Aus der Chirurgischen Klinik
der Medizinischen Fakultät Mannheim
(Direktor: Prof. Dr. med. Christoph Reißfelder)
und
Deutsches Krebsforschungszentrum
Nachwuchs-Klinische Kooperationseinheit Translationale Chirurgische Onkologie
(Leiter: Apl. Prof. Dr. med. Sebastian Schölch)

Molecular characterization of circulating tumor cells in pancreatic cancer

Inauguraldissertation
zur Erlangung des medizinischen Doktorgrades
der
Medizinischen Fakultät Mannheim
der Ruprecht-Karls-Universität
zu
Heidelberg

vorgelegt von
Lei Zhu

aus
Jiangsu
2021

Dekan: Prof. Dr. med. Sergij Goerdt
Referent: Apl. Prof. Dr. med. Sebastian Schölch

This cumulative dissertation includes:

Zhu L, Kan KJ, Grün JL, Hissa B, Yang C, Győrffy B, Loges S, Reißfelder C, Schölch S. GAS2L1 Is a Potential Biomarker of Circulating Tumor Cells in Pancreatic Cancer. Cancers (Basel). 2020 Dec 15;12(12):3774. doi: 10.3390/cancers12123774.

Zhu L, Hissa B, Győrffy B, Jann JC, Yang C, Reissfelder C, Schölch S. Characterization of Stem-like Circulating Tumor Cells in Pancreatic Cancer. Diagnostics (Basel). 2020 May 15;10(5):305. doi: 10.3390/diagnostics10050305

Other publications during promotion not included in this dissertation:

Maier CF*, **Zhu L***, Nanduri LK*, Kühn D, Kochall S, Thepkaysone M-L, William D, Grützmann K, Klink B, Betge J, Weitz J, Rahbari NN, Reißfelder C, Schölch S. Patient-Derived Organoids of Cholangiocarcinoma. International Journal of Molecular Sciences. 2021; 22(16):8675. doi: 10.3390/ijms22168675

Kan KJ*, Guo F*, **Zhu L**, Pallavi P, Sigl M, Keese M. Weighted Gene Co-Expression Network Analysis Reveals Key Genes and Potential Drugs in Abdominal Aortic Aneurysm. Biomedicines. 2021 May 13;9(5):546. doi: 10.3390/biomedicines9050546

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

CARLIN	CRISPR array repair lineage tracing
cDNA	complementary DNA
CI	Confidence interval
CK	Cytokeratins
CSC	Cancer stem cell
CTC	Circulating tumor cell
DAPI	4',6-diamidino-2-phenylindole
DEP	Dielectrophoresis
DFS	Disease free survival
DNA	Deoxyribonucleic acid
ECM	Extracellular matrix
EMT	Epithelial-mesenchymal transition
EPCAM	Epithelial cell adhesion molecule
GEMM	Genetically engineered mouse model
GESTALT	Genome editing of synthetic target arrays for lineage tracing
HR	Hazard ratio
ISET	isolation by size of epithelial tumor cells
MEMS	Micro-electro-mechanical system
ORA	Over-representation analysis
OS	Overall survival
PanIN	Pancreatic intraepithelial neoplasia
PBMC	peripheral blood mononuclear cell
PCA	Principal component analysis
PDAC	Pancreatic ductal adenocarcinoma
PFS	Progression free survival
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
RR	Relative risk
scRNA-seq	Single-cell RNA sequencing
TCGA	The Cancer Genome Atlas
t-SNE	t-Distributed stochastic neighbor embedding
WGCNA	Weighted gene co-expression network analysis

1 INTRODUCTION

1.1 Pancreatic cancer

Pancreatic cancer (pancreatic ductal adenocarcinoma, PDAC) is a highly fatal malignancy with a dismal prognosis. According to the Robert Koch Institute, 18,687 people were diagnosed with pancreatic cancer in Germany in 2017.[1] With 18,005 deaths, the mortality approaches the incidence, leading to PDAC being the fourth common cause of death from malignant neoplasms[1]. Only 9% of patients survive 5 years after diagnosis.[1,2] Furthermore, both the age-standardized incidence and mortality have slightly increased since 2000, especially in the elderly.[2]

Down to the histopathologic subtypes, pancreatic adenocarcinoma is the dominant diagnosis. Pancreatic adenocarcinoma occupies 85.8% (97,923/114,166) of all malignant pancreatic tumors.[3] It is widely believed that pancreatic adenocarcinoma originates from cells lining the pancreatic duct.[4] Therefore, pancreatic ductal adenocarcinoma (PDAC) is the most commonly used acronym.[4] Of all PDAC patients, only 7.6% (7,436/97,923) are diagnosed in localized stages, and 15.6% (14,830/95,172) eventually undergo radical or palliative surgery.[3] The 61.4% (60,169/97,923) of patients initially present with metastasis and have a median overall survival (OS) as low as 4.0 months.[3] Even for those patients who are eligible for potentially curative resection, 54.7% (158/289) suffer from recurrence or distant metastases in the following months in a large, multicenter randomized trial.[5]

Kamisawa et al.[6] investigated 130 necropsy cases of PDAC and found that 80% (104/130) developed hematogenous metastases. Among the prevalent metastatic lesions, the liver is the principal target (62%, 81/130), followed by the lung (55%, 71/130) and bone (25%, 32/130).[6] This phenomenon is probably because the liver is the first line of defense—all the peripancreatic veins draining into the portal vein. Millions of CTCs are continuously shed into the liver microvasculature. Meanwhile, only 12.3% (16/130) of patients showed macroscopic lymph node metastases without hematogenous metastases, microscopically peripancreatic lymph node metastases are are found in 52-78%.[7,8] However, there are indications that the peripancreatic (“sentinel”) lymph nodes filter the lymphatic fluid and represent a cul-de-sac rather than a temporary intermediate station, thus not contributing to the development of distant metastases.[9–11]

Although metastasis is one of the cardinal characteristics of pancreatic cancer and has a high lethality, its cascade remains incompletely understood. This dissertation will thus focus on hematogenous metastases in pancreatic cancer.

1.2 Circulating tumor cells

Multiple studies investigate the molecular backgrounds of different steps of hematogenous dissemination, including tumor cell invasion[12], migration[13], and intravasation[14]. Traditionally, circulating tumor cells (CTC) are defined as tumor cells that detach from the primary tumor, intravasate, and spread in the bloodstream. It is also possible that tumor cells enter hematogenous circulation by infiltrating into the lymph node blood vessels.[15–17] Independently of the mechanism of intravasation, CTCs are integral to the metastatic process in malignant tumors.[18–22]

Solid tumors exhibit different degrees of metastatic organotropism. The mechanisms behind this are still incompletely understood. However, it is clear that the next capillary bed in the tumor-draining circulation (i.e., the liver in gastrointestinal malignancies) often acts as a filter for CTCs and tumor microemboli and thus is a predilection site for distant metastases. Consequently, CTCs in portal vein samples are found more frequently[23] and in greater numbers[24] than in the peripheral venous blood of patients with pancreatic cancer. In addition to the spatial distribution, researchers reported a circadian rhythm in CTC release by longitudinal in vivo monitoring.[25]

Although the number of blood cells exceeds the number of CTCs by about 1:10.000.000 on average, there are still far more CTCs than clinically relevant metastases, reflecting that metastasis is an inefficient process. CTCs are faced with numerous challenges once they have entered circulation, including anoikis, fluid shear stress, and immune surveillance.[9,26] Hence, CTCs have a short half-life of 1 – 2.4 hours.[27] The next capillary bed in the tumor-draining blood compartment is of higher risk of metastatic colonization.[28] CTCs or circulating tumor microemboli can get lodged in the microvasculature, intraluminally expand, and ultimately penetrate the vessels[29] or extravasate by breaching vascular walls[30]. After extravasation, more obstacles are waiting for CTCs before developing into clinically relevant metastatic lesions,[30] most prominently the requirement to adapt to and thrive in a foreign microenvironment.[18] Another crucial factor of macroscopic metastasis is angiogenesis, which is indispensable in tumorigenesis[31] as implanted tumor cells cannot outgrow to lesions > ~ 3 mm diameter without neoangiogenesis.[32]

1.3 Mouse models for pancreatic CTCs research

Due to the rarity of CTCs in the peripheral blood of PDAC patients, in vivo investigations of pancreatic CTCs with mouse models are often used to gain insights into the cellular biology of this cell population. Immunodeficient mice and genetically engineered mouse models are widely used.

1.3.1 Immunodeficient mouse models

To date, several studies have attempted to monitor pancreatic CTCs in vivo through inoculating fluorescent protein-labeled commercial human PDAC cell lines into immunodeficient mice.[33–35] Yu et al.[33] found more CTCs in their orthotopic than subcutaneous xenograft tumor model using in vivo flow cytometric (IVFC) inspection. In another study[34], researchers observed the peak of CTC burden at the fifth-week post-inoculation with multi-photo microscopy, Intravital Imaging System (IVIS). Rivera-Báez and colleagues[36] successfully established 3 CTC cell lines from their pancreatic cancer patients in a seminal study. To investigate the tumorigenicity and morphologic features recapitulation of the novel CTC cell lines, they injected cells subcutaneously into NSG mice and observed 100% tumor formation and widespread metastasis.[36]

There are limited publications of patient-derived xenografts to investigate CTCs in pancreatic cancer.[37,38] Dimitrov-Markov et al.[38] generated extremely metastatic PDAC patient-derived xenograft (PDX) (subcutaneous) mouse models, and performed single-cell RNA sequencing (scRNA-seq) of 10 CTC successfully.[38] Torphy et al.[37] found serial monitoring of CTC could predict and monitor therapy response to guide therapeutic regimens.

While the above research reported positive results, accumulative evidence suggests that orthotopic xenograft models are superior to ectopic (i.e., subcutaneous) xenografts in metastasis research.[39–41] In 1999, Killlion et al.[40] reviewed the literature and demonstrated that subcutaneous xenografts lack distant metastases. This conclusion was supported by Dai et al.[39], who established two orthotopic and one subcutaneous xenograft models. Thirty-six days after implantation, they showed that 80% (16/20) of the mice developed metastasis in the orthotopic models. In contrast, metastasis was absent in the subcutaneous model. This data suggests that the orthotopic environment can simulate the process of metastasis more accurately.

1.3.2 Genetically engineered mouse models (GEMMs)

In genetically engineered mouse models (GEMMs), the genetic hallmarks of malignant disease are recapitulated in the respective organ of the mouse. The animals develop “their own” tumors, which is has been proven to mimic the human disease in multiple entities realistically.[42–44] Cre/LoxP is the most commonly used tool for genetic manipulations in the mouse.[45–48] Cell type-specific[44,45,49,50] and/or tamoxifen-inducible Cre (“Cre^{ERT2}”)[51–53] promoters allow spatial and temporal control of the expression of the target genes. As the host organisms are immunocompetent, GEMMs are ideal models to investigate host-tumor interactions.

In the PDAC cohort from the cancer genome atlas program (TCGA), *Kras* (93.3%, 140/150) and *TP53* (72.0%, 108/150) are the two most frequent genes harboring oncogenic mutations. As a result, the most prominent GEMM of PDAC incorporates mutant *Kras* and *Trp53* to obtain a compound mutant strain.[44,54,55] Hingorani et al.[44,55] established this mouse model of pancreatic intraepithelial neoplasia (PanIN) and histopathologically confirmed that the lesions eventually progressed to invasive and metastatic tumors. This mouse model is known as KPC (LSL-*Kras*$^{G12D/+}$, LSL-*Trp53*$^{R172H/+}$, Pdx1-Cre). A few years later, Bardeesy et al.[56] crossed *Trp53*$^{lox/lox\ or\ lox/+}$rather than *Trp53*$^{R172H/+}$ with Pdx1-Cre and LSL-Kras$^{G12D/+}$ (KP_fC) and found a more rapid progression of PDAC.

In subsequent studies, researchers introduced the reporter allele RosaYFP into KPfC/KPC (KPfCY/KPCY) mice for lineage labeling.[57,58] Rhim et al.[57] explored the circulating pancreatic cells by flow cytometry and found YFP+ circulating cells even before tumor formation. Ko and colleagues[59] employed the KP_fCY mouse model to validate the efficiency of CTC isolation in a novel microfluidic chip. A pioneering study was conducted by Simeonov and colleagues[60], who developed a method (macsGESTALT, multiplexed, activatable, clonal, and subclonal; genome editing of synthetic target arrays for lineage tracing) and applied it to an orthotopic allograft mouse model with KPCY cells. They were able to reconstruct the dissemination of single cells derived from multiple metastatic sites, including CTCs, using single-cell barcoding by the macsGESTALT construct.[60]

1.4 CTC capture technologies

Due to unresolved biological and technical issues induced by the rarity of these cells, our understanding of CTCs and their fates are incomplete.

Although up to 4 x 10^6 CTCs per gram of primary tumor are shed into the systemic circulation daily[61], CTCs are extremely rare compared to billions of hematocytes per milliliter of blood. As a result, various CTC enrichment/capture technologies with a different selection and isolation principles were developed in the past decades, mainly based on biological or physical features distinguishing CTCs from other mononucleated blood cells.[62] Since a series of reviews[62–65] have summarized the existing CTC isolation technologies very well, this dissertation will introduce this part only briefly.

1.4.1 Biological properties

Generally, CTCs are a heterogeneous cell population, and there is currently no uniform

marker available which identifies all CTCs while excluding all other cells present in the blood.[66–69]

Epithelial cell surface markers, most prominently Epithelial Cell Adhesion Molecule (*EPCAM*)) are widely used to positively select CTCs.[70–75] However, to acquire the invasive features required for escape from the primary tumor bulk, CTCs experience epithelial-mesenchymal transition (EMT). As a result, carcinoma cells undergoing partial or complete EMT acquires mesenchymal properties while downregulating epithelial traits.[76–78] In fact, recent research demonstrated that metastatic lesions are predominantly derived from CTCs in a late-hybrid EMT state and were *EPCAM* negative.[60] Furthermore, previous studies confirmed that *EPCAM*-negative CTCs are also involved in the metastatic cascade.[79–81]

CD45 (*PTPRC*) is a popular marker for the negative selection of CTCs via the depletion of hematopoietic cells.[82] CD45 depletion of whole blood samples leaves a diverse cell population that includes most CTCs but is not a pure CTC sample. In addition, this method might be limited by losing CTC-leukocyte aggregates, which would be depleted from such a sample.

Generally, combinations of multiple markers promise more selective CTC capture[83,84], but at the same time only isolate increasingly selected subgroups of CTCs while losing the other CTCs not included in the marker combination.

1.4.2 Physical properties

Techniques that rely on physical properties to distinguish CTCs from hematopoietic cells by density, size, or elasticity have unique advantages. They are unbiased and independent of cell surface markers, which might be present only in subfraction CTCs depending on their physical state.

Density gradient centrifugation is the most widely used strategy to enrich CTCs. Typically, the density of a density gradient medium is 1.077 g/ml, like Ficoll-paque™ or Lymphoprep™. After centrifugation, CTCs should be theoretically enriched in the peripheral blood mononuclear cell (PBMC) layer. However, this method is prone to losing CTC clusters as a result of their higher density.[85] Other isolation methods based on density, such as OncoQuick[86] and AccuCyte[87] were also reported.

Size-based enrichment assays developers postulate that CTCs are larger than other blood cells. Filtration is label-free, fast, and convenient. However, the size overlap of CTCs and other blood cells limits the purity of the resulting samples. Hao et al.[88] reviewed size-based technologies for CTC isolation in 2018. Based on the same principle, techniques such as ISET (isolation by size of epithelial tumor cells)[89], MEMS

(micro-electro-mechanical system)-based microfilter approaches[90], ScreenCell™ [91] have similar limitations.

DEPArray™ [92,93] is an image-based cell sorter based on dielectrophoresis. It utilizes the dielectric charge across the cell membrane of CTCs to arrest and manipulate them without physical touch.[77]

1.4.3 Microfluidics

In CTC research, microfluidic techniques enrich corpuscular objects of different sizes or densities in biological fluids. A multitude of different microfluidic devices for CTC enrichment or quantification have been developed.[94] As an example, Stott et al.[95] developed a herringbone-chip, which generates micro vortices that direct CTCs toward capture on the EpCAM antibody-coated walls. Another team optimized imaging and characterization functions in the CTC-iChip microfluidic system.[96] This strategy integrates nucleated cell separation, size-based sorting, and immunoaffinity selection.[96] A more popular platform, Parsortix, was reported in 2015.[97] However, it requires further identification and isolation methods (usually surface-marker based) to obtain relatively pure CTC samples like most other microfluidic devices solely based on physical characteristics of CTCs.[97]

1.5 Clinical value of CTC in pancreatic cancer

Several groups have shown that CTCs occur in any stage of pancreatic cancer.[70,98] While most research validated the prognostic function of CTCs in pancreatic cancer[23,24,70,99,100], some articles support opposite conclusions[101,102]. This controversy may result from different CTC identification techniques, leading to different CTC subpopulations to be investigated. Since there is no consensus isolation strategy for pancreatic CTCs, the results can be compared only to a minimal degree.

1.5.1 Overall survival (OS)

Pancreatic cancer patients with detectable CTCs in the peripheral blood have the significantly worse OS. The translational arm of an international randomized phase III trial published in 2013 reported that CTC positivity of 79 pancreatic patients at baseline and/or 2 months correlated with shorter OS in both univariate (log-rank test) and multivariable analysis (Cox proportional hazards model) (RR = 2.5, p = 0.01).[103] A similar result was seen in another prospective study with 69 patients. In this study, the presence of CTCs lead to decreased survival in univariate (p = 0.030) and multivariate analysis (HR = 2.093; CI, 1.081–4.050; p = 0.028).[74] Poruk et al.[104] attempted to assess

CTCs with different markers and separately analyzed the influence of epithelial (CK+, DAPI+, CD45-) and mesenchymal (CK+, Vimentin+, DAPI+, CD45-) CTCs on survival. They found that only the detection of epithelial CTCs was associated with worse OS ($p = 0.008$), while the presence of mesenchymal-like CTCs did not ($p = 0.39$).[104]
Not all research groups chose the qualitative detection of CTCs (positivity vs. negativity) as cut-offs. Other authors used a semi-quantitative scale in order to classify different groups. A pilot study with 20 PDAC patients revealed that increased CTC numbers (> 10 CTCs/ml) predict dismal OS ($p < 0.001$).[100]

1.5.2 Disease/recurrence free survival (DFS/RFS)

In addition to OS, disease/recurrence free survival (DFS/RFS) is another popular outcome in clinical trials involved in cancer metastasis.[74,75,104,105]
In a prospective cohort with long-term following-up, 98 patients with resectable pancreatic cancer were recruited.[75] Although this trial only detected CTCs in 7 cases with the widely used CellSearch ® system, all patients with detectable CTCs developed metastasis in the postoperative follow-up with significantly shorter DFS than patients without CTCs (HR = 2.8; $p = 0.008$).[75] Effenberger et al.[74] illustrated similar results in 69 PDAC patients, CTC positivity lead to shorter PFS in univariate ($p = 0.009$) and multivariate analysis (HR = 4.543; CI, 1.549–13.329; $p = 0.006$).[74]
Although it did not influence OS, the detection of vimentin-positive CTCs increased the risk of cancer recurrence significantly (HR = 2.78, CI, 1.31–5.88, $p = 0.01$).[104]
While most researchers isolated CTCs from peripheral venous blood, Tien et al.[24] also quantified CTCs in the portal vein. In this study, 84.6% (11/13) of patients with a high portal CTC burden developed hepatic metastases within 6 months. The metastatic incidence in patients with lower CTC counts was significantly lower (12.8%, 6/47, $p < 0.001$).

1.6 Single-cell RNA sequencing (scRNA-seq)

1.6.1 Technologies

In 1990, researchers used PCR technology to achieve exponential amplification of cDNA molecules, and for the first time, showed that it is feasible to perform transcriptome analysis on single cells.[106] Two decades later, Tang et al.[107] made technical adjustments to make single cells compatible with high-throughput technologies, and for the first time, conducted a completely unbiased transcriptome-wide study of mRNA in a single cell.[107]
Early single-cell research aims to conduct in-depth research on a small number of

precious cells.[107–109] Later technological evolution proved that different cell types could be identified without pre-sorting cells.[110]

Since single-cell transcriptome sequencing technology was first reported in 2009, transcriptome sequencing technology has evolved rapidly.[111–114] Groups from Harvard University combined microfluidic technology with scRNA-seq and developed Drop-seq[115] and inDrop[116], respectively. Both technologies use microfluidics to generate drops of single cells and barcoded microbeads, and the reverse transcription-amplification library is realized in the droplets. This way, multiple transcriptomes can be sequenced and assigned to individual cells using single-cell barcode information. This technology has made it possible to analyze thousands of single cells' gene expressions in a single sequencing run.

scRNA-seq seems like an ideal technology for CTC research, as each cell can be tracked and analyzed. Transcriptome information can be extracted from the data, and subsequent analysis can be performed through bioinformatics strategies. Even though the vast majority of sequenced cells are mere blood cells and only a few CTC transcriptomes are contained in the large datasets, the CTC expression data can be extracted and analyzed from the datasets.

1.6.2 Bioinformatics

The immense potential of scRNA-seq sparked the development of corresponding analysis tools[117]. Powerful toolkits were published in the past years, including Seurat[118], Scater[119], or Scanpy[120]. Satija and colleagues[118] developed Seurat in 2015, which provides integrated pipelines to analyze single-cell sequencing data with in situ RNA patterns and is widely used.[121–123]

Generally, cells are classified into clusters according to the similarity of transcriptomic profiles via distance metrics. This approach allows researchers to identify different cell types from one merging sample, which means technical feasibility for pointing CTCs out from the hematogeneous cells.

Another integral procedure is trajectory analysis, which can investigate intrinsic differentiation or transition processes and biological function dynamics.[124] In brief, the single-cell sequencing map resembles a snapshot of a continuous process, and transcriptional distances from a root cell are portrayed according to the pseudotime, a proxy for developmental time.[125,126] The emerging of Monocle[127] and Wanderlust[128] are important tools for trajectory analysis.

1.6.3 Frontiers

The eruption of scRNA-seq technologies and bioinformatic strategies has fueled CTC biology research considerably.[129–132] However, there is only very limited data on pancreatic CTCs.[60,133]

In 2014, Ting et al.[133] applied CTC-iChip[96] to the KPfC mouse model, enabling antigen-agnostic isolation of CTCs and the following scRNA-seq. Notably, they revealed that various extracellular matrix (ECM) transcripts were enriched in pancreatic CTCs and concluded that the expression of cell-autonomous ECM genes might contribute to blood-borne dissemination.[133]

Recently, single-cell barcoding technology was expanded to in vivo lineage tracing, including but not limited to GESTALT (genome editing of synthetic target arrays for lineage tracing)[134], PolyloxExpress[135], and CARLIN (CRISPR array repair lineage tracing)[136]. These inducible lineage recorders allow the identification of metastasis-initiating subclones.[60]

1.7 Cancer stem cells in pancreatic circulating tumor cells

Although intensive research efforts focus on CTCs, the metastasis-initiating subpopulation of CTCs is still essentially unknown.[137] Therefore, no therapeutically addressable target has been identified on CTCs to date. A complex set of remarkable cellular properties is required to complete the metastatic cascade, including migration and invasion capabilities, immune escape strategies, and, ultimately, colony-forming capability.[137] As only a small subpopulation of CTCs can perform all necessary steps of metastasis, the number of CTCs shed over time is exponentially more than the number of overt metastases.[138] As a result, CTCs are highly heterogeneous and usually non-tumorigenic.[139]

While myriad CTCs are shed into the bloodstream, the sum of macroscopic metastases is less by several orders of magnitude.[84] Therefore, it can be assumed that merely a small population of CTCs metastasis-initiating, i.e., colonizing distant organs and forming overt metastases.[20,137] Hence, this metastatic cascade requires self-renewal and tumor-initiating capacities of CTCs, which are hallmarks of cancer stem cells (CSC).[10] A subfraction of CTCs must therefore have intrinsic pluripotency.

In a study with 24 metastatic PDAC patients, investigators found that 70.8% of patient blood samples contained CTCs with stem cell properties (CD133+, CK+, CD45-, DAPI+).[73] In contrary, Semaan et al.[78] found a significantly lower number of stem-like CTCs (CD133+, DAPI+, CD45-) (2.4 ± 0.5 cells/7.5 ml), accounting for 8.2% of CTCs from 74 PDAC patients. Franses et al.[140] evaluated a panel of stemness genes in CTCs

from PDAC patients, including *CD133*, *CD24*, *CD44*, *KLF4*, and *LIN28B*. They knocked out *LIN28B* through CRISPR both in vitro and in vivo and found a less metastatic phenotype.[140]

1.8 Rationale and objectives

To date, the discovery and validation of universal and specific markers for pancreatic CTCs are still in their infancy, while immunoaffinity assays are, despite their limitations, the most popular CTC harvesting strategies. The biology of CTCs and their role in the metastatic cascade are still unclear. As only a rare fraction of CTCs carries cancer stem cells features and can initiate metastases, the identification of this population is crucial for the subsequent development of anti-metastatic targeted therapies.

Therefore, the objectives of this work are:

1) To identify novel biomarkers of pancreatic CTCs.
2) To characterize pancreatic CTCs with stem-like features.

2 RESULTS

2.1 *GAS2L1* is a latent biomarker of pancreatic CTC.

Dissertation relevant publication:

Zhu L, Kan KJ, Grün JL, Hissa B, Yang C, Győrffy B, Loges S, Reißfelder C, Schölch S. GAS2L1 Is a Potential Biomarker of Circulating Tumor Cells in Pancreatic Cancer. **Cancers (Basel)**. 2020 Dec 15;12(12):3774. doi: 10.3390/cancers12123774.([141])

Summary: The research presented here aimed to identify latent biomarkers of pancreatic CTCs.

Therefore, the publicly available dataset GSE51372[133] was re-analyzed with Seurat. 157 murine cells were assigned to 5 clusters after data cleaning, normalization, and dimensional reductions (PCA and t-Distributed stochastic neighbor embedding (t-SNE)). Three heterogeneous clusters of CTCs were observed. Clusters 1 and 3 are pure CTCs, while cluster 0 is a mixture of murine cells.

The following functional enrichment analysis revealed distinct biologic functions of CTC clusters. While the proteoglycans in the cancer pathway was enriched in cluster 1, cluster 3 exhibited the enrichment of the platelet activation pathway. However, both clusters showed enrichment of the focal adhesion pathway.

Nine transcripts were up-regulated in pancreatic CTCs, and two (*Clic4* and *Gas2l1*) were identified as cluster markers. However, only *Gas2l1* could be shown to better identify CTCs than the current gold standard *Epcam* (positive control).

Additional receiver operating characteristic (ROC) examination demonstrated the synergistic function of CTC identification through combining *Gas2l1* and *Epcam*. Moreover, no transcriptomic and proteomic expression overlap was observed between these two markers.

Lastly, to evaluate the prognostic value of *GAS2L1*, we explored it in the TCGA PDAC cohort. Notably, *GAS2L1* shows significant over-expression in PDAC compared to normal pancreatic tissue and a prognostic value for RFS but not OS.

cancers

MDPI

Article

GAS2L1 Is a Potential Biomarker of Circulating Tumor Cells in Pancreatic Cancer

Lei Zhu [1,2], Ke-Jia Kan [2], Johanna L. Grün [2], Barbara Hissa [2,†], Cui Yang [2], Balázs Győrffy [3,4], Sonja Loges [5,6], Christoph Reißfelder [2] and Sebastian Schölch [1,2,*]

1 Junior Clinical Cooperation Unit Translational Surgical Oncology, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; lei.zhu@dkfz.de
2 Department of Surgery, Universitäts Medizin Mannheim, Medical Faculty Mannheim, Heidelberg University, 68167 Mannheim, Germany; Kejia.Kan@medma.uni-heidelberg.de (K.-J.K.); johanna.gruen@umm.de (J.L.G.); bhissa@beilstein-institut.de (B.H.); cui.yang@umm.de (C.Y.); christoph.reissfelder@umm.de (C.R.)
3 Department of Bioinformatics and 2nd Department of Pediatrics, Semmelweis University, H-1094 Budapest, Hungary; gyorffy.balazs@med.semmelweis-univ.hu
4 TTK Cancer Biomarker Research Group, Institute of Enzymology, H-1117 Budapest, Hungary
5 Division of Personalized Medical Oncology (A420), German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany; s.loges@dkfz.de
6 Department of Personalized Oncology, University Hospital Mannheim, University of Heidelberg, 68167 Mannheim, Germany
* Correspondence: s.schoelch@dkfz.de; Tel.: +49-621-383-5152
† Current address: Beilstein-Institut zur Förderung der Chemischen Wissenschaften, 60487 Frankfurt am Main, Germany.

Received: 17 November 2020; Accepted: 13 December 2020; Published: 15 December 2020

Simple Summary: The analysis of circulating tumor cells (CTC) is a mainstay of liquid biopsy of solid malignancies. However, research to date has not yet determined a universal and specific marker for CTCs of pancreatic cancer. Genetically engineered mouse models (GEMMs) of pancreatic cancer, can mimic the human disease very closely. This study aimed to identify potential biomarkers for CTCs in a GEMM of pancreatic cancer and further validate markers in human samples. Therefore, we analyzed single-cell RNA sequencing data of murine pancreatic CTCs and performed advanced bioinformatic analyses. We demonstrated that the focal adhesion pathway is functionally enriched in pancreatic CTCs. In addition, we suggest Gas2l1/GAS2L1 as a potential surface marker of pancreatic CTCs. In combination with *Epcam/EPCAM*, *Gas2l1/GAS2L1* identify the majority of pancreatic CTCs. Furthermore, pancreatic cancer patients with overexpression of *GAS2L1* have an unfavorable prognostic outcome.

Abstract: Pancreatic cancer is a malignant disease with high mortality and a dismal prognosis. Circulating tumor cell (CTC) detection and characterization have emerged as essential techniques for early detection, prognostication, and liquid biopsy in many solid malignancies. Unfortunately, due to the low EPCAM expression in pancreatic cancer CTCs, no specific marker is available to identify and isolate this rare cell population. This study analyzed single-cell RNA sequencing profiles of pancreatic CTCs from a genetically engineered mouse model (GEMM) and pancreatic cancer patients. Through dimensionality reduction analysis, murine pancreatic CTCs were grouped into three clusters with different biological functions. *CLIC4* and *GAS2L1* were shown to be overexpressed in pancreatic CTCs in comparison with peripheral blood mononuclear cells (PBMCs). Further analyses of PBMCs and RNA-sequencing datasets of enriched pancreatic CTCs were used to validate the overexpression of *GAS2L1* in pancreatic CTCs. A combinatorial approach using both *GAS2L1* and *EPCAM* expression leads to an increased detection rate of CTCs in PDAC in both GEMM and patient samples. *GAS2L1* is thus proposed as a novel biomarker of pancreatic cancer CTCs.

