Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

Kristin Decker, M.Sc.

born in: Bayreuth (Germany)

Oral examination: 04.09.2019

Characterization of disseminated breast cancer cells in kidney, pancreas and spleen

Referees: Prof Dr Andreas Trumpp

Prof Dr med. Claus R. Bartram

Curiosity

- the rover and the concept –

is what science is all about: the quest to reveal the unknown.

Ahmed Zewail

Declaration

The work presented in this thesis was performed from February 2015 until June 2019 in the junior research group ‘Metastatic Niches’ at the German Cancer Research Center (DKFZ), Heidelberg as well as in the Heidelberg Institute for Stem Cell Research and Experimental Medicine (HI-STEM gGmbH) under the supervision of Dr Thordur Oskarsson.

Declarations according to § 8 (3) c), d) and h) of the doctoral degree regulations:

c) I hereby declare that I have written the submitted dissertation myself and in this process have used no other sources or materials than those explicitly indicated.

d) I hereby declare that I have not applied to be examined at any other institution, nor have I used the dissertation in this way or any other form at any other institution as an examination paper, nor submitted it to any other faculty as dissertation.

h) I hereby consent to the verification of the dissertation by means of electronic data processing programs against standing scientific standards

Table of contents

1 Summary

In 2018 about 2 million people died from cancer in Europe alone, with 90% of cancer related deaths being due to metastatic spread of the disease. Despite the high impact of metastasis for patients` survival, the molecular mechanisms of metastatic progression are still poorly understood and therapeutic opportunities are limited. Metastasis can occur months to years after removal of the primary tumor, rendering potential therapeutic interventions even more difficult. This suggests that therapy resistant cancer cells can survive in the body, able to initiate metastasis after latency periods. Our knowledge of the molecular mechanisms enabling these cells to survive in unfavorable microenvironments is limited and specific therapeutic targeting of these cells is currently not possible.

In this study breast cancer dissemination was investigated in orthotopic *in vivo* models using different human breast cancer cell lines as well as patient derived xenograft (PDX) models. Breast cancer cells were not only detected in lung, liver and bone, organs that are prone to develop breast cancer metastases, but also in kidney, pancreas and spleen. These organs rarely develop metastases in patients and show no metastatic growth in our models. The integrity and viability of those cells was confirmed by *in vitro* cultivation of cancer cells, isolated from different organs. Thus, it was hypothesized that disseminated cancer cells (DCCs) in kidney, pancreas and spleen may resemble a subpopulation of cancer cells in a metastatic latency period.

To further investigate this hypothesis the *in vivo* location of DCCs in different organs, their survival capability as well as growth potential was further analyzed. It was confirmed that the majority of DCCs in the kidney had extravasated, but stayed non-proliferative and were located in close proximity to blood vessels. Furthermore, DCCs from PDX models, located in kidney, pancreas and spleen, survive for a prolonged time period without significant growth after resection of the mammary tumor. DCCs isolated from kidney, pancreas and spleen maintain their growth potential and were able to initiate metastatic growth in the lung upon intravenous injection. While being able to form lung metastases, injected DCCs still do not grow in their organ, originally isolated from. This suggests that not an aggressive subpopulation was selected but that the microenvironment influenced the growth pattern of these cells. Taken together, it was demonstrated that breast cancer DCCs survive in the non-supportive organ environments of kidney, pancreas and spleen without growth and maintain their metastatic potential.

To identify survival mechanisms applied by DCCs, gene expression profiling of DCCs isolated from kidney and pancreas compared to cancer cells from the mammary tumor was performed, analyzing the transcriptional profile of DCCs being distinctly different from cells of the mammary tumor. However, substantial similarities were identified between DCCs in kidney and pancreas. These observations suggested that DCCs despite their different organ of isolation may use similar survival mechanisms in unfavorable microenvironments. Based on the similarities of DDCs from kidney and pancreas, it was hypothesized that mutual transcriptomic changes in kidney and pancreas DCCs may be regulated intrinsically. A second gene expression profile of cancer cells from kidney, pancreas, lung and mammary tumor, having been cultured *in vitro* for 48h, revealed that a significant percentage of the transcriptomic changes in DCCs were maintained upon short-term culture, supporting the hypothesis that these changes are regulated cancer cell intrinsically. Furthermore, DCCs from kidney and pancreas have a different

transcriptomic profile compared to lung derived metastatic cells, suggesting that DCCs resemble a latent subpopulation in the metastatic cascade.

As the aim was to identify survival cues of DCCs that may be used therapeutically, the common signaling pathways and functions in kidney and pancreas DCCs were analyzed in more detail. Using Gene Set Enrichment Analysis (GSEA) and Gene Ontology (GO) analysis molecular mechanisms and pathways were identified. Mechanisms of epithelial to mesenchymal transition (EMT) as well as cell death mechanisms were repressed in DCCs from both organs. A reduced rate of apoptotic events in DCCs of the kidney was confirmed by cleaved caspase 3 staining compared to lung metastasis and cells from the mammary tumor.

Furthermore, based on the gene expression profiles, pronounced metabolic changes were observed in DCCs. Genes involved in metabolic pathways including energy, glucose and nucleotide metabolism were downregulated in DCCs. A closer look revealed that DCCs underwent a metabolic switch from glycolysis to oxidative phosphorylation. Genes involved in glucose import into the cell as well as enzymes of the glycolytic pathway were repressed at the transcriptional level. In contrast, the expression of genes, encoding for proteins of the electron transport chain in mitochondria, was upregulated in DCCs. These metabolic changes were accompanied by a downregulation of hypoxia response genes. This downregulation may happen in context of the metabolic switch from glycolysis to oxidative phosphorylation as hypoxia leads to induction of glycolysis as several of the glycolytic enzymes are direct target genes of the hypoxia induced transcription factor HIF1.

Antigen presentation mediated by major histocompatibility complex II (MHCII) molecules was repressed on mRNA and on protein level in DCCs as well. The repression of antigen presentation is most likely caused by cancer cell autonomous mechanisms as the experiments were performed in immune compromised mouse models. Further investigations suggested that YAP1 regulated genes, which were upregulated in DCCs, may negatively regulate MHC II gene expression. *In vitro* experiments confirmed that YAP1 can regulate MHC II transcript as well as protein levels.

Moreover, DNA repair mechanisms and cell cycle checkpoint genes were enriched in DCCs from kidney and pancreas.

Lastly, the transcriptomic profile of DCCs from kidney and pancreas showed enrichment for chemotherapy resistance signatures. Following up on this enrichment, the therapy resistance of DCCs was further investigated *in vivo* using two different chemotherapeutic treatments, a combination of Doxorubicin and Cyclophosphamide as well as monotherapy with Paclitaxel. While cancer cells in the mammary tumor and metastatic cells in the lung responded well to the treatment, DCCs in kidney, pancreas and spleen were not affected and the number of cancer cell in these organs remained practically unchanged. Thus, in this model breast cancer DCCs in kidney, pancreas and spleen resemble a pool of cancer cells able to survive therapy.

Due to the central role of chemotherapy resistance for the outcome of therapeutic interventions, the chemotherapy resistance mechanisms of DCCs were analyzed in more detail. The chemotherapy resistance signature, enriched in DCCs, was compared to the genes upregulated in kidney and pancreas DCCs to identify overlaps. The comparison revealed an overlap of the tetraspanin (TSPAN) gene family members TSPAN8 and TSPAN1 in all three datasets. To investigate the importance of the identified molecules, the fractions of TSPAN expressing cancer cells were analyzed in a mouse model as well as in patient samples. The TSPAN8 single as well

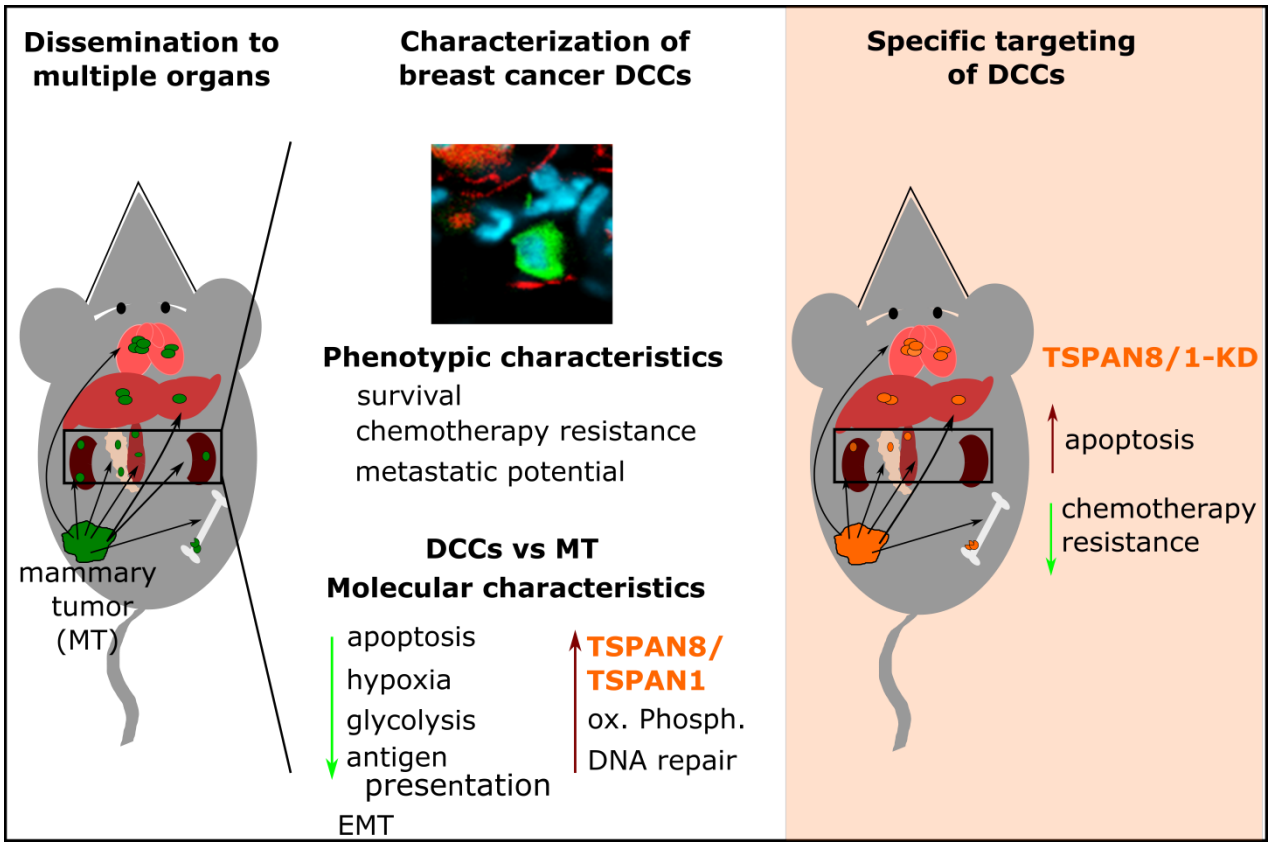
as TSPAN8 and TSPAN1 double-positive cell population was increased in DCCs in kidney, pancreas and spleen compared to the mammary tumor. Furthermore, TSPAN8 single positive as well as TSPAN8 and TPSAN1 co-expressing cancer cells were detected in ascites and pleural effusion samples of three out of four breast cancer patients by FACS analysis.

As survival without growth was identified as an essential characteristic of DCCs from kidney, pancreas and spleen, the role of TSPAN8 and TSPAN1 for survival was investigated. Under sphere forming conditions *in vitro,* knockdown of TSPAN8 and TSPAN1 resulted in an increased rate of apoptosis. Furthermore, TSPAN8 and TSPAN1 co-expressing cells were enriched in the label retaining, non-proliferative population of spheres. Due to the context dependency of TSPAN functions, the role of TSPAN8 and TSPAN1 was also analyzed in DCCs. TSPAN8 and TSPAN1 knockdown were injected in the mammary fat pad, DCC were isolated and transcriptomic analysis was performed compared to control cells. This experiment confirmed an association of TSPAN8 and TSPAN1 with chemotherapy resistance of DCCs as the therapy resistance signature enriched in DCCs was lost upon knockdown of TSPANs. The transcriptional data from TSPAN8 and TSPAN1 knockdown DCCs compared to control DCCs indicated further that TSPAN8 and TSPAN1 may also regulate stem cell properties in DCCs *in vivo*.

Functionally, *in vivo* injections of TSPAN8 and TSPAN1 double knockdown cells confirmed the crucial role of these two molecules for survival and chemotherapy resistance of DCCs. The knockdown significantly reduced the number of cancer cells in pancreas and spleen. In addition, knockdown of TSPAN8 and TSPAN1 sensitized DCCs to chemotherapy. A combination of the knockdown with Doxorubicin and Cyclophosphamide treatment resulted in a significant reduction of the cancer cell numbers in kidney, pancreas and spleen *in vivo* compared to chemotherapy alone.

In conclusion, this study provides insights into the biology of dormant disseminated breast cancer cells and may reveal novel opportunities to develop therapeutic strategies against dormant DCCs during metastatic latency periods. Depletion of TSPAN1 and TSPAN8 rendered DCCs sensitive to chemotherapy. Thus, the combination of TSPAN8 and TSPAN1 inhibition with chemotherapy might be a therapeutic strategy worth considering further investigations.

2 Graphical abstract



3 Zusammenfassung

In Europa sind 2018 etwa 2 Millionen Menschen an Krebs gestorben, wobei 90% dieser Krebs assoziierten Todesfälle von Metastasen verursacht wurden. Obwohl die Signifikanz von Metastasen für das Überleben von Krebspatienten bekannt ist, ist unser Verständnis der Entstehung von Mestastasen immer noch unvollständig und die therapeutischen Möglichkeiten daher stark begrenzt. Klinische Observationen zeigen, dass Metastasen selbst Monate bis Jahre nach Entfernung des Primärtumors entstehen können. Dies verkomplizieren den Behandlungsprozess, denn diese Beobachtungen lassen vermuten, dass therapieresistente Krebszellen, die Metastasen nach Latenzperioden induzieren können, lange im Körper überleben.

In dieser Studie wurde die Disseminierung von Brustkrebs in orthotopischen *in vivo* Modellen unter Verwendung von humanen Zelllinien sowie von Patienten-abgeleiteten Xenografts (PDX) untersucht. Dabei wurden Krebszellen nicht nur in Organen detektiert, die permissiv für das Wachstum von Brustkrebsmetastasen sind, Lunge, Leber und Knochen, sondern auch in Pankreas, Niere und Milz. In letzteren Organen entwickeln sich weder in Patienten noch in unseren Mausmodellen Metastasen. Die Integrität und Viabilität der Krebszellen wurde mit *in vitro* Kultivierung der Zellen, welche aus verschiedenen Organen isoliert wurden, bestätigt. Basierend auf diesen Ergebnissen wurde die Hypothese aufgestellt, dass es sich bei den disseminierten Brustkrebszellen (DCCs) in Niere, Pankreas und Milz um eine Subpopulation von Krebszellen handeln könnte, welche sich in einer metastatischen Latenzphase befindet.

Um diese Hypothese näher zu beleuchten wurden die DCCs *in vivo* auf ihre Lokalisation innnerhalb der Organe, ihre Überlebensfähigkeit sowie ihr Wachstumspotential weiterführend untersucht. Es konnte gezeigt werden, dass die Mehrheit der DCCs in der Niere extravasiert hatte, aber nicht proliferierend in direktem Kontakt zu Blutgefäßen blieb. Weiterhin konnten DCCs von PDX Modellen ohne Wachstum mehrere Wochen nach Resektion des Tumors in Niere, Pankreas und Milz überleben. DCCs, welche aus Niere, Pankreas oder Milz isoliert wurden, behalten ihr Wachstumspotential und konnten nach intravenöser Injektion Metastasen in der Lunge induzieren. Obwohl DCCs in der Lunge Metastasen gebildet haben, konnten sie in ihrem Ursprungsorgan weiterhin nicht wachsen. Diese Daten deuten darauf hin, dass keine aggressive Subpopulation selektiert wurde, sondern dass hauptsächlich die Mikroumgebung das Wachstumsmuster der Zellen beeinflusst. Zusammenfassend konnte gezeigt werden, dass Brustkrebs DCCs in den nachteiligen Mikroumgebungen der Niere, der Pankreas und der Milz ohne Wachstum überleben und ihr metastatisches Potential erhalten.

Um die molekularen Mechanismen für das Überleben der DCCs zu untersuchen, wurden umfassende Genexpressionsprofile von, aus Niere oder Pankreas isolierten, DCCs erstellt, und diese mit Krebszellen aus dem Mammakarzinom verglichen. Auf Transkriptionsebene zeigen disseminierte Krebszellen klare Unterschiede zu den Zellen des Mammakarzinoms. Die Unterschiede zwischen Krebszellen in Niere und Pancreas sind weniger stark ausgeprägt und die beiden Populationen weisen essentielle Ähnlichkeiten auf. Diese Beobachtungen deuten darauf hin, dass in DCCs aus unterschiedlichen Organen ähnliche Überlebensmechanismen in unvorteilhaften Mikroumgebungen zur Anwendung kommen könnten. Auf Grund dieser Ähnlichkeiten, wurde die Hypothese aufgestellt, dass gemeinsame transkriptionelle Änderungen in DCCs aus beiden Organen, krebszellintrinsisch reguliert sein könnten. Ein zweites Genexpressionsprofil von Krebszellen, welche aus Niere, Pankreas, Lunge und Mammakarzinom

isoliert und 48h *in vitro* kultiviert wurden, hat gezeigt, dass ein signifikanter Anteil der transkriptionellen Änderungen in DCCs in Kurzzeitkultur stabil war und somit intrinsisch reguliert zu sein scheint. Des Weiteren, unterscheiden sich DCCs aus Niere und Pankreas in ihrem Expressionprofil von Lungenmetastatsen. Dies deutet daraufhin, dass DCCs eine latente Population in der metastatischen Kaskade darstellen.

Die therapeutische Intervention zur Beseitigung der latenten DCCs ist von entscheidender Bedeutung um spätere Mestasen zu verhindern. Daher ist die Detektion molekularer Mechanismen, welche therapeutisch genuzt werden könnten, der nächste Schritt der Untersuchung. Hier wurde der Fokus auf die Gemeinsamkeiten in DCCs aus Niere und Pankreas gelegt. Mit „Gene Set Enrichment Analysis“ (GSEA) und „Gene Ontology“ (GO) Analysen wurden molekulare Mechanismen und Signalwege identifiziert, die im Folgenden genauer erläutert werden. Es wurde festgestellt, dass Mechanismen des epithelialen zu mesenchymal Übergangs (EMT) in DCCs in Niere und Pancreas inhibiert und Zelltod Mechanismen herunterregulierten sind. Eine geringere Apoptoserate in DCCs wurde mit einer cleaved Cascapase 3 Färbungen bestätigt.

Weiterhin wurden in den Genexpressionsdaten ausgeprägte metabolische Änderungen in DCCs beobachtet. Gene in metabolischen Signalwegen, einschließlich Energie-, Glukose und Nukleotidmetabolismus, waren herunterreguliert. Ein genauerer Blick hat gezeigt, dass DCCs einen metabolischen Wechsel von Glykolyse zu oxidativer Phosphorylierung unterliefen. Gene, die in Glukoseimport involviert sind sowie Enzyme, die an der Glykolyse beteiligt sind, waren in DCCs auf transkriptioneller Ebene herunterreguliert. Im Gegensatz dazu war die Expression von Genen erhöht, die für Proteine der Elektronentransportkette der oxidativen Phosphorylierung in Mitochondrien kodieren. Gene der Hypoxieantwort waren ebenfalls herunterreguliert in DCCs. Diese Regulation könnte mit der Verschiebung von Glykolyse zu oxidativer Phosphorylierung in Verbindung stehen, da Hypoxie zur Aktivierung von Glykolyse führt, da glykolytische Enzyme direkte Zielgene des Hypoxie induzierten Transkriptionsfaktors HIF1 sind.

Außerdem war Antigenpräsentation, durch „Major histocompatibility complex II“ (MHCII) Moleküle, in DCCs auf Transkript- sowie Proteinlevel herunterreguliert. Die Experimente weisen auf eine krebszellautonome Regulierung hin, da immunkompremierte Mausmodelle verwendet wurden. Ein möglicher Mechanismus hierfür ist eine YAP1 basierende Regulation von Genen, welche MHCII Expression negativ regulieren können. YAP1 Gensignaturen waren in .DCCs angereichert und in vitro konnte gezeigt werden, dass YAP1 MHCII Moleküle auf Transkript- und Proteinlevel reguliert.

Weiterhin waren DNA Reparaturmechanismen sowie Zellzyklus Checkpoint Gene im Genexpressionprofil der DCCs induziert.

Letzlich waren Chemotherapieresistenz Gensignaturen, die mit Zelllinien oder mit Patientenmaterial erstellt wurden, bei GSEA angereichert in DCCs. Die Chemotherapieresistenz der DCCs wurde *in vivo* unter Verwendung zwei verschiedener Chemotherapeutika weiter untersucht. Während das Mammakarzinom sowie die Metastasen in der Lunge gut auf die Behandlung angesprochen haben, wurden die DCCs in Niere, Pankreas und Milz von der Behandlung nicht beeinflusst und die Anzahl der Krebszellen in den Organen blieb unverändert. Somit stellen DCCs in Niere, Pankreas und Milz in diesem Modell eine therapieresistenten Pool an Brustkrebszellen darstellen.

Auf Grund der zentralen Bedeutung der Chemotherapieresistenz für den Erfolg der therapeutischen Intervention, wurden diese Ergebnisse tiefergehend analysiert Die, in DCCs angereicherte, Resistenzsignatur wurde mit den hochregulierten Genen in DCCs in Niere und Pankreas verglichen um Überschneidungen zu identifizieren. Dieser Verleich hat zur Identifizierung zweier Moleküle der Tetraspanin (TSPAN) Familie, TSPAN8 und TSPAN1, geführt. Um die Bedeutung der identifizierten Tetraspanine weiter zu beleuchten, wurden die Anteile der TSPAN8 und TSPAN1 exprimierenden DCCs in den Organen des Mausmodels sowie in Patientenproben untersucht. Die TSPAN8 einfach positive sowie TSPAN8 und TSPAN1 doppelpositive Zellpopulation war erhöht in DCCs in Niere, Pankreas und Milz im Vergleich zum Mammakarzinom. Außerdem wurden TSPAN8 einfach positive sowie TSPAN8 und TSPAN1 doppelpositive Krebszellen in Aszites Proben und pleuralen Effusionen von Brustkrebspatienten in drei von vier Patientenproben mittels FACS Analysen detektiert.