Keywords: biomarkers; circulating tumor cells; single-cell RNA sequencing; liquid biopsy; pancreatic neoplasms; pancreatic ductal adenocarcinoma; mice; genetically engineered mouse model; computational biology

1. Introduction

Pancreatic ductal adenocarcinoma (PDAC) is the fourth most frequent cause of cancer-related death in western countries [1]. Fewer than one in five patients are eligible for potentially curative therapy since the tumor is unresectable in most cases at the time of diagnosis [2]. Even after resection with curative intent, most patients develop local or distant recurrence and ultimately succumb to the disease. Liquid biopsy, i.e., the detection and characterization of tumor-related molecules or intact tumor cells in a blood sample, is a tool that may enable clinicians to detect metastatic disease early on and thus avoid surgery in patients who may not benefit from resection due to occult metastatic disease.

Liquid biopsy is defined as investigating tumor biology in different blood-borne sources such as circulating tumor DNA (ctDNA) or circulating tumor cells (CTCs). CTCs are shed from the tumor, survive in circulation, and ultimately form distant metastases. In contrast to ctDNA, CTCs are viable tumor cells and, as such, they might represent a more accurate reflection of the current tumor biology compared to ctDNA also originating from dead cells. Furthermore, ex vivo culture of CTCs allows the downstream function experiments or even personalized translational research [3–5]. CTCs have been shown to have prognostic and predictive value in pancreatic cancer [6].

The extreme rarity of CTCs in many epithelial tumors has led to numerous CTC isolation technologies, which can detect CTCs based on physical properties or biologic markers [7]. Techniques that rely on physical properties to distinguish CTCs from blood cells by size, density, or even elasticity have unique advantages because they are unbiased and independent of cell surface markers, which might be present only in a subfraction CTCs depending on their biological state. However, to date, most techniques to detect CTC still depend on surface markers. The most common biomarker used to identify and isolate epithelial tumor-derived CTCs is the Epithelial Cell Adhesion Molecule (EPCAM). However, recent studies found that EPCAM-negative CTCs are also involved in the metastatic cascade [8–10], demonstrating the need for other markers to identify and isolate CTCs. Generally, CTCs are a heterogeneous cell population, and there is currently no uniform marker available which identifies all CTCs while excluding all other cells present in the blood [11–14]. As a way to overcome this problem, combinations of multiple markers have been used [15,16]; however, the currently available means to identify and characterize CTCs are still unsatisfactory.

This study aims to combine and analyze publicly available, single-cell RNAseq expression profiles of human and murine PDAC-derived CTCs to discover new biomarkers that could be potentially used to identify pancreatic cancer-derived CTCs.

2. Results

The workflow of the analysis is depicted in Figure 1.

2.1. Pancreatic CTC Heterogeneity Leads to Distinct Clustering

157 cells of the GSE51372 dataset were taken into the principal component analysis (PCA). Among these cells, there were 75 CTCs from the GEMM (henceforth referred to as CTC group), 34 primary tumor cell samples from the GEMM (TuRNA group) (from which 10 or 100 pg RNA of diluted bulk RNA were taken), 20 EGFP-positive primary tumor cells (bulk tumor cells (BTC) group), 16 single tumor cells from the NB508 cell line (NB508 group) and 12 white blood cells (WBCs) from a control mouse (WBC group). As shown in Supplementary Figure S1A,B, there are significant sequencing depth and gene detection differences among the different sample types, which could impair the reliability of the traditional sequencing data analysis. Upon PCA, CTCs formed two distinct clusters,

one of which clustered with both BTC and, surprisingly, with WBC, while the second cluster was separate, indicating substantial heterogeneity among murine pancreatic CTCs (Figure 2A). The groups nb508 and TuRNA formed distinct clusters.

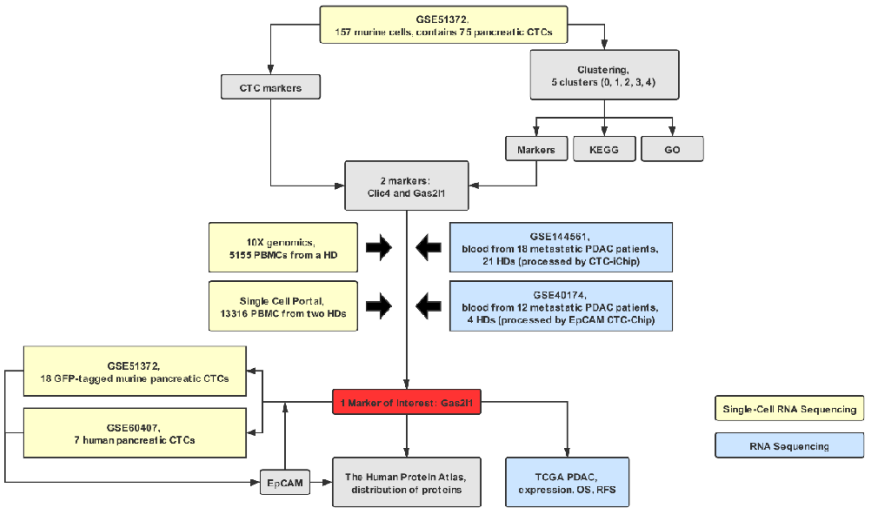

Figure 1. Study flow diagram. CTCs, circulating tumor cells; EPCAM, epithelial cell adhesion molecule; GO, gene ontology; HD, healthy donor; KEGG, Kyoto Encyclopedia of Genes and Genomes; PBMC, peripheral blood mononuclear cells; PDAC, pancreatic ductal adenocarcinoma; OS, overall survival; RFS, relapse-free survival.

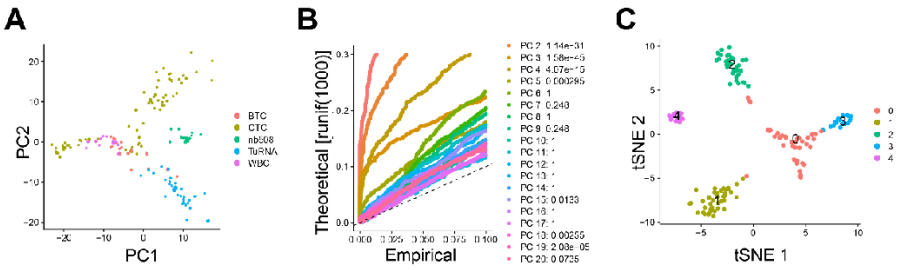

Figure 2. Clustering of murine cells after single-cell sequencing. There are 157 cells from the GSE51372 dataset that were taken into the analysis, including 75 circulating tumor cells (CTCs) from the genetically engineered mouse model ("CTC"), 34 primary tumor cells ("TuRNA") (from which 10 or 100 pg RNA of diluted bulk RNA were taken), 20 EGFP-positive primary bulk tumor cells ("BTC"), 16 single tumor cells from NB508 cell line ("NB508"), and 12 WBCs from a control mouse ("WBC"). (**A**) Principal component analysis (PCA) of the cells. (**B**) JackStraw plot of the first 20 principal components (PCs). The distribution of p values for each PC with a uniform distribution (dashed line). The first five PCs show strong enrichment of features with low p values (solid curve above the dashed line). (**C**) t-Distributed Stochastic Neighbor Embedding (t-SNE) plot of 157 cells. All cells can be distributed into five clusters (cluster 0, 1, 2, 3, and 4). Cluster 0 consists of murine pancreatic CTCs, bulk tumor cells (BTCs), and white blood cells (WBCs). Cluster 1 only contains CTCs. Cluster 2 consists of diluted bulk RNA (TuRNA) and BTCs. Cluster 3 only contains CTCs and cluster 4 represents NB508 cells. BTCs, bulk tumor cells; CTCs, circulating tumor cells; PC, principal component; t-SNE, t-Distributed Stochastic Neighbor Embedding; TuRNA, RNA of the bulk tumor; WBC, white blood cells.

As depicted in Figure 2B, there is a drop of significance levels after the first five primary components (PCs). These five PCs were taken into the subsequent t-Distributed Stochastic Neighbor Embedding (t-SNE) analysis. Here, all 157 samples were grouped into five distinct clusters (Figure 2C).

Murine pancreatic CTCs were identified as three distinct clusters (clusters 0, 1, and 3), which is in line with the original study results [17]. While cluster 0 was a mixture of CTCs, BTCs, and WBCs, clusters 1 and 3 contained only CTCs and were therefore included in the functional enrichment analysis.

Cluster 1 comprised 39 of 75 CTCs (52.0%). Compared to the other four clusters, there are 139 genes upregulated and 15 transcripts with reduced expression in cluster 1. Further analysis (Figure 3A) showed that over-expressed genes were enriched in the proteoglycans in cancer pathway (KEGG: 05205; FDR = 2.79×10^{-3}, fold change = 5.62). Consistently, the enriched gene ontology term proteoglycan metabolic process was significantly upregulated (GO: 0006029; FDR = 9.98×10^{-3}, fold change = 9.28).

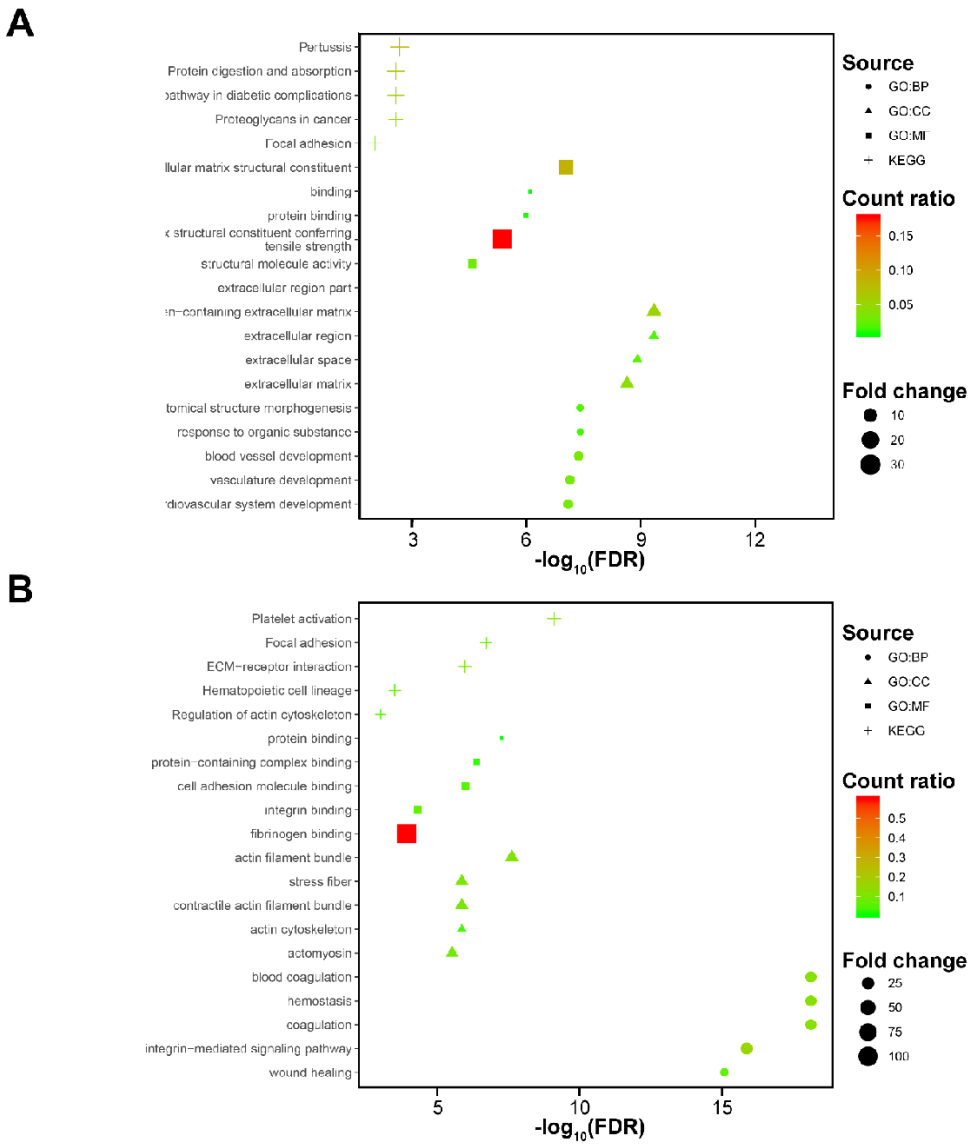

Figure 3. Functional enrichment analysis of cluster-specific markers. Functional enrichment analysis of genes in the clusters 1 (**A**) and 3 (**B**), top five enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, Gene Ontology (GO) terms for biological processes (BP), cell components (CC) and molecular functions (MF) [fold enriched > 2 and false discovery rate (FDR) < 0.001]. BP, biological process; CC, cellular component; FDR, false discovery rate; KEGG, Kyoto Encyclopedia of Genes and Genomes; MF, molecular function; GO, gene ontology.

Cluster 3 contains 26.7% (20/75) of pancreatic CTCs. Pathway analysis of cluster 3 signatures resulted in enrichment of the platelet activation pathway (KEGG: 04611; FDR = 7.74×10^{-10}, fold changed = 13.88) (Figure 3B). Abundant terms associated with platelets were also enriched in the GO analysis, including platelet activation (GO: 0030168; FDR = 1.62×10^{-13}, fold changed = 30.12), platelet aggregation (GO: 0070527; FDR = 7.78×10^{-10}, fold changed = 35.94), regulation of platelet activation (GO: 0010543; FDR = 1.05×10^{-3}, fold changed = 23.46 (Figure 3B). This strong platelet signal may be a result of platelets adhering to the CTCs.

Interestingly, the focal adhesion pathway was enriched significantly in both clusters ((KEGG: 04510; cluster 1: FDR = 9.82×10^{-3}, fold changed = 4.97; cluster 3: FDR = 4.90×10^{-7}, fold changed = 7.81). Further enriched GO annotations also support this result, such as actin cytoskeleton organization (GO: 0030036) and cell-matrix adhesion (GO: 0007160).

2.2. *Clic4 and Gas2l1 Are Overexpressed in Pancreatic CTCs*

As the study's primary aim was to identify biomarkers that can distinguish CTCs from blood cells, the expression levels of CTCs (group CTC) were compared to those of WBCs (group WBC). Interestingly, only nine genes were significantly over-expressed in CTCs, *Capns1*, *Csrp1*, *Rpl41*, *Bsg*, *Ppp2ca*, *Clic4*, *Gas2l1*, *Aldh2*, and *Fkbp8*. Two of them, *Clic4* and *Gas2l1*, were identified as markers of cluster 3.

The traditional leukocyte marker *Cd45 (Ptprc)* and the epithelial marker *Epcam* were used as negative and positive control markers, respectively. We also included *Sparc*, which was reported in the original study [17], into our analysis. While *Clic4*, *Gas2l1*, and *Sparc* were highly enriched in pancreatic CTCs (also surpassing the expression of *Epcam*), *Cd45* exhibited a low expression pattern in CTCs (Figure 4). To confirm the expression of candidate markers PBMCs from healthy humans, we use a single-cell RNA-seq dataset with 5155 PBMCs from 10x Genomics to perform the same analysis procedure. As expected, *CLIC4*, *GAS2L1*, and *SPARC* are rarely expressed in PBMC, while *CD45* (*PTPRC*) was ubiquitously expressed. This result was reproducible in another PBMC single-cell RNA sequencing dataset (Supplementary Figure S2).

2.3. *Gas2l1 and Epcam Expression Identify Distinct CTCs Subpopulations and Synergize in CTC Identification*

Further calculation of the AUC value and Mann–Whitney test of *Clic4*, *Gas2l1*, *Sparc* and *Epcam* were implemented. As depicted in Figure 5A, *Clic4* and *Gas2l1* exhibit significantly higher AUC values ($AUC_{Clic4} = 0.963$, $p < 0.001$; $AUC_{Gas2l1} = 0.938$, $p < 0.001$) than *Sparc* and *Epcam* ($AUC_{Sparc} = 0.695$, $p = 0.031$; $AUC_{Epcam} = 0.514$, $p = 0.873$).

To validate whether the potential markers could work in enriched CTC samples, we also explored two other datasets, GSE40174 and GSE144561, which represent blood samples of PDAC patients processed by microfluidic CTC chips (EpCAM ^{Hb}CTC-Chip [18] and CTC-iChip [19]). Remarkably, only *GAS2L1* is significantly overexpressed in the blood of metastatic PDAC patients in both datasets, while the differential expression of *CLIC4* and *SPARC* between the two groups is inconsistent (Figure 5B). As *GAS2L1* overexpression was found in these samples, we selected *Gas2l1* for the following analysis.

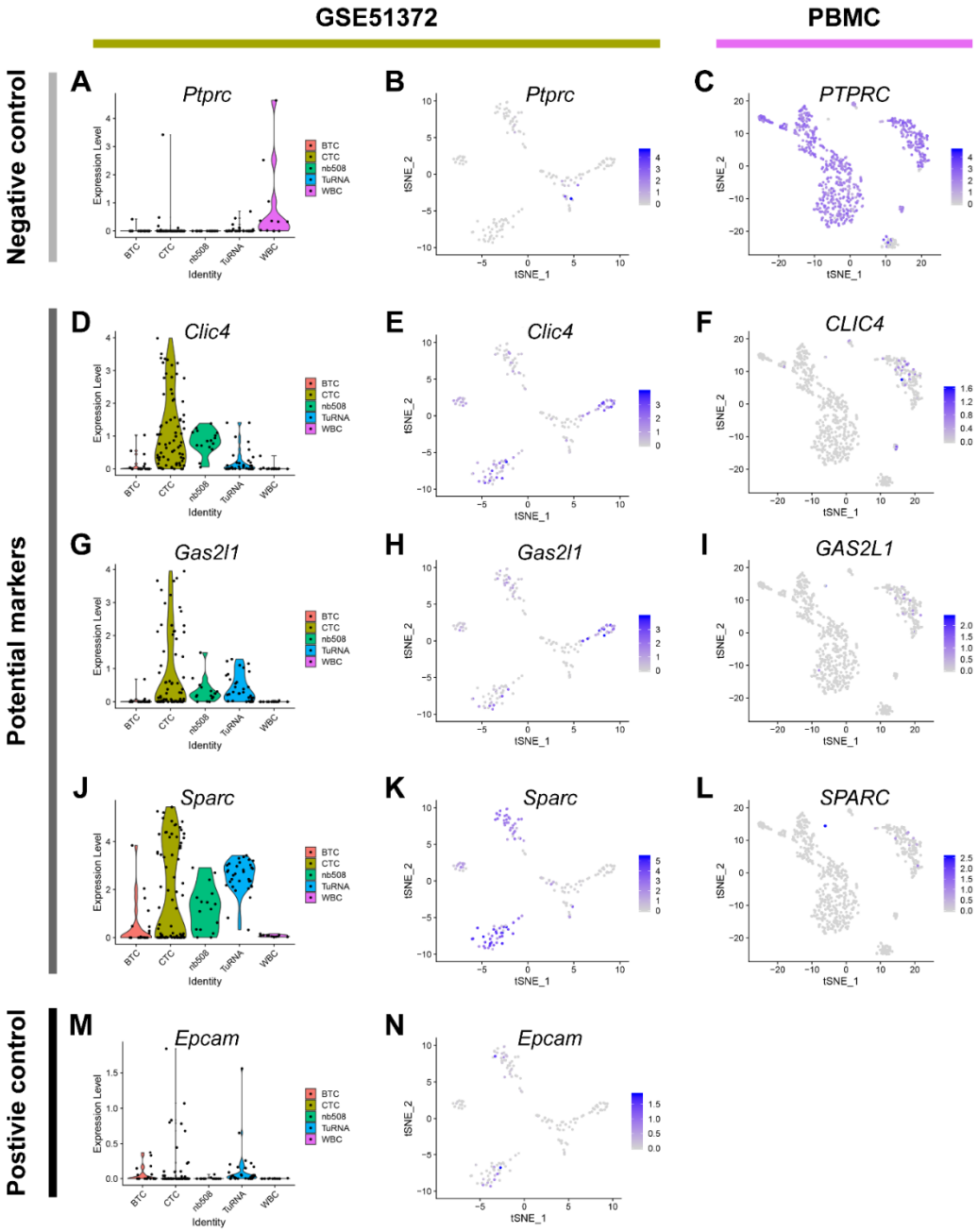

Figure 4. Expression of markers in circulating tumor cells (CTCs) and peripheral blood mononuclear cells (PBMCs). There are a violin plot and scatter plots for each marker to show the expression level in the CTC and PBMC datasets. *Ptprc/Cd45* (**A–C**) is a negative control, *Clic4* (**D–F**), *Gas2l1* (**G–I**) and *Sparc* (**J–L**) are potential CTC markers, *Epcam* (**M**,**N**) is used as the positive control. *Epcam* is not expressed in the PBMC dataset. BTCs, bulk tumor cells; CTCs, circulating tumor cells; t-SNE, t-Distributed Stochastic Neighbor Embedding; TuRNA, RNA of the bulk tumor; WBC, white blood cells.

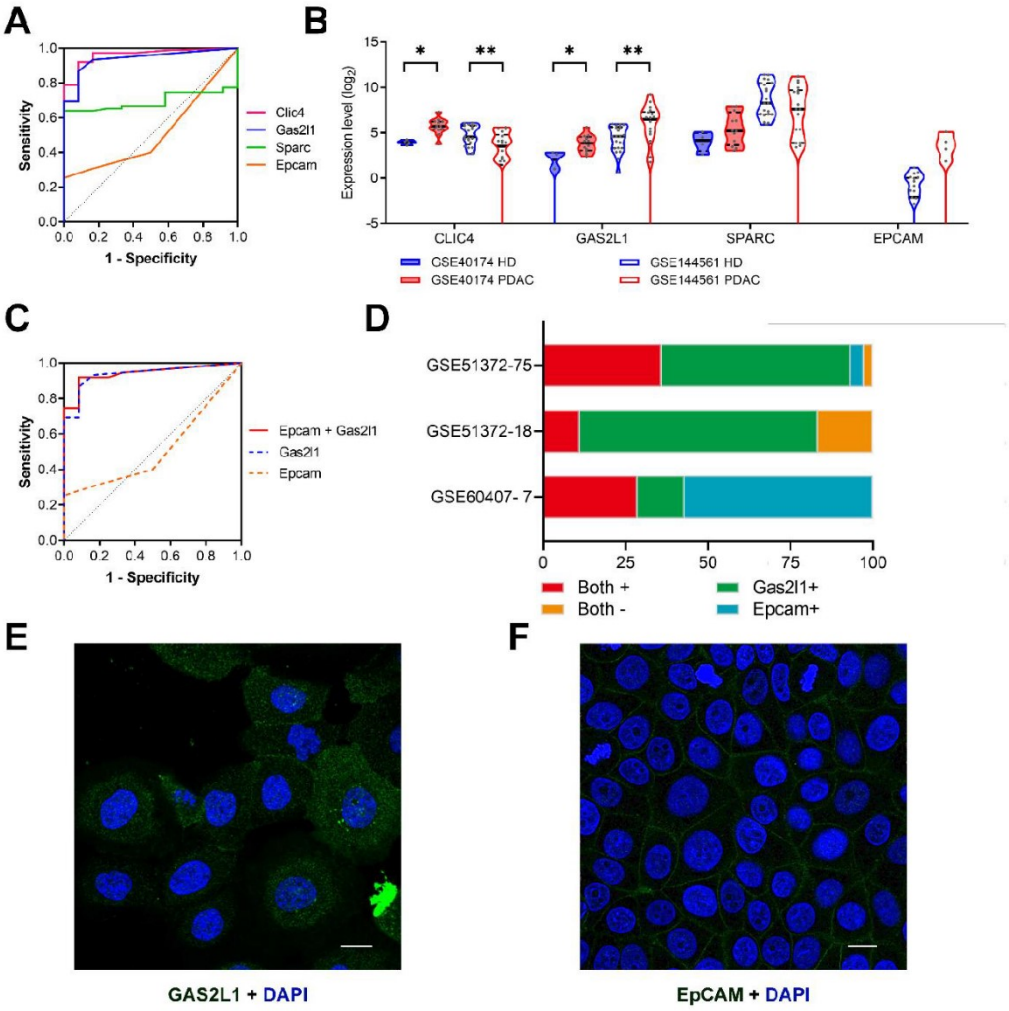

Figure 5. T *Gas2l1* is a potential marker for pancreatic circulating tumor cells (CTCs). (**A**) Receiver operator characteristic (ROC) analysis of markers in murine samples (GSE51372). (**B**) The violin plots reflect the expression level of *CLIC4*, *GAS2L1*, *SPARC*, and *EPCAM* in human blood samples (GSE40174 and GSE144561). *, $p < 0.050$; **, $p < 0.001$. (**C**) The ROC curve reflects the combination of *Gas2l1* and *Epcam* in murine samples (GSE51372). (**D**) The stacked bars show the *Gas2l1* and *Epcam* expression ratio in 75 murine pancreatic CTCs (GSE51372) and two validation datasets (18 GFP-Tagged murine CTCs from GSE51372 and 7 human pancreatic CTCs from GSE60407). The subcellular location of GAS2L1 (**E**) and EPCAM (**F**) in human cell lines (A-431 and MCF7, respectively); scale bar, 20 µm. Blue, nucleus; Green, antibody. These two pictures were obtained from the Human Protein Atlas (https://www.proteinatlas.org/). AUC, area under curve; *EPCAM*, epithelial cell adhesion molecule; HD, healthy donor; PDAC, pancreatic ductal adenocarcinoma; ROC, receiver operator characteristic.

The expression of *Gas2l1* was analyzed in 18 GFP-tagged murine pancreatic CTCs (GSE51372) and seven human pancreatic CTCs. While *Gas2l1* is expressed in most murine CTCs (83.3%, 15/18), three of seven human pancreatic CTCs (GSE60407) lack *GAS2L1* expression. This result indicates that similar to other known CTC markers, *GAS2L1* cannot identify all but only a subset of CTCs. In studies aiming to quantify all tumor-derived cells in the bloodstream, *GAS2L1* should therefore be combined with other markers such as *EPCAM* (Figure 5C). At least, *GAS2L1* is significantly overexpressed in CTC enriched cell population after the EpCAM ^{Hb}CTC-Chip enrichment. In fact,

the positivity of one or both markers identifies the majority of murine pancreatic CTCs (GSE51372; 73 of 75 CTCs (97.3%); 15 of 18 GFP-tagged CTCs (83.3%)) and all seven CTCs in the human pancreatic CTC dataset (GSE60407) (Figure 5D; Table 1). Interestingly, there is no statistically significant overlap in *Gas2l1*$^{+}$ and *Epcam*$^{+}$ murine pancreatic CTC populations (Figure 5D; Spearman r = −0.119, $p = 0.310$, Supplementary Figure S3), suggesting their complementary potential. The GAS2L1 protein is located both in the cytoplasm and the plasma membrane and the EPCAM protein is located in the plasma membrane (Figure 5E,F). Therefore, antibodies binding to GAS2L1 and EPCAM can be used to identify this CTC subpopulation without prior permeabilization of the cells.

Table 1. Expression of *GAS2L1* in CTC samples.

Dataset–Sample Size	*Gas2l1*$^{+}$ [1]	*Epcam*$^{+}$	*Gas2l1*$^{+}$ and/or *Epcam*$^{+}$	Total
GSE51372-75	70 (93.3%)	30 (40.0%)	73 (97.3%)	75
GSE51372-18	15 (83.3%)	2 (11.1%)	15 (83.3%)	18
GSE60407-7	3 (42.9%)	6 (85.7%)	7 (100.0%)	7

[1] +, positive expression.

2.4. Intratumoral GAS2L1 Negatively Correlates with Recurrence-Free Survival

To evaluate the expression of *GAS2L1* in pancreatic cancer and normal tissues, we utilized The Gene Expression Profiling Interactive Analysis 2 (GEPIA 2) (http://gepia2.cancer-pku.cn/#analysis) tool [20]. *GAS2L1* was significantly overexpressed ($p < 0.001$) in pancreatic adenocarcinoma as compared to matched normal tissue and the normal pancreatic data from the Broad Genotype-Tissue Expression (GTEx) portal (Figure 6A). We further investigated whether the expression levels of *GAS2L1* correlated with clinical prognosis (OS and RFS) in the pancreatic adenocarcinoma cohort of TCGA. Although OS was not influenced by *GAS2L1* [$p = 0.937$; HR = 0.98 (0.65–1.48)], patients with higher *GAS2L1* expression [$p = 0.006$; HR = 2.75 (1.21–6.24)] have significantly worse RFS (Figure 6B,C). These results point towards a prognostic value of *GAS2L1* in resected, non-metastatic pancreatic cancer.