Vorhergehend wurde die Fähigkeit der DCCs ohne Wachstum zu Überleben als Kernmerkmal dieser Zellpopulation identifiziert. Daher wurde die Bedeutung von TSPAN8 und TSPAN1 für das Überleben in Spherekulturen von MDA-MB-231 Zellen untersucht. In TSPAN8 und TSPAN1 knockdown Zellen war die Apoptoserate erhöht. Außerdem war die TSPAN8 und TSPAN1 ko-expremierende Zellpopulation in der „label-retaining“, nicht-proliferirenden Fraktion der Spheres angereichert. Auf Grund der Kontextabhängigkeit der TSPAN Funktionen, wurde die Rolle von TSPAN8 und TSPAN1 auch direkt in DCCs analysiert. Hierzu wurden TSPAN8 und TSPAN1 doppel-knockdown Zellen *in vivo* injizieret, DCCs isoliert und das Genexpressionsprofil im Vergleich zu Kontroll-DCCs analysiert. Dieser Versuch hat die essenzielle Rolle der beiden Moleküle für die Chemotherapieresistenz der DCCs bestätigt. Die Resistenzsignatur, die in den DCCs angereichert war, ist durch den knockdown der Tetraspanine verloren gegangen. Die Transkriptionsdaten von TSPAN8 und TSPAN1 knockdown DCCs im Vergleich zu Kontroll-DCCs zeigen, dass die Tetraspanine Stammzelleigenschaften in DCCs in vivo regulieren.

Auch funktional konnten i*n vivo* Injektionen von TSPAN8 und TSPAN1 doppel-knockdown Zellen die essenzielle Rolle der beiden Moleküle für das Überleben und die Chemotherapieresistenz der DCCs bestätigen. Durch den doppel-knockdown wurde die Anzahl an DCCs in Pankreas und Milz signifikant reduziert sowie DCCs für Chemotherapie sensitiviert. Die Kombination aus TSPAN8 und TSPAN1 knockdown mit Doxorubicin und Cyclophosphamid Behandlung hat zu einer signifikanten Reduktion der Anzahl der Tumorzellen in Niere, Pankreas und Milz im Vergleich zu Chemotherapie ohne knockdown geführt.

Zusammenfassend stellt diese Studie wertvolle Informationen zur Biologie inaktiver disseminierter Brustkrebszellen zur Verfügung, die neue Möglichkeiten zur Entwicklung therapeutischer Strategien gegen Brustkrebs im latenten metastatischen Stadium eröffnen könnten. TSPAN8 und TSPAN1 wurden als Vermittler von Chemotherapieresistenz in disseminierten Brustkrebszellen identifiziert. Durch Depletion von TSPAN8 und TSPAN1 wurden DCCs für Chemotherapie sensitiviert. Die Kombination aus TSPAN Inhibierung und Chemotherapie stellt eine neue therapeutische Strategie dar, die es wert ist für weitere Analysen in Betracht gezogen zu werden.

4 Introduction

4.1 Biology of metastasis

In the last decades, research has improved patients' outcomes through the means of better diagnosis and treatment. While in 1973, 145 of 100,000 cancer patients in Germany died from the disease, this number was reduced by 27% to 106 cancer deaths in 2013 [1] (**Figure 1 A**). Despite these improvements, about two million people have died from cancer in Europe in 2018 and the risk for cancer associated death at the age of 75 or younger is about 29% [1]. In most cancer entities the reason for this still high mortality rate is metastatic spread. 90% of cancer-related deaths can be associated to metastatic disease [2, 3]. Metastasis is the spread of cancer cells from the primary site to distant organs, where they expand to a secondary tumor. This expansion can ultimately lead to failure of organ functions, which is a common cause of death [4] or to secondary malignancies including dyspnea initiated by cancer cells in lung and pleura [5], cachexia leading to extensive muscle wasting [6] or thrombosis [4, 7].

The high impact of metastases on disease outcome is supported by the 1- and 5-year survival rates of cancer patients depending on the stage of the disease. Cancer stages are assigned using the TNM system. T stands for the size of the primary tumor, N for the lymph node status and M for the presence of distant metastases. Using this system, cancers are grouped into five stages. Stage 0 represents a cancer in situ, stage I, II and III describe primary tumors with increasing size and severity of local invasion into the surrounding tissue and stage IV refers to cancers with distant metastases [8]. Carcinomas are cancers that are initiated in epithelial cells and account for about 85% of all cancer cases [9]. **Figure 1 B** shows the survival rates of the two most common carcinomas in Europe, breast and lung cancer [1], depending on the stage of the disease at diagnosis. Among breast cancer patients, about 99% survive five year or more when their disease is diagnosed at stage I. 88% survive five years with a stage II disease at diagnosis and the 5-year survival rate is about 55% when the patient is diagnosed with a stage III disease. When the disease is diagnosed at the metastatic stage (stage IV), the survival rate drops to only 14% [10]. Lung cancer progresses much faster leading to low survival rates five years after diagnosis even in patients with early stage cancers [11]. Thus, the 1-year survival rates were analyzed. In stage I cancer lung cancer patients, the net 1-year survival is about 80% and drops to 40% at stage III of the disease. When the disease is diagnosed at stage IV, the 1-year survival probability is further reduced to only 17% in both sexes (**Figure 1 B**) [11]. Similar results were also obtained for kidney, bladder and ovarian cancer [9] stressing that metastases are the most common cause of cancer associated deaths.

Despite the high impact of metastases for disease outcome, the complex molecular process is still not fully understood. The lack of knowledge translates into limited therapeutic opportunities once the tumor has seeded to distant sites [12, 13]. This emphasizes the need for better a understanding of metastases development and progression to find novel treatment options and to improve patients´ outcome.

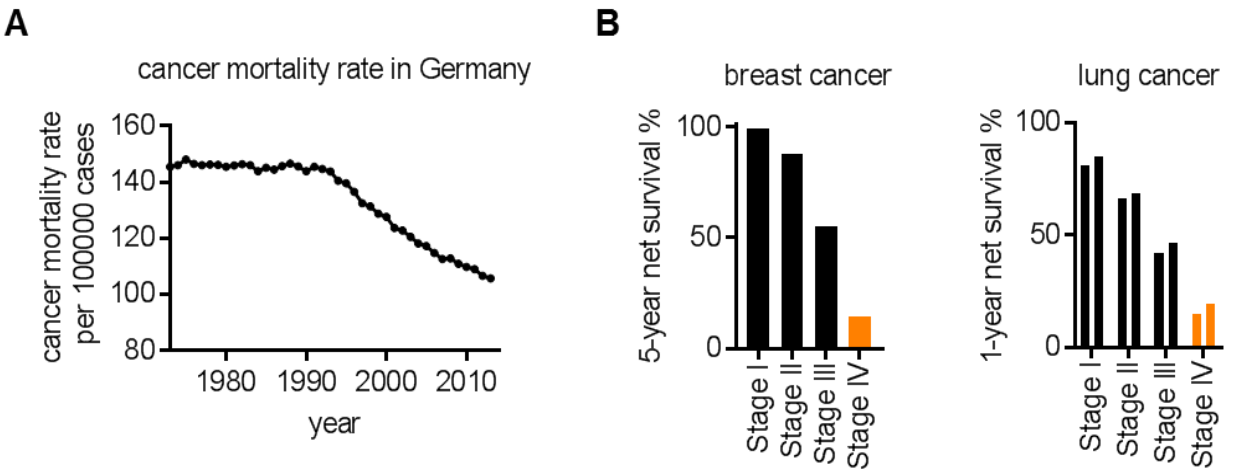


Figure 1 Most cancer related deaths are caused by metastasis today

A The cancer mortality rate in Germany over time: The world health organization (WHO) accessed the mortality rate of 36 cancer types in 185 countries over a time span of 40 years. In Germany, the mortality rate for all cancer entities was 145 deaths per 100,000 cases in 1973. Over the years the mortality rate decreased by about 27%. In 2013, 106 people died of cancer per 100,000 cases. The graph was plotted with data obtained from [1].

B Survival percentages of the two most common carcinomas in Europe, breast and lung cancer [1], significantly drop with progression of the disease. In breast cancer, more than 90% of patients survive five year when their disease is diagnosed at stage I or II. When breast cancer is diagnosed at stage III the survival rate is about 55%. However, the survival drops to 14% when the disease has reached the metastatic stage (Stage IV) [10]. In lung cancer, more than 80% survive one year when the disease is diagnosed at stage I. This decreases to 43% survival at stage III and further to 16% at stage IV [11]. The data was assessed by Cancer Research UK between 2003 and 2006 for lung cancer and 2002 and 2006 for breast cancer. For breast cancer only data for female patients are shown. For lung cancer, the left bar of each stage represents male cases, the right one female subjects. The metastatic stage of the disease (stage IV) is shown in orange.

4.1.1 Metastasis seeding – an early or late event?

Based on the findings that early surgery improves cancer patients' outcome [14] and that repetitive rounds of *in vivo* selection increase the metastatic potential of cancer cells [15, 16], metastasis was viewed as late event during cancer progression. However, more and more studies suggest that cancer cell seeding may already happen early during cancer development [14, 17-19]. Podsypania *et al.* have shown that even untransformed mouse mammary cells expressing the oncogenic transgenes Myc and $Kras^{D12}$ can bypass transformation in the mammary gland and directly seed the lung when brought into systemic circulation [20]. Therefore, two different models for metastatic progression exist: the linear and the parallel progression model [21]. The linear progression model assumes that cancer cells undergo several rounds of mutations in the primary tumor generating clones of high proliferation potential and aggressiveness. These more aggressive clones are then shed late during cancer development and establish metastases at distant sites (**Figure 2 A**) [21]. Building up on this concept, macroscopic metastases may also be able to seed further metastases themselves [22]. The parallel progression model that has been proposed by Christoph Klein [21] is based on the

experimental data showing that cancer cell dissemination occurs early during tumor development [14, 20]. Early disseminated cells further evolve in the secondary organ independent of the primary tumor (**Figure 2 B**). Microenvironmental cues selecting for site-specific advantages in disseminated cancer cells play an essential role in this model [21]. In summary, the parallel progression model assumes parallel seeding of cancer cells to several different organs and independent accumulation of genetic and epigenetic alterations in cancer cells from primary tumor and metastases [21]. The clinical observation that early removal of the primary tumor positively impacts patient survival [23, 24], contradicts the parallel model. It may be explained by a potential influence of the primary tumor on the metastatic site from the distance. Several studies indicate that factors secreted from the primary tumor can pave the way for metastatic growth [25, 26]. Removal of the primary tumor diminishes also these stimuli, potentially explaining why tumor removal is essential for metastasis even in the parallel progression model [21].

Important evidence for one or the other model comes from genetic analysis of primary tumors and their metastases. While several studies have identified only low divergence between primary tumor and metastases samples [27-30], other studies have detected more genetic differences between primary tumor and metastasis pointing to the parallel progression modelb[31-33]. However, Turajlic and Swanton have emphasized that, based on the genetic heterogeneity of tumors, multiple sampling of primary tumors may be necessary and could change our view of the divergence of primary tumor and metastasis [34]. They have proposed that one of several biopsies may contain the cells that seeded the metastasis thereby decreasing the divergence and supporting the linear rather than the parallel progression model [34]. Additional research is needed to enhance our understanding of the dynamics of metastatic progression [21].

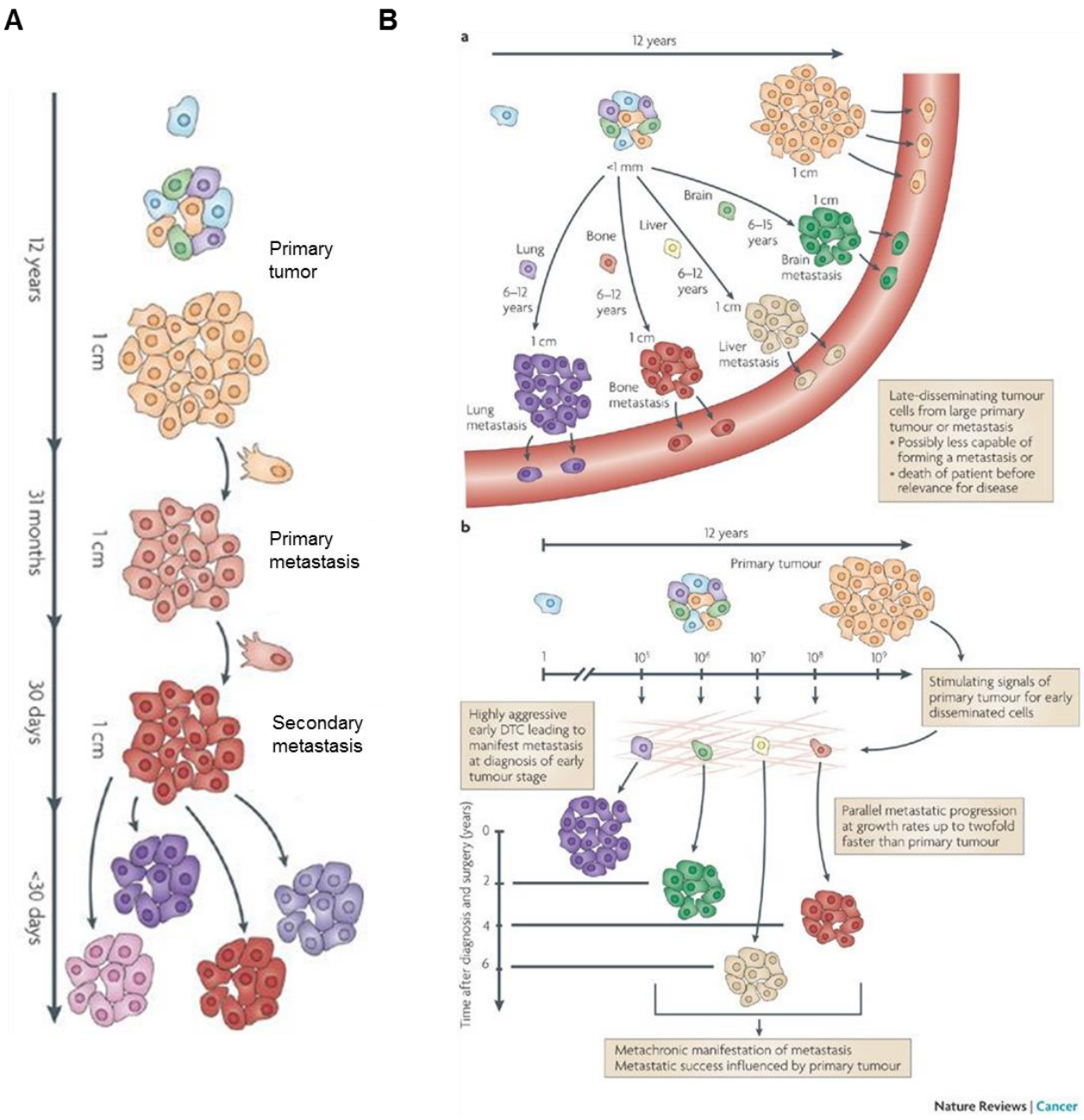


Figure 2 Theoretical models of metastases formation: linear versus parallel progression

Based on experimental findings, two different models for metastasis progression have been proposed: linear (**A**) and parallel (**B**) progression models. The figure was adapted from [21].

A The linear progression model assumes that dissemination of fully metastatic cancer cells is a late event in tumor progression. Therefore, cancer cells from the primary tumor undergo several rounds of genetic selection to yield more aggressive metastatic progenies. The time of co-evolution of metastatic and non-metastatic cells in the primary tumor results in smaller genetic differences between cells from metastases and tumors. The model also implies seeding from one metastasis to another.

B The parallel progression model in contrast suggests that cancer cell seeding occurs early in tumor development. Cancer cells in different organs would develop individually and in parallel to the primary tumor. Thereby, this model assumes that metastases acquire genetic variations that are rather distinct from the primary tumor. However, the primary tumor can influence metastatic growth from the distance e.g. by secretion of soluble factors.

4.1.2 Metastatic cascade

Metastasis formation is a complex multistep process that can be divided into two parts: (1) the physical translocation of cancer cells with an invasive phenotype from the primary tumor and (2) the colonization of the secondary organ (**Figure 3)** [2]. During translocation of cancer cells, a subpopulation within the primary tumor establishes an invasive phenotype that enables the cells to invade the surrounding tissue and intravasate into vessels. These invasive cells get thereby distributed through the body [2]. This process is occurring frequently: About 1 million breast cancer cells are shed into the vasculature per gram of primary tumor every day [35, 36].

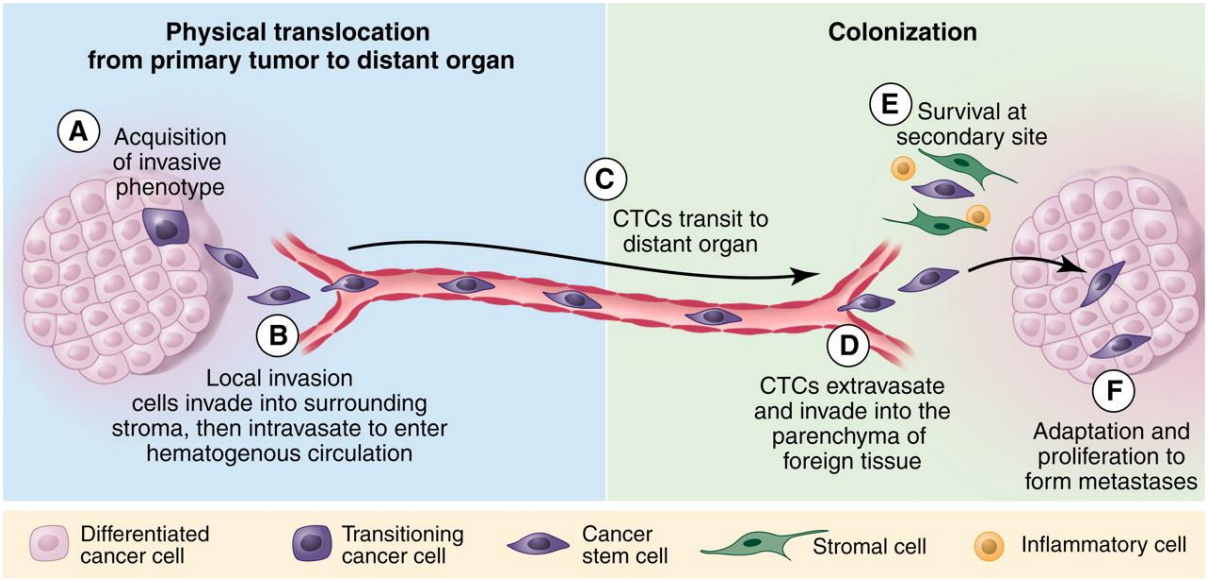


Figure 3 The metastatic cascade – a complex multi-step process

The metastatic process is a complex multistep process that can be divided into two parts: physical translocation of cancer cells from the primary tumor to distant organs (left; blue) and colonization of the secondary site (right; green). The translocation of cancer cells from the primary tumor involves acquisition of an invasive phenotype (A), the local spreading of these invasive cells into the surrounding stroma and the intravasation into vessels (B), by which they get transported through the body as circulating tumor cells (CTCs) (C). The colonization spans the extravasation of CTCs from blood vessels and the invasion of the parenchyma of the distant organ (D), the survival in the foreign microenvironment (E) and finally the adaptation allowing active proliferation leading to macrometastasis (F). [2]

The second step of the metastatic cascade, the colonization, is the rate limiting step of the process [37]. During the colonization phase, cancer cells extravasate from the vessels into the organ parenchyma, where they need to survive and finally start proliferating, forming metastasis (**Figure 3**) [2, 3]. In 1950, Zeidman *et al.* have reported that only a very small proportion of injected cancer cells is able to initiate metastasis [38]. The quantification of cancer cell numbers has revealed that only 1% of intravenous injected melanoma cells survived 24h after injection. Of those cells, alive in the lung 14 days after injection (0.2% of the starting population), only 20% were able to initiate macrometastasis [39]. In line with the low probability of metastatic growth, Fidler and Kripke have demonstrated in 1977 that only a minor subpopulation of cancer cells

from the primary tumor can seed metastases. Single clones established from murine melanoma cells *in vitro* showed greatly different capacities to metastasize and invade [40], suggesting that tumors are heterogeneous and only a small proportion of cancer cells from the primary tumor is able to conclude the complex process and initiate metastasis.

A subpopulation within the primary tumor that has tumor initiating capacity is referred to as cancer stem cell (CSC). The concept of CSCs goes back to the hierarchical organization observed in healthy tissues [41-43], where only a subset of cells – adult stem cells – is able to repopulate the organ. Adult stem cells are long lived, undifferentiated and are able to undergo asymmetric cell divisions to give rise to several differentiated cell types within the tissue [41, 44]. A similar hierarchy has been postulated for tumors, harboring a subpopulation of cancer stem cells (CSC) with tumor initiating capacity. Cancer stem cells have for example been experimentally identified in breast [45] and colorectal cancer [46, 47] based on linage tracing and tumor repopulation assays and have been transcriptionally profiled among others in colorectal cancer [44]. It has been shown that CSCs maintain some characteristics of healthy stem cells e.g. LGR5 expression [48-50], which has also been detected in mammary gland stem cells [51]. Several lines of evidence suggest that cells with stem cell properties also play an important role in metastasis initiation. The expression of LGR5 for examples has been detected in lung metastasis initiating cells in breast cancer [52]. Recently, de Sousa e Melo *et al.* have demonstrated that Lgr5+ CSCs in colon cancer are also crucial for liver metastasis growth [53]. Furthermore, stem cell signaling such as Wnt and Notch has been shown to be essential for metastatic progression of cancer cells [52, 54, 55].

Recent gene expression analysis of primary tumors and matching metastasis has detected a metastatic gene signature that is able to predict patient outcome, already in the bulk primary tumor [56, 57]. These data suggest that the subpopulation of cancer cells harboring these alterations may not be as small as originally assumed. This further emphasizes the complexity of the process and the impact of further factors such as the microenvironment. Given these and other studies [15, 45, 58], Weigelt *et al.* have proposed an “integrative model of metastasis” using breast cancer as tumor entity (**Figure 4**) [59]. Based on gene expression data, this model assumes that tumors with good and poor prognosis can already be distinguished by the gene expression profile of the primary tumor. Only tumors with poor prognosis harbor a subpopulation of CSCs, which can acquire metastasis-initiating characteristics under the influence of the microenvironment. Only this subpopulation of metastatic CSCs is finally able to conclude the metastatic cascade leading to the low frequency of metastasis initiating cells observed experimentally [40, 59, 60].