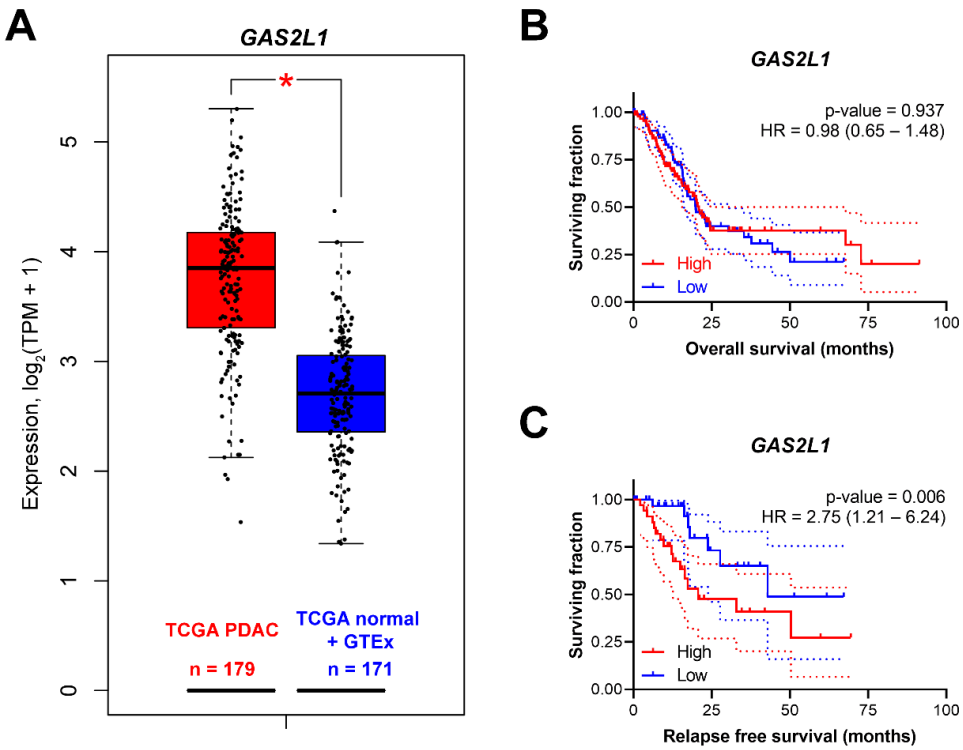

Figure 6. Prognostic value of *GAS2L1*. (**A**) *GAS2L1* expression in the TCGA pancreatic tumors with corresponding matched normal tissue and GTEx pancreatic tissue. *, $p < 0.001$. Overall survival (OS) (**B**) and relapse-free survival (RFS) (**C**) of *GAS2L1*. GTEx, The Genotype-Tissue Expression; HR, hazard ratio; PDAC, pancreatic ductal adenocarcinoma; TCGA, The Cancer Genome Atlas; TPM, transcripts per million.

3. Discussion

A growing body of evidence suggests that CTCs contribute to the development of metastases [5,21,22]. Besides, it is well established that CTCs can be found early in PDAC development [23–26]. In a genetically engineered mouse model (GEMM) of PDAC, pancreatic cells were detected in the bloodstream even before malignancy could be detected by histologic examination of the pancreas [27]. These observations encourage the hypothesis that CTC could be used as an early indicator of pancreatic malignancy. Furthermore, several research groups demonstrated that CTCs have prognostic relevance in pancreatic cancer [26,28–30].

It is still a technical challenge to distinguish CTCs from the surrounding blood components as a rare and heterogeneous population. Various technologies have been developed to isolate CTCs based on the fact that their physical properties (i.e., size, density, elasticity) differ slightly from those of leukocytes [7,31,32]. Alternatively, biologic differences such as protein expression on the cell surface can also be used to distinguish and isolate CTCs using fluorophore- or magnetic beat-coupled antibodies [33].

This study demonstrates that pancreatic CTCs are highly heterogeneous and can be divided into three distinct clusters, two of which are pure CTC clusters. Interestingly, both pure CTC clusters showed increased expression of the focal adhesion pathway and several relevant gene ontology terms, including actin cytoskeleton organization (GO: 0030036) and cell-matrix adhesion (GO: 0007160).

The presented data also suggests that *Gas2l1* may be used as a potential identification marker for pancreatic CTCs. The role of *Gas2l1* (growth arrest specific 2 like 1) in cancer is largely unknown, and even more so in CTCs. Prior studies have noted that *Gas2l1* encodes a member of the growth

arrest-specific 2 (GAS2) protein family, which guides microtubules towards focal adhesions through physical crosslinking of growing microtubules to actin stress fibers [34–36]. Only a few studies are investigating the role of GAS2 in oncogenesis and there is no consensus on whether GAS2 acts as a tumor suppressor or oncogene. GAS2 is upregulated in malignant glioma [37]. In colorectal cancer (CRC), fecal GAS2 was proposed as a non-invasive marker for early recurrence as it can be found in the feces of patients with recurrent CRC [38]. Besides, GAS2 expression is associated with proliferative activity in CRC [39]. In contrast, GAS2 seems to act as a tumor suppressor by inhibiting cell growth in hepatocellular carcinoma [40,41]. There is currently no data regarding the role of GAS2 in PDAC.

GAS2L1 is required for centrosome and microtubule dynamics [34], which are required for cell polarization and migration [42,43]. Microtubules have a pivotal role in regulating cell protrusion and forming focal adhesions at the anterior migration margin [42,44]. As a result, microtubules are indispensable for CTC attachment to the capillary endothelium [45,46]. All of the above mechanisms are critically required for the successful completion of CTC seeding and, ultimately, the process of metastasis, suggesting a role of Gas2l1 in this process.

Gas2l1 is also reported to be expressed in platelets [46]. Multiple platelet pathways are overexpressed in cluster 3, most likely due to platelets adhering to the CTCs in this cluster. Therefore, it is reasonable to assume that the *Gas2l1* transcripts found in cluster 3 are at least partly derived from platelets. The CTC–platelet interaction may activate integrins to form a fibrin-based protective envelope for CTCs [46–48]. This may serve as a possible explanation for the link between worse RFS and overexpression of *GAS2L1* in human PDAC. However, this hypothesis is limited by the missing correlation between *GAS2L1* and OS in the clinical dataset, as well as the comparison between *GAS2L1* expression in tumor tissue and CTCs. In addition, only four of seven human pancreatic single CTCs expressed *GAS2L1*. This may reflect CTC heterogeneity or unknown confounding factors such as aspirin medication of the study participants. Generally, the available data from only seven human CTCs need to be put on a broader basis before drawing definitive conclusions.

The current literature recognizes the critical role of epithelial cell adhesion molecule (EPCAM) in identifying CTCs due to its absence in normal blood cells [49,50]. However, there is increasing concern that CTCs may at least partially lose epithelial traits, including EPCAM expression during epithelial–mesenchymal transition (EMT) [51–55]. For instance, only 40% (30/75) of the murine pancreatic CTCs in GSE51372 express *EPCAM*. Therefore, the usage of EPCAM as an identification marker in the majority of CTC studies leads to a severe bias toward epithelial CTCs. Several studies suggest that EPCAM-negative CTCs also have aggressive metastatic potential [56–58]. The addition of *GAS2L1* expression as a criterium to identify CTCs may increase the number of positively identified CTCs. As there is no significant overlap in the *Gas2l1*$^+$ and *Epcam*$^+$ CTC populations, GAS2L1 may be complementary to EPCAM as a selection marker in CTCs. This is supported by the fact that 7/7 human and 73/75 (97.3%) murine pancreatic CTCs were positive for either *GAS2L1* and/or *EPCAM*. Importantly, GAS2L1 protein can be found on the cell surface, therefore enabling the use of GAS2L1 antibodies for the identification and isolation of live CTCs without the need for prior permeabilization, which severely limits the use of subsequent RNA-based assays. This makes the combination of surface EPCAM/GAS2L1 a promising option for the identification of pancreatic CTCs.

4. Materials and Methods

4.1. Data Collection

All datasets analyzed in this study are publicly available. Gene expression datasets and clinical information profiles for human PDAC were obtained from the Cancer Genome Atlas (TCGA) data portal (https://tcga-data.nci.nih.gov/tcga/) [59]. The RNAseq-based expression normalization fragments per kilobase of transcript per million mapped reads upper quartile (FPKM-UQ) normalization method was used.

Two single-cell sequencing datasets of pancreatic CTCs were downloaded from the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). GSE51372 contains CTCs from LSL-*Kras*^{G12D}, LSL-*Trp53*$^{flox/flox\ or\ +}$, *Pdx1-Cre* mice [17]. Another dataset (GSE60407) consists of seven expression profiles of CTCs from three PDAC patients [17]. Both datasets were generated on an AB 5500xl Genetic Analyzer on GPL15907 (*Mus musculus*) and GPL16288 (*Homo sapiens*) platforms, respectively. Another two RNA-seq datasets of enriched pancreatic CTC samples are also accessible through GEO Series accession number GSE40174 [60] and GES144561 [61].

Additionally, a single-cell gene expression dataset of 5155 peripheral blood mononuclear cells (PBMCs) of a healthy donor was obtained from 10X genomics (https://support.10xgenomics.com/single-cell-gene-expression/datasets/3.0.2/5k_pbmc_v3_nextgem). The dataset was established on the Illumina NovaSeq platform, and sample demultiplexed, barcode processed, and single-cell 3′ gene counted by Cell Ranger 3.0.2. Single-cell expression profiles of PBMCs from healthy donors were downloaded from the Broad Single Cell Portal (BETA) (https://singlecell.broadinstitute.org/single_cell) and used for validation.

4.2. Data Quality Check, Pre-Processing, and Clustering

Seurat, a specific R toolkit for quality control (QC) and exploration of single-cell transcriptomic data, was employed according to the instructions provided by the Satija lab (https://satijalab.org/) [62].

After removing cells with fewer than 200 unique genes per cell from the dataset, we performed a global-scaling normalization, multiplied all remaining gene expression by 10,000, followed by log transformation. After normalization, the top 3000 highly variable genes were extracted from each dataset based on mean variance.

Both linear (Principal Component Analysis (PCA)) and non-linear dimensional reduction (t-Distributed Stochastic Neighbor Embedding (t-SNE)) were performed after the scaling (linear transformation). The JackStrawPlot function was employed to calculate the top principal components of the dataset. Principal components (PCs) with low p-values after random permutation and PCA recalculation were identified as significant and submitted to the following clustering using the FindClusters function to iteratively group cells together based on modularity optimization techniques.

4.3. Statistical Analysis and Visualization

The cluster of interest was compared to all other clusters to identify the markers of the target clusters. $p < 0.050$ and |log(Fold Change, FC)| > 1 were chosen as cutoffs to define significant markers.

The functional enrichment analysis, which consists of the pathway analysis of the Kyoto Encyclopedia of Genes and Genomes (KEGG) [63] and the functional interpretations of Gene Ontology (GO) [64], was completed by the g:Profiler (https://biit.cs.ut.ee/gprofiler/gost) [65]. All three aspects—biological processes (BP), cellular components (CC), and molecular functions (MF)—were included in the GO analysis. Fold enrichment > 2.0 and Benjamini–Hochberg false discovery rate (FDR) < 0.001 were defined as statistically significant.

The prognostic value of GAS2L1 was explored in the TCGA pancreatic cancer cohort. Overall survival (OS) and relapse-free survival (RFS) curves were displayed with p-values (log-rank test) and 95% confidence intervals (CI) of hazard ratios (HR). The Kaplan–Meier plots were generated by GraphPad Prism 8 to visualize the differences and comparisons with $p < 0.050$ were considered statistically significant.

The expression in different groups was compared using the Receiver Operating Characteristic (ROC) and Mann–Whitney tests in the R statistical environment using Bioconductor libraries (www.bioconductor.org). Markers with an area under the curve (AUC) value of the ROC check > 0.70 and $p < 0.050$ were considered statistically significant. GraphPad Prism 8 was used to plot the figures.

5. Conclusions

In summary, the here presented data suggests Gas2l1/GAS2L1 as a potential biomarker for pancreatic CTCs. The combination of EPCAM/GAS2L1 surface protein stainings may increase the detection rate and comprehensiveness of CTC studies in pancreatic cancer without limiting the availability of the isolated CTCs for downstream analyses. As the available single-cell RNA-seq datasets of pancreatic CTCs are very limited in sample size, we utilized several complementary datasets to validate our findings. However, further experiments including murine and human studies, are necessary to evaluate the viability of the proposed GAS2L1 and EPCAM combination strategy in identifying pancreatic CTCs.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-6694/12/12/3774/s1, Figure S1: Visualization of the quality check metrics of 157 single-cell sequencing profiles; Figure S2: Expression of markers in peripheral blood mononuclear cells (PBMCs) from two healthy donors; Figure S3: *Gas2l1* and *Epcam* expression do not correlate in murine pancreatic CTCs.

Author Contributions: Conceptualization, L.Z.; Methodology, L.Z. and K.-J.K.; Validation, L.Z., K.-J.K., J.L.G., B.H., S.L., and S.S.; Formal analysis, L.Z., K.-J.K., and B.G.; Investigation, L.Z., K.-J.K., and B.G.; Data curation, L.Z., K.-J.K., and B.G.; Visualization, L.Z., K.-J.K., B.G., and S.S.; Funding acquisition, L.Z. and K.-J.K.; Project administration, S.S.; Supervision, C.R. and S.S.; Writing—original draft, L.Z., K.-J.K., B.H., and S.S.; Writing—review and editing, L.Z., K.-J.K., J.L.G., B.H., C.Y., B.G., S.L., C.R., and S.S. All authors have read and agreed to the published version of the manuscript.

Funding: Lei Zhu and Ke-Jia Kan were supported by the China Scholarship Council (CSC), no. 201908080072 and no. 201706230257.

Acknowledgments: Lei Zhu acknowledges Jian-Ming Zeng (Faculty of Health Sciences, University of Macau, China) and his team for their bioinformatic support. Balázs Győrffy was supported by the Higher Education Institutional Excellence Programme (2020-4.1.1.-TKP2020) of the Ministry for Innovation and Technology in Hungary, within the framework of the Bionic thematic programme of the Semmelweis University. The Hector Stiftung II supports Sebastian Schölch and Sonja Loges. This paper is dedicated to the memory of our friend Professor Moritz Koch.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2019. *CA Cancer J. Clin.* **2019**, *69*, 7–34. [CrossRef]
2. Vincent, A.; Herman, J.; Schulick, R.; Hruban, R.H.; Goggins, M. Pancreatic cancer. *Lancet* **2011**, *378*, 607–620. [CrossRef]
3. Jeong, K.-Y.; Kim, E.K.; Park, M.H.; Kim, H.M. Perspective on Cancer Therapeutics Utilizing Analysis of Circulating Tumor Cells. *Diagnostics* **2018**, *8*, 23. [CrossRef]
4. Yu, M.; Bardia, A.; Aceto, N.; Bersani, F.; Madden, M.W.; Donaldson, M.C.; Desai, R.; Zhu, H.; Comaills, V.; Zheng, Z.; et al. Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. *Science* **2014**, *345*, 216–220. [CrossRef]
5. Klotz, R.; Thomas, A.; Teng, T.; Han, S.M.; Iriondo, O.; Li, L.; Restrepo-Vassalli, S.; Wang, A.; Izadian, N.; MacKay, M.; et al. Circulating Tumor Cells Exhibit Metastatic Tropism and Reveal Brain Metastasis Drivers. *Cancer Discov.* **2020**, *10*, 86–103. [CrossRef]
6. Martini, V.; Timme-Bronsert, S.; Fichtner-Feigl, S.; Hoeppner, J.; Kulemann, B. Circulating Tumor Cells in Pancreatic Cancer: Current Perspectives. *Cancers* **2019**, *11*, 1659. [CrossRef]
7. Sharma, S.; Zhuang, R.; Long, M.; Pavlovic, M.; Kang, Y.; Ilyas, A.; Asghar, W. Circulating tumor cell isolation, culture, and downstream molecular analysis. *Biotechnol. Adv.* **2018**, *36*, 1063–1078. [CrossRef]
8. Lustberg, M.B.; Balasubramanian, P.; Miller, B.; Garcia-Villa, A.; Deighan, C.; Wu, Y.; Carothers, S.; Berger, M.; Ramaswamy, B.; Macrae, E.R.; et al. Heterogeneous atypical cell populations are present in blood of metastatic breast cancer patients. *Breast Cancer Res.* **2014**, *16*, R23. [CrossRef] [PubMed]

9. Bertolini, G.; D'Amico, L.; Moro, M.; Landoni, E.; Perego, P.; Miceli, R.; Gatti, L.; Andriani, F.; Wong, D.; Caserini, R.; et al. Microenvironment-Modulated Metastatic CD133+/CXCR4+/EpCAM- Lung Cancer-Initiating Cells Sustain Tumor Dissemination and Correlate with Poor Prognosis. *Cancer Res.* **2015**, *75*, 3636–3649. [CrossRef] [PubMed]
10. Wen, K.-C.; Sung, P.-L.; Chou, Y.-T.; Pan, C.-M.; Wang, P.-H.; Lee, O.K.-S.; Wu, C.-W. The role of EpCAM in tumor progression and the clinical prognosis of endometrial carcinoma. *Gynecol. Oncol.* **2018**, *148*, 383–392. [CrossRef] [PubMed]
11. Poudineh, M.; Aldridge, P.M.; Ahmed, S.; Green, B.J.; Kermanshah, L.; Nguyen, V.; Tu, C.; Mohamadi, R.M.; Nam, R.K.; Hansen, A.; et al. Tracking the dynamics of circulating tumour cell phenotypes using nanoparticle-mediated magnetic ranking. *Nat. Nanotechnol.* **2017**, *12*, 274–281. [CrossRef] [PubMed]
12. Pestrin, M.; Salvianti, F.; Galardi, F.; De Luca, F.; Turner, N.; Malorni, L.; Pazzagli, M.; Di Leo, A.; Pinzani, P. Heterogeneity of PIK3CA mutational status at the single cell level in circulating tumor cells from metastatic breast cancer patients. *Mol. Oncol.* **2015**, *9*, 749–757. [CrossRef] [PubMed]
13. Zeinali, M.; Lee, M.; Nadhan, A.; Mathur, A.; Hedman, C.; Lin, E.; Harouaka, R.; Wicha, M.S.; Zhao, L.; Palanisamy, N.; et al. High-Throughput Label-Free Isolation of Heterogeneous Circulating Tumor Cells and CTC Clusters from Non-Small-Cell Lung Cancer Patients. *Cancers* **2020**, *12*, 127. [CrossRef] [PubMed]
14. Nanduri, L.K.; Hissa, B.; Weitz, J.; Schölch, S.; Bork, U. The prognostic role of circulating tumor cells in colorectal cancer. *Expert Rev. Anticancer Ther.* **2019**, *19*, 1077–1088. [CrossRef] [PubMed]
15. Pecot, C.V.; Bischoff, F.Z.; Mayer, J.A.; Wong, K.L.; Pham, T.; Bottsford-Miller, J.; Stone, R.L.; Lin, Y.G.; Jaladurgam, P.; Roh, J.W.; et al. A novel platform for detection of CK+ and CK− CTCs. *Cancer Discov.* **2011**, *1*, 580–586. [CrossRef] [PubMed]
16. Baccelli, I.; Schneeweiss, A.; Riethdorf, S.; Stenzinger, A.; Schillert, A.; Vogel, V.; Klein, C.; Saini, M.; Bäuerle, T.; Wallwiener, M.; et al. Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay. *Nat. Biotechnol.* **2013**, *31*, 539–544. [CrossRef]
17. Ting, D.T.; Wittner, B.S.; Ligorio, M.; Vincent Jordan, N.; Shah, A.M.; Miyamoto, D.T.; Aceto, N.; Bersani, F.; Brannigan, B.W.; Xega, K.; et al. Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells. *Cell Rep.* **2014**, *8*, 1905–1918. [CrossRef]
18. Yu, M.; Stott, S.; Toner, M.; Maheswaran, S.; Haber, D.A. Circulating tumor cells: Approaches to isolation and characterization. *J. Cell Biol.* **2011**, *192*, 373–382. [CrossRef]
19. Fachin, F.; Spuhler, P.; Martel-Foley, J.M.; Edd, J.F.; Barber, T.A.; Walsh, J.; Karabacak, M.; Pai, V.; Yu, M.; Smith, K.; et al. Monolithic Chip for High-throughput Blood Cell Depletion to Sort Rare Circulating Tumor Cells. *Sci. Rep.* **2017**, *7*. [CrossRef]
20. Tang, Z.; Kang, B.; Li, C.; Chen, T.; Zhang, Z. GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **2019**, *47*, W556–W560. [CrossRef]
21. Gkountela, S.; Castro-Giner, F.; Szczerba, B.M.; Vetter, M.; Landin, J.; Scherrer, R.; Krol, I.; Scheidmann, M.C.; Beisel, C.; Stirnimann, C.U.; et al. Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding. *Cell* **2019**, *176*, 98–112.e14. [CrossRef] [PubMed]
22. Liu, X.; Taftaf, R.; Kawaguchi, M.; Chang, Y.-F.; Chen, W.; Entenberg, D.; Zhang, Y.; Gerratana, L.; Huang, S.; Patel, D.B.; et al. Homophilic CD44 Interactions Mediate Tumor Cell Aggregation and Polyclonal Metastasis in Patient-Derived Breast Cancer Models. *Cancer Discov.* **2019**, *9*, 96–113. [CrossRef] [PubMed]
23. Kulemann, B.; Pitman, M.B.; Liss, A.S.; Valsangkar, N.; Fernández-Del Castillo, C.; Lillemoe, K.D.; Hoeppner, J.; Mino-Kenudson, M.; Warshaw, A.L.; Thayer, S.P. Circulating tumor cells found in patients with localized and advanced pancreatic cancer. *Pancreas* **2015**, *44*, 547–550. [CrossRef] [PubMed]
24. Gao, Y.; Zhu, Y.; Zhang, Z.; Zhang, C.; Huang, X.; Yuan, Z. Clinical significance of pancreatic circulating tumor cells using combined negative enrichment and immunostaining-fluorescence in situ hybridization. *J. Exp. Clin. Cancer Res.* **2016**, *35*. [CrossRef]
25. Rhim, A.D.; Thege, F.I.; Santana, S.M.; Lannin, T.B.; Saha, T.N.; Tsai, S.; Maggs, L.R.; Kochman, M.L.; Ginsberg, G.G.; Lieb, J.G.; et al. Detection of circulating pancreas epithelial cells in patients with pancreatic cystic lesions. *Gastroenterology* **2014**, *146*, 647–651. [CrossRef]

26. Court, C.M.; Ankeny, J.S.; Sho, S.; Winograd, P.; Hou, S.; Song, M.; Wainberg, Z.A.; Girgis, M.D.; Graeber, T.G.; Agopian, V.G.; et al. Circulating Tumor Cells Predict Occult Metastatic Disease and Prognosis in Pancreatic Cancer. *Ann. Surg. Oncol.* **2018**, *25*, 1000–1008. [CrossRef]
27. Rhim, A.D.; Mirek, E.T.; Aiello, N.M.; Maitra, A.; Bailey, J.M.; McAllister, F.; Reichert, M.; Beatty, G.L.; Rustgi, A.K.; Vonderheide, R.H.; et al. EMT and dissemination precede pancreatic tumor formation. *Cell* **2012**, *148*, 349–361. [CrossRef]
28. Kurihara, T.; Itoi, T.; Sofuni, A.; Itokawa, F.; Tsuchiya, T.; Tsuji, S.; Ishii, K.; Ikeuchi, N.; Tsuchida, A.; Kasuya, K.; et al. Detection of circulating tumor cells in patients with pancreatic cancer: A preliminary result. *J. Hepatobiliary Pancreat. Surg.* **2008**, *15*, 189–195. [CrossRef]
29. Bidard, F.C.; Huguet, F.; Louvet, C.; Mineur, L.; Bouché, O.; Chibaudel, B.; Artru, P.; Desseigne, F.; Bachet, J.B.; Mathiot, C.; et al. Circulating tumor cells in locally advanced pancreatic adenocarcinoma: The ancillary CirCe 07 study to the LAP 07 trial. *Ann. Oncol.* **2013**, *24*, 2057–2061. [CrossRef]
30. Chang, M.-C.; Chang, Y.-T.; Chen, J.-Y.; Jeng, Y.-M.; Yang, C.-Y.; Tien, Y.-W.; Yang, S.-H.; Chen, H.-L.; Liang, T.-Y.; Wang, C.-F.; et al. Clinical Significance of Circulating Tumor Microemboli as a Prognostic Marker in Patients with Pancreatic Ductal Adenocarcinoma. *Clin. Chem.* **2016**, *62*, 505–513. [CrossRef]
31. Lee, J.-S.; Park, S.S.; Lee, Y.K.; Norton, J.A.; Jeffrey, S.S. Liquid biopsy in pancreatic ductal adenocarcinoma: Current status of circulating tumor cells and circulating tumor DNA. *Mol. Oncol.* **2019**, *13*, 1623–1650. [CrossRef] [PubMed]
32. Rossi, E.; Zamarchi, R. Single-Cell Analysis of Circulating Tumor Cells: How Far Have We Come in the -Omics Era? *Front. Genet.* **2019**, *10*, 958. [CrossRef] [PubMed]
33. Tsavellas, G.; Huang, A.; McCullough, T.; Patel, H.; Araia, R.; Allen-Mersh, T.G. Flow cytometry correlates with RT-PCR for detection of spiked but not circulating colorectal cancer cells. *Clin. Exp. Metastasis* **2002**, *19*, 495–502. [CrossRef]
34. Au, F.K.C.; Jia, Y.; Jiang, K.; Grigoriev, I.; Hau, B.K.T.; Shen, Y.; Du, S.; Akhmanova, A.; Qi, R.Z. GAS2L1 Is a Centriole-Associated Protein Required for Centrosome Dynamics and Disjunction. *Dev. Cell* **2017**, *40*, 81–94. [CrossRef] [PubMed]
35. Stroud, M.J.; Nazgiewicz, A.; McKenzie, E.A.; Wang, Y.; Kammerer, R.A.; Ballestrem, C. GAS2-like proteins mediate communication between microtubules and actin through interactions with end-binding proteins. *J. Cell. Sci.* **2014**, *127*, 2672–2682. [CrossRef] [PubMed]
36. van de Willige, D.; Hummel, J.J.; Alkemade, C.; Kahn, O.I.; Au, F.K.; Qi, R.Z.; Dogterom, M.; Koenderink, G.H.; Hoogenraad, C.C.; Akhmanova, A. Cytolinker Gas2L1 regulates axon morphology through microtubule-modulated actin stabilization. *EMBO Rep.* **2019**, *20*. [CrossRef]
37. Sui, R.; Piao, H.-Z. UCHL1 enhances the malignant development of glioma via targeting GAS2. *Eur. Rev. Med. Pharmacol. Sci.* **2020**, *24*, 6195–6203. [CrossRef]
38. Huang, C.-J.; Lee, C.-L.; Yang, S.-H.; Chien, C.-C.; Huang, C.-C.; Yang, R.-N.; Chang, C.-C. Upregulation of the growth arrest-specific-2 in recurrent colorectal cancers, and its susceptibility to chemotherapy in a model cell system. *Biochim. Biophys. Acta (BBA)-Mol. Basis Dis.* **2016**, *1862*, 1345–1353. [CrossRef]
39. Chang, C.-C.; Huang, C.-C.; Yang, S.-H.; Chien, C.-C.; Lee, C.-L.; Huang, C.-J. Data on clinical significance of GAS2 in colorectal cancer cells. *Data Brief* **2016**, *8*, 82–86. [CrossRef]
40. Zhu, R.; Mok, M.T.S.; Kang, W.; Lau, S.S.K.; Yip, W.-K.; Chen, Y.; Lai, P.B.S.; Wong, V.W.S.; To, K.-F.; Sung, J.J.Y.; et al. Truncated HBx-dependent silencing of GAS2 promotes hepatocarcinogenesis through deregulation of cell cycle, senescence and p53-mediated apoptosis. *J. Pathol.* **2015**, *237*, 38–49. [CrossRef]
41. Zhu, R.-X.; Cheng, A.S.L.; Chan, H.L.Y.; Yang, D.-Y.; Seto, W.-K. Growth arrest-specific gene 2 suppresses hepatocarcinogenesis by intervention of cell cycle and p53-dependent apoptosis. *World J. Gastroenterol.* **2019**, *25*, 4715–4726. [CrossRef] [PubMed]
42. Heikenwalder, M.; Lorentzen, A. The role of polarisation of circulating tumour cells in cancer metastasis. *Cell. Mol. Life Sci.* **2019**, *76*, 3765–3781. [CrossRef] [PubMed]
43. Luxton, G.W.G.; Gundersen, G.G. Orientation and Function of the Nuclear-Centrosomal Axis During Cell Migration. *Curr. Opin. Cell Biol.* **2011**, *23*, 579–588. [CrossRef] [PubMed]
44. Etienne-Manneville, S. Microtubules in cell migration. *Annu. Rev. Cell Dev. Biol.* **2013**, *29*, 471–499. [CrossRef]
45. Killilea, A.N.; Csencsits, R.; Le, E.; Patel, A.M.; Kenny, S.J.; Xu, K.; Downing, K.H. Cytoskeletal Organization in Microtentacles. *Exp. Cell Res.* **2017**, *357*, 291–298. [CrossRef]

46. Matrone, M.A.; Whipple, R.A.; Balzer, E.M.; Martin, S.S. Microtentacles tip the balance of cytoskeletal forces in circulating tumor cells. *Cancer Res.* **2010**, *70*, 7737–7741. [CrossRef]
47. Lonsdorf, A.S.; Krämer, B.F.; Fahrleitner, M.; Schönberger, T.; Gnerlich, S.; Ring, S.; Gehring, S.; Schneider, S.W.; Kruhlak, M.J.; Meuth, S.G.; et al. Engagement of αIIbβ3 (GPIIb/IIIa) with αvβ3 integrin mediates interaction of melanoma cells with platelets: A connection to hematogenous metastasis. *J. Biol. Chem.* **2012**, *287*, 2168–2178. [CrossRef]
48. Uppal, A.; Wightman, S.C.; Ganai, S.; Weichselbaum, R.R.; An, G. Investigation of the essential role of platelet-tumor cell interactions in metastasis progression using an agent-based model. *Theor. Biol. Med. Model.* **2014**, *11*, 17. [CrossRef]
49. Cohen, S.J.; Alpaugh, R.K.; Gross, S.; O'Hara, S.M.; Smirnov, D.A.; Terstappen, L.W.M.M.; Allard, W.J.; Bilbee, M.; Cheng, J.D.; Hoffman, J.P.; et al. Isolation and characterization of circulating tumor cells in patients with metastatic colorectal cancer. *Clin. Colorectal Cancer* **2006**, *6*, 125–132. [CrossRef]
50. Spurr, N.K.; Durbin, H.; Sheer, D.; Parkar, M.; Bobrow, L.; Bodmer, W.F. Characterization and chromosomal assignment of a human cell surface antigen defined by the monoclonal antibody AUAI. *Int. J. Cancer* **1986**, *38*, 631–636. [CrossRef]
51. Datar, R.H.; Zheng, A.; Cote, R.J. Significance of Studying Circulating Tumor Cells. In *Circulating Tumor Cells*; Springer: New York, NY, USA, 2016; p. 8.
52. Krebs, M.G.; Metcalf, R.L.; Carter, L.; Brady, G.; Blackhall, F.H.; Dive, C. Molecular analysis of circulating tumour cells—biology and biomarkers. *Nat. Rev. Clin. Oncol.* **2014**, *11*, 129–144. [CrossRef] [PubMed]
53. Rao, C.G.; Chianese, D.; Doyle, G.V.; Miller, M.C.; Russell, T.; Sanders, R.A.; Terstappen, L.W.M.M. Expression of epithelial cell adhesion molecule in carcinoma cells present in blood and primary and metastatic tumors. *Int. J. Oncol.* **2005**, *27*, 49–57. [CrossRef] [PubMed]
54. Steinert, G.; Schölch, S.; Niemietz, T.; Iwata, N.; García, S.A.; Behrens, B.; Voigt, A.; Kloor, M.; Benner, A.; Bork, U.; et al. Immune escape and survival mechanisms in circulating tumor cells of colorectal cancer. *Cancer Res.* **2014**, *74*, 1694–1704. [CrossRef] [PubMed]
55. Zhu, L.; Hissa, B.; Győrffy, B.; Jann, J.-C.; Yang, C.; Reissfelder, C.; Schölch, S. Characterization of Stem-like Circulating Tumor Cells in Pancreatic Cancer. *Diagnostics* **2020**, *10*, 305. [CrossRef]
56. Nicolazzo, C.; Gradilone, A.; Loreni, F.; Raimondi, C.; Gazzaniga, P. EpCAMlow Circulating Tumor Cells: Gold in the Waste. *Dis. Markers* **2019**, *2019*, 1718920. [CrossRef]
57. Schölch, S.; García, S.A.; Iwata, N.; Niemietz, T.; Betzler, A.M.; Nanduri1, L.K.; Bork, U.; Kahlert, C.; Thepkaysone, M.-L.; Swiersy, A.; et al. Circulating tumor cells exhibit stem cell characteristics in an orthotopic mouse model of colorectal cancer. *Oncotarget* **2016**, *7*, 27232–27242. [CrossRef]
58. Zhang, L.; Ridgway, L.D.; Wetzel, M.D.; Ngo, J.; Yin, W.; Kumar, D.; Goodman, J.C.; Groves, M.D.; Marchetti, D. The identification and characterization of breast cancer CTCs competent for brain metastasis. *Sci. Transl. Med.* **2013**, *5*, 180ra48. [CrossRef]
59. Cancer Genome Atlas Research Network Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* **2017**, *32*, 185–203.e13. [CrossRef]
60. Yu, M.; Ting, D.T.; Stott, S.L.; Wittner, B.S.; Ozsolak, F.; Paul, S.; Ciciliano, J.C.; Smas, M.E.; Winokur, D.; Gilman, A.J.; et al. RNA sequencing of pancreatic circulating tumour cells implicates WNT signaling in metastasis. *Nature* **2012**, *487*, 510–513. [CrossRef]
61. Franses, J.W.; Philipp, J.; Missios, P.; Bhan, I.; Liu, A.; Yashaswini, C.; Tai, E.; Zhu, H.; Ligorio, M.; Nicholson, B.; et al. Pancreatic circulating tumor cell profiling identifies LIN28B as a metastasis driver and drug target. *Nat. Commun.* **2020**, *11*, 3303. [CrossRef]
62. Stuart, T.; Butler, A.; Hoffman, P.; Hafemeister, C.; Papalexi, E.; Mauck, W.M.; Hao, Y.; Stoeckius, M.; Smibert, P.; Satija, R. Comprehensive Integration of Single-Cell Data. *Cell* **2019**, *177*, 1888–1902.e21. [CrossRef] [PubMed]
63. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **2019**, *28*, 1947–1951. [CrossRef] [PubMed]
64. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25–29. [CrossRef] [PubMed]

65. Raudvere, U.; Kolberg, L.; Kuzmin, I.; Arak, T.; Adler, P.; Peterson, H.; Vilo, J. g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **2019**, *47*, W191–W198. [CrossRef]

2.2 The adherens junction pathway in pancreatic CTC with stem features

Dissertation relevant publication:

Zhu L, Hissa B, Győrffy B, Jann JC, Yang C, Reissfelder C, Schölch S. Characterization of Stem-like Circulating Tumor Cells in Pancreatic Cancer. **Diagnostics (Basel)**. 2020 May 15;10(5):305. doi: 10.3390/diagnostics10050305 ([142])

Summary: The objective of this study was to characterize pancreatic CTCs with stem-like features.

A pool of phenotypic markers was utilized to characterize the murine pancreatic CTC dataset GSE51372.[142] The marker panel including EMT-transcription factors, epithelial, mesenchymal, stem-like/pluripotency, and proliferation markers.

The principal component analysis (PCA) of 72 pancreatic CTCs displayed a heterogeneous distribution. Depending on cells with (CTC-S) or without stem-like features (CTC-N), CTCs were divided into two groups. The following heatmap with hierarchical clustering analysis supported this classification.

We found hybrid EMT markers expression in the CTC-S group. Not surprisingly, severe heterogeneity was found for most markers. Surprisingly, the commonly used pancreatic stem cell marker *Cd44* was significantly reduced in the CTC-S group while other stem markers (*Klf4*, *Aldh1a2*) and mesenchymal markers (*Fn1*, *Vim*) were significantly over-expressed.

Both over-representation analysis (ORA) and weighted gene co-expression network analysis (WGCNA) revealed that the adherens junction pathway is significantly enriched in CTC-S. *Ctnnb1* is the vital consensus component of the adherens junction pathway in integrated bioinformatics strategies. Further correlation analysis between *Ctnnb1* and stem cell markers indicated a significant positive relationship with *Klf4*. The prognostic value of *CTNNB1* and *KLF4* was examined in the TCGA PDAC cohort. Higher *CTNNB1* predicts worse progress-free survival and relapse-free survival in this cohort.

diagnostics

MDPI

Article

Characterization of Stem-like Circulating Tumor Cells in Pancreatic Cancer

Lei Zhu [1], Barbara Hissa [1], Balázs Győrffy [2,3], Johann-Christoph Jann [4], Cui Yang [1], Christoph Reissfelder [1,5] and Sebastian Schölch [1,5,*]

1 Department of Surgery, Universitätsmedizin Mannheim, Medical Faculty Mannheim, Heidelberg University, 68167 Mannheim, Germany; Lei.Zhu@medma.uni-heidelberg.de (L.Z.); Barbara.Hissa@medma.uni-heidelberg.de (B.H.); Cui.Yang@umm.de (C.Y.); christoph.reissfelder@umm.de (C.R.)
2 2nd Department of Pediatrics, Semmelweis University, H-1094 Budapest, Hungary; gyorffy.balazs@yahoo.com
3 TTK Cancer Biomarker Research Group, Institute of Enzymology, H-1117 Budapest, Hungary
4 Department of Medicine III, Universitätsmedizin Mannheim, Medical Faculty Mannheim, Heidelberg University, 68167 Mannheim, Germany; Johann-Christoph.Jann@medma.uni-heidelberg.de
5 German Cancer Consortium (DKTK) & German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany
* Correspondence: sebastian.schoelch@medma.uni-heidelberg.de; Tel.: +49-621-383-5152

Received: 17 February 2020; Accepted: 13 May 2020; Published: 15 May 2020

Abstract: Pancreatic ductal adenocarcinoma (PDAC) is the fourth most frequent cause of death from cancer. Circulating tumor cells (CTCs) with stem-like characteristics lead to distant metastases and thus contribute to the dismal prognosis of PDAC. Our purpose is to investigate the role of stemness in CTCs derived from a genetically engineered mouse model of PDAC and to further explore the potential molecular mechanisms. The publically available RNA sequencing dataset GSE51372 was analyzed, and CTCs with (CTC-S) or without (CTC-N) stem-like features were discriminated based on a principal component analysis (PCA). Differentially expressed genes, weighted gene co-expression network analysis (WGCNA), and further functional enrichment analyses were performed. The prognostic role of the candidate gene (*CTNNB1*) was assessed in a clinical PDAC patient cohort. Overexpression of the pluripotency marker *Klf4* (Krüppel-like factor 4) in CTC-S cells positively correlates with *Ctnnb1* (β-Catenin) expression, and their interaction presumably happens via protein–protein binding in the nucleus. As a result, the adherens junction pathway is significantly enriched in CTC-S. Furthermore, the overexpression of *Ctnnb1* is a negative prognostic factor for progression-free survival (PFS) and relapse-free survival (RFS) in human PDAC cohort. Overexpression of *Ctnnb1* may thus promote the metastatic capabilities of CTCs with stem-like properties via adherens junctions in murine PDAC.

Keywords: pancreatic cancer; PDAC; circulating tumor cells; CTC; stem cells; stem-like; stemness; adherens junctions; epithelial-mesenchymal transition; EMT

1. Introduction

With 57,000 new cases and 46,000 deaths annually, pancreatic ductal adenocarcinoma (PDAC) is the 9th most frequent malignant disease in males and 8th most prevalent in females and the 4th most frequent cause of death in both genders [1]. At the time of diagnosis, only a fraction of patients are amenable to surgical resection of the tumor. However, even patients who undergo complete surgical resection are at a high risk of either local or systemic recurrence [2].

In order to study the molecular mechanisms of pathogenesis and metastasis of pancreatic cancer, several genetically induced mouse models (GEMMs) have been established in recent decades [3–6].

The most prominent and best-characterized model is the LSL-*Kras*^{G12D}, LSL-*Trp53*^{R172H}, *Pdx1-Cre* (KPC) mouse model [6].

It is now well established that circulating tumor cells (CTCs) are an integral part of the metastatic cascade in malignant disease [7–18]. CTCs are shed from the primary tumor, survive in circulation, and ultimately colonize distant organs where they establish clinically overt metastases [9,16]. However, while thousands or millions of CTCs are shed into the bloodstream over time, the number of metastases is several orders of magnitude lower [19]. Therefore, it can be assumed that only a small fraction of CTCs are actually tumorigenic and thus clinically relevant. In order to achieve the invasive phenotype required to leave the primary tumor bulk and enter circulation, CTCs from epithelial tumors undergo a process called epithelial–mesenchymal transition (EMT). During this process, the cells acquire mesenchymal properties (e.g., migratory capability) while downregulating epithelial traits [20,21]. This process is reverted during mesenchymal-epithelial transition (MET). As both EMT and MET dynamically fluctuate, CTCs in circulation exhibit a high degree of plasticity and represent a heterogeneous population consisting of epithelial, mesenchymal, and intermediary cells [20,22].

In a previous study, CTCs from the GEMM (LSL-*Kras*^{G12D}, LSL-*Trp53*$^{flox/flox\ or\ flox/+}$, *Pdx1-Cre*) were isolated and submitted to single-cell RNA sequencing. The raw data of this experiment was released to the Gene Expression Omnibus (GEO) database, and it is available for analysis [23]. In previous studies from our own group, we were able to demonstrate that, apart from immune evasive capabilities, CTCs also possess stem cell properties [13,17]. The main goal of this study was therefore to investigate the role of stem cell properties in murine CTCs of PDAC and to further characterize CTCs with stem cell properties.

2. Results

Details regarding the study workflow are depicted in Figure 1.

2.1. Pancreatic CTCs Can Be Clustered into Stem-Like and Non-Stem like Categories

After pre-processing, we kept 72 samples for the following study, and 11,931 genes were obtained (Figure S1A,B). In order to classify CTCs into subgroups, the top 3000 genes that presented the highest variability in expression were chosen for the principal component analysis (PCA) [19,24,25]. From these 3000 genes, epithelial (Epcam, Krt7, Krt8, Krt18, and Krt19) [23], mesenchymal (Fn1, S100a4 and Vim) [26], stemness/pluripotency markers (Aldh1a1, Aldh1a2, Cd24a, Cd44, and Klf4) [25,27,28] and the proliferation marker Mki67 [29] were included in subsequent analyses.

All 72 individual CTCs could be divided into three clusters (Figure S2A, B). However, this study focuses on whether the CTCs have stem-like features or not, a dichotomous classification; therefore, the CTCs were divided into clusters based on stem-like features. According to the PCA loading plot (Figure 2A), the cluster located in the first quadrant (blue dots) exhibited stemness markers and was defined as CTC-S (CTCs with stem-like features) (Figure 2B). The other two clusters were merged (red dots) and named CTC-N (CTC without stemness features. The correlation heatmap with hierarchical clustering showed a similar result (Figure 2C).

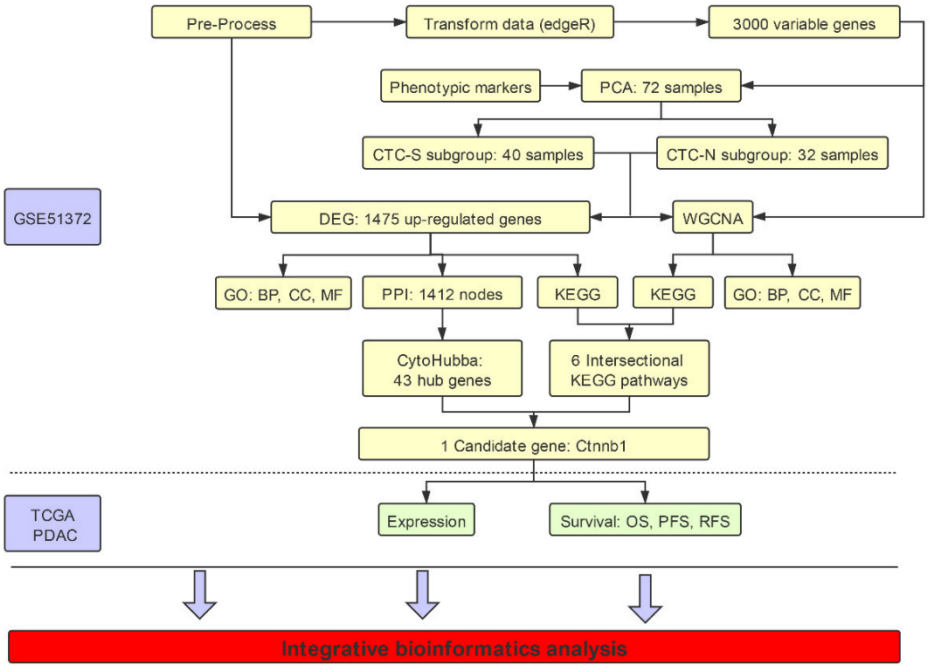

Figure 1. Illustration of the workflow in the integrative bioinformatics analysis. Data set GSE51372 from the Gene Expression Omnibus (GEO) database was used, which contains 75 murine pancreatic CTCs single-cell sequencing datasets. After filtering out three samples with low quality, we further excluded low expression genes, and EdgR transformed the remaining 72 samples. CTCs with stem-like features (CTC-S) and CTC with non-stem-like features (CTC-N) were defined based on the principal component analysis (PCA) of the top 3000 most variable genes. We performed weighted gene co-expression network analysis (WGCNA) on these 3000 genes, while all read counts of 11,931 genes were used to analyze the differentially expressed genes. Further functional enrichment analysis, including the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, Gene Ontology (GO) aspects on both differentially expressed genes and WGCNA results was performed. The cytoHubba plugin of Cytoscape 3.5.1 was employed to identify hub genes. Candidate pathways (adherens junction) and genes (*Ctnnb1*) were identified. The clinical value of *KLF4* and *CTNNB1* was validated in the TCGA (The Cancer Genome Atlas) pancreatic adenocarcinoma (PDAC) cohort. The workflow was plotted using ProcessOn (available online: https://www.processon.com/diagrams, accessed on 23 March 2020).

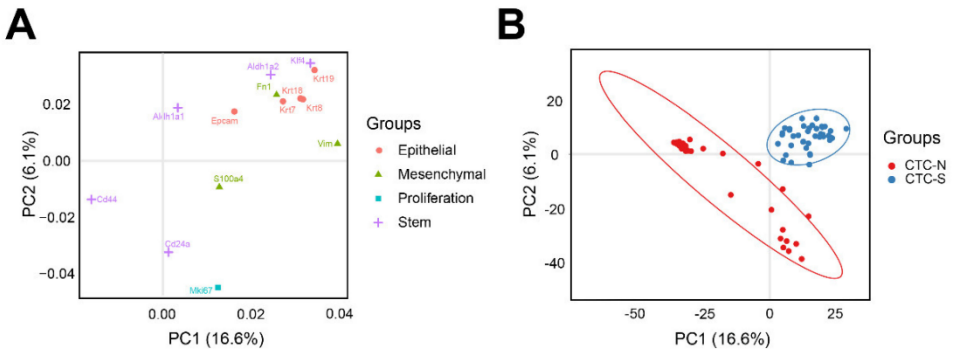

Figure 2. *Cont.*

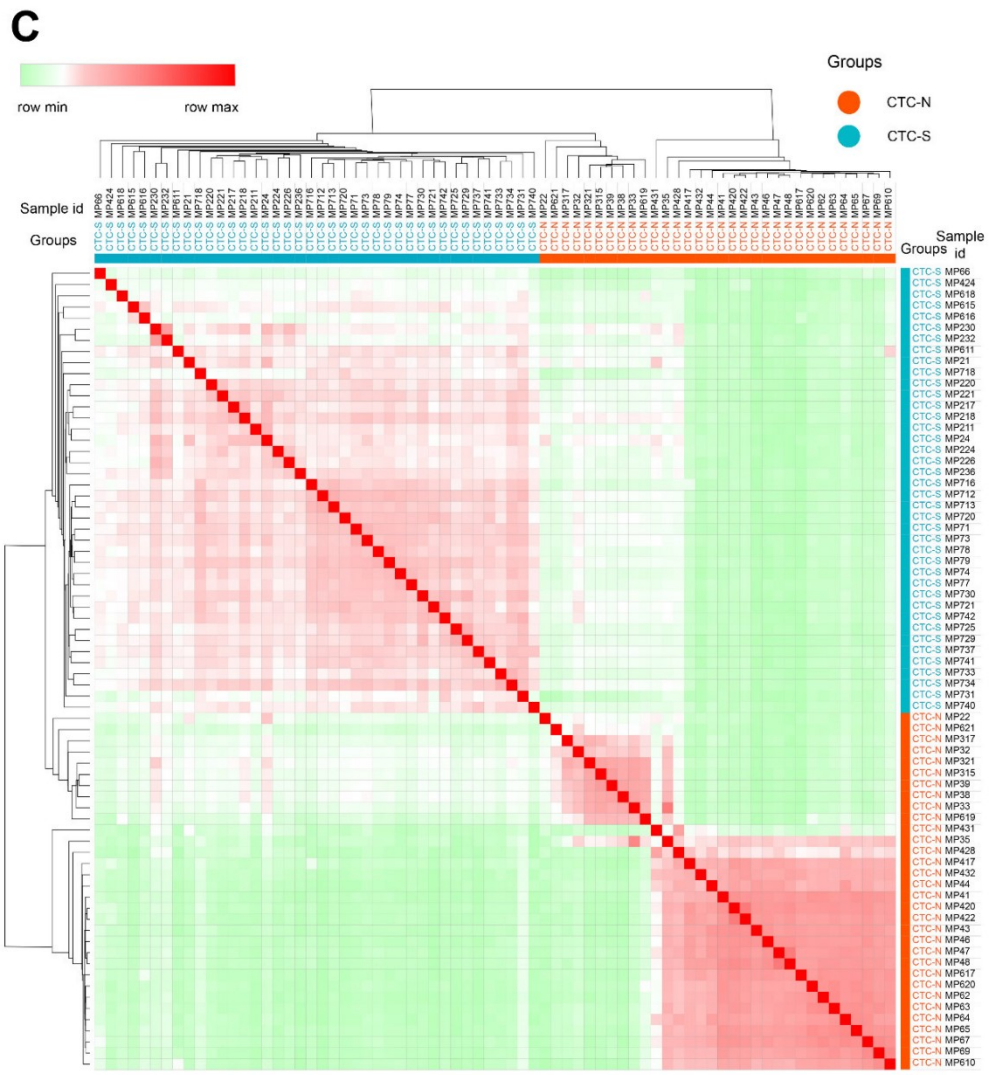

Figure 2. Defining the CTC-S and CTC-N subgroups. (**A**) Principal component analysis (PCA) loading plot of the marker set. The *x* and *y* axes represent the principal components 1 (PC1) = 16.6% variance and PC2 = 6.1% variance, respectively. 3 of the 5 stem markers (Aldh1a1, Aldh1a2, and Klf4) are located in the first quadrant, indicating that stem markers tend to correlate with both PC1 and PC2 positively. The PCA loading data were download from the ClustVis and visualized by the imageGP (available online: http://www.ehbio.com/ImageGP/index.php/Home/Index/index.html,accessed on 23 April 2020). (**B**) PCA scores plot of 72 samples. All samples were divided into three clusters. The cluster located in the first quadrant was defined as CTCs with stem-like features (CTC-S) since they present with stem markers PCA loadings in (A), as the other two clusters were combined and defined as CTC with non-stem-like features (CTC-N). The corresponding ellipses were plotted based on a 95% probability from the same group. (**C**) The correlation heatmap was visualized by MORPHEUS (available online: https://software.broadinstitute.org/morpheus/, accessed on 7 March 2020). We chose the average linkage method to perform the hierarchical clustering. The heatmap demonstrates the distinct subgroups of CTC-S and CTC-N. The colors of the square matrices illustrate the Pearson's correlation coefficient, with red indicating a strong correlation and green a weak correlation. All samples are listed in the same order in both horizontal and vertical axes.

2.2. Both Epithelial and Mesenchymal Markers are Expressed on CTC-S

To evaluate each cell within their respective subgroups in more detail, we expanded the marker pool, as not all markers of interest were listed in the 3000 most variable genes. EMT markers (*Snai1*, *Snai2*, *Twist1*, *Zeb1*, and *Zeb2*), epithelial markers (*Cdh1*, *Egfr*, *Epcam*, *Krt7*, *Krt8*, *Krt18*, and *Krt19*) [23], mesenchymal markers (*Cdh2*, *Fn1*, *Itga5*, *Sdc1*, *S100a4*, and *Vim*) [20,23], pancreatic cancer stemness markers (*Abcg2*, *Aldh1a1*, *Aldh1a2*, *Cd24a*, *Cd44*, *Cxcr4*, *Met*, *Prom1*, *Klf4*, *Nanog* and *Sox2*) [30,31], and a proliferation marker (*Mki67*) were included to plot the heatmap (Figure 3A).

A high degree of expression heterogeneity was observed for most markers. Compared with CTC-N, CTC-S cells demonstrated an increased expression of epithelial markers such as *Egfr*, *Epcam*, and *keratins*. Parallelly, the mesenchymal markers *Fn1* [$\log_2$(fold change) (FC) = 4.07, false discovery rate (FDR) = 1.42×10^{-5}] and *Vim* ($\log_2$FC = 6.37, FDR = 5.89×10^{-10}) were up-regulated in the CTC-S group, resulting in a phenotype with both epithelial and mesenchymal characteristics. Notably, stemness transcripts showed variable expression, while *Aldh1a2* ($\log_2$FC = 4.98, FDR = 1.61×10^{-8}) and *Klf4* ($\log_2$FC = 9.41, FDR = 1.45×10^{-21}) were enriched in CTC-S, *Cd44* was significantly down-regulated ($\log_2$FC = −2.57, FDR = 0.012) (Figure 3A).

2.3. The Adherens Junction Pathway is Functionally Enriched in CTC-S

Compared with CTC-N, 1475 genes were significantly up- and 247 significantly down-regulated in the CTC-S group (Figure S3A, B).

We next constructed a protein–protein interaction (PPI) network of up-regulated genes by STRING (available online: https://string-db.org/, accessed on 19 August 2019) [32], establishing 1412 nodes and 6695 interactions (Figure not shown). We used the cytoHubba plugin in Cytoscape to narrow this number down to 43 genes, which were subsequently defined as hub genes.

The biological domains of biological process (BP), cellular component (CC), and molecular function (MF) were utilized to describe the biological domain of genes enriched in the CTC-S group. The cellular components' anchoring junction [Gene Ontology (GO): 0070161, fold enriched = 3.56, FDR = 1.55×10^{-40}] and adherens junction (GO: 0005912, fold enriched = 3.53, FDR = 9.73×10^{-39}) were significantly functionally enriched. The enrichment of cell adhesion molecule binding (GO: 0050839, fold enriched = 3.34, FDR = 6.48×10^{-31}) was observed in the MF domain (Figure 3B).

While GO enrichment analysis focuses on the functions of the genes, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis mainly reflects the interaction network and the signaling pathways of the requested genes. In this analysis, several signaling pathways were significantly enriched, including adherens' junction (KEGG: 04520, fold enriched = 3.42, FDR = 1.47×10^{-5}) and fluid shear stress and atherosclerosis pathways (KEGG: 05418, fold enriched = 2.61, FDR = 2.05×10^{-5}) (Figure 3B).

In order to gain deeper insight into the molecular biology of the CTC subsets, we used a weighted gene co-expression network analysis (WGCNA) in this study.

After network construction (Figure S4A, B), we identified five gene co-expression modules (blue, brown, green, turquoise, and yellow) with sizes between 67 and 1384 genes in the 72 samples (Figure 4A). Both brown and turquoise modules correlated with stemness (correlation coefficient > 0.7, adjusted p-value <0.001), but the results of the brown module were more significant (correlation coefficient = 0.95, adjusted p-value = 3.6×10^{-37}).

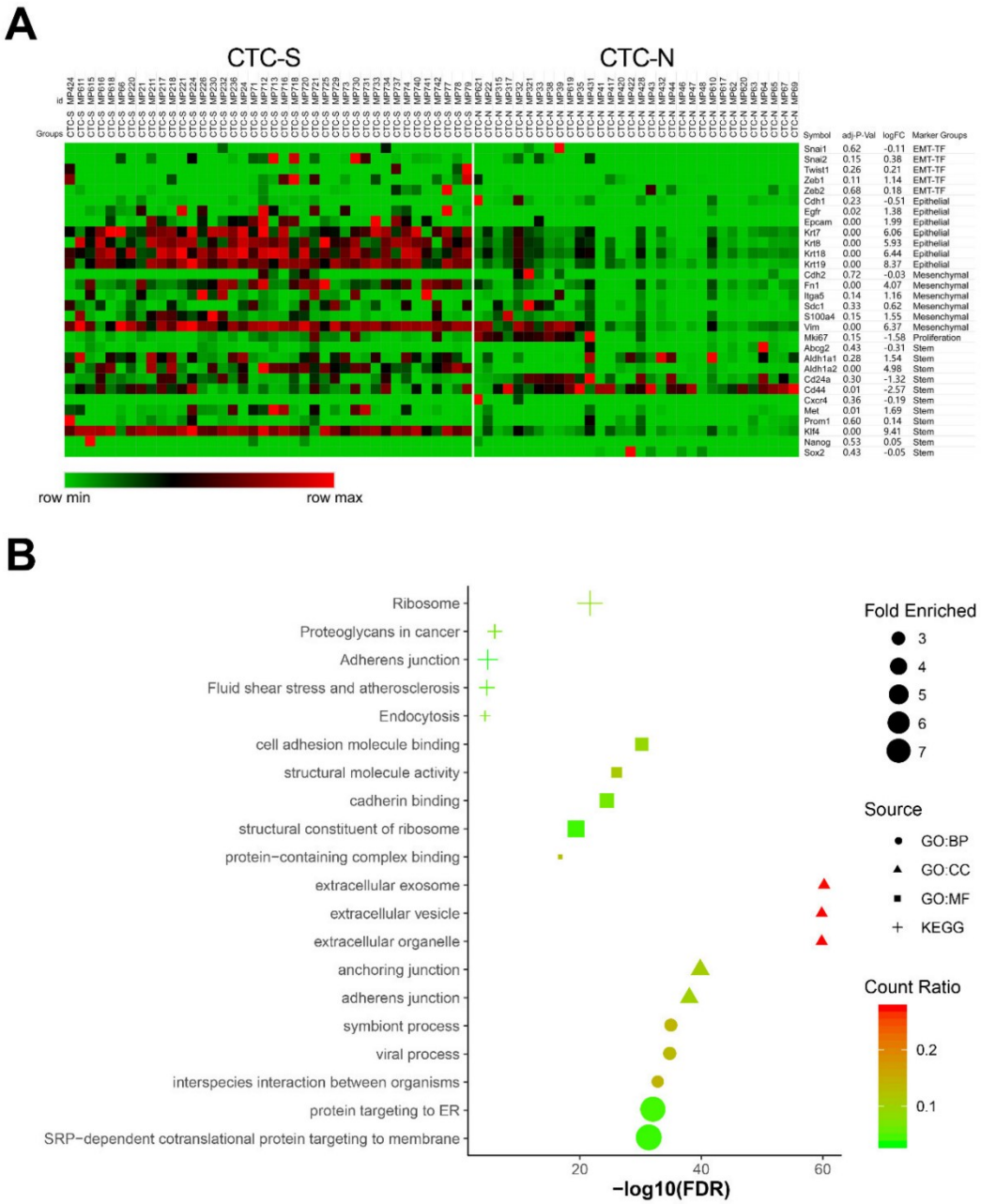

Figure 3. Differentially expressed genes between CTC-S and CTC-N groups. (**A**) MORPHEUS-generated heatmap showing the expression of the marker set in CTC-S and CTC-N groups. The maximum and the minimum values in each row are displayed red and green, respectively. The marker names, groups, $\log_2$(Fold Change (FC)) values, and false discovery rate (FDR) values were listed on the right side. (**B**) The functional enrichment analysis of up-regulated genes in the CTC-S subgroup was visualized by imageGP (available online: http://www.ehbio.com/ImageGP/index.php/Home/Index/index.html, 23 April 2020), top 5 enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, Gene Ontology (GO) terms for biologic processes (BP), cell components (CC) and molecular functions (MF), each aspect was shown (fold enriched >2 and FDR <0.010).

We therefore used the brown module in the functional enrichment analysis. In the brown module, 645 genes were functionally enriched in the adherens junction pathway (KEGG: 04520, fold enriched = 3.71, FDR = 3.59×10^{-3}), which is consistent with the results of the analysis of the differentially expressed genes. In addition, anchoring junction (GO: 0070161, fold enriched = 3.27, FDR = 2.22×10^{-13}), adherens junction organization (GO: 0034332, fold enriched = 4.29, FDR = 8.11×10^{-3}) and regulation of adherens junction organization (GO: 1903391, fold enriched = 13.11, FDR = 8.12×10^{-3}) were all significantly enriched again (Figure 4B).