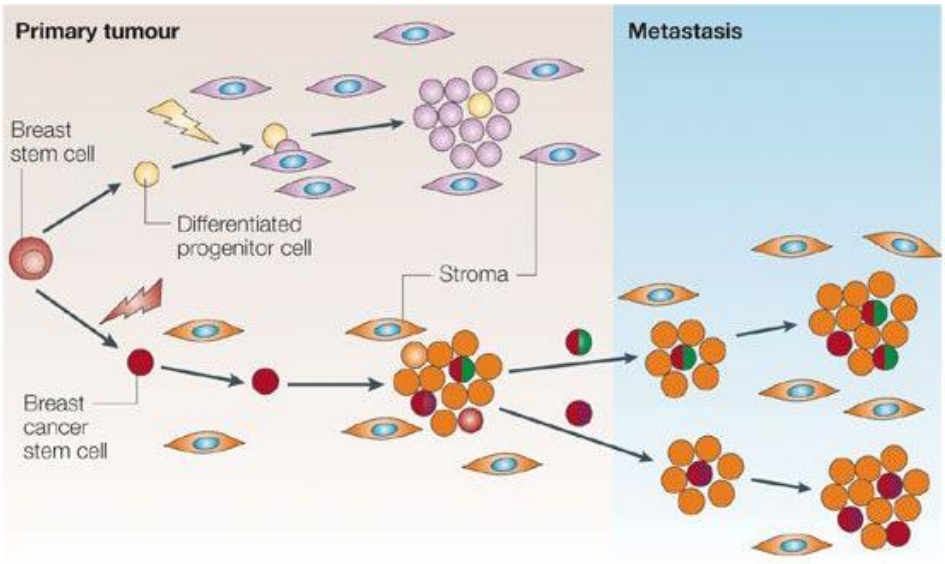


Figure 4 Integrative model of cancer metastasis

Weigelt *et al.* have proposed an integrative model of cancer metastasis based on several decades of metastasis research. This model summarizes the data obtained by *in vitro* experiments [40] as well as by gene expression profiling of primary tumor and metastasis samples [56, 57]. It suggests that metastases occur via collaboration of a poor prognostic phenotype already present in the primary tumor (orange; lower panel) and the effect of the stroma on cancer stem cells (red). Thereby a rare subpopulation of metastatic cancer stem cells (red/green or red/purple) is generated that is able to cope with the non-permissive microenvironment at the secondary site. Good prognostic tumors (pink; upper part) are generated from cancer progenitor cells (yellow) and do not harbor cancer stem cells. [59]

The afore mentioned studies and the model predicted by Weigelt *et al.*, stress that the oncogenic transformation of cells happening in the primary tumor and leading to cellular properties, such as unlimited proliferation, resistance to cell death and an instable genome [61], is not sufficient for metastatic outgrowth [60, 62]. In the following chapter, the current understanding of the individual steps cancer cells need to pass to establish overt metastases at the secondary site are described.

4.1.2.1 Invasion

The first hurdle on the way to metastasis is the acquisition of molecular traits that allow a subpopulation of cancer cells to spread to distant organs [3]. Thus, cancer cells must leave the primary tumor, invade the basement membrane and acquire a migratory phenotype [63]. An essential molecular mechanism in this context is the epithelial to mesenchymal transition (EMT), which refers to the acquisition of mesenchymal features and the loss of epithelial properties [36, 64]. Epithelial cells are characterized by close contact to neighboring cells, an apicobasal polarity axis and are separated from other cell types via a basal membrane [65]. Epithelial cells function as barrier and enable exchange e.g. of nutrients. Mesenchymal cells in contrast are only loosely connected via a 3D-extracellular matrix and make up connective tissue [65]. The

interconversion of these cell types via EMT - and its reversion called mesenchymal to epithelial transition (MET), are molecular mechanisms that are occurring for example during critical stages of the embryonic development [36, 64], during wound healing e.g. in keratinocytes [66] and in the ovarian surface epithelium during the menstrual cycle [67]. In 1982, Greenburg and Hay have described for the first time the process of EMT. They have observed that fully differentiated epithelial cells plated in a collagen matrix can acquire a spindle shape, were able to detach from the explant and migrate via filopodia through a 3D matrix. These transformed cells had no longer the epithelium typical apicobasal polarity but were indistinguishable from mesenchymal cells *in vivo* [68].

EMT is tightly controlled by a network of microenvironmental triggers regulating a circuit of molecular signals ultimately controlling EMT transcription factors (TF) (**Figure 5**) [69]. Sensing of environmental cues such as oxygen levels, inflammation and stresses allows the integration of the microenvironmental situation into EMT decision making (**Figure 5**) [69]. These cues are translated into intracellular signaling via a complex network of molecular pathways including NOTCH, HIF1/2, WNT, NF-κB and TGFβ (**Figure 5**). One of the most prominent molecular changes during EMT is the loss of E-cadherin [36, 70]. Thus, this molecular alteration has been applied as marker for EMT and was used to identify TFs associated with EMT [69]. Three main families of EMT-TFs have been identified: Snail (SNAI1 and SNAI2) [71, 72], Zeb (ZEB1 and ZEB2) [73, 74] and basic helix-loop-helix TF (E47 and TWIST) [75, 76].

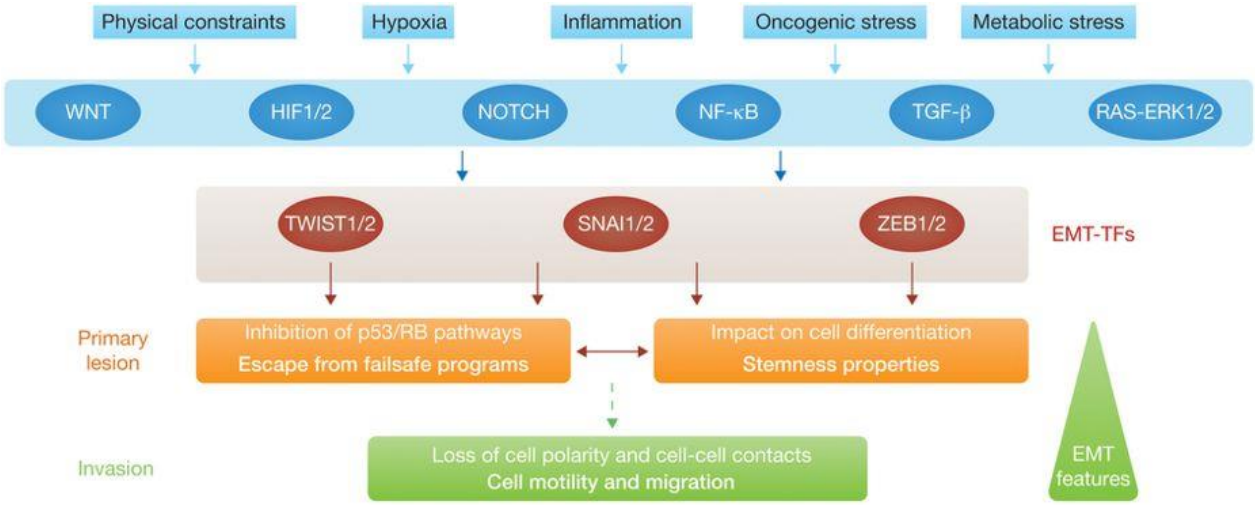


Figure 5 Epithelial to mesenchymal transition (EMT) is tightly regulated

Cancer cells hijack the epithelial to mesenchymal transition (EMT) mechanism, which is crucial during embryonic development. Cancer cells in the primary lesion undergo EMT, lose cell polarity and cell-cell contacts and gain a migratory phenotype enabling invasion. EMT is controlled by a network of signaling pathways. These pathways include WNT, HIF1/2, NOTCH, NF-κB, TGFβ and the RAS-ERK pathway that are regulated by microenvironmental triggers such as hypoxia, inflammation, oncogenic or metabolic stress. Downstream of these pathways are transcription factors (TF) that are crucial for EMT: members of the ZEB, SNAIL and TWIST family mediate the phenotypic changes [69].

In addition to its role in development and homeostasis, EMT also plays an important part in certain diseases such as fibrosis [77-79] and cancer [65, 80-82]. One of the first indications of EMT being an important event during tumor progression has been published by Gabbert *et al.* [81, 83]. The authors have observed that differentiated colonic adenocarcinomas contained single tumor cells or small aggregates at the invasive front that did not form proper junctions, had only a rudimentary basal membrane and pseudopodia-like protrusions [83]. In their publication they have described their findings as dedifferentiation at the invasive front. In today's context, this study can be classified as an early indication for the existence of EMT in tumors [81, 83]. Further evidence for the presence of EMT during cancer metastases was provided by Brabletz and colleagues [80]. In patient tissues of colorectal primary tumors and their matched lymph-node metastases, they have identified epithelial like, E-cadherin positive and differentiated growth patterns in the primary tumor as well as in metastasis. The invasive front however, contained mesenchymal E-cadherin negative cells [80]. This study has suggested the generation of invasive cells in the primary tumor via EMT, which was reversed in the lymph nodes through MET to allow metastatic growth [80]. Since then, the importance of reversal of the mesenchymal invasive phenotype by MET has been confirmed in several studies [84, 85]. A direct functional link between the EMT-TF Twist, the induction of EMT and metastatic potential was provided by the group of Weinberg [76]. They have demonstrated that ablation of Twist in breast cancer cells inhibited lung metastasis [76]. These studies together have led to the working model that cancer cells need to undergo EMT to obtain an invasive, migratory phenotype that must be reversed at the secondary site to allow metastatic outgrowth.

Supplementary to the acquisition of the migratory phenotype of mesenchymal cells, the EMT program also promotes metastasis through other mechanisms. EMT promotes degradation of the basement membrane through upregulation of matrix-metallo proteases (MMP) by EMT-TFs [86, 87], resistance to cell death through Snail [88], increased stem cell properties [89, 90] and decreased proliferation [88, 91]. In the context of stemness, experimental evidence suggests that certain cancer cells have the ability to switch between CSC and non-CSC phenotypes, which is dependent on an open chromatin confirmation of the EMT-TF ZEB1 [92]. Results by Schmidt *et al.* have suggested that CSC properties are initiated in invasive cancer cells by activation of the EMT-TF TWIST but are stable even upon inactivation of TWIST during MET [93]. Through intensive research in the field of EMT, we know today that the process is not an on-off switch between fully epithelial and mesenchymal states but a continuous transition with intermediate states allowing cells to express epithelial and mesenchymal markers at once [94-96]. Pastushenko *et al.* have recently identified several intermediate EMT stages in different microenvironments within the primary tumor. These stages were equally tumorigenic but had different capacities to invade and metastasize with the hybrid stages being the most metastatic [94]. Together, the evidence suggests that the plasticity of cancer cells may be important for metastatic progression. However, the clinical importance of EMT for cancer is still being discussed as evidence in patients is largely missing due to technical difficulties [65].

Importantly, EMT independent invasion mechanisms have also been identified in different cancer entities [97-99]. The expression of podoplanin at the invasive front for example has been reported in carcinomas to drive filopodia generation and migration [97]. Another crucial signaling axis that is involved in invasive properties of cancer cells is integrin signaling. This very complex signaling network has amongst other functions been reported to promote EMT via activation of transforming growth factor beta (TGF-β) signaling [100, 101]. Independent of EMT, integrin

signaling activates focal adhesion kinase, which initiates migration through several downstream pathways [63] and remodels the basement membrane via activation of MMPs [100].