To identify the pathways with the strongest influence on the CTC-S phenotype, we intersected functional enrichment pathways from the results of differentially expressed genes and WGCNA. This resulted in six pathways, including focal adhesion (KEGG: 04510), adherens junction (KEGG: 04520), ECM-receptor interaction (KEGG: 04512), AGE-RAGE signaling pathway in diabetic complications (KEGG: 04933), proteoglycans in cancer (KEGG: 05205) and fluid shear stress and atherosclerosis (KEGG: 05418) (Figure 5A). Since adherens junction organization (GO: 0034332) was also significantly enriched in the GO: BP aspect of functional enrichment analysis of both differentially expressed genes and WGCNA, we identified the adherens junction as a pathway of interest in CTC-S. In addition, double enriched GO terms, such as the regulation of adherens junction organization (GO: 1903391), cell–cell junction (GO: 0005911), and cell–cell junction assembly (GO: 0007043) provide further evidence that adherens junctions are biologically relevant in the CTC-S group.

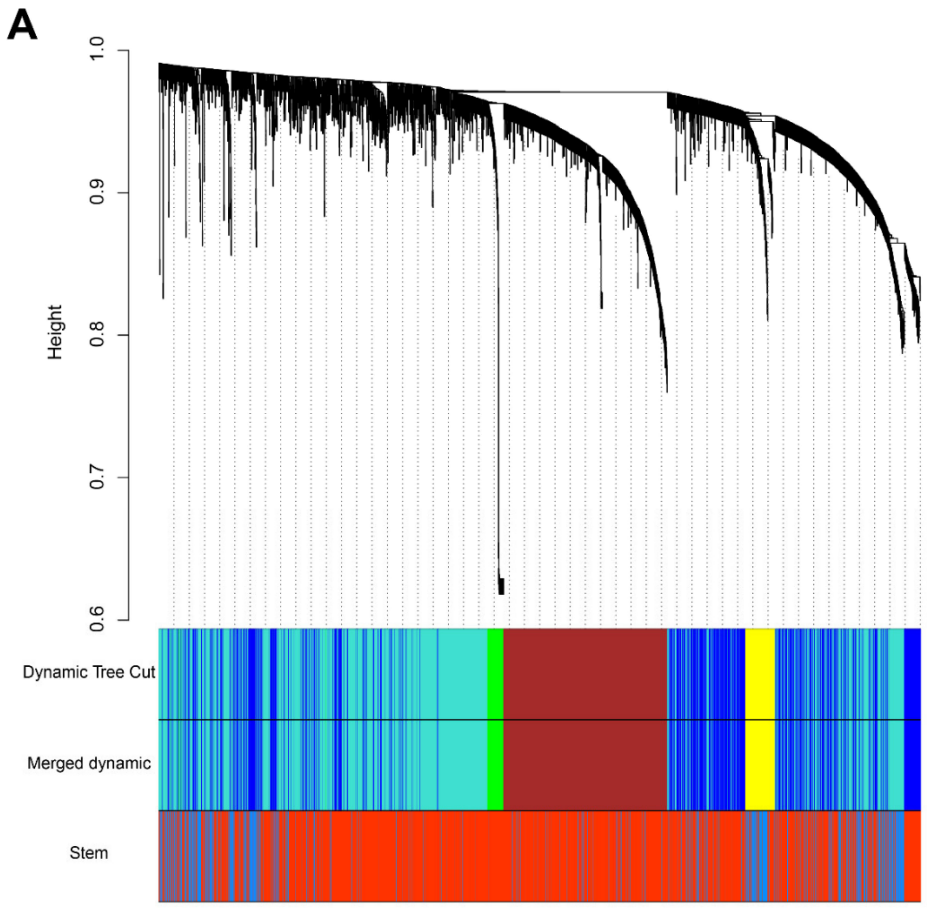

Figure 4. *Cont.*

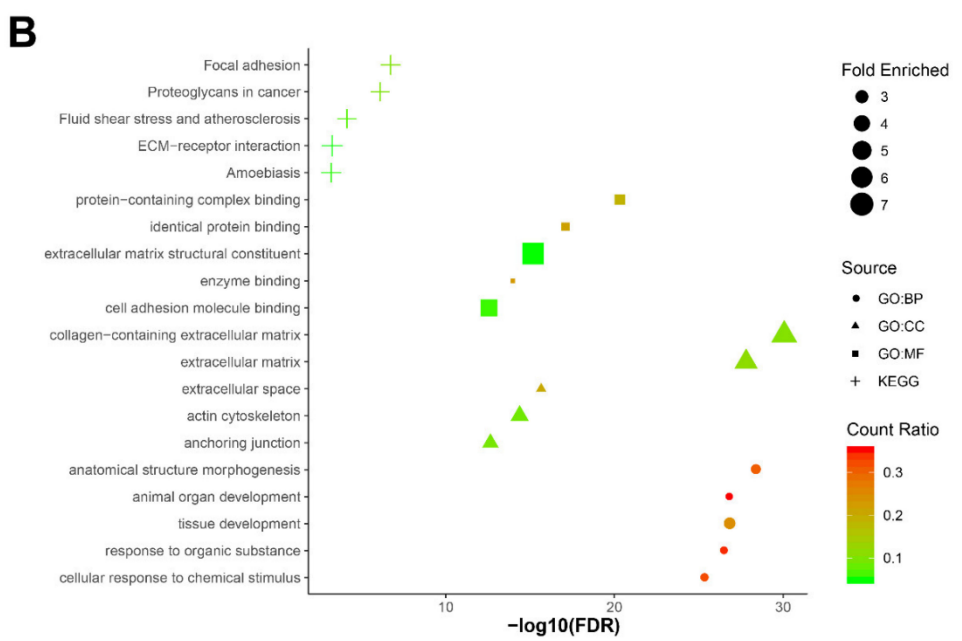

Figure 4. Weighted gene co-expression network analysis (WGCNA) of 72 pancreatic CTCs. (**A**) The cluster dendrogram of 3000 genes, five modules (blue, brown, green, turquoise, and yellow), were identified. The brown module correlates with the stem trait most significantly (r = 0.95, adjusted p-value = 3.6×10^{-3}). (**B**) The functional enrichment analysis of genes in the brown module (generated by imageGP, available online: http://www.ehbio.com/ImageGP/index.php/Home/Index/index.html, accessed on 23 April 2020), top 5 enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, Gene Ontology (GO) terms for biological processes (BP), cell components (CC) and molecular functions (MF) are shown [fold enriched > 2 and false discovery rate (FDR) <0.010].

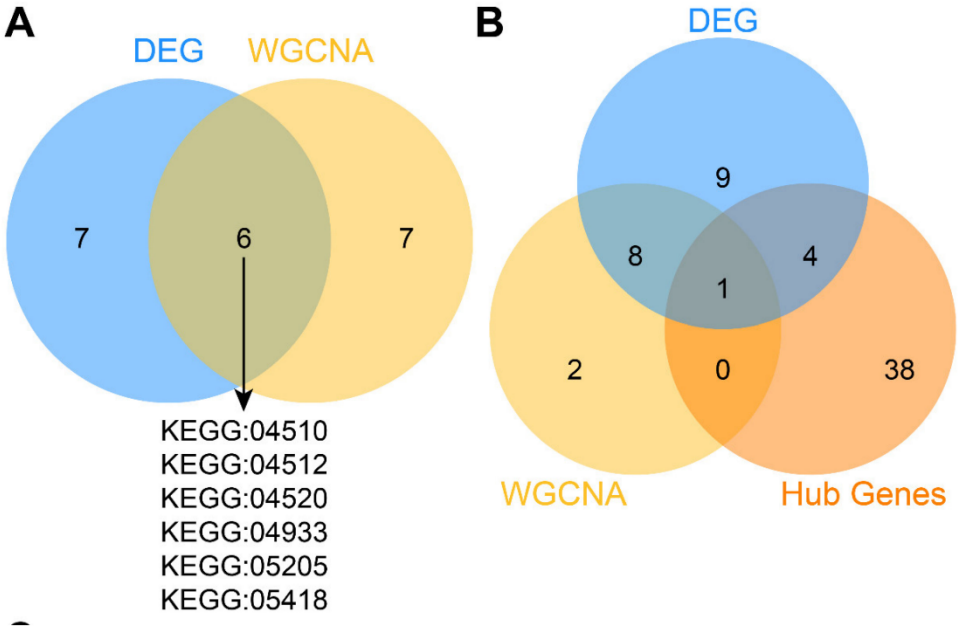

Figure 5. *Cont.*

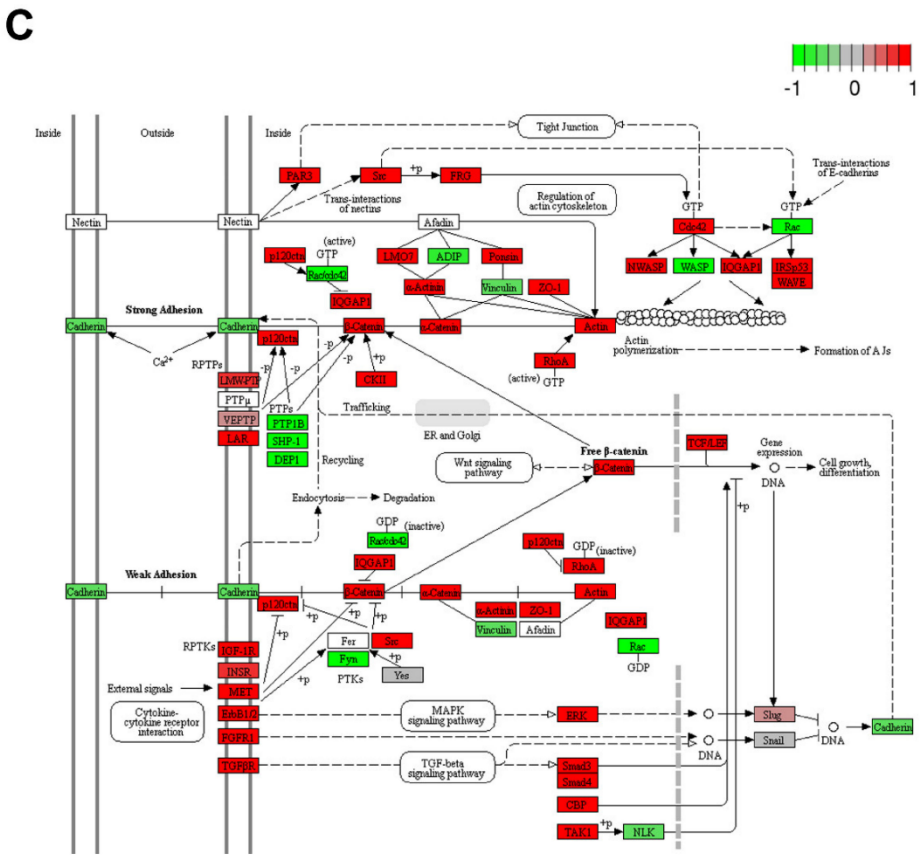

Figure 5. The adherens junctions pathway is significantly enriched in the CTC-S group. (**A**) Venn plot of enriched pathways in differentially expressed genes and weighted gene co-expression network analysis (WGCNA). Six pathways included focal adhesion (KEGG: 04510), ECM-receptor interaction (KEGG: 04512), adherens junction (KEGG: 04520), AGE-RAGE signaling pathway in diabetic complications (KEGG: 04933), proteoglycans in cancer (KEGG: 05205) and fluid shear stress and atherosclerosis (KEGG: 05418) were significantly enriched in both differentially expressed genes and WGCNA. (**B**) The intersection of enriched genes in the adherens junction pathway of differentially expressed genes, enriched genes in the adherens junction pathway of the brown module in WGCNA, and hub genes. *Ctnnb1* is at the intersection of all three lists. (**C**) Gene expression in the adherens junction pathway was plotted by PATHVIEW [33]. Red represents up-regulated genes, while green represents down-regulated genes in the CTC-S group. The Venn plots in this study were generated by imageGP (available online: http://www.ehbio.com/ImageGP/index.php/Home/Index/index.html, accessed on 23 April 2020).

In order to identify the key genes in the adherens junction pathway and the entire PPI network., hub genes of up-regulated genes in the PPI, genes enriched in the adherens junction pathway of differentially expressed genes and the WGCNA were intersected. The gene *Ctnnb1* ($\log_2$FC = 3.24, FDR = 2.40×10^{-4}) was found significant in all three analyses (Figure 5B). The adherens junction pathway map and gene expression heatmap are plotted in Figure 5C and Figure S5.

2.4. Ctnnb1 and Klf4 Positively Correlate

Next, we evaluated the correlation between *Ctnnb1* and stemness/pluripotency markers (*Aldh1a2*, *Met*, and *Klf4*), which were significantly up-regulated in the CTC-S group (Figure 6). Among these, we found that *Ctnnb1* overexpression most significantly correlates with *Klf4* upregulation (r = 0.50, FDR = 4.99×10^{-5}).

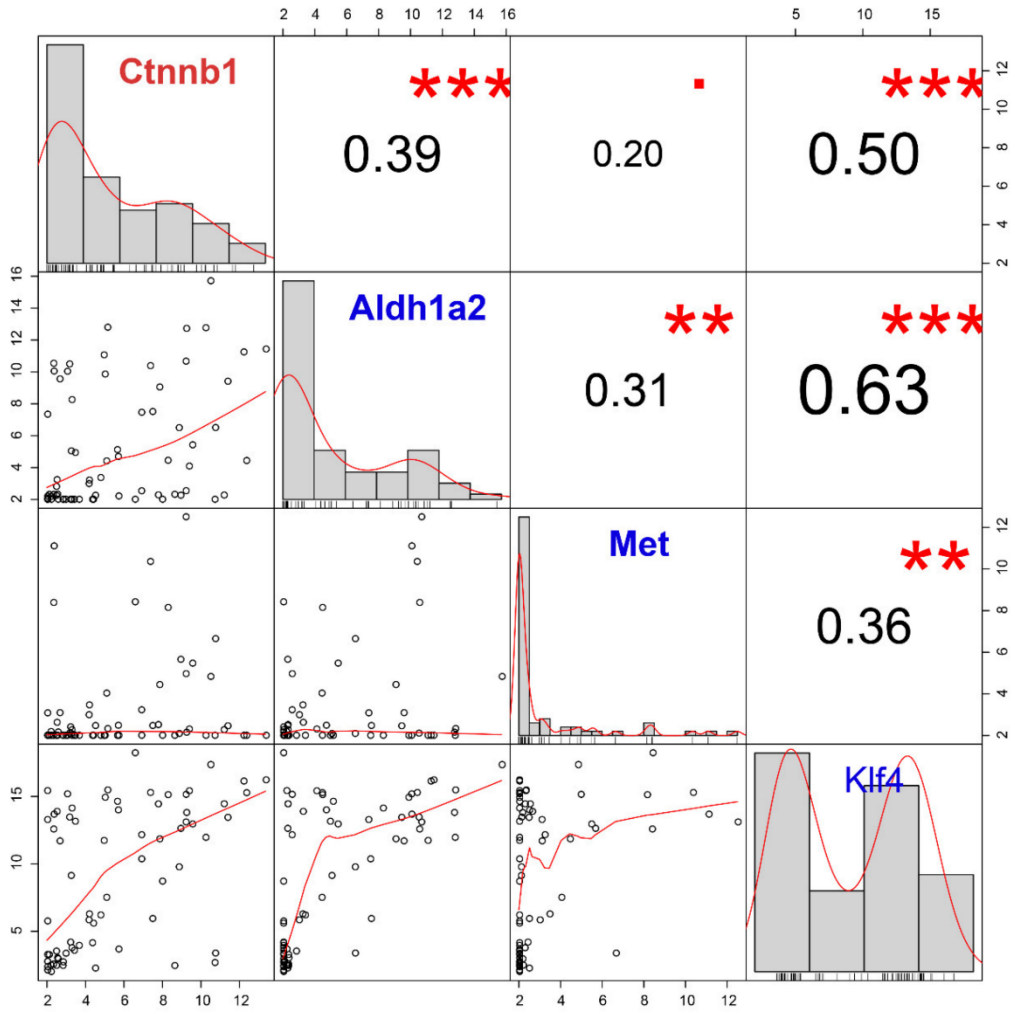

Figure 6. The correlation analysis of the candidate gene (*Ctnnb1*) and stem markers (*Aldh1a2*, *Met* and *Klf4*) demonstrates the positive correlation between *Klf4* and *Ctnnb1*. Gene expression histograms are represented on the diagonal. The lower left part is the scatter plots and the numbers in the upper right part correspond to the Pearson's correlation coefficient. *** FDR < 0.001; **, 0.001 < FDR < 0.010;. 0.050 < FDR < 0.100.

2.5. CTNNB1 Is a Negative Prognostic Factor in Human PDAC

As reported above, *Klf4* and *Ctnnb1* were significantly up-regulated in murine pancreatic CTCs with stem-like features. To explore whether *KLF4* and *CTNNB1* are differentially expressed between pancreatic tumor and normal tissues, the visualization web tool Gene Expression Profiling Interactive Analysis 2 (GEPIA 2) (available online: http://gepia2.cancer-pku.cn, accessed on 7 March 2020) was utilized [34]. When evaluating The Cancer Genome Atlas (TCGA) data, both *KLF4* and *CTNNB1* were up-regulated in PDAC compared to the matched normal tissues and the Genotype–Tissue Expression (GTEx) data [35] (Figure 7A,B).

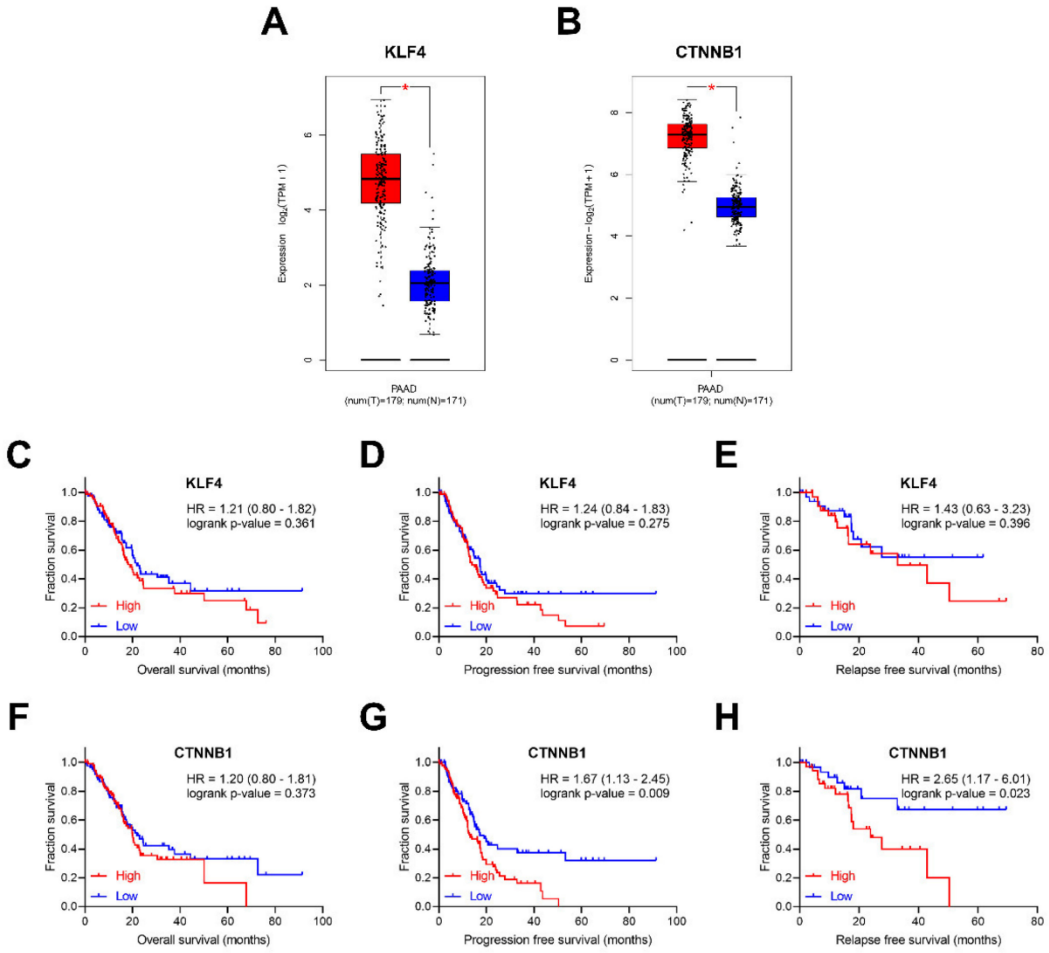

Figure 7. The clinical value of *KLF4* and *CTNNB1*. The box-plots with data overlaid dot plots for *KLF4* (**A**) and *CTNNB1* (**B**), showing expression in TCGA PAAD tumors with correspondent match PAAD normal tissue and GTEx pancreatic tissue. * p-value < 0.001. Clinical prognostic value of *KLF4*, overall survival (OS) (**C**), progression-free survival (PFS) (**D**), and relapse-free survival (RFS) (**E**). Clinical prognostic value of *CTNNB1*, OS (**F**), PFS (**G**) and RFS (**H**). Box plots were generated using GEPIA2 (available online: http://gepia2.cancer-pku.cn/#index, accessed on 7 March 2020).

We further investigated whether candidate genes expression levels correlate with clinical prognosis, including overall survival (OS), progression-free survival (PFS) and relapse-free survival (RFS) (Figure 7C–H) in the TCGA cohort [36]. No significant influence of *KLF4* on survival was observed. In contrast, although *CTNNB1* expression does not influence OS (Hazard Ratio (HR) = 1.20, log-rank p-value = 0.373), patients with higher *CTNNB1* expression have significantly worse PFS (HR = 1.67, log-rank p-value = 0.009) and RFS (HR = 2.65, log-rank p-value = 0.023).

3. Discussion

The aim of this study was to identify and characterize pancreatic CTCs with stem-like characteristics and explore potential underlying mechanisms through integrated bioinformatics analysis.

Numerous molecular markers, including *Cd24*, *Cd44*, *Prom1* (*Cd133*), and other genes, have been applied in order to define pancreatic cancer stem cells (CSCs) [30,37,38]. The CTC subgroup identified in this study as CTC-S demonstrated heterogeneous expression of stemness and pluripotency markers in comparison to CTC-N, with some markers showing increased expression (*Aldh1a2* and *Klf4*),

some with reduced levels (*Abcg2*, *Cd44*, *Cxcr4*, *Nanog* and *Sox2*) and others with no significant change in expression (*Aldh1a1*, *Cd24a*, *Met* and *Prom1*). The mutual exclusivity phenomenon may be due to the heterogeneous nature of stem-like cells and stemness markers [39]. Similar results have been published for breast cancer, where $CD44^{high}$ / $CD24^{low}$ defines a cluster of tumor cells with stem-like features [40]. This mutual exclusivity is confirmed in the here-presented results.

The invasive capabilities (e.g., migration and invasion) of tumor cells are a prerequisite for successful initiation and completion of metastasis. Accumulating evidence suggests that these features can be acquired through epithelial–mesenchymal transition (EMT), which is considered as a hallmark of tumor cell dissemination [31,41]. The CTC-S subpopulation exhibits a mixed epithelial/mesenchymal phenotype, while the CTC-N group exhibits neither epithelial nor mesenchymal markers. This is consistent with previous studies relating EMT to stemness [42,43]. Furthermore, CTCs with partial EMT phenotype or stem-like features may predict unfavorable survival in cancers independently [44].

Epithelial markers (most prominently *EpCAM*) have been widely used as identification and isolation markers of CTCs, and the prognostic value of *EpCAM*-positive CTCs is evident [15,45–48], but previous data have pointed towards a relative down-regulation of epithelial markers on CTCs [13]. The here-presented results draw a more complex image of *EpCAM* expression in CTCs—it can be speculated that *EpCAM* expression on CTCs is a continuum from *EpCAM*-negative to highly *EpCAM*-positive cells. While, in this work, *EpCAM* seems to be relatively overexpressed on cells with presumably higher metastatic potential, more research on the role of *EpCAM* on CTCs is clearly required.

In theory, the loss of cell adhesion molecules during EMT should also coincide with the down-regulation of genes associated with adherens junctions. While both *Cdh1* and *Cdh2* (generally considered a mesenchymal marker) are downregulated in the CTC-S subgroup, most of the genes in the adherens junction pathway are up-regulated in comparison to CTC-N, resulting in a mixed epithelial/mesenchymal phenotype of CTC-S. This may represent a transitory state allowing CTC-S to undergo functional adaptation during the metastatic cascade [49]. Additionally, this overexpression of adherens junctions may enable CTC-S to bind to endothelial cells prior to extravasation.

Intercellular adhesion has been demonstrated to be crucial for successful metastasis of epithelial cancers [50]. Among the different types of intercellular adhesions, cell–cell adherens junctions are the most common and contribute to cell polarity, tissue architecture maintenance, cell movement limitation, and proliferation [50]. Strong adherens junctions are mediated by the cadherin–catenin complex [51–53]. The here-presented data show that the CTC-S cells express low levels of E-cadherin (*Cdh1*), while the other genes from the adherens junction pathway were up-regulated, such as *Rhoa* and *Cdc42*, both of which are necessary for adherens junction maintenance [54,55]. We hypothesized that CTC-S cells are in the process of forming the adherens junction through actin cytoskeleton remodeling [56].

Besides, another critical component of adherens junctions, plakoglobin (also known as γ-catenin), was found to be necessary for the formation of CTC clusters and further contributed to the metastatic cascade [57,58]. In summary, adherens junctions are critical to the formation of CTC clusters, which can form tumor-microemboli, ultimately outgrowing to overt metastases [57]. These cellular aggregates have been detected in 81% of PDAC patients with unfavorable OS and PFS [59]. However, it must be taken into account that all 75 pancreatic CTCs analyzed in this study were isolated individually, which means there are no data derived from CTC clusters in this dataset; the validity of the here-presented data concerning CTC clusters is therefore limited.

Numerous studies have explored the *Wnt* signaling pathway, mainly due to the fact that canonical *Wnt* signaling regulates, stabilizes, and promotes the accumulation of the β-catenin by inhibiting its degradation [60,61]. A few studies have demonstrated that *Wnt*/β-catenin signaling plays a crucial role in the plasticity of stem cells [62,63]. Furthermore, *Wnt* signaling is activated in pancreatic CTCs according to the original study of GSE51372 [64]. Unexpectedly, we found no significant correlation between *Wnt* family members and *Ctnnb1* in the pancreatic CTC-S group. In fact, according to our results, these *Wnt* family members are even expressed at low levels (Figure S6). Given that *Klf4* is overexpressed in the CTC-S group, we believe that *Klf4* inhibits *Wnt* signaling in this context [65].

Besides, the β-catenin has been determined to be necessary for cell adhesion and pluripotency, even without Wnt signaling [66].

Klf4 (Krüppel-Like Factor 4) is enriched in CTC-S, which positively correlates with *Ctnnb1*. As reported, *Klf4* is a versatile marker that can both suppress [67] or support [68] PDAC. Several studies have investigated *Klf4*, which encodes a protein that belongs to the Krüppel family of transcription factors and plays a vital role in maintaining embryonic or induced pluripotent stem cells as well as in preventing their differentiation [28,69,70]. In a recent study, Zheng et al. [71] found that *Klf4* promotes CTC survival by increasing intracellular reactive oxygen species. Intriguingly, *Klf4* can interact with *Ctnnb1* through transcription regulation or direct protein interaction. As Tiwari et al. [72] reported, *Ctnnb1* is one of the direct repression targets of the transcription factor *Klf4*. On the opposite side, β-catenin could also bind with the promoter of *Klf4* through activating the canonical *Wnt* pathway, as reported [73]. However, we found that the canonical *Wnt* family members are expressed at low levels, as previously mentioned. The evidence suggests that the interaction between *Klf4* and β-catenin might be happening through the formation of a protein complex rather than via the transcriptional regulation in the pancreatic CTC-S group. Even though there are studies claiming that *Klf4* inhibits β-catenin [74–76], there are other reports showing that the *Klf4* / β-catenin complex is necessary for the self-renewal capacity of stem/cancer cells [68]. The consensus is that the protein–protein interaction between *Klf4* and β-catenin happens in the nucleus [68,74–76]. According to previously published results [76,77], the cytoplasmic β-catenin could migrate to the nucleus freely, and there promote the transcription and translation of β-catenin targeted genes in conjunction with transcription factors. Among the β-catenin binding proteins extracted from epithelial cells nuclei, a 55-kDa protein could be identified as being *KLF4* [76]. Furthermore, the de novo synthesis of both E-cadherin and β-catenin increased at a similar rate in order to reconstruct an adherens junction after the proteolytic disruption of this extracellular interaction in epithelial cancer cells [77].

Clinically, *KLF4* expression in the primary tumor does significantly influence survival. *CTNNB1* is a negative prognostic marker for RFS and PFS. For OS, a trend can be seen, but it fails to reach statistical significance. This supports the hypothesis that the high expression of *Ctnnb1* in CTC-S contributes to the formation of adherens junctions, which in turn may promote the survival of CTC-S in circulation [78]. However, it must be taken into account that the clinical validation was performed using RNA-seq expression data of bulk tumor cells from the primary tumor rather than single-cell RNA sequencing of CTCs, which limits its validity.

It is important to reinforce that this study is an in silico analysis of a single pre-existing dataset from a single pancreatic cancer model, since no other high-throughput sequencing datasets of pancreatic CTCs are publically available. To reduce the bias from limited study resources, we combined the analysis of differentially expressed genes, WGCNA, correlation analysis, and added clinical survival data. The limited sample size, and the absence of prognostic data directly associated with CTC expression data, may partially limit the results. With all that under consideration, more preclinical and clinical studies are necessary to confirm our findings. However, despite its exploratory nature, this study offers new insight into pancreatic CTCs with stem-like characteristics.

4. Materials and Methods

4.1. Data Collection, Data Quality Check, and Normalization

Single-cell RNA-sequencing data of 75 murine pancreatic CTCs with sufficient quality in dataset GSE51372 were downloaded from the Gene Expression Omnibus (GEO) database (available online: http://www.ncbi.nlm.nih.gov/geo/, accessed on 19 August 2019). Gene expression and clinical information profiles for Pancreas Adenocarcinoma (PAAD) patients were obtained from The Cancer Genome Atlas (TCGA) data portal (available online: https://tcga-data.nci.nih.gov/tcga/, accessed on 19 August 2019) [36].

In the pre-process part, an integrated web tool iDEP 9.0 (available online: http://bioinformatics.sdstate.edu/idep/, accessed on 7 March 2020) was employed to analyze the 75 CTCs sequencing data [79]. Three samples with a sequencing depth of less than two counts per million (CPM) were excluded, genes with less than 0.5 CPM in all the samples were also filtered out, and then transformed the remaining 72 samples with EdgeR: $\log_2$(CPM + c), pseudo count c = 4 [79]. (Figure S1) Mitochondrial RNA (mtRNA) was not used for QC as mtRNA is not contained in the GSE51372 read counts file.

4.2. Principal Component Analysis (PCA)

The top 3000 (approximately equal to 25%) [80] most variable genes were identified by the iDEP 9.0. Then, we used an open-source web tool named ClustVis (available online: https://biit.cs.ut.ee/clustvis/, accessed on 7 March 2020) to perform the PCA and generate the PCA score plot [81]. Imputation was deemed unnecessary since the data were normalized in the aforementioned part. PCA loading data were downloaded from ClustVis and visualized using imageGP (available online: http://www.ehbio.com/ImageGP/index.php/, accessed on 23 April 2020). Next, the corresponding stem markers (PCA loadings) were used to define the distinct PCA score cluster with stem-like features.

4.3. Identification of Differentially Expressed Genes

After pre-processing, read counts of 11,931 genes were used to identify the differentially expressed genes. Data analysis was performed using the R package DESeq2 [82]. |Fold change (FC) | > 2 and a corrected *p*-value, false discovery rate (FDR) < 0.050 (Benjamini–Hochberg procedure) were set as cutoffs.

4.4. Weighted Gene Co-Expression Network Analysis (WGCNA)

We constructed a gene co-expression network of the 3000 most variable genes using the *WGCNA* package [83] in R x64 3.6.1 based on Euclidean distance. Before the topological overlap matrix construction, we estimated the soft threshold β. The details are described in Figure S4. To identify the module correlated with stem-like traits, we used the dynamic tree-cut algorithm (automatic single block method) to cluster dendrogram branches into several modules and assigned them colors. The minimum module size is set at 30, and the modules with larger than 0.9 pairwise correlation were merged. Only the module which significantly positively correlated with stem-like traits was included in the subsequent study.

4.5. Definition of Hub Genes

The up-regulated genes were submitted to the STRING 11.0 database [32], and the Protein–Protein Interaction (PPI) network was reconstructed via the Cytoscape software, version 3.5.1 or higher [84]. In addition to text mining, other basic settings were selected, such as co-expression, co-occurrence, databases, experiments, gene fusion, and neighborhood as active interaction sources. Besides that, only interaction pairs that scored >0.9 in the network were selected.

We used the cytoHubba plugin in Cytoscape, which provides 12 topological analysis methods including cluster coefficient, degree, the density of maximum neighborhood component, edge percolated component, maximal clique centrality, maximum neighborhood component and six other centralities (betweenness, bottleneck, closeness, eccentricity, radiality, and stress) for ranking nodes [85]. If the analyzed genes ranked in the top 10% in more than eight of the aforementioned methods, they were defined as hub genes.

4.6. Functional Enrichment Analysis

We used g:Profiler (available online: https://biit.cs.ut.ee/gprofiler/gost, accessed on 23 March 2020) [86] to perform functional enrichment analysis. Several enriched pathways of the Kyoto Encyclopedia of Genes and Genomes (KEGG) [87] and functional interpretations of gene ontology (GO) [88],

including biological processes (BP), cellular components (CC) and molecular functions (MF), were identified [28]. The enrichment of fold >2.0 and FDR <0.010 were set as cutoffs.

4.7. Kaplan-Meier Survival Plot

Pathologically confirmed PDAC samples and their corresponding RNA-seq data from TCGA were included in this experiment [20]. Pre-processed level 3 data generated using Illumina HiSeq 2000 RNA Sequencing V2 and gene symbols *KLF4* and *CTNNB1* were used. For each tumor sample of the patient, the expression level was determined using a combination of MapSplice and RSEM. The individual sample files were merged in R using the *plyr* package [89]. In the next visualization step, GraphPad Prism 8.2.1 was utilized for the Log-rank (Mantel–Cox) test and to generate the Kaplan–Meier plots.

4.8. Statistical Analysis

Correlations of gene expression were evaluated with *cor* function (Pearson method by default) in R with the *PerformanceAnalytics* package. We calculated the Benjamini–Hochberg FDR. Only FDR <0.050 was considered statistically significant.

Additional tools used in this study are listed in Table S1.

5. Conclusions

In summary, we performed an integrated bioinformatic analysis of single-cell RNA sequencing data of pancreatic CTCs derived from the GEMM. We identified two distinct cell populations with (CTC-S) and without (CTC-N) stem cell-like properties. Various markers, including EMT-transcription factors, epithelial, mesenchymal, stemness, pluripotency, and proliferation markers, were used to characterize the CTC-S population. The adherens junction pathway was found to be significantly enriched CTC-S. This pathway may be activated by up-regulated *Ctnnb1* (β-catenin) through its conjunction with transcription factor *Klf4*, thus enabling CTC-S to survive in the bloodstream and promote distant metastasis.

To conclude, this study suggests that pancreatic CTCs with stem-like features might survive in the bloodstream and reach target organs due to evaluated *Ctnnb1* expression and the activation of intracellular adherens junctions pathway.

Supplementary Materials: The following are available online at http://www.mdpi.com/2075-4418/10/5/305/s1, Figure S1: Pre-processing the raw data of the pancreatic circulating tumor cells (CTCs); Figure S2: The classification of 72 samples; Figure S3: The differential expressed genes; Figure S4: The construction of the weighted gene co-expression network of 72 pancreatic CTCs; Figure S5: The heatmap of genes involved in adherens junction pathway; Figure S6: Correlation analysis of *Ctnnb1* and *Wnt* family members in pancreatic CTC-S group. Table S1: Online web tools utilized in this study.

Author Contributions: Conceptualization, L.Z.; Data curation, L.Z. and B.G.; Formal analysis, L.Z., B.G. and J.-C.J.; Funding acquisition, L.Z.; Investigation, L.Z., B.G. and J.-C.J.; Methodology, L.Z.; Project administration, S.S.; Validation, L.Z., B.H. and S.S.; Visualization, L.Z, B.G. and S.S.; Writing—original draft, L.Z., B.H. and S.S.; Writing—review and editing, L.Z., B.H., B.G., J.-C.J., C.Y., C.R. and S.S. All authors have read and agreed to the published version of the manuscript.

Funding: L.Zhu is supported by the China Scholarship Council (CSC) for his doctoral degree, 201908080072.

Acknowledgments: L. Zhu would like to thank Xijin Ge, Department of Mathematics and Statistics, South Dakota State University, USA, for technical support on the iDEP.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2019. *CA Cancer J. Clin.* **2019**, *69*, 7–34. [CrossRef] [PubMed]
2. Strobel, O.; Neoptolemos, J.; Jäger, D.; Büchler, M.W. Optimizing the outcomes of pancreatic cancer surgery. *Nat. Rev. Clin. Oncol.* **2019**, *16*, 11–26. [CrossRef] [PubMed]
3. Ding, Y.; Cravero, J.D.; Adrian, K.; Grippo, P. Modeling Pancreatic Cancer In Vivo, From Xenograft and Carcinogen-Induced Systems to Genetically Engineered Mice. *Pancreas* **2010**, *39*, 283–292. [CrossRef] [PubMed]
4. Eferl, R.; Casanova, E. *Mouse Models of Cancer: Methods and Protocols*; Humana Press: New York, NY, USA, 2015; Volume 1267, ISBN 978-1-4939-2297-0.
5. Aguirre, A.J.; Bardeesy, N.; Sinha, M.; Lopez, L.; Tuveson, D.A.; Horner, J.; Redston, M.S.; Depinho, R.A. Activated Kras and Ink4a/Arf deficiency cooperate to produce metastatic pancreatic ductal adenocarcinoma. *Genes Dev.* **2003**, *17*, 3112–3126. [CrossRef]
6. Hingorani, S.R.; Wang, L.; Multani, A.S.; Combs, C.; Deramaudt, T.B.; Hruban, R.H.; Rustgi, A.K.; Chang, S.; Tuveson, D.A. Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. *Cancer Cell* **2005**, *7*, 469–483. [CrossRef]
7. Massagué, J.; Obenauf, A.C. Metastatic colonization by circulating tumour cells. *Nature* **2016**, *529*, 298–306. [CrossRef]
8. Court, C.M.; Ankeny, J.S.; Sho, S.; Winograd, P.; Hou, S.; Song, M.; Wainberg, Z.A.; Girgis, M.D.; Graeber, T.G.; Agopian, V.G.; et al. Circulating Tumor Cells Predict Occult Metastatic Disease and Prognosis in Pancreatic Cancer. *Ann. Surg. Oncol.* **2018**, *25*, 1000–1008. [CrossRef]
9. Steinert, G.; Schölch, S.; Koch, M.; Weitz, J. Biology and significance of circulating and disseminated tumour cells in colorectal cancer. *Langenbecks Arch. Surg.* **2012**, *397*, 535–542. [CrossRef]
10. Rahbari, N.N.; Bork, U.; Kircher, A.; Nimitz, T.; Schölch, S.; Kahlert, C.; Schmidt, T.; Steinert, G.; Ulrich, A.B.; Reissfelder, C.; et al. Compartmental differences of circulating tumor cells in colorectal cancer. *Ann. Surg. Oncol.* **2012**, *19*, 2195–2202. [CrossRef]
11. Bork, U.; Grützmann, R.; Rahbari, N.N.; Schölch, S.; Distler, M.; Reissfelder, C.; Koch, M.; Weitz, J. Prognostic relevance of minimal residual disease in colorectal cancer. *World J. Gastroenterol* **2014**, *20*, 10296–10304. [CrossRef]
12. Schölch, S.; Bork, U.; Rahbari, N.N.; García, S.; Swiersy, A.; Betzler, A.M.; Weitz, J.; Koch, M. Circulating tumor cells of colorectal cancer. *Cancer Cell Microenviron.* **2014**, *1*, e323.
13. Steinert, G.; Schölch, S.; Niemietz, T.; Iwata, N.; García, S.A.; Behrens, B.; Voigt, A.; Kloor, M.; Benner, A.; Bork, U.; et al. Immune Escape and Survival Mechanisms in Circulating Tumor Cells of Colorectal Cancer. *Cancer Res.* **2014**, *74*, 1694–1704. [CrossRef]
14. Van Noort, V.; Schölch, S.; Iskar, M.; Zeller, G.; Ostertag, K.; Schweitzer, C.; Werner, K.; Weitz, J.; Koch, M.; Bork, P. Novel Drug Candidates for the Treatment of Metastatic Colorectal Cancer through Global Inverse Gene-Expression Profiling. *Cancer Res.* **2014**, *74*, 5690–5699. [CrossRef] [PubMed]
15. Bork, U.; Rahbari, N.N.; Schölch, S.; Reissfelder, C.; Kahlert, C.; Büchler, M.W.; Weitz, J.; Koch, M. Circulating tumour cells and outcome in non-metastatic colorectal cancer: A prospective study. *Br. J. Cancer* **2015**, *112*, 1306–1313. [CrossRef] [PubMed]
16. Rahbari, N.N.; Bork, U.; Schölch, S.; Reissfelder, C.; Thorlund, K.; Betzler, A.; Kahlert, C.; Schneider, M.; Ulrich, A.B.; Büchler, M.W.; et al. Metastatic Spread Emerging From Liver Metastases of Colorectal Cancer: Does the Seed Leave the Soil Again? *Ann. Surg.* **2016**, *263*, 345–352. [CrossRef] [PubMed]
17. Schölch, S.; García, S.A.; Iwata, N.; Niemietz, T.; Betzler, A.M.; Nanduri, L.K.; Bork, U.; Kahlert, C.; Thepkaysone, M.-L.; Swiersy, A.; et al. Circulating tumor cells exhibit stem cell characteristics in an orthotopic mouse model of colorectal cancer. *Oncotarget* **2016**, *7*, 27232–27242. [CrossRef]
18. Nanduri, L.K.; Hissa, B.; Weitz, J.; Schölch, S.; Bork, U. The prognostic role of circulating tumor cells in colorectal cancer. *Expert Rev. Anticancer Ther.* **2019**, *19*, 1077–1088. [CrossRef]
19. Baccelli, I.; Schneeweiss, A.; Riethdorf, S.; Stenzinger, A.; Schillert, A.; Vogel, V.; Klein, C.; Saini, M.; Bäuerle, T.; Wallwiener, M.; et al. Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay. *Nat. Biotechnol.* **2013**, *31*, 539–544. [CrossRef]

20. Barriere, G.; Fici, P.; Gallerani, G.; Fabbri, F.; Zoli, W.; Rigaud, M. Circulating tumor cells and epithelial, mesenchymal and stemness markers: Characterization of cell subpopulations. *Ann. Transl. Med.* **2014**, *2*, 109.
21. Nagrath, S.; Jack, R.M.; Sahai, V.; Simeone, D.M. Opportunities and Challenges for Pancreatic Circulating Tumor Cells. *Gastroenterology* **2016**, *151*, 412–426. [CrossRef]
22. Agnoletto, C.; Corrà, F.; Minotti, L.; Baldassari, F.; Crudele, F.; Cook, W.J.J.; Di Leva, G.; D'Adamo, A.P.; Gasparini, P.; Volinia, S. Heterogeneity in circulating tumor cells: The relevance of the stem-cell subset. *Cancers* **2019**, *11*, 483. [CrossRef] [PubMed]
23. Ting, D.T.; Wittner, B.S.; Ligorio, M.; Jordan, N.V.; Shah, A.M.; Miyamoto, D.T.; Aceto, N.; Bersani, F.; Brannigan, B.W.; Xega, K.; et al. Single-Cell RNA Sequencing Identifies Extracellular Matrix Gene Expression by Pancreatic Circulating Tumor Cells. *Cell Rep.* **2014**, *8*, 1905–1918. [CrossRef]
24. Akrap, N.; Andersson, D.; Bom, E.; Gregersson, P.; Ståhlberg, A.; Landberg, G. Identification of Distinct Breast Cancer Stem Cell Populations Based on Single-Cell Analyses of Functionally Enriched Stem and Progenitor Pools. *Stem Cell Rep.* **2016**, *6*, 121–136. [CrossRef] [PubMed]
25. Lapin, M.; Tjensvoll, K.; Oltedal, S.; Javle, M.; Smaaland, R.; Gilje, B.; Nordgård, O. Single-cell mRNA profiling reveals transcriptional heterogeneity among pancreatic circulating tumour cells. *BMC Cancer* **2017**, *17*, 1–10.
26. Shibue, T.; Weinberg, R.A. EMT, CSCs, and drug resistance: The mechanistic link and clinical implications. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 611–629. [CrossRef]
27. Singh, S.; Arcaroli, J.; Messersmith, W.; Vasiliou, V. Acetaldehyde and retinaldehyde-metabolizing enzymes in colon and pancreatic cancers. *Adv. Exp. Med. Biol.* **2015**, *815*, 281–294.
28. Yamane, M.; Ohtsuka, S.; Matsuura, K.; Nakamura, A.; Niwa, H. Overlapping functions of Krüppel-like factor family members: Targeting multiple transcription factors to maintain the naïve pluripotency of mouse embryonic stem cells. *Development* **2018**, *145*, 1–12. [CrossRef]
29. Scholzen, T.; Gerdes, J. The Ki-67 Protein: From the Known and the Unknown. *J. Cell. Phys.* **2000**, *322*, 311–322. [CrossRef]
30. Ishiwata, T.; Matsuda, Y.; Yoshimura, H.; Sasaki, N.; Ishiwata, S.; Ishikawa, N.; Takubo, K.; Arai, T.; Aida, J. Pancreatic cancer stem cells: Features and detection methods. *Pathol. Oncol. Res.* **2018**, *24*, 797–805. [CrossRef]
31. Zhou, P.; Li, B.; Liu, F.; Zhang, M.; Wang, Q.; Liu, Y.; Yao, Y.; Li, D. The epithelial to mesenchymal transition (EMT) and cancer stem cells: Implication for treatment resistance in pancreatic cancer. *Mol. Cancer* **2017**, *16*, 1–11. [CrossRef]
32. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [CrossRef] [PubMed]
33. Luo, W.; Brouwer, C. Pathview: An R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics* **2013**, *29*, 1830–1831. [CrossRef] [PubMed]
34. Tang, Z.; Kang, B.; Li, C.; Chen, T.; Zhang, Z. GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **2019**, *47*, W556–W560. [CrossRef]
35. Lonsdale, J.; Thomas, J.; Salvatore, M.; Phillips, R.; Lo, E.; Shad, S.; Hasz, R.; Walters, G.; Garcia, F.; Young, N.; et al. GTEx Consortium The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **2013**, *45*, 580–585. [CrossRef] [PubMed]
36. Raphael, B.J.; Hruban, R.H.; Aguirre, A.J.; Moffitt, R.A.; Yeh, J.J.; Stewart, C.; Robertson, A.G.; Cherniack, A.D.; Gupta, M.; Getz, G.; et al. Cancer Genome Atlas Research Network Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* **2017**, *32*, 185–203.e13. [CrossRef]
37. Lytle, N.K.; Ferguson, L.P.; Rajbhandari, N.; Gilroy, K.; Fox, R.G.; Deshpande, A.; Schürch, C.M.; Hamilton, M.; Robertson, N.; Lin, W.; et al. A Multiscale Map of the Stem Cell State in Pancreatic Adenocarcinoma. *Cell* **2019**, *177*, 572–586.e22. [CrossRef]
38. Varillas, J.I.; Zhang, J.; Chen, K.; Barnes, I.I.; Liu, C.; George, T.J.; Fan, Z.H. Microfluidic isolation of circulating tumor cells and cancer stem-like cells from patients with pancreatic ductal adenocarcinoma. *Theranostics* **2019**, *9*, 1417–1425. [CrossRef]
39. Islam, F.; Gopalan, V.; Lam, A.K.-Y. *Cancer Stem Cells, Methods and Protocols*, 1st ed.; Papaccio, G., Desiderio, V., Eds.; Humana Press: New York, NY, USA, 2018; ISBN 978-1-4939-7400-9.

40. Al-Hajj, M.; Wicha, M.S.; Benito-Hernandez, A.; Morrison, S.J.; Clarke, M.F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3983–3988. [CrossRef]
41. Brabletz, T.; Hlubek, F.; Spaderna, S.; Schmalhofer, O.; Hiendlmeyer, E.; Jung, A.; Kirchner, T. Invasion and Metastasis in Colorectal Cancer: Epithelial-Mesenchymal Transition, Mesenchymal-Epithelial Transition, Stem Cells and β-Catenin. *CTO* **2005**, *179*, 56–65. [CrossRef]
42. Mani, S.A.; Guo, W.; Liao, M.-J.; Eaton, E.N.; Ayyanan, A.; Zhou, A.Y.; Brooks, M.; Reinhard, F.; Zhang, C.C.; Shipitsin, M.; et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **2008**, *133*, 704–715. [CrossRef]
43. Morel, A.-P.; Lièvre, M.; Thomas, C.; Hinkal, G.; Ansieau, S.; Puisieux, A. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS ONE* **2008**, *3*, e2888. [CrossRef] [PubMed]
44. Papadaki, M.A.; Stoupis, G.; Theodoropoulos, P.A.; Mavroudis, D.; Georgoulias, V.; Agelaki, S. Circulating Tumor Cells with Stemness and Epithelial-to-Mesenchymal Transition Features Are Chemoresistant and Predictive of Poor Outcome in Metastatic Breast Cancer. *Mol. Cancer Ther.* **2019**, *18*, 437–447. [CrossRef] [PubMed]
45. Rahbari, N.N.; Aigner, M.; Thorlund, K.; Mollberg, N.; Motschall, E.; Jensen, K.; Diener, M.K.; Büchler, M.W.; Koch, M.; Weitz, J. Meta-analysis shows that detection of circulating tumor cells indicates poor prognosis in patients with colorectal cancer. *Gastroenterology* **2010**, *138*, 1714–1726. [CrossRef] [PubMed]
46. Cohen, S.J.; Alpaugh, R.K.; Gross, S.; O'Hara, S.M.; Smirnov, D.A.; Terstappen, L.W.M.M.; Allard, W.J.; Bilbee, M.; Cheng, J.D.; Hoffman, J.P.; et al. Isolation and characterization of circulating tumor cells in patients with metastatic colorectal cancer. *Clin. Colorectal Cancer* **2006**, *6*, 125–132. [CrossRef]
47. Cohen, S.J.; Punt, C.J.A.; Iannotti, N.; Saidman, B.H.; Sabbath, K.D.; Gabrail, N.Y.; Picus, J.; Morse, M.A.; Mitchell, E.; Miller, M.C.; et al. Prognostic significance of circulating tumor cells in patients with metastatic colorectal cancer. *Ann. Oncol.* **2009**, *20*, 1223–1229. [CrossRef]
48. Cohen, S.J.; Punt, C.J.A.; Iannotti, N.; Saidman, B.H.; Sabbath, K.D.; Gabrail, N.Y.; Picus, J.; Morse, M.; Mitchell, E.; Miller, M.C.; et al. Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. *J. Clin. Oncol.* **2008**, *26*, 3213–3221. [CrossRef]
49. Tam, W.L.; Weinberg, R.A. The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat. Med.* **2013**, *19*, 1438–1449. [CrossRef]
50. Conacci-Sorrell, M.; Zhurinsky, J.; Ben-Ze'ev, A. The cadherin-catenin adhesion system in signaling and cancer. *J. Clin. Investig.* **2002**, *109*, 987–991. [CrossRef]
51. Hinck, L.; Näthke, I.S.; Papkoff, J.; Nelson, W.J. Beta-catenin: A common target for the regulation of cell adhesion by Wnt-1 and Src signaling pathways. *Trends Biochem. Sci.* **1994**, *19*, 538–542. [CrossRef]
52. Yamada, S.; Pokutta, S.; Drees, F.; Weis, W.I.; Nelson, W.J. Deconstructing the Cadherin-Catenin-Actin Complex. *Cell* **2005**, *123*, 889–901. [CrossRef]
53. Drees, F.; Pokutta, S.; Yamada, S.; Nelson, W.J.; Weis, W.I. α-Catenin Is a Molecular Switch that Binds E-Cadherin-β-Catenin and Regulates Actin-Filament Assembly. *Cell* **2005**, *123*, 903–915. [CrossRef] [PubMed]
54. Sahai, E.; Marshall, C.J. ROCK and Dia have opposing effects on adherens junctions downstream of Rho. *Nat. Cell Biol.* **2002**, *4*, 408–415. [CrossRef] [PubMed]
55. Harris, T.J.C.; Tepass, U. Adherens junctions: From molecules to morphogenesis. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 502–514. [CrossRef] [PubMed]
56. Menke, A.; Giehl, K. Regulation of adherens junctions by Rho GTPases and p120-catenin. *Arch. Biochem. Biophys.* **2012**, *524*, 48–55. [CrossRef]
57. Aceto, N.; Bardia, A.; Miyamoto, D.T.; Donaldson, M.C.; Wittner, B.S.; Spencer, J.A.; Yu, M.; Pely, A.; Engstrom, A.; Zhu, H.; et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **2014**, *158*, 1110–1122. [CrossRef]
58. Kemler, R. From cadherins to catenins: Cytoplasmic protein interactions and regulation of cell adhesion. *Trends Genet.* **1993**, *9*, 317–321. [CrossRef]
59. Chang, M.-C.; Chang, Y.-T.; Chen, J.-Y.; Jeng, Y.-M.; Yang, C.-Y.; Tien, Y.-W.; Yang, S.-H.; Chen, H.-L.; Liang, T.-Y.; Wang, C.-F.; et al. Clinical Significance of Circulating Tumor Microemboli as a Prognostic Marker in Patients with Pancreatic Ductal Adenocarcinoma. *Clin. Chem.* **2016**, *62*, 505–513. [CrossRef]

60. Janda, C.Y.; Dang, L.T.; You, C.; Chang, J.; de Lau, W.; Zhong, Z.A.; Yan, K.S.; Marecic, O.; Siepe, D.; Li, X.; et al. Surrogate Wnt agonists that phenocopy canonical Wnt/β-catenin signaling. *Nature* **2017**, *545*, 234–237. [CrossRef]
61. Gerlach, J.P.; Emmink, B.L.; Nojima, H.; Kranenburg, O.; Maurice, M.M. Wnt signalling induces accumulation of phosphorylated β-catenin in two distinct cytosolic complexes. *Open Biology* **2014**, *4*, 1–11. [CrossRef]
62. Woo, D.; Yu, M. Circulating tumor cells as "liquid biopsies" to understand cancer metastasis. *Transl. Res.* **2018**, *201*, 128–135. [CrossRef]
63. Zhang, P.; Chang, W.-H.; Fong, B.; Gao, F.; Liu, C.; Al Alam, D.; Bellusci, S.; Lu, W. Regulation of induced pluripotent stem (iPS) cell induction by Wnt/β-catenin signaling. *J. Biol. Chem.* **2014**, *289*, 9221–9232. [CrossRef] [PubMed]
64. Yu, M.; David, T.T.; Stott, S.L.; Wittner, B.S.; Ozsolak, F.; Paul, S.; Ciciliano, J.C.; Smas, M.E.; Winokur, D.; Gilman, A.J.; et al. RNA sequencing of pancreatic circulating tumour cells implicates WNT signaling in metastasis. *Nature* **2012**, *487*, 510–513. [CrossRef] [PubMed]
65. Evans, P.M.; Chen, X.; Zhang, W.; Liu, C. KLF4 interacts with beta-catenin/TCF4 and blocks p300/CBP recruitment by beta-catenin. *Mol. Cell. Biol.* **2010**, *30*, 372–381. [CrossRef]
66. Rudloff, S.; Kemler, R. Differential requirements for β-catenin during mouse development. *Development* **2012**, *139*, 3711–3721. [CrossRef] [PubMed]
67. Zammarchi, F.; Morelli, M.; Menicagli, M.; Di Cristofano, C.; Zavaglia, K.; Paolucci, A.; Campani, D.; Aretini, P.; Boggi, U.; Mosca, F.; et al. KLF4 is a novel candidate tumor suppressor gene in pancreatic ductal carcinoma. *Am. J. Pathol.* **2011**, *178*, 361–372. [CrossRef] [PubMed]
68. Hoffmeyer, K.; Raggioli, A.; Rudloff, S.; Anton, R.; Hierholzer, A.; Del Valle, I.; Hein, K.; Vogt, R.; Kemler, R. Wnt/β-catenin signaling regulates telomerase in stem cells and cancer cells. *Science* **2012**, *336*, 1549–1554. [CrossRef]
69. Karagiannis, P.; Takahashi, K.; Saito, M.; Yoshida, Y.; Okita, K.; Watanabe, A.; Inoue, H.; Yamashita, J.K.; Todani, M.; Nakagawa, M.; et al. Induced Pluripotent Stem Cells and Their Use in Human Models of Disease and Development. *Phys. Rev.* **2018**, *99*, 79–114. [CrossRef]
70. Chronis, C.; Fiziev, P.; Papp, B.; Butz, S.; Bonora, G.; Sabri, S.; Ernst, J.; Plath, K. Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell* **2017**, *168*, 442–459.e20. [CrossRef]
71. Zheng, Y.; Miyamoto, D.T.; Wittner, B.S.; Sullivan, J.P.; Aceto, N.; Jordan, N.V.; Yu, M.; Karabacak, N.M.; Comaills, V.; Morris, R.; et al. Expression of β-globin by cancer cells promotes cell survival during blood-borne dissemination. *Nat. Commun.* **2017**, *8*, 1–12. [CrossRef]
72. Tiwari, N.; Meyer-Schaller, N.; Arnold, P.; Antoniadis, H.; Pachkov, M.; van Nimwegen, E.; Christofori, G. Klf4 is a transcriptional regulator of genes critical for EMT, including Jnk1 (Mapk8). *PLoS ONE* **2013**, *8*, e57329. [CrossRef]
73. Ai, Z.; Shao, J.; Wu, Y.; Yu, M.; Du, J.; Shi, X.; Shi, X.; Zhang, Y.; Guo, Z. CHIR99021 enhances Klf4 Expression through β-Catenin Signaling and miR-7a Regulation in J1 Mouse Embryonic Stem Cells. *PLoS ONE* **2016**, *11*, e0150936. [CrossRef] [PubMed]
74. Parenti, S.; Montorsi, L.; Fantini, S.; Mammoli, F.; Gemelli, C.; Atene, C.G.; Losi, L.; Frassineti, C.; Calabretta, B.; Tagliafico, E.; et al. KLF4 Mediates the Effect of 5-ASA on the β-Catenin Pathway in Colon Cancer Cells. *Cancer Prev. Res.* **2018**, *11*, 503–510. [CrossRef] [PubMed]
75. Sellak, H.; Wu, S.; Lincoln, T.M. KLF4 and SOX9 transcription factors antagonize β-catenin and inhibit TCF-activity in cancer cells. *Biochim. Biophys. Acta* **2012**, *1823*, 1666–1675. [CrossRef] [PubMed]
76. Zhang, W.; Chen, X.; Kato, Y.; Evans, P.M.; Yuan, S.; Yang, J.; Rychahou, P.G.; Yang, V.W.; He, X.; Evers, B.M.; et al. Novel cross talk of Kruppel-like factor 4 and beta-catenin regulates normal intestinal homeostasis and tumor repression. *Mol. Cell. Biol.* **2006**, *26*, 2055–2064. [CrossRef] [PubMed]
77. Klinke, D.J.; Horvath, N.; Cuppett, V.; Wu, Y.; Deng, W.; Kanj, R. Interlocked positive and negative feedback network motifs regulate β-catenin activity in the adherens junction pathway. *Mol. Biol. Cell* **2015**, *26*, 4135–4148. [CrossRef]
78. Hofmann, C.; Obermeier, F.; Artinger, M.; Hausmann, M.; Falk, W.; Schoelmerich, J.; Rogler, G.; Grossmann, J. Cell-Cell Contacts Prevent Anoikis in Primary Human Colonic Epithelial Cells. *Gastroenterology* **2007**, *132*, 587–600. [CrossRef]
79. Ge, S.X.; Son, E.W.; Yao, R. iDEP: An integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **2018**, *19*, 1–24. [CrossRef]

80. Blischak, J. PCA of Highly Variable Genes. Available online: https://jdblischak.github.io/singlecell-qtl/pca-variable.html (accessed on 6 March 2020).
81. Metsalu, T.; Vilo, J. ClustVis: A web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **2015**, *43*, W566–W570. [CrossRef]
82. 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]
83. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [CrossRef]
84. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef] [PubMed]
85. Chin1, C.-H.; Chen, S.-H.; Wu, H.-H.; Ho, C.-W.; Ko, M.-T.; Lin, C.-Y. From cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **2014**, *8*, 1–7. [CrossRef]
86. Raudvere, U.; Kolberg, L.; Kuzmin, I.; Arak, T.; Adler, P.; Peterson, H.; Vilo, J. g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **2019**, *47*, W191–W198. [CrossRef]
87. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **2019**, *28*, 1947–1951. [CrossRef] [PubMed]
88. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **2000**, *25*, 25–29. [CrossRef] [PubMed]
89. Wickham, H. The split-apply-combine strategy for data analysis. *J. Stat. Softw.* **2009**, *40*, 174473. [CrossRef]

3 DISCUSSION AND CONCLUSION

As early as 2014, pioneering research attempted to establish scRNA-seq of pancreatic CTCs from KPfC mice.[133] A stromal protein, SPARC, was identified to be enriched abundantly in pancreatic CTCs.[133] However, this result has not been confirmed in the following years, and only limited related data was published.[143]

On the other hand, it is widely accepted that disseminated tumor cells require stem cell features to establish overt distant colonization[10,144]. Some researchers established CTC in vitro culture with hypoxia or serum-free conditions, believing that only stem cells could survive and probably represent the progenitor population of metastasis.[145,146] It is unclear whether the resulting CTC cell lines still exhibit the same biology as their parental CTCs in circulation. Regardless, using stem markers to define CTCs with stem cell phenotype is a popular strategy.[34,73,140]

This cumulative dissertation attempted to address the biomarker and stem fraction characterization based on one central dataset (GSE521372) of CTC single-cell RNA sequencing.[147]

In the first part, *Gas2l1* was found to be a useful marker for CTC identification, especially when combined with *Epcam* as biomarkers for pancreatic CTCs. Previous literature introduced *GAS2L1* as an essential protein for cell polarization and migration.[148,149] Therefore, it is reasonable that pancreatic CTCs rely on *Gas2l1* and its guiding microtubules to focal adhesions for migration. Notably, Rowley et al.[150] demonstrated that *Gas2l1* is expressed in murine platelets abundantly. The gene ontology analysis of pancreatic CTCs in this study also confirmed the enrichment of platelets markers. This reflects the alternative possibility that *Gas2l1* expression on pancreatic CTCs results from platelets covering CTCs and thus leading to transcriptomic contamination.