An additional EMT independent migration mechanism that was adapted by cancer cells from healthy organ morphogenesis is collective cell migration. Collective cell migration refers to the movement of a cluster of cells that stays connected via cell-cell junctions [102]. The group of Peter Friedl has elucidated that this mechanism may play an essential role during cancer metastasis. They have demonstrated that oligoclonal clusters of cancer cells can be detected in the blood of breast cancer patients and these clusters have a higher metastatic potential than single disseminated cells [103]. Several other studies have confirmed that cell clusters are more efficient in initiating metastasis [103-106]. Only recently, Zajac and colleagues have elucidated cluster dissemination and its impact on survival and disease progression in patient samples using colorectal cancer material [105]. This and other studies [103] have also shown that clusters do not form in circulation but that polyclonal colonies of up to 250 cells [105] are already shed from the primary tumor. They have demonstrated that these clusters outperform single cells in invasive properties and metastasis formation capabilities in mouse models [103, 105].

In addition to cancer cell intrinsic cues leading to an invasive phenotype, several groups have demonstrated that stromal cells can also trigger cancer cell invasion [98, 99]. Macrophages for example can increase cancer cell migration, invasion and intravasation [98, 107, 108]. Wyckoff *et al.* have shown that inhibition of macrophage infiltration was linked to less cancer cells in the blood of mice [98]. Macrophages, recruited to cancer cells via cancer cell secreted colony stimulating factor (CSF1), increase cancer cell invasion and migration via secretion of epidermal growth factor [107, 108]. Moreover, fibroblasts can influence cancer cell invasion. In a study by Dumont *et al.* the effect of fibroblasts was mediated by EMT regulation. Activated fibroblasts promoted EMT thereby triggering metastasis [109]. Further studies have shown that fibroblasts may influence EMT via MMP based processes [110, 111]. Production of MMP3, a matrix metalloproteinase that is expressed by activated fibroblasts in the tumor context [111], has been linked to cleavage of E-cadherin and thereby loosening of epithelial cell-cell junctions as well as upregulation of mesenchymal markers such as vimentin (VIM) [110]. The group of Eric Sahai has revealed that fibroblasts also actively pull cancer cells at the invasive front of the primary tumor by establishing an N-cadherin – E-cadherin junctions. These junctions that were mechanically active were also present in patient material [99]. Thereby, fibroblasts lead cancer cells away from the primary tumor and enable invasion. While pulling cancer cells, fibroblasts also generate tracks via force-mediated matrix remodeling underlying integrin signaling [112]. These tracks enable collective invasion of tumor cells as cancer cells can follow these tracks using Cdc42 regulated myosin light chains [112]. These studies have highlighted that fibroblasts can on one hand directly interact with cancer cells, which leads to transmission of force and active pulling of cancer cells at the invasive front. On the other hand fibroblasts can remodel the stroma to create tracks cancer cells can collectively follow to enhance invasion.

4.1.2.2 Intravasation

Upon obtaining an invasive phenotype, cancer cells need to intravasate into the vasculature to be transported throughout the body. For body-wide distribution, cancer cells can theoretically take two different paths: the lymph system or the blood vasculature. For a long time it was thought that metastatic cells are mainly distributed through the hematogenous system and that the lymphatic system is a dead end for cancer cells [2, 113]. Lymph nodes have been viewed as bridgeheads containing a pool of cancer cells that has successfully metastasized and is therefore more likely to metastasize also to other organs [113, 114]. Only recently, Pereira *et al.* and Brown *et al.* have demonstrated directly that cancer cells from lymph vessels are able seed lung metastasis [115, 116]. These studies confirm that also lymph node metastasis can be a source for distant metastases [116]. Cancer cell intravasation of the lymph system is considered a passive process [113] as lymph vessels do not have tight endothelial junctions as present in blood vessels [117]. Thereby, cancer cells only need to invade the connective tissue and the invasion of the lymphatics is considered rather efficient [117]. The accessibility of the lymphatic system may explain the high frequency of cancer cell spreading to the sentinel lymph node [117].

Intravasation into blood vessels is an active process. Cancer cells must cross the endothelial barrier by disrupting cellular junctions and degrade the vascular basement membrane [118]. This process is facilitated by contact and crosstalk between cancer cells and endothelial cells. Proteases, such as MMPs, cathepsins and serine proteases, play an essential role during intravasation by cleavage of cell adhesion molecules, degradation of extracellular matrix and activation of cytokines and growth factors [113]. One example for a crosstalk mechanism is the secretion of Angiopoietin-like 4 (ANGTL4) by cancer cells, which binds and activates endothelial integrins and weakens endothelial cell-cell contacts [118]. In addition to tumor derived factors, stromal cells can also impact intravasation. A role of macrophages in the secretion of factors that promote intravasation has been reported by several groups [113, 119, 120].

Two models exist for the intravasation of cancer cells into vessels. The conventional model assumes that primarily cancer cells, invading the local stroma, migrate to vessels and intravasate. Deryugina *et al.* have suggested an alternative model based on their data generated with an unbiased intravasation-scoring method in live animals [121]. The authors have shown that intravasation in tumors expressing epidermal growth factor receptor (EGFR), which was essential for intratumoral vasculature formation, occurred primarily within the core of the primary tumor independent of stromal invasion [121]. The role of intra-tumor intravasation has also been confirmed by other studies, showing a correlation between intra-tumor vessel density and the number of tumor cells in circulation [98, 117, 122, 123]. Abnormalities of vessels, such as discontinued basement membrane and incomplete cell junctions, inside tumors further ease the vascular entry of cancer cells [63, 124].

4.1.2.3 Transport through circulation

Having entered circulation, cancer cells face a completely different microenvironment: the shear stress of the blood flow and the immune cells present in the blood stream [113]. To survive in this unfavorable environment, cancer cells can use blood platelets as shields. Cancer cell expression of coagulation factors leads to aggregation of platelets around the respective cancer cells. These clots have been shown to protect cancer cells from immune cell mediated killing, primarily by natural killer (NK) cells [113, 125, 126]. Furthermore, the association with platelets also reduces the shear stress acting on cancer cells in circulation [113]. In accordance with these findings, the number of activated platelets present in cancer patients has been linked to prognosis [127-129] and treatment with anti-coagulants has been shown to reduce metastatic spread [130].

Another hurdle is the survival of cancer cells without anchorage, which in endothelial cells triggers anoikis, a specific form of apoptosis [131]. Therefore, tumor cells have developed several mechanisms to circumvent apoptosis induction triggered by lack of adhesion. One example is activation of integrin signaling, which promotes survival via activation of PI3K signaling and induction of anti-apoptotic proteins such as BCL2 [63]. In addition, cancer cells remain in circulation only for short time periods [132] and may be trapped in capillaries quickly due to their size (20-30 µm compared to the capillary diameter of 8 µm). Thus, leaving not enough time for anoikis induction [12].

The detection of cancer cells inside the vasculature – referred to as circulating tumor cells (CTCs) has diagnostic value [133-136]. In breast cancer for example, a threshold of five CTCs in 7.5 ml of blood was linked to poor overall and progression-free survival [134]. In esophageal squamous cell carcinoma the number of CTCs correlated with disease progression. Higher numbers of CTCs reflected tumor invasion, lymph node spread and distant metastasis [137]. Furthermore, CTCs are also used for therapy response predictions as it was observed that the number of CTCs decreases upon treatment success and increased during resistance [138]. Based on these results several methods and devices for the detection of CTCs have been developed. [139, 140] [141, 142].

4.1.2.4 Extravasation

Upon arrival at the secondary site, cancer cells need to leave the vessels and enter the parenchyma of the secondary organ. This can be done by two mechanisms: Firstly, Cancer cells that got stuck in capillaries due to their size [63] can start proliferating in the vessel bed ultimately leading to bursting of the capillary [143]. Secondly, cancer cells can actively leave the vasculature, and extravasate from vessels by crossing the endothelial cell layer [12]. The extravasation process is promoted by the secretion of cancer cell derived factors such as COX2, ANGTL4 and MMPs that induce vascular hyper-permeability [144, 145]. In addition, cancer cells can also recruit stromal cells such as inflammatory monocytes that support extravasation [146].

4.1.2.5 Survival at the distant site

Cancer cells that managed to disseminate from the primary tumor and extravasate from vessels are again challenged by a unfavorable microenvironment in the tissue parenchyma to which they are poorly adapted [12]. For survival of cancer cells at the secondary site, they may adapt intrinsically to the requirements of the tissue microenvironment or reprogram the microenvironment to a less hostile state. An example of breast cancer cells` intrinsic signaling, improving survival selectively in the bone microenvironment, is Src tyrosine kinase signaling [147]. Zhang *et al.* have shown that Src signaling, activated in cancer cells, allows them to respond to the pro-survival factor CXCL12 provided by the bone microenvironment while at the same time inhibiting the pro-apoptotic effect of bone secreted TRAIL [147]. This complex crosstalk of cancer cells and bone stroma promotes cancer cell survival [147]. Oskarsson *et al.* have demonstrated the interplay between cancer cell intrinsic adaptations and cancer mediated modulations of the microenvironment at the secondary site of the lung. They have shown that breast cancer cells in the lung produce the extracellular matrix (ECM) protein Tenasin C (TNC) in early stages of the metastatic cascade enhancing survival and stemness properties of cancer cells. In the later stages, when cancer cell induced stromal modulations have taken place, the production of TNC is taken over by lung stromal cells [52].

An important component of the non-permissive microenvironment at the secondary organ is the immune system. Upon arrival of single disseminated cells at the distant site, immune cells display anti-cancer immune surveillance and can attack cancer cells [148, 149]. Inhibition of this initial anti-metastatic effect of T-cells or NK cells has been shown to increase metastasis in mouse models [149, 150]. To circumvent immune clearing, cancer cells have different strategies. Cancer cells can downregulate antigen presentation by inhibition or mutation of major histocompatibility complex (MHC), which binds antigens and presents them to the cell surface. Repression of MHC molecules themselves or their processing machinery is a frequently oberved immune evasion mechanism [151, 152]. Furthermore, increased expression of immune checkpoint molecules such as programmed death receptor ligand 1 (PD-L1) has been reported for example in breast cancer [153] and is associated with increased T-cell exhaustion and apoptosis [154].

In recent years, evidence has accumulated, suggesting that the primary tumor may influence cancer cell survival and growth at the secondary site already prior to cancer cell arrival. The finding that the primary tumor may prepare the “soil” for metastasis has been summarized in the concept of the pre-metastatic niche [155]. First evidence for the presence of a pre-metastatic niche was published by Kaplan *et al.* They have demonstrated that hematopoietic progenitor cells, expressing vascular epithelial growth factor receptor 1 (VEGFR1), were recruited to the metastatic site already before arrival of cancer cells and that this recruitment was crucial for metastasis formation [156]. They have further suggested that this recruitment is mediated by fibronectin secreted by cells in the primary tumor [156]. Since, a variety of different factors, promoting pre-metastatic niche formation, have been identified [155]. Examples for tumor secreted factors establishing a pre-metastatic niche in the lung are tumor necrosis factor alpha (TNFα) and TGFβ [157]. Microenvironmental cues present in the primary tumor such as inflammation and hypoxia have been shown to trigger secretion of factors leading to recruitment of myeloid cells to the lung, priming the microenvironment [158, 159].

Exosomes, shed from the primary tumor, were also indicated to be able to initiate a pre-metastatic niche [160-162]. Costa-Silva *et al.* have demonstrated that exosomes released from pancreatic cancer cells were taken up by Kupffer cells in the liver, leading to recruitment of macrophages, which was critical for liver metastasis formation [160]. Independent of the way of acquisition, hallmarks of pre-metastatic niches have been formulated by Liu and Cao [155]: immunosuppression, [163], inflammation [157], vascular permeability [164], lymph angiogenesis (in the pre-metastatic niche of lymph nodes) [165], organ tropism [166] as well as reprogramming of metabolism [167] and stroma [168].