The second publication demonstrated that the adherence junction pathway was enriched in murine pancreatic CTCs with stem cell characteristics. It is conceivable that this phenomenon reflects CTC-S preparing for attachment to the vessel wall in the target organ. The downregulation of the canonical *Wnt* pathway found in the study suggests that Klf4 and β-catenin interact on the protein rather than the transcriptomic level in the pancreatic CTC-S group. The Klf4 / β-catenin interaction may help CTCs acquire stemness, regain intercellular adhesion capabilities, and form cell clusters through actin cytoskeleton remodeling.[151]

As both publications represent *in silico* analysis, their translational value needs further experimental validation. Immunocytochemistry or single-cell RT-PCR in lineage-traced

pancreatic CTCs would be suitable methods to validate Gas2l1 expression. In vivo genetic manipulation (e.g., using the cre/loxP system), knocking out individual stem cell markers followed by the analysis of the metastatic phenotype of the model would constitute a stem-feature verification experiment.

This dissertation has shown that the combination of *GAS2L1* and *EPCAM* may be a highly sensitive marker combination for identifying CTCs in pancreatic cancer patients. The second major finding of this work is that pancreatic CTCs with stem-like features exhibit elevated *Ctnnb1* expression, which may contribute to the metastatic properties of CTCs through activation of intracellular adherens junctions pathway. However, despite the advances in our understanding of CTCs in pancreatic cancer, major steps of the metastatic cascade are still incompletely understood and require further investigations before clinically relevant breakthroughs can be made.

4 SUMMARY

Pancreatic ductal adenocarcinoma (PDAC) is a highly lethal malignancy in humans due to high recurrence rates and early systemic dissemination.[1–3] This major health burden leads to tremendous research efforts to decipher the metastatic cascade of PDAC.[11,152–156] CTCs, as a well-known intermediary step between primary and secondary tumors, have been a lasting research topic for decades.[60,102,157–160]

Although exponentially increasing numbers of publications are investigating CTCs, there has been no consensus marker for CTCs in the past decades, most likely as a result of CTC heterogeneity. Universally, researchers applied EPCAM for immunoaffinity isolation of pancreatic CTCs to harvest a particular CTC; however, this subpopulation is controversial and may not represent the metastasis-initiating subfraction of CTCs.[60]

GEMMs are valuable tools to investigate pancreatic CTCs.[60,133] Since they retain histopathological features, consistent transcriptomics alterations, and the mutational landscape of human PDAC,[44,55] the CTCs derived from GEMMs should resemble human pancreatic CTCs closely.

Previous pioneering research already attempted scRNA-seq of pancreatic CTCs from KPfC mice. The re-analysis was performed on their dataset (GSE51372) with Seurat. Three heterogeneous clusters of CTCs with distinct biologic phenotypes were identified. In this study, the *Gas2l1* exhibited better CTC identification than *Epcam*, *Cd45*, *Sparc*, and absence expression in peripheral blood mononuclear cells. Probably, pancreatic CTCs realize migration through *Gas2l1* guiding microtubules to focal adhesions. Notably, few pieces of the literature demonstrated that *Gas2l1* is in murine platelets. This reflects the possibility that pancreatic CTCs cloak with platelets. Regardless, *GAS2L1* shows significant over-expression in PDAC compared to normal pancreatic tissue and a remarkable indicator value of relapse-free survival in the PDAC cohort (The Cancer Genome Atlas Program). Future immunocytochemistry and single-cell RT-PCR in lineage-traced pancreatic CTCs are necessary to validate its efficiency.

Generally, stem markers were used to identify CTCs with stem-like features. In case a single marker distorted the stem-related classification, a pool of phenotypic markers was applied on the same murine pancreatic CTC dataset (GSE51372). Combined with the following principal component analysis and hierarchical clustering, Murine pancreatic CTCs were divided into with and without stem-like features groups. Both over-representation analysis and weighted gene co-expression network analysis revealed that the adherens junction pathway is significantly enriched in pancreatic

stem-like CTCs. Transcription factor, *Klf4* was found to have a significant association with stem-like population. Reasonable speculation is that the stem fraction CTCs regain intercellular junction and form cell clusters through actin cytoskeleton remodeling. However, further transcriptomic comparisons between CTCs and CTC clusters, genetic manipulations are alternative solutions for validations.

In summary, pancreatic CTCs are vital components of hematogeneous metastasis. The emerging scRNA-seq techniques and corresponding bioinformatic toolkits accelerated the development of CTC research. Despite limitations existing, these two in silico analyses of scRNA-seq data of pancreatic CTCs derived from GEMM and explored biomarkers and stem-like populations of these CTCs. As scRNA-seq technology is evolving rapidly, the dawn of uncovering more secrets of CTCs is breaking.

5 LITERATURE LIST

1. *Krebs - Pancreatic cancer*, *Krebs - Pancreatic cancer*. Online: https://www.krebsdaten.de/Krebs/EN/Content/Cancer_sites/Pancreatic_cancer/pancreatic_cancer_node.html, Stand: 25.08.2021.

2. Robert Koch Insitute: *Cancer in Germany 2015/2016*. Berlin, 2020.

3. Luo, G, Fan, Z, Gong, Y, Jin, K, Yang, C, Cheng, H, Huang, D, Ni, Q, Liu, C, Yu, X: Characteristics and Outcomes of Pancreatic Cancer by Histological Subtypes. *Pancreas*, 48: 817–822, 2019.

4. Bardeesy, N, DePinho, RA: Pancreatic cancer biology and genetics (Bandiera_abtest: aCg_type: Nature Research Journalsnumber: 12Primary_atype: Reviewspublisher: Nature Publishing Group). *Nat Rev Cancer*, 2: 897–909, 2002.

5. Neoptolemos, JP, Stocken, DD, Friess, H, Bassi, C, Dunn, JA, Hickey, H, Beger, H, Fernandez-Cruz, L, Dervenis, C, Lacaine, F, Falconi, M, Pederzoli, P, Pap, A, Spooner, D, Kerr, DJ, Büchler, MW, European Study Group for Pancreatic Cancer: A randomized trial of chemoradiotherapy and chemotherapy after resection of pancreatic cancer. *N Engl J Med*, 350: 1200–1210, 2004.

6. Kamisawa, T, Isawa, T, Koike, M, Tsuruta, K, Okamoto, A: Hematogenous metastases of pancreatic ductal carcinoma. *Pancreas*, 11: 345–349, 1995.

7. Winter, JM, Cameron, JL, Campbell, KA, Arnold, MA, Chang, DC, Coleman, J, Hodgin, MB, Sauter, PK, Hruban, RH, Riall, TS, Schulick, RD, Choti, MA, Lillemoe, KD, Yeo, CJ: 1423 pancreaticoduodenectomies for pancreatic cancer: A single-institution experience. *J Gastrointest Surg*, 10: 1199–1210; discussion 1210-1211, 2006.

8. Raut, CP, Tseng, JF, Sun, CC, Wang, H, Wolff, RA, Crane, CH, Hwang, R, Vauthey, J-N, Abdalla, EK, Lee, JE, Pisters, PWT, Evans, DB: Impact of Resection Status on Pattern of Failure and Survival After Pancreaticoduodenectomy for Pancreatic Adenocarcinoma. *Ann Surg*, 246: 52–60, 2007.

9. Joyce, JA, Pollard, JW: Microenvironmental regulation of metastasis. *Nat Rev Cancer*, 9: 239–252, 2009.

10. Chaffer, CL, Weinberg, RA: A perspective on cancer cell metastasis. *Science*, 331: 1559–1564, 2011.

11. Wong, SY, Hynes, RO: Lymphatic or hematogenous dissemination: how does a metastatic tumor cell decide? *Cell Cycle*, 5: 812–817, 2006.

12. Thiery, JP, Acloque, H, Huang, RYJ, Nieto, MA: Epithelial-Mesenchymal Transitions in Development and Disease. *Cell*, 139: 871–890, 2009.

13. Hamidi, H, Ivaska, J: Every step of the way: integrins in cancer progression and metastasis. *Nat Rev Cancer*, 18: 533–548, 2018.

14. Gupta, GP, Massagué, J: Cancer metastasis: building a framework. *Cell*, 127: 679–695, 2006.

15. Brown, M, Assen, FP, Leithner, A, Abe, J, Schachner, H, Asfour, G, Bago-Horvath, Z, Stein, JV, Uhrin, P, Sixt, M, Kerjaschki, D: Lymph node blood vessels provide exit routes for metastatic tumor cell dissemination in mice. *Science*, 359: 1408–1411, 2018.

16. Pereira, ER, Kedrin, D, Seano, G, Gautier, O, Meijer, EFJ, Jones, D, Chin, S-M, Kitahara, S, Bouta, EM, Chang, J, Beech, E, Jeong, H-S, Carroll, MC, Taghian, AG, Padera, TP: Lymph node metastases can invade local blood vessels, exit the node, and colonize distant organs in mice. *Science*, 359: 1403–1407, 2018.

17. Ubellacker, JM, Tasdogan, A, Ramesh, V, Shen, B, Mitchell, EC, Martin-Sandoval, MS, Gu, Z, McCormick, ML, Durham, AB, Spitz, DR, Zhao, Z, Mathews, TP, Morrison, SJ: Lymph protects metastasizing melanoma cells from ferroptosis. *Nature*, 585: 113–118, 2020.

18. Massagué, J, Obenauf, AC: Metastatic colonization by circulating tumour cells. *Nature*, 529: 298–306, 2016.

19. Court, CM, Ankeny, JS, Sho, S, Winograd, P, Hou, S, Song, M, Wainberg, ZA, Girgis, MD, Graeber, TG, Agopian, VG, Tseng, H-R, Tomlinson, JS: Circulating Tumor Cells Predict Occult Metastatic Disease and Prognosis in Pancreatic Cancer. *Ann. Surg. Oncol.*, 25: 1000–1008, 2018.

20. Steinert, G, Schölch, S, Koch, M, Weitz, J: Biology and significance of circulating and disseminated tumour cells in colorectal cancer. *Langenbecks Arch Surg*, 397: 535–542, 2012.

21. Rahbari, NN, Bork, U, Kircher, A, Nimitz, T, Schölch, S, Kahlert, C, Schmidt, T, Steinert, G, Ulrich, AB, Reissfelder, C, Büchler, MW, Koch, M, Weitz, J: Compartmental differences of circulating tumor cells in colorectal cancer. *Ann. Surg. Oncol.*, 19: 2195–2202, 2012.

22. Steinert, G, Schölch, S, Niemietz, T, Iwata, N, García, SA, Behrens, B, Voigt, A, Kloor, M, Benner, A, Bork, U, Rahbari, NN, Büchler, MW, Stoecklein, NH, Weitz, J, Koch, M: Immune escape and survival mechanisms in circulating tumor cells of colorectal cancer. *Cancer Research*, 74: 1694–1704, 2014.

23. White, MG, Lee, A, Vicente, D, Hall, C, Kim, MP, Katz, MHG, Lee, JE, Ikoma, N, Lucci, A, Tzeng, C-WD: Measurement of Portal Vein Blood Circulating Tumor Cells is Safe and May Correlate With Outcomes in Resected Pancreatic Ductal Adenocarcinoma. *Ann Surg Oncol*, 28: 4615–4622, 2021.

24. Tien, YW, Kuo, H-C, Ho, B-I, Chang, M-C, Chang, Y-T, Cheng, M-F, Chen, H-L, Liang, T-Y, Wang, C-F, Huang, C-Y, Shew, J-Y, Chang, YC, Lee, EY, Lee, W-H: A High Circulating Tumor Cell Count in Portal Vein Predicts Liver Metastasis From Periampullary or Pancreatic Cancer: A High Portal Venous CTC Count Predicts Liver Metastases. *Medicine (Baltimore)*, 95: e3407, 2016.

25. Zhu, X, Suo, Y, Fu, Y, Zhang, F, Ding, N, Pang, K, Xie, C, Weng, X, Tian, M, He, H, Wei, X: In vivo flow cytometry reveals a circadian rhythm of circulating tumor cells. *Light Sci Appl*, 10: 110, 2021.

26. Strilic, B, Offermanns, S: Intravascular Survival and Extravasation of Tumor Cells. *Cancer Cell*, 32: 282–293, 2017.

27. Meng, S, Tripathy, D, Frenkel, EP, Shete, S, Naftalis, EZ, Huth, JF, Beitsch, PD, Leitch, M, Hoover, S, Euhus, D, Haley, B, Morrison, L, Fleming, TP, Herlyn, D, Terstappen, LWMM, Fehm, T, Tucker, TF, Lane, N, Wang, J, Uhr, JW: Circulating Tumor Cells in Patients with Breast Cancer Dormancy. *Clin Cancer Res*, 10: 8152–8162, 2004.

28. Denève, E, Riethdorf, S, Ramos, J, Nocca, D, Coffy, A, Daurès, J-P, Maudelonde, T, Fabre, J-M, Pantel, K, Alix-Panabières, C: Capture of viable circulating tumor cells in the liver of colorectal cancer patients. *Clin Chem*, 59: 1384–1392, 2013.

29. Al-Mehdi, AB, Tozawa, K, Fisher, AB, Shientag, L, Lee, A, Muschel, RJ: Intravascular origin of metastasis from the proliferation of endothelium-attached tumor cells: a new model for metastasis. *Nat Med*, 6: 100–102, 2000.

30. Luzzi, KJ, MacDonald, IC, Schmidt, EE, Kerkvliet, N, Morris, VL, Chambers, AF, Groom, AC: Multistep nature of metastatic inefficiency: dormancy of solitary cells after successful extravasation and limited survival of early micrometastases. *Am J Pathol*, 153: 865–873, 1998.

31. Ferrara, N: VEGF and the quest for tumour angiogenesis factors. *Nat Rev Cancer*, 2: 795–803, 2002.

32. Gimbrone, MA, Leapman, SB, Cotran, RS, Folkman, J: Tumor dormancy in vivo by prevention of neovascularization. *J Exp Med*, 136: 261–276, 1972.

33. Yu, Q, Yao, Y, Zhu, X, Gao, Y, Chen, Y, Wang, R, Xu, P, Wei, X, Jiang, L: In Vivo Flow Cytometric Evaluation of Circulating Metastatic Pancreatic Tumor Cells after High-Intensity Focused Ultrasound Therapy. *Cytometry A*, 97: 900–908, 2020.

34. Kuo, CW, Chueh, D-Y, Chen, P: Real-time in vivo imaging of subpopulations of circulating tumor cells using antibody conjugated quantum dots. *J Nanobiotechnology*, 17: 26, 2019.

35. Ito, Y, Inoue, E, Matsui, Y, Kobuchi, S, Moyama, C, Amagase, K, Yoshimura, M, Ikehara, Y, Nakata, S, Nakanishi, H: Cytology-based Detection of Circulating Tumour Cells in Human Pancreatic Cancer Xenograft Models With KRAS Mutation. *Anticancer Res*, 40: 6781–6789, 2020.

36. Rivera-Báez, L, Lohse, I, Lin, E, Raghavan, S, Owen, S, Harouaka, R, Herman, K, Mehta, G, Lawrence, TS, Morgan, MA, Cuneo, KC, Nagrath, S: Expansion of Circulating Tumor Cells from Patients with Locally Advanced Pancreatic Cancer Enable Patient Derived Xenografts and Functional Studies for Personalized Medicine. *Cancers (Basel)*, 12: 2020.

37. Torphy, RJ, Tignanelli, CJ, Kamande, JW, Moffitt, RA, Herrera Loeza, SG, Soper, SA, Yeh, JJ: Circulating tumor cells as a biomarker of response to treatment in patient-derived xenograft mouse models of pancreatic adenocarcinoma. *PLoS One*, 9: e89474, 2014.

38. Dimitrov-Markov, S, Perales-Patón, J, Bockorny, B, Dopazo, A, Muñoz, M, Baños, N, Bonilla, V, Menendez, C, Duran, Y, Huang, L, Perea, S, Muthuswamy, SK, Al-Shahrour, F, Lopez-Casas, PP, Hidalgo, M: Discovery of New Targets to Control Metastasis in Pancreatic Cancer by Single-cell Transcriptomics Analysis of Circulating Tumor Cells. *Mol Cancer Ther*, 19: 1751–1760, 2020.

39. Dai, L, Lu, C, Yu, XI, Dai, L-J, Zhou, JX: Construction of orthotopic xenograft mouse models for human pancreatic cancer. *Exp Ther Med*, 10: 1033–1038, 2015.

40. Killion, JJ, Radinsky, R, Fidler, IJ: Orthotopic models are necessary to predict therapy of transplantable tumors in mice. *Cancer Metastasis Rev*, 17: 279–284, 1998.

41. Loukopoulos, P, Kanetaka, K, Takamura, M, Shibata, T, Sakamoto, M, Hirohashi, S: Orthotopic transplantation models of pancreatic adenocarcinoma derived from cell lines and primary tumors and displaying varying metastatic activity. *Pancreas*, 29: 193–203, 2004.

42. McFadden, DG, Papagiannakopoulos, T, Taylor-Weiner, A, Stewart, C, Carter, SL, Cibulskis, K, Bhutkar, A, McKenna, A, Dooley, A, Vernon, A, Sougnez, C, Malstrom, S, Heimann, M, Park, J, Chen, F, Farago, AF, Dayton, T, Shefler, E, Gabriel, S, Getz, G, Jacks, T: Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. *Cell*, 156: 1298–1311, 2014.

43. Betzler, AM, Nanduri, LK, Hissa, B, Blickensdörfer, L, Muders, MH, Roy, J, Jesinghaus, M, Steiger, K, Weichert, W, Kloor, M, Klink, B, Schroeder, M, Mazzone, M, Weitz, J, Reissfelder, C, Rahbari, NN, Schölch, S: Differential Effects of Trp53 Alterations in Murine Colorectal Cancer. *Cancers (Basel)*, 13: 808, 2021.

44. Hingorani, SR, Wang, L, Multani, AS, Combs, C, Deramaudt, TB, Hruban, RH, Rustgi, AK, Chang, S, Tuveson, DA: Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. *Cancer Cell*, 7: 469–483, 2005.

45. Magnuson, MA, Osipovich, AB: Pancreas-specific Cre driver lines and considerations for their prudent use. *Cell Metab*, 18: 9–20, 2013.

46. Shibata, H, Toyama, K, Shioya, H, Ito, M, Hirota, M, Hasegawa, S, Matsumoto, H, Takano, H, Akiyama, T, Toyoshima, K, Kanamaru, R, Kanegae, Y, Saito, I, Nakamura, Y, Shiba, K, Noda, T: Rapid colorectal adenoma formation initiated by conditional targeting of the Apc gene. *Science (New York, N.Y.)*, 278: 120–123, 1997.

47. Zheng, B, Sage, M, Cai, WW, Thompson, DM, Tavsanli, BC, Cheah, YC, Bradley, A: Engineering a mouse balancer chromosome. *Nat Genet*, 22: 375–378, 1999.

48. Ventura, A, Kirsch, DG, McLaughlin, ME, Tuveson, DA, Grimm, J, Lintault, L, Newman, J, Reczek, EE, Weissleder, R, Jacks, T: Restoration of p53 function leads to tumour regression in vivo. *Nature*, 445: 661–665, 2007.

49. Takehara, T, Tatsumi, T, Suzuki, T, Rucker, EB, Hennighausen, L, Jinushi, M, Miyagi, T, Kanazawa, Y, Hayashi, N: Hepatocyte-specific disruption of Bcl-xL leads to continuous hepatocyte apoptosis and liver fibrotic responses. *Gastroenterology*, 127: 1189–1197, 2004.

50. Wiles, KN, Alioto, CM, Hodge, NB, Clevenger, MH, Tsikretsis, LE, Lin, FTJ, Tétreault, M-P: IκB Kinase-β Regulates Neutrophil Recruitment Through Activation of STAT3 Signaling in the Esophagus. *Cell Mol Gastroenterol Hepatol*, S2352-345X(21)00152–1, 2021 (zur Publikation angenommen).

51. Quante, M, Marrache, F, Goldenring, JR, Wang, TC: TFF2 mRNA transcript expression marks a gland progenitor cell of the gastric oxyntic mucosa. *Gastroenterology*, 139: 2018-2027.e2, 2010.

52. Moro-Sibilot, L, Blanc, P, Taillardet, M, Bardel, E, Couillault, C, Boschetti, G, Traverse-Glehen, A, Defrance, T, Kaiserlian, D, Dubois, B: Mouse and Human Liver Contain Immunoglobulin A-Secreting Cells Originating From Peyer's Patches and Directed Against Intestinal Antigens. *Gastroenterology*, 151: 311–323, 2016.

53. Chen, L, Vasoya, RP, Toke, NH, Parthasarathy, A, Luo, S, Chiles, E, Flores, J, Gao, N, Bonder, EM, Su, X, Verzi, MP: HNF4 Regulates Fatty Acid Oxidation and Is Required for Renewal of Intestinal Stem Cells in Mice. *Gastroenterology*, 158: 985-999.e9, 2020.

54. Lee, JW, Komar, CA, Bengsch, F, Graham, K, Beatty, GL: Genetically Engineered Mouse Models of Pancreatic Cancer: The KPC Model (LSL-KrasG12D/+;LSL-Trp53R172H/+;Pdx-1-Cre), Its Variants and Their Application in Immuno-oncology Drug Discovery. *Curr Protoc Pharmacol*, 73: 14.39.1-14.39.20, 2016.

55. Hingorani, SR, Petricoin, EF, Maitra, A, Rajapakse, V, King, C, Jacobetz, MA, Ross, S, Conrads, TP, Veenstra, TD, Hitt, BA, Kawaguchi, Y, Johann, D, Liotta, LA, Crawford, HC, Putt, ME, Jacks, T, Wright, CVE, Hruban, RH, Lowy, AM, Tuveson, DA: Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer Cell*, 4: 437–450, 2003.

56. Bardeesy, N, Aguirre, AJ, Chu, GC, Cheng, K, Lopez, LV, Hezel, AF, Feng, B, Brennan, C, Weissleder, R, Mahmood, U, Hanahan, D, Redston, MS, Chin, L, DePinho, RA: Both p16Ink4a and the p19Arf-p53 pathway constrain progression of pancreatic adenocarcinoma in the mouse. *PNAS*, 103: 5947–5952, 2006.

57. Rhim, AD, Mirek, ET, Aiello, NM, Maitra, A, Bailey, JM, McAllister, F, Reichert, M, Beatty, GL, Rustgi, AK, Vonderheide, RH, Leach, SD, Stanger, BZ: EMT and dissemination precede pancreatic tumor formation. *Cell*, 148: 349–361, 2012.

58. Li, J, Byrne, KT, Yan, F, Yamazoe, T, Chen, Z, Baslan, T, Richman, LP, Lin, J, Sun, YH, Rech, AJ, Balli, D, Hay, CA, Sela, Y, Merrell, AJ, Liudahl, SM, Gordon, N, Norgard, RJ, Yuan, S, Yu, S, Chao, T, Ye, S, Eisinger-Mathason, TSK, Faryabi, RB, Tobias, JW, Lowe, S, Coussens, LM, Wherry, EJ, Vonderheide, RH, Stanger, BZ: Tumor cell-intrinsic factors underlie heterogeneity of immune cell infiltration and response to immunotherapy. *Immunity*, 49: 178-193.e7, 2018.

59. Ko, J, Bhagwat, N, Yee, SS, Black, T, Redlinger, C, Romeo, J, O'Hara, M, Raj, A, Carpenter, EL, Stanger, BZ, Issadore, D: A magnetic micropore chip for rapid (<1 hour) unbiased circulating tumor cell isolation and in situ RNA analysis. *Lab Chip*, 17: 3086–3096, 2017.

60. Simeonov, KP, Byrns, CN, Clark, ML, Norgard, RJ, Martin, B, Stanger, BZ, Shendure, J, McKenna, A, Lengner, CJ: Single-cell lineage tracing of metastatic cancer reveals selection of hybrid EMT states. *Cancer Cell*, S1535-6108(21)00271–3, 2021 (zur Publikation angenommen).

61. Butler, TP, Gullino, PM: Quantitation of cell shedding into efferent blood of mammary adenocarcinoma. *Cancer Res*, 35: 512–516, 1975.

62. Sharma, S, Zhuang, R, Long, M, Pavlovic, M, Kang, Y, Ilyas, A, Asghar, W: Circulating tumor cell isolation, culture, and downstream molecular analysis. *Biotechnol. Adv.*, 36: 1063–1078, 2018.

63. Yang, Y-P, Giret, TM, Cote, RJ: Circulating Tumor Cells from Enumeration to Analysis: Current Challenges and Future Opportunities. *Cancers (Basel)*, 13: 2723, 2021.

64. Jackson, JM, Witek, MA, Kamande, JW, Soper, SA: Materials and microfluidics: enabling the efficient isolation and analysis of circulating tumour cells. *Chem Soc Rev*, 46: 4245–4280, 2017.

65. Harouaka, R, Kang, Z, Zheng, S-Y, Cao, L: Circulating tumor cells: advances in isolation and analysis, and challenges for clinical applications. *Pharmacol Ther*, 141: 209–221, 2014.

66. Poudineh, M, Aldridge, PM, Ahmed, S, Green, BJ, Kermanshah, L, Nguyen, V, Tu, C, Mohamadi, RM, Nam, RK, Hansen, A, Sridhar, SS, Finelli, A, Fleshner, NE, Joshua, AM, Sargent, EH, Kelley, SO: Tracking the dynamics of circulating tumour cell phenotypes using nanoparticle-mediated magnetic ranking. *Nat Nanotechnol*, 12: 274–281, 2017.

67. Pestrin, M, Salvianti, F, Galardi, F, De Luca, F, Turner, N, Malorni, L, Pazzagli, M, Di Leo, A, Pinzani, P: Heterogeneity of PIK3CA mutational status at the single cell level in circulating tumor cells from metastatic breast cancer patients. *Mol Oncol*, 9: 749–757, 2015.

68. Zeinali, M, Lee, M, Nadhan, A, Mathur, A, Hedman, C, Lin, E, Harouaka, R, Wicha, MS, Zhao, L, Palanisamy, N, Hafner, M, Reddy, R, Kalemkerian, GP, Schneider, BJ, Hassan, KA, Ramnath, N, Nagrath, S: High-Throughput Label-Free Isolation of Heterogeneous Circulating Tumor Cells and CTC Clusters from Non-Small-Cell Lung Cancer Patients. *Cancers (Basel)*, 12: 2020.

69. Nanduri, LK, Hissa, B, Weitz, J, Schölch, S, Bork, U: The prognostic role of circulating tumor cells in colorectal cancer. *Expert Review of Anticancer Therapy*, 0: 1–12, 2019.

70. Song, BG, Kwon, W, Kim, H, Lee, EM, Han, YM, Kim, H, Byun, Y, Lee, KB, Lee, KH, Lee, KT, Lee, JK, Jang, J-Y, Park, JK: Detection of Circulating Tumor Cells in Resectable Pancreatic Ductal Adenocarcinoma: A Prospective Evaluation as a Prognostic Marker. *Front Oncol*, 10: 616440, 2021.

71. Pantel, K, Brakenhoff, RH, Brandt, B: Detection, clinical relevance and specific biological properties of disseminating tumour cells. *Nat Rev Cancer*, 8: 329–340, 2008.

72. Nagrath, S, Sequist, LV, Maheswaran, S, Bell, DW, Irimia, D, Ulkus, L, Smith, MR, Kwak, EL, Digumarthy, S, Muzikansky, A, Ryan, P, Balis, UJ, Tompkins, RG, Haber, DA, Toner, M: Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature*, 450: 1235–1239, 2007.

73. Varillas, JI, Zhang, J, Chen, K, Barnes, II, Liu, C, George, TJ, Fan, ZH: Microfluidic isolation of circulating tumor cells and cancer stem-like cells from patients with pancreatic ductal adenocarcinoma. *Theranostics*, 9: 1417–1425, 2019.