4.1.2.6 Tumor latency – dormancy

Even those cancer cells that managed to reach the parenchyma of in the distant organ and survive in the non-permissive microenvironment are not guaranteed to grow [12]. Based on experimental findings that cancer cells can already disseminate very early during primary tumor development [14, 18, 19] and the clinical observation that metastasis in patients of some tumor types can occur even years or decades after removal of the primary tumor [19], it is apparent that a subpopulation of cancer cells can survive in the body without macroscopic growth. This subpopulation may nevertheless initiate metastasis later in life. The observation that cancer cell can survive in the body without growth was already made early by the pathologist Rupert Willis [169, 170]. He has coined the term tumor dormancy, which refers to the time span between arrival of cancer cells at the secondary site and metastatic outgrowth, in which no measurable disease progression is observed [169, 170]. [62]. Dependent on the tumor type these latency periods can last years or even decades e.g. in breast cancer [19, 171] or only few weeks or month as in lung or colon cancer [62, 172]. In light of the parallel and linear metastatic progression model (page 10), short latency periods imply that the cancer cells are already equipped with molecular functions that allow rapid outgrowth without elaborate adaptations to the microenvironment at the distant site [62]. Whereas, long latency periods may be explained by early dissemination of immature cancer cells that need to evolve at the distant site [21, 62] as well as by the non-permissive microenvironment at the distant side that restrict metastatic growth [62]. The interplay of these factors leads to a net growth output of zero during latency periods resulting in tumor dormancy.

Today, three dormancy mechanisms are recognized: dormancy on a cellular level, dormancy being controlled by the immune system and angiogenic dormancy on colony level mediated by a balance between growth and cell death [62, 173, 174]. Angiogenic dormancy has been linked to a lack of vascularization [175]. The critical role of the vasculature for cancer cell growth is well established for the primary tumor [61]. If the distance between tumor cells and blood vessels exceeds the diffusion rate of nutrients (100-500 µm), the lack of oxygen and nutrients leads to apoptosis of cancer cells [176-178]. Similar to the primary tumor also metastatic growth depends on blood supply at the distant site [175, 179]. Kienast *et al.* have shown that inhibition of the angiogenic switch in lung cancer micrometastasis using vascular endothelial growth factor (VEGF) inhibitors repressed metastatic growth and resulted in long time dormancy [179]. Holmgren *et al.* have reported that the proliferation rate between dormant and growing lung metastasis was not changed. However, in dormant colonies, cell death was significantly more

frequent. They further linked this balance of death and growth to angiogenic inhibition [175]. Omitting this inhibition triggered metastatic growth [175].

More recently, the immune system has been linked to tumor dormancy. Eyles *et al.* have revealed in a spontaneous melanoma mouse model, that disseminated cancer cells (DCCs) in the lung were kept dormant by CD8+ T-cells [149]. Depletion of CD8+ T-cells promoted metastatic growth, suggesting that T-cells restrict outgrowth of melanoma cells in the lung [149]. However, clinical evidence confirming a role of the immune system for dormancy of single disseminated cells is missing [174]. Beside a direct cytotoxic effect on DCCs mediated by T-cells, immune cells may also contribute to tumor dormancy by modelling the microenvironment non-permissive for metastatic growth [174, 180]. One example is the secretion of angiogenesis-inducing cytokines by CD4+ T-cells through interferon-gamma signaling. This secretion resulted in growth arrest of DCCs in the pancreas without cell killing. Inhibition of this pathway reversed the effects and rendered CD4+ T-cell supportive of tumor growth [180].

Cellular dormancy is mediated by single cells entering a quiescent state, which is characterized by reversibility and lack of proliferation [174]. Dasgupta *et al.* have defined three characteristic properties that discriminate dormant cancer cells from other cancer subpopulations: cell survival, therapy resistance and growth arrest [173]. Tumor dormancy may not only be a passive process due to lack of molecular features and mitogenic signals, enabling metastatic growth, as rather an active program initiated in a subset of cancer cells [181-183]. Evidence for this assumption comes from dormancy signatures in breast cancer that were not restricted to genes that are repressed upon growth factor starvation [174, 184]. Also dormancy gene expression signatures, generated in fibroblasts *in vitro* have suggested an active component for dormancy induction [185]. In this study, dormancy has been induced by different stimuli and each stimulus has led to a different expression profile emphasizing that dormancy is context dependent and not only a consequence of cell cycle exit [185].

Furthermore, there is evidence that tumor dormancy may recapitulate a quiescence program of normal stem cells [174]. A significant overlap of more than 60% between the genes induced in dormant head and neck cancer cells and quiescent healthy stem cells from the hair follicle, muscle and blood has supported this idea [174]. In addition, signals in the healthy stem cell niche such as BMP7 signaling have been shown to play a role in maintenance of cancer cell dormancy [186] and to overwrite oncogenic signaling [187]. Another essential characteristic of dormant cells is their plasticity to uncouple signaling pathways when appropriate. One example for molecular plasticity is the rewiring of metabolic and survival pathways that is often observed during tumor dormancy [174]. Several lines of evidence have suggested that reduction of AKT signaling is linked to growth arrest and survival in dormant cells [188-190]. Although AKT signaling is closely linked to metabolism via mTOR activation, dormant cells are able to uncouple the pathways maintained mTOR signaling independent of AKT [191].

The microenvironment also plays an essential role in the regulation of cancer cells dormancy. For example, it regulates p38 activity in cancer cells, which is well described to maintain dormancy of cancer cells [181]. Aguirre-Ghiso and colleagues have revealed that the balance between ERK and p38 activation regulates growth or dormancy of DCCs and that this balance is regulated by contact with the microenvironment via the urokinase plasminogen activator receptor (uPAR). High uPAR activity leads to fibronectin fibril formation and activation of ERK via integrin signaling. If the balance is shifted towards ERK, cancer cells actively proliferate, whereas higher levels of p38 induce dormancy in cancer cells [181, 192]. Endothelial cells, lining up vessel

walls, are another example for a stromal cell type, actively supporting dormant disseminated cancer cells [193] [179, 194]. Ghajar *et al.* have demonstrated that disseminated breast cancer cells in bone, brain and lung preferentially reside upon the tissue microvasculature in mouse models [193]. Using *in vitro* models they have been able to show that endothelial cells of stable microvasculature keep DCCs dormant through production of thrombospondin-1 (THBS1). Sprouting neovasculature has the obverse effect and induces metastatic growth via secretion of TGFβ and periostin (POSTN) [193].

These examples emphasize the important role of the microenvironment for cancer cells dormancy. Dependent on their ability to support metastatic growth, organs have been classified into dormancy-permissive or –restrictive [174]. This classification is based on the finding that some organs such as the bone are more often carriers of DCCs than others [195]. However, the presence of DCCs in an organ does not always correlate with the development of metastasis. For example in gastric cancer, DCCs in the bone marrow are frequently detected in patients but the occurrence of bone metastasis is nevertheless rare [174]..

4.1.2.7 Metastatic colonization

During the course of the disease, dormant DCCs may get reactivated leading to macroscopic metastasis. Several studies have investigated the mechanisms leading to reawakening of dormant cancer cells. Intrinsic properties of secondary organs were identified early as important factors for cancer cell growth [196]. A study by Huesemann *et al.* has confirmed that the microenvironment has a crucial impact. Transplantation of pre-malignant bone DCCs into the bone-marrow of irradiated mice, led to growth of previously dormant cells [14]. This study emphasizes that the microenvironment determines metastatic growth to a significant proportion [14], which is in concordance with “the seed and soil hypothesis” postulated by the 19th century surgeon Stephen Paget [197]. The analysis of 900 autopsy records of cancer patients, made him realize that metastases do not occur randomly in all organs with the same probability. Thereafter, he has hypothesized that only some cancer cells, the seeds, are able to grow in a friendly microenvironment, the soil. For successful metastasis seed and soil have to be compatible [197, 198]. Intensive research has been conducted to unravel the role of the microenvironment during cancer metastasis. These finding have led to the “integrative metastasis model” postulated by Weigelt *et al.* (**Figure 4**), taking the impact of the microenvironment on metastatic growth into account [59] (discussed in more detail on page 13).

Also changes in the composition of the microenvironment for example through inflammation or aging may lead to reactivation of dormant cells [2, 199]. Albrengues *et al.* have shown recently that tobacco smoke can reactivate dormant breast cancer cells in the lung, via induction of chronical inflammation [199]. In this study, the authors linked inflammation to the generation of neutrophil extracellular traps (NETs) that led to remodeling of proliferative signaling in cancer cells [199]. Moreover, Barkan *et al.* have suggested that cytoskeletal organization plays a critical role for dormancy and reactivation. They have shown that fibronectin production of cancer cells activates integrin signaling and stress fiber production, which is linked to reactivation of growth in dormant cancer cells [200]. Alternatively, cancer cells can also educate the stroma, being initially

hostile, to become increasingly supportive with proceeding metastatic progression [12, 201, 202]. One way is to trigger an inflammatory response by recruitment of leucocytes that support tumorigenesis [201, 203, 204]. Inflammation further leads to recruitment of innate immune cells, such as macrophages, dendritic cells, NK cells, MDSC, neutrophils and mast cells, as well as adaptive immune cells, such as T- and B-cells, that can have pro-tumorigenic as well as anti-tumorigenic effects [113, 201].

Moreover, non-immune stromal cells such as fibroblasts can influence metastatic colonization. Several studies have demonstrated that a complex interaction between cancer cells and fibroblasts helps maintaining a cancer stem cell (CSC) population and thereby enables metastasis. Malanchi *et al.* have observed that infiltrating cancer cells induced POSTN expression in fibroblasts, which is essential for maintenance of the CSC pool. POSTN influences CSCs by recruitment of Wnt ligands thereby increasing Wnt signaling [205]. Depletion of POSTN reduced the CSC pool and inhibited metastasis formation [205]. Another mechanism for the enrichment of CSCs by fibroblasts has been presented by Chen *et al.* Fibroblast mediated secretion of insulin growth factor II (IGF-II) activates IGF receptor (IGFR) signaling in cancer cells leading to expression of NANOG and promotion of stem cell properties [206]. Activation of this signaling axis has been linked to reduced relapse-free survival in stage 1 non-small cell lung cancer patients [206].

In addition to modifications within and by the microenvironment that help cancer cells to adapt, cancer cells themselves may also further evolve in the secondary organ leading to generation of clones that can cope better with the situation at the distant site [2, 62]. Interleukin 11 and NF-kB expression by cancer cells in the bone marrow has been associated with induction of osteolytic bone metastasis [15, 207]. Genes enabling metastatic growth of breast cancer cells in the lung [16], liver [208] and brain [209] have been reported. The little overlap between the identified genes again emphasizes the interplay between cell intrinsic properties and extrinsic signals [12].