74. Effenberger, KE, Schroeder, C, Hanssen, A, Wolter, S, Eulenburg, C, Tachezy, M, Gebauer, F, Izbicki, JR, Pantel, K, Bockhorn, M: Improved Risk Stratification by Circulating Tumor Cell Counts in Pancreatic Cancer. *Clin Cancer Res*, 24: 2844–2850, 2018.

75. Hugenschmidt, H, Labori, KJ, Borgen, E, Brunborg, C, Schirmer, CB, Seeberg, LT, Naume, B, Wiedswang, G: Preoperative CTC-Detection by CellSearch® Is Associated with Early Distant Metastasis and Impaired Survival in Resected Pancreatic Cancer. *Cancers (Basel)*, 13: 485, 2021.

76. Barriere, G, Fici, P, Gallerani, G, Fabbri, F, Zoli, W, Rigaud, M: Circulating tumor cells and epithelial, mesenchymal and stemness markers: Characterization of cell subpopulations. *Annals of Translational Medicine*, 2: 109, 2014.

77. Nagrath, S, Jack, RM, Sahai, V, Simeone, DM: Opportunities and Challenges for Pancreatic Circulating Tumor Cells. *Gastroenterology*, 151: 412–426, 2016.

78. Semaan, A, Bernard, V, Kim, DU, Lee, JJ, Huang, J, Kamyabi, N, Stephens, BM, Qiao, W, Varadhachary, GR, Katz, MH, Shen, Y, San Lucas, FA, Gascoyne, P, Alvarez, HA, Maitra, A, Guerrero, PA: Characterisation of circulating tumour cell phenotypes identifies a partial-EMT sub-population for clinical stratification of pancreatic cancer. *Br J Cancer*, 124: 1970–1977, 2021.

79. Lustberg, MB, Balasubramanian, P, Miller, B, Garcia-Villa, A, Deighan, C, Wu, Y, Carothers, S, Berger, M, Ramaswamy, B, Macrae, ER, Wesolowski, R, Layman, RM, Mrozek, E, Pan, X, Summers, TA, Shapiro, CL, Chalmers, JJ: Heterogeneous atypical cell populations are present in blood of metastatic breast cancer patients. *Breast Cancer Res.*, 16: R23, 2014.

80. Bertolini, G, D'Amico, L, Moro, M, Landoni, E, Perego, P, Miceli, R, Gatti, L, Andriani, F, Wong, D, Caserini, R, Tortoreto, M, Milione, M, Ferracini, R, Mariani, L, Pastorino, U, Roato, I, Sozzi, G, Roz, L: Microenvironment-Modulated Metastatic CD133+/CXCR4+/EpCAM- Lung Cancer-Initiating Cells Sustain Tumor Dissemination and Correlate with Poor Prognosis. *Cancer Res.*, 75: 3636–3649, 2015.

81. Wen, K-C, Sung, P-L, Chou, Y-T, Pan, C-M, Wang, P-H, Lee, OK-S, Wu, C-W: The role of EpCAM in tumor progression and the clinical prognosis of endometrial carcinoma. *Gynecol. Oncol.*, 148: 383–392, 2018.

82. Pantel, K, Alix-Panabières, C, Riethdorf, S: Cancer micrometastases. *Nat Rev Clin Oncol*, 6: 339–351, 2009.

83. Pecot, CV, Bischoff, FZ, Mayer, JA, Wong, KL, Pham, T, Bottsford-Miller, J, Stone, RL, Lin, YG, Jaladurgam, P, Roh, JW, Goodman, BW, Merritt, WM, Pircher, TJ, Mikolajczyk, SD, Nick, AM, Celestino, J, Eng, C, Ellis, LM, Deavers, MT, Sood, AK: A novel platform for detection of CK+ and CK− CTCs. *Cancer Discov*, 1: 580–586, 2011.

84. Baccelli, I, Schneeweiss, A, Riethdorf, S, Stenzinger, A, Schillert, A, Vogel, V, Klein, C, Saini, M, Bäuerle, T, Wallwiener, M, Holland-Letz, T, Höfner, T, Sprick, M, Scharpff, M, Marmé, F, Sinn, HP, Pantel, K, Weichert, W, Trumpp, A: Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay. *Nature Biotechnology*, 31: 539–544, 2013.

85. Woelfle, U, Breit, E, Zafrakas, K, Otte, M, Schubert, F, Müller, V, Izbicki, JR, Löning, T, Pantel, K: Bi-specific immunomagnetic enrichment of micrometastatic tumour cell clusters from bone marrow of cancer patients. *J Immunol Methods*, 300: 136–145, 2005.

86. Rosenberg, R, Gertler, R, Friederichs, J, Fuehrer, K, Dahm, M, Phelps, R, Thorban, S, Nekarda, H, Siewert, JR: Comparison of two density gradient centrifugation systems for the enrichment of disseminated tumor cells in blood. *Cytometry*, 49: 150–158, 2002.

87. Campton, DE, Ramirez, AB, Nordberg, JJ, Drovetto, N, Clein, AC, Varshavskaya, P, Friemel, BH, Quarre, S, Breman, A, Dorschner, M, Blau, S, Blau, CA, Sabath, DE, Stilwell, JL, Kaldjian, EP: High-recovery visual identification and single-cell retrieval of circulating tumor cells for genomic analysis using a dual-technology platform integrated with automated immunofluorescence staining. *BMC Cancer*, 15: 360, 2015.

88. Hao, S-J, Wan, Y, Xia, Y-Q, Zou, X, Zheng, S-Y: Size-based separation methods of circulating tumor cells. *Adv Drug Deliv Rev*, 125: 3–20, 2018.

89. Pinzani, P, Salvadori, B, Simi, L, Bianchi, S, Distante, V, Cataliotti, L, Pazzagli, M, Orlando, C: Isolation by size of epithelial tumor cells in peripheral blood of patients with breast cancer: correlation with real-time reverse transcriptase-polymerase chain reaction results and feasibility of molecular analysis by laser microdissection. *Hum Pathol*, 37: 711–718, 2006.

90. Zheng, S, Lin, H, Liu, J-Q, Balic, M, Datar, R, Cote, RJ, Tai, Y-C: Membrane microfilter device for selective capture, electrolysis and genomic analysis of human circulating tumor cells. *J Chromatogr A*, 1162: 154–161, 2007.

91. Desitter, I, Guerrouahen, BS, Benali-Furet, N, Wechsler, J, Jänne, PA, Kuang, Y, Yanagita, M, Wang, L, Berkowitz, JA, Distel, RJ, Cayre, YE: A New Device for Rapid Isolation by Size and Characterization of Rare Circulating Tumor Cells. *Anticancer Research*, 31: 427–441, 2011.

92. Gambari, R, Borgatti, M, Altomare, L, Manaresi, N, Medoro, G, Romani, A, Tartagni, M, Guerrieri, R: Applications to cancer research of „lab-on-a-chip“ devices based on dielectrophoresis (DEP). *Technol Cancer Res Treat*, 2: 31–40, 2003.

93. Trapani, MD, Manaresi, N, Medoro, G: DEPArray™ system: An automatic image-based sorter for isolation of pure circulating tumor cells. *Cytometry Part A*, 93: 1260–1266, 2018.

94. Dong, Y, Skelley, AM, Merdek, KD, Sprott, KM, Jiang, C, Pierceall, WE, Lin, J, Stocum, M, Carney, WP, Smirnov, DA: Microfluidics and circulating tumor cells. *J Mol Diagn*, 15: 149–157, 2013.

95. Stott, SL, Hsu, C-H, Tsukrov, DI, Yu, M, Miyamoto, DT, Waltman, BA, Rothenberg, SM, Shah, AM, Smas, ME, Korir, GK, Floyd, FP, Gilman, AJ, Lord, JB, Winokur, D, Springer, S, Irimia, D, Nagrath, S, Sequist, LV, Lee, RJ, Isselbacher, KJ, Maheswaran, S, Haber, DA, Toner, M: Isolation of circulating tumor cells using a microvortex-generating herringbone-chip. *PNAS*, 107: 18392–18397, 2010.

96. Ozkumur, E, Shah, AM, Ciciliano, JC, Emmink, BL, Miyamoto, DT, Brachtel, E, Yu, M, Chen, P, Morgan, B, Trautwein, J, Kimura, A, Sengupta, S, Stott, SL, Karabacak, NM, Barber, TA, Walsh, JR, Smith, K, Spuhler, PS, Sullivan, JP, Lee, RJ, Ting, DT, Luo, X, Shaw, AT, Bardia, A, Sequist, LV, Louis, DN, Maheswaran, S, Kapur, R, Haber, DA, Toner, M: Inertial focusing for tumor antigen-dependent and -independent sorting of rare circulating tumor cells. *Sci Transl Med*, 5: 179ra47, 2013.

97. Xu, L, Mao, X, Imrali, A, Syed, F, Mutsvangwa, K, Berney, D, Cathcart, P, Hines, J, Shamash, J, Lu, Y-J: Optimization and Evaluation of a Novel Size Based Circulating Tumor Cell Isolation System. *PLoS One*, 10: e0138032, 2015.

98. Kulemann, B, Pitman, MB, Liss, AS, Valsangkar, N, Fernández-Del Castillo, C, Lillemoe, KD, Hoeppner, J, Mino-Kenudson, M, Warshaw, AL, Thayer, SP: Circulating tumor cells found in patients with localized and advanced pancreatic cancer. *Pancreas*, 44: 547–550, 2015.

99. Martini, V, Timme-Bronsert, S, Fichtner-Feigl, S, Hoeppner, J, Kulemann, B: Circulating Tumor Cells in Pancreatic Cancer: Current Perspectives. *Cancers (Basel)*, 11: 2019.

100. Amantini, C, Morelli, MB, Nabissi, M, Piva, F, Marinelli, O, Maggi, F, Bianchi, F, Bittoni, A, Berardi, R, Giampieri, R, Santoni, G: Expression Profiling of Circulating Tumor Cells in Pancreatic Ductal Adenocarcinoma Patients: Biomarkers Predicting Overall Survival. *Front Oncol*, 9: 874, 2019.

101. Hugenschmidt, H, Labori, KJ, Brunborg, C, Verbeke, CS, Seeberg, LT, Bendigtsen Schirmer, C, Renolen, A, Borgen, E, Naume, B, Wiedswang, G: Cytokeratin-positive cells in the bone marrow from patients with pancreatic, periampullary malignancy and benign pancreatic disease show no prognostic information. *BMC Cancer*, 20: 1107, 2020.

102. Khoja, L, Backen, A, Sloane, R, Menasce, L, Ryder, D, Krebs, M, Board, R, Clack, G, Hughes, A, Blackhall, F, Valle, JW, Dive, C: A pilot study to explore circulating tumour cells in pancreatic cancer as a novel biomarker. *Br J Cancer*, 106: 508–516, 2012.

103. Bidard, FC, Huguet, F, Louvet, C, Mineur, L, Bouché, O, Chibaudel, B, Artru, P, Desseigne, F, Bachet, JB, Mathiot, C, Pierga, JY, Hammel, P: Circulating tumor

cells in locally advanced pancreatic adenocarcinoma: the ancillary CirCe 07 study to the LAP 07 trial. *Ann. Oncol.*, 24: 2057–2061, 2013.

104. Poruk, KE, Valero, V, Saunders, T, Blackford, AL, Griffin, JF, Poling, J, Hruban, RH, Anders, RA, Herman, J, Zheng, L, Rasheed, ZA, Laheru, DA, Ahuja, N, Weiss, MJ, Cameron, JL, Goggins, M, Iacobuzio-Donahue, CA, Wood, LD, Wolfgang, CL: Circulating Tumor Cell Phenotype Predicts Recurrence and Survival in Pancreatic Adenocarcinoma. *Ann Surg*, 264: 1073–1081, 2016.

105. Park, Y, Jun, HR, Choi, HW, Hwang, DW, Lee, JH, Song, KB, Lee, W, Kwon, J, Ha, SH, Jun, E, Kim, SC: Circulating tumour cells as an indicator of early and systemic recurrence after surgical resection in pancreatic ductal adenocarcinoma. *Sci Rep*, 11: 1644, 2021.

106. Brady, G, Barbara, M, Iscove, N: Representative in Vitro cDNA Amplification From Individual Hemopoietic Cells and Colonies. *Methods Mol Biol*, 2: 17–25, 1990.

107. Tang, F, Barbacioru, C, Wang, Y, Nordman, E, Lee, C, Xu, N, Wang, X, Bodeau, J, Tuch, BB, Siddiqui, A, Lao, K, Surani, MA: mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*, 6: 377–382, 2009.

108. Tang, F, Barbacioru, C, Bao, S, Lee, C, Nordman, E, Wang, X, Lao, K, Surani, MA: Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell*, 6: 468–478, 2010.

109. Ramsköld, D, Luo, S, Wang, Y-C, Li, R, Deng, Q, Faridani, OR, Daniels, GA, Khrebtukova, I, Loring, JF, Laurent, LC, Schroth, GP, Sandberg, R: Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol*, 30: 777–782, 2012.

110. Guo, G, Huss, M, Tong, GQ, Wang, C, Li Sun, L, Clarke, ND, Robson, P: Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev Cell*, 18: 675–685, 2010.

111. Sasagawa, Y, Nikaido, I, Hayashi, T, Danno, H, Uno, KD, Imai, T, Ueda, HR: Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol*, 14: R31, 2013.

112. Hashimshony, T, Wagner, F, Sher, N, Yanai, I: CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep*, 2: 666–673, 2012.

113. Jaitin, DA, Kenigsberg, E, Keren-Shaul, H, Elefant, N, Paul, F, Zaretsky, I, Mildner, A, Cohen, N, Jung, S, Tanay, A, Amit, I: Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science*, 343: 776–779, 2014.

114. Picelli, S, Faridani, OR, Björklund, AK, Winberg, G, Sagasser, S, Sandberg, R: Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc*, 9: 171–181, 2014.

115. Macosko, EZ, Basu, A, Satija, R, Nemesh, J, Shekhar, K, Goldman, M, Tirosh, I, Bialas, AR, Kamitaki, N, Martersteck, EM, Trombetta, JJ, Weitz, DA, Sanes, JR, Shalek, AK, Regev, A, McCarroll, SA: Highly Parallel Genome-wide

Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, 161: 1202–1214, 2015.

116. Klein, AM, Mazutis, L, Akartuna, I, Tallapragada, N, Veres, A, Li, V, Peshkin, L, Weitz, DA, Kirschner, MW: Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, 161: 1187–1201, 2015.

117. Rostom, R, Svensson, V, Teichmann, SA, Kar, G: Computational approaches for interpreting scRNA-seq data. *FEBS Lett*, 591: 2213–2225, 2017.

118. Satija, R, Farrell, JA, Gennert, D, Schier, AF, Regev, A: Spatial reconstruction of single-cell gene expression data (Bandiera_abtest: aCg_type: Nature Research Journalsnumber: 5Primary_atype: Researchpublisher: Nature Publishing GroupSubject_term: Gastrulation;Machine learning;Statistical methodsSubject_term_id: gastrulation;machine-learning;statistical-methods). *Nat Biotechnol*, 33: 495–502, 2015.

119. McCarthy, DJ, Campbell, KR, Lun, ATL, Wills, QF: Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics*, 33: 1179–1186, 2017.

120. Wolf, FA, Angerer, P, Theis, FJ: SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*, 19: 15, 2018.

121. Danopoulos, S, Bhattacharya, S, Mariani, TJ, Al Alam, D: Transcriptional characterisation of human lung cells identifies novel mesenchymal lineage markers. *Eur Respir J*, 55: 1900746, 2020.

122. Zhao, X, Wu, S, Fang, N, Sun, X, Fan, J: Evaluation of single-cell classifiers for single-cell RNA sequencing data sets. *Brief Bioinform*, 21: 1581–1595, 2020.

123. Zhao, G, Lu, H, Chang, Z, Zhao, Y, Zhu, T, Chang, L, Guo, Y, Garcia-Barrio, MT, Chen, YE, Zhang, J: Single-cell RNA sequencing reveals the cellular heterogeneity of aneurysmal infrarenal abdominal aorta. *Cardiovascular Research*, cvaa214, 2020 (zur Publikation angenommen).

124. Luecken, MD, Theis, FJ: Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol*, 15: e8746, 2019.

125. Moignard, V, Woodhouse, S, Haghverdi, L, Lilly, AJ, Tanaka, Y, Wilkinson, AC, Buettner, F, Macaulay, IC, Jawaid, W, Diamanti, E, Nishikawa, S-I, Piterman, N, Kouskoff, V, Theis, FJ, Fisher, J, Göttgens, B: Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nat Biotechnol*, 33: 269–276, 2015.

126. Haghverdi, L, Büttner, M, Wolf, FA, Buettner, F, Theis, FJ: Diffusion pseudotime robustly reconstructs lineage branching. *Nat Methods*, 13: 845–848, 2016.

127. Trapnell, C, Cacchiarelli, D, Grimsby, J, Pokharel, P, Li, S, Morse, M, Lennon, NJ, Livak, KJ, Mikkelsen, TS, Rinn, JL: The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*, 32: 381–386, 2014.

128. Bendall, SC, Davis, KL, Amir, E-AD, Tadmor, MD, Simonds, EF, Chen, TJ, Shenfeld, DK, Nolan, GP, Pe'er, D: Single-cell trajectory detection uncovers

progression and regulatory coordination in human B cell development. *Cell*, 157: 714–725, 2014.

129. Aceto, N, Bardia, A, Miyamoto, DT, Donaldson, MC, Wittner, BS, Spencer, JA, Yu, M, Pely, A, Engstrom, A, Zhu, H, Brannigan, BW, Kapur, R, Stott, SL, Shioda, T, Ramaswamy, S, Ting, DT, Lin, CP, Toner, M, Haber, DA, Maheswaran, S: Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell*, 158: 1110–1122, 2014.

130. Szczerba, BM, Castro-Giner, F, Vetter, M, Krol, I, Gkountela, S, Landin, J, Scheidmann, MC, Donato, C, Scherrer, R, Singer, J, Beisel, C, Kurzeder, C, Heinzelmann-Schwarz, V, Rochlitz, C, Weber, WP, Beerenwinkel, N, Aceto, N: Neutrophils escort circulating tumour cells to enable cell cycle progression. *Nature*, 566: 553–557, 2019.

131. Cheng, Y-H, Chen, Y-C, Lin, E, Brien, R, Jung, S, Chen, Y-T, Lee, W, Hao, Z, Sahoo, S, Min Kang, H, Cong, J, Burness, M, Nagrath, S, S. Wicha, M, Yoon, E: Hydro-Seq enables contamination-free high-throughput single-cell RNA-sequencing for circulating tumor cells (Bandiera_abtest: aCc_license_type: cc_byCg_type: Nature Research Journalsnumber: 1Primary_atype: Researchpublisher: Nature Publishing GroupSubject_term: Cancer stem cells;Gene expression analysis;Lab-on-a-chip;Metastasis;SequencingSubject_term_id: cancer-stem-cells;gene-expression-analysis;lab-on-a-chip;metastases;sequencing). *Nat Commun*, 10: 2163, 2019.

132. Ruan, H, Zhou, Y, Shen, J, Zhai, Y, Xu, Y, Pi, L, Huang, R, Chen, K, Li, X, Ma, W, Wu, Z, Deng, X, Wang, X, Zhang, C, Guan, M: Circulating tumor cell characterization of lung cancer brain metastases in the cerebrospinal fluid through single-cell transcriptome analysis. *Clin Transl Med*, 10: e246, 2020.

133. Ting, DT, Wittner, BS, Ligorio, M, Jordan, NV, Shah, AM, Miyamoto, DT, Aceto, N, Bersani, F, Brannigan, BW, Xega, K, Ciciliano, JC, Zhu, H, MacKenzie, OC, Trautwein, J, Arora, KS, Shahid, M, Ellis, HL, Qu, N, Bardeesy, N, Rivera, MN, Deshpande, V, Ferrone, CR, Kapur, R, Ramaswamy, S, Shioda, T, Toner, M, Maheswaran, S, Haber, DA: Single-Cell RNA Sequencing Identifies Extracellular Matrix Gene Expression by Pancreatic Circulating Tumor Cells. *Cell Rep*, 8: 1905–1918, 2014.

134. McKenna, A, Findlay, GM, Gagnon, JA, Horwitz, MS, Schier, AF, Shendure, J: Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science*, 353: aaf7907, 2016.

135. Pei, W, Shang, F, Wang, X, Fanti, A-K, Greco, A, Busch, K, Klapproth, K, Zhang, Q, Quedenau, C, Sauer, S, Feyerabend, TB, Höfer, T, Rodewald, H-R: Resolving Fates and Single-Cell Transcriptomes of Hematopoietic Stem Cell Clones by PolyloxExpress Barcoding. *Cell Stem Cell*, 27: 383-395.e8, 2020.

136. Bowling, S, Sritharan, D, Osorio, FG, Nguyen, M, Cheung, P, Rodriguez-Fraticelli, A, Patel, S, Yuan, W-C, Fujiwara, Y, Li, BE, Orkin, SH, Hormoz, S, Camargo, FD: An Engineered CRISPR-Cas9 Mouse Line for Simultaneous Readout of Lineage Histories and Gene Expression Profiles in Single Cells. *Cell*, 181: 1410-1422.e27, 2020.

137. Rahbari, NN, Bork, U, Schölch, S, Reissfelder, C, Thorlund, K, Betzler, A, Kahlert, C, Schneider, M, Ulrich, AB, Büchler, MW, Weitz, J, Koch, M: Metastatic Spread Emerging From Liver Metastases of Colorectal Cancer: Does the Seed Leave the Soil Again? *Ann. Surg.*, 263: 345–352, 2016.

138. Kochall, S, Thepkaysone, M-L, García, SA, Betzler, AM, Weitz, J, Reissfelder, C, Schölch, S: Isolation of Circulating Tumor Cells in an Orthotopic Mouse Model of Colorectal Cancer. *J Vis Exp*, 125: 2017.

139. García, SA, Weitz, J, Schölch, S: Circulating Tumor Cells. *Methods Mol Biol*, 1692: 213–219, 2018.

140. Franses, JW, Philipp, J, Missios, P, Bhan, I, Liu, A, Yashaswini, C, Tai, E, Zhu, H, Ligorio, M, Nicholson, B, Tassoni, EM, Desai, N, Kulkarni, AS, Szabolcs, A, Hong, TS, Liss, AS, Fernandez-Del Castillo, C, Ryan, DP, Maheswaran, S, Haber, DA, Daley, GQ, Ting, DT: Pancreatic circulating tumor cell profiling identifies LIN28B as a metastasis driver and drug target. *Nat Commun*, 11: 3303, 2020.

141. Zhu, L, Kan, K-J, Grün, JL, Hissa, B, Yang, C, Győrffy, B, Loges, S, Reißfelder, C, Schölch, S: GAS2L1 Is a Potential Biomarker of Circulating Tumor Cells in Pancreatic Cancer. *Cancers (Basel)*, 12: E3774, 2020.

142. Zhu, L, Hissa, B, Győrffy, B, Jann, J-C, Yang, C, Reissfelder, C, Schölch, S: Characterization of Stem-like Circulating Tumor Cells in Pancreatic Cancer. *Diagnostics (Basel)*, 10: 2020.

143. Lapin, M, Tjensvoll, K, Oltedal, S, Javle, M, Smaaland, R, Gilje, B, Nordg\a ard, O: Single-cell mRNA profiling reveals transcriptional heterogeneity among pancreatic circulating tumour cells. *BMC Cancer*, 17: 1–10, 2017.

144. Massagué, J, Obenauf, AC: Metastatic colonization by circulating tumour cells. *Nature*, 529: 298–306, 2016.

145. Maheswaran, S, Haber, DA: Ex Vivo Culture of CTCs: An Emerging Resource to Guide Cancer Therapy. *Cancer Res*, 75: 2411–2415, 2015.

146. Yu, M, Bardia, A, Aceto, N, Bersani, F, Madden, MW, Donaldson, MC, Desai, R, Zhu, H, Comaills, V, Zheng, Z, Wittner, BS, Stojanov, P, Brachtel, E, Sgroi, D, Kapur, R, Shioda, T, Ting, DT, Ramaswamy, S, Getz, G, Iafrate, AJ, Benes, C, Toner, M, Maheswaran, S, Haber, DA: Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. *Science*, 345: 216–220, 2014.

147. Ting, DT, Wittner, BS, Ligorio, M, Vincent Jordan, N, Shah, AM, Miyamoto, DT, Aceto, N, Bersani, F, Brannigan, BW, Xega, K, Ciciliano, JC, Zhu, H, MacKenzie, OC, Trautwein, J, Arora, KS, Shahid, M, Ellis, HL, Qu, N, Bardeesy, N, Rivera, MN, Deshpande, V, Ferrone, CR, Kapur, R, Ramaswamy, S, Shioda, T, Toner, M, Maheswaran, S, Haber, DA: Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells. *Cell Rep*, 8: 1905–1918, 2014.

148. Heikenwalder, M, Lorentzen, A: The role of polarisation of circulating tumour cells in cancer metastasis. *Cell. Mol. Life Sci.*, 76: 3765–3781, 2019.

149. Luxton, GWG, Gundersen, GG: Orientation and Function of the Nuclear-Centrosomal Axis During Cell Migration. *Curr Opin Cell Biol*, 23: 579–588, 2011.

150. Rowley, JW, Oler, AJ, Tolley, ND, Hunter, BN, Low, EN, Nix, DA, Yost, CC, Zimmerman, GA, Weyrich, AS: Genome-wide RNA-seq analysis of human and mouse platelet transcriptomes. *Blood*, 118: e101-111, 2011.

151. Hoffmeyer, K, Raggioli, A, Rudloff, S, Anton, R, Hierholzer, A, Del Valle, I, Hein, K, Vogt, R, Kemler, R: Wnt/β-catenin signaling regulates telomerase in stem cells and cancer cells. *Science*, 336: 1549–1554, 2012.

152. Yachida, S, Jones, S, Bozic, I, Antal, T, Leary, R, Fu, B, Kamiyama, M, Hruban, RH, Eshleman, JR, Nowak, MA, Velculescu, VE, Kinzler, KW, Vogelstein, B, Iacobuzio-Donahue, CA: Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature*, 467: 1114–1117, 2010.

153. Keleg, S, Büchler, P, Ludwig, R, Büchler, MW, Friess, H: Invasion and metastasis in pancreatic cancer. *Mol Cancer*, 2: 14, 2003.

154. Collisson, EA, Maitra, A: Pancreatic Cancer Genomics 2.0: Profiling Metastases. *Cancer Cell*, 31: 309–310, 2017.

155. Vakoc, CR, Tuveson, DA: Soils and Seeds That Initiate Pancreatic Cancer Metastasis. *Cancer Discov*, 7: 1067–1068, 2017.

156. Aiello, NM, Brabletz, T, Kang, Y, Nieto, MA, Weinberg, RA, Stanger, BZ: Upholding a role for EMT in pancreatic cancer metastasis. *Nature*, 547: E7–E8, 2017.

157. Yu, J, Gemenetzis, G, Kinny-Köster, B, Habib, JR, Groot, VP, Teinor, J, Yin, L, Pu, N, Hasanain, A, van Oosten, F, Javed, AA, Weiss, MJ, Burkhart, RA, Burns, WR, Goggins, M, He, J, Wolfgang, CL: Pancreatic circulating tumor cell detection by targeted single-cell next-generation sequencing. *Cancer Lett*, 493: 245–253, 2020.

158. Charles Jacob, HK, Charles Richard, JL, Signorelli, R, Kashuv, T, Lavania, S, Vaish, U, Boopathy, R, Middleton, A, Boone, MM, Sundaram, R, Dudeja, V, Saluja, AK: Modulation of Early Neutrophil Granulation: The Circulating Tumor Cell-Extravesicular Connection in Pancreatic Ductal Adenocarcinoma. *Cancers (Basel)*, 13: 2727, 2021.

159. Uchikura, K, Takao, S, Nakajo, A, Miyazono, F, Nakashima, S, Tokuda, K, Matsumoto, M, Shinchi, H, Natsugoe, S, Aikou, T: Intraoperative molecular detection of circulating tumor cells by reverse transcription-polymerase chain reaction in patients with biliary-pancreatic cancer is associated with hematogenous metastasis. *Ann Surg Oncol*, 9: 364–370, 2002.

160. Lukyanchuk, VV, Friess, H, Kleeff, J, Osinsky, SP, Ayuni, E, Candinas, D, Roggo, A: Detection of circulating tumor cells by cytokeratin 20 and prostate stem cell antigen RT-PCR in blood of patients with gastrointestinal cancers. *Anticancer Res*, 23: 2711–2716, 2003.

6 CURRICULUM VITAE

PERSONAL INFORMATION

Last name, first name:	Zhu, Lei
Date of birth:	02. August 1992
Place of birth:	Jiangsu (PR China)

SCHOOL CAREER

2007 – 2010	Jiangsu Yancheng Middle School

UNIVERSITY CAREER

2010 – 2015	Nanjing Medical University
	Bachelor of Medicine
2015 – 2018	Fudan University
	Master of Clinical Medicine
02.12.2016	Physician license (China)
15.11.2018	Certificate of resident standardization training (Surgery) (China)
Since WS2018/19	Doctoral Student, University of Heidelberg, Germany

7 ACKNOWLEDGMENTS

"Promotion" is a German word used to define a doctoral career in Germany, and I believe in and perform it firmly. As a preamble of scientific research, it requests me to promote myself to think about scientific topics independently; knowing why, what, and how.

I sincerely thank my supervisor, Prof. Dr. med. Sebastian Schölch. He gave me a chance and opened the gate of research for me. It's impressive that he reviewed my paper for eleven rounds with myriad corrections and recommendations until I published my first paper. My projects cannot move on successfully without his sustained support. He spared his time, even on holidays or vacations, to meet with me, discuss the projects. It's my great honor to follow him in the past years. I believe this experience will be a priceless treasure in my scientific career.

I really appreciate Marzena Knyssok-Sypniewski and Lukas Schmitt for their thoughtful technical support. I also would like to thank Kejia Kan, Qianwen Zhou, and other Chinese colleagues in the laboratory. They gave me help with various troubleshooting.

My gratitude to my wife, Nina Wang, is beyond words. She made a huge personal sacrifice and enormous contributions to our family. Due to the COVID-19 pandemic and inconvenient international travel, I was absent during all of her pregnancy and nursing. Thanks to my parents, they took care of my wife and child carefully and supported me abidingly.

Last but not least, I would like to express my sincere thanks to the China Scholarship Council (CSC), which supported me uninterruptedly for three years. I am so proud that I have a strong country behind me.

Thank you very much to all the people I met here!