4.1.3 Organ tropism of metastases

The complex link between the gene expression pattern of cancer cells and signals coming from the microenvironment, results in a distinct organ colonization pattern for each cancer type [16, 173, 210]. The observation that different cancer entities develop metastasis preferably in specific organs is referred to as organ tropism of metastases [62, 173]. A summary of the most common metastatic sites in different cancer entities is provided in **Table 1**. But not only the organs affected by metastasis in different tumor entities vary but also the range of organs a cancer type can seed to. Some cancers as e.g. prostate cancer seed to one organ, namely the bone [62, 211]. Other types such as breast and lung cancer can initiate metastasis in a broad range of different organs [212, 213]. This suggests that beside organ specific cues as it has been demonstrated for brain [214] and bone [15], some molecular properties are beneficial for cancer cells in several distant sites. Another level of complexity is added by the observation that, even though two cancer types show the same organ tropism, the clinical outcome can be very different due to different latency times (p22) of individual cancer entities [62]. Prominent examples are lung and breast cancer. Both tumor types metastasize to lung, liver, bone and

brain. Breast cancer metastases develop often after years or decades of tumor remission [19], whereas lung cancer metastases occur fast, often only months after tumor removal [172].

Table 1 Organ tropisms of common tumor types

Different tumor entities metastasize to different organs. This observation is referred to as organ tropism of metastases. (adapted from [62])

Tumor type	Principal sites of metastasis
Breast	Bone, lung, liver and brain
Lung adenocarcinoma	Brain, bone, adrenal gland and liver
Skin melanoma	Lung, brain, skin and liver
Colorectal	Liver and lung
Pancreatic	Liver and lung
Prostate	Bone
Sarcoma	Lung
Uveal melanoma	Liver

Although the general steps, cancer cells have to undergo to establish metastases are the same in all organs, differences in structure and microenvironment of different body compartments require specific molecular functions for each organ [62]. This hypothesis is strengthened by results from Minn *et al*. demonstrating the presence of several subpopulations with metastatic potential for different organs in pleural effusions of breast cancer patients [215]. This concept is further underlined by different studies in the last decade that have identified gene signatures associated with metastatic progression of cancer [15, 16, 57, 216, 217]. The comparison of the signatures generated in different cancer entities reveals only a small overlap between metastasis associated genes in different cancers or even different metastatic sites from one cancer type [4]. It is likely that a subset of the necessary genes is already expressed in the primary tumor but is only crucial at the distant site [62]. Examples are LOX and ANGPTL4, which are expressed by the primary tumor but are dispensable. However, they play crucial roles for lung metastasis in breast cancer [145, 218]. Other genes may not be expressed in the primary tumor (or only in a very small subpopulation of cells), but induced at the metastatic site [62]. Interleukin 11 is an example for this class of genes [15].

In addition to the cancer cells intrinsic properties, the cellular composition of the microenvironment, which is distinct in different organs, plays a crucial role for cancer cell survival and growth [52, 147]. For example, in the bone marrow niche, osteoblasts interact with cancer cells leading to a survival advantage of DCCs. On one hand, direct contact of cancer cells with osteogenic cells via E-cadherin - N-cadherin junctions has been shown to result in activation of mTOR signaling in cancer cells, promoting metastatic growth [219]. On the other hand, Jagged1 expressed on cancer cells, has been shown to activate Notch signaling in osteoblasts resulting in secretion of IL-6, which feeds back to cancer cells promoting proliferation and survival [220]. Moreover, the immune system has been shown to influence the organ tropism of metastasis for

example via cytokine secretion [221]. Mueller *et al.* have observed a relationship between the chemokine receptor expression of breast cancer and melanoma cells and metastatic growth in organs that express high levels of the corresponding ligands [221].

Another determining factor for organ tropism can be the blood circulation pattern and the vessel architecture in different organs [62, 222, 223]. As the endothelia in different organs vary significantly [224-226], tumor cells also have to possess specific capabilities for each target organ. Colonization of the lung for example requires invasion of the basement membrane, which is mediated by interaction of integrins on endothelial cells with cancer cell-expressed laminin 5 [227].

4.2 Metastasis and therapy

A major problem of anti-cancer therapy is that the metastatic stage of the disease can only been targeted when it has reached an aggressively growing macrometastatic state [13] Wide-spread and early dissemination of cancer cells in combination with the possibility for existence of a therapy resistant subpopulation that can serve as a reservoir for relapse [62, 228], limits our options for therapeutic intervention early during metastatic progression. In breast cancer, therapy resistant DCCs, detected in the bone marrow even three years after adjuvant therapy, have been correlated with disease relapse [229], stressing the impact of minimal residual disease. In this chapter the mechanisms of therapy resistance of metastatic cells as well as possible treatment strategies for early targeting of disseminated cancer cells will be discussed.

Therapy resistance of metastasis and DCCs can be acquired in response to treatment [230] or can be an intrinsic mechanism [62] caused by survival signaling, interaction with the microenvironment or dormancy. One example for a cancer subtype prone for acquired resistance is a lung adenocarcinoma with gain-of-function in the endothelial growth factor receptor (EGFR). This tumor type can become resistant to EGFR inhibition by amplification of MET, leading to PI3K activation and resulting in the same downstream signaling as EGFR activation [231]. This resistance strategy has the advantage that MET signaling can additionally promote pro-metastatic cues such as migration [62]. This example demonstrates that therapy resistance mechanisms can at the same time promote metastatic progression.

Intrinsic characteristics of cancer cells acquired during the process of metastasis or even before can also render them resistant to therapy. The treatment response of metastatic cells can be affected by the presence of a dormant subpopulation (page 22) or cancer stem cells (page 13). The lack of proliferation in combination with intrinsic characteristics may explain the therapy resistance of dormant cells [173]. Chemotherapeutic agents are designed to target actively proliferating cells and will thereby not affect dormant disseminated cancer cells or CSCs [232, 233]. This selection for non-proliferative cancer cells can lead to the presence of minimal residual disease after the treatment [233].

The microenvironment such as the bone marrow can also confer therapy resistance to cancer cells [62]. Recently, a protective role of the microenvironment was demonstrated for breast cancer DCCs in the bone marrow. These cells were protected from chemotherapy independent of their proliferation status by interactions with perivascular endothelial cells via a yet unknown mechanism [234]. However, an interruption of the integrin mediated interaction between endothelial cells and DCCs rendered the cancer cells sensitive to chemotherapy [234].

Independent of therapy resistances, the metastatic state is difficult to target due to the early dissemination of cancer cells from the mammary tumor [14, 19, 21]. Therapeutic agents would therefore need to target survival and outgrowth of metastatic cells at the secondary site rather than dissemination from the primary tumor [12]. This approach is however, difficult as drugs addressing the primary tumor may only have limited effects against not actively proliferating metastatic cells [12]. This discrepancy between therapy effectiveness in the primary tumor and in early metastatic lesions may be caused by limited drug availability at the metastatic site due to poor vascularization or organ specific limitation e.g. the blood brain barrier as well as by a protective role of the metastatic niche as discussed above [12]. Furthermore, metastases can acquire genetic and transcriptomic changes independent of the primary tumor resulting in

different drug vulnerabilities and dependencies that affect the response to therapy [235]. These examples stress the need for identification of metastasis specific therapeutic strategies. Studies investigating the molecular mechanisms responsible for survival and growth of DCCs have led to identification of targets on metastatic cancer cells themselves [147, 236] as well as in the microenvironment [237, 238], potentially affecting metastatic disease onset. Recently therapeutic strategies targeting cancer stem cells have come into the focus of anti-cancer and metastasis treatment. As CSCs are viewed as initiators of cancer and metastasis growth, therapeutic targeting could have long lasting benefits [44]. One example for CSC targeting is the use of cytotoxic drugs conjugated to cancer stem cells markers such as LGR5 [239].

These results encourage further studies on the survival and growth of DCCs to bring specific anti-metastatic therapies into the clinic and enable targeting of the early steps of metastatic progression. This will help to overcome today's obstacle of only being able to treat late stage, actively growing macroscopic metastasis [13].

4.3 Breast cancer metastasis

In this study metastatic dissemination was investigated in the context of breast cancer. Breast cancer is the most common cancer type in women with more than 500,000 new cases in Europe in 2018 and a 30% mortality rate [1]. The vast majority of breast cancer cases (more than 99%) are carcinomas that arise from epithelial cells within the breast, lining the ducts [240] (**Figure 6**). Dependent on the invasion status of the disease, breast cancer is classified as ductal carcinoma in situ (DCIS), invasive or metastatic breast cancer. DCIS is characterized by locally restricted growth inside of preexisting ducts and is non-invasive. Importantly, DCIS can over time break the basement membrane and develop into invasive breast cancer [240, 241]. (**Figure 6**). Similar to most other cancer entities, the metastatic state of breast cancer, stage IV, is linked to significantly decreased 5-year patient survival rates (**Figure 1**) [10].

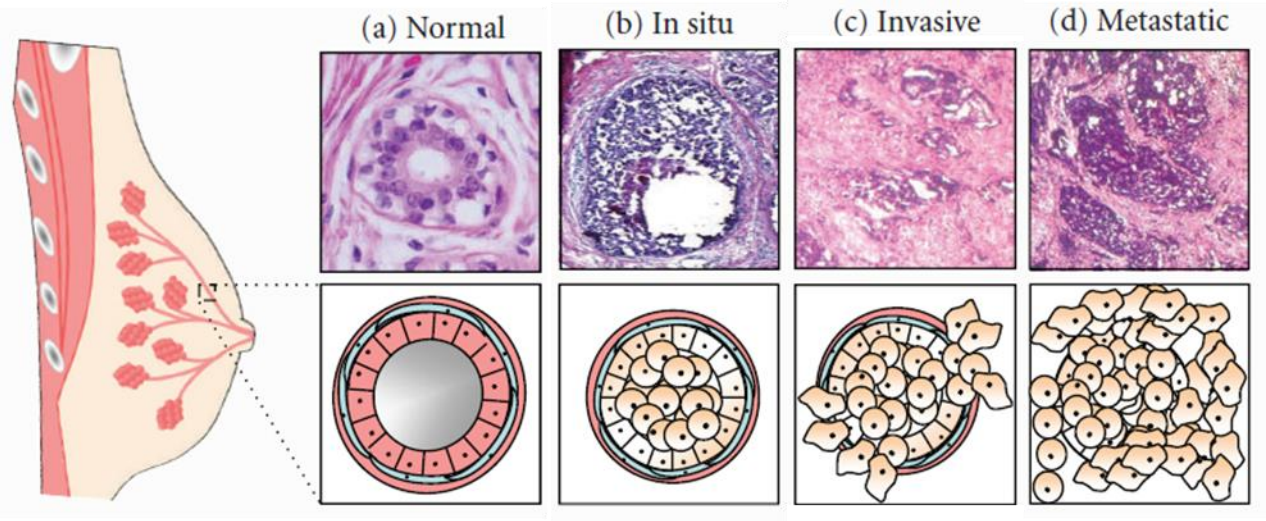


Figure 6 Stages of breast carcinoma development and progression

Breast carcinomas, which account for 99% of breast malignancies, arise in epithelial cells in the ducts of the breast (a). In early stages of the disease – referred to as ductal cancer in situ, growth is restricted to the lumen of ducts (b). In the course of the disease breast carcinomas can break the basement membrane and thereby become invasive (c). Metastatic breast cancer is the end stage of the disease with cancer cells invading the vessels and inducing metastatic growth at the distant site (d). The image was adapted from [242]

Even breast cancer of a single stage is not one homogenous disease but can be further distinguished based on its molecular characteristics. Based on gene expression profiles and hormone receptor expressions, breast cancer is commonly divided into four main subgroups: luminal A, luminal B, Human epidermal growth factor receptor 2 positive (HER2+) and basal like breast cancer [243]. Breast cancer classification in that way adds significant clinical information on top of the TNM staging applied for all cancer entities. Luminal A and B tumors are estrogen receptor (ER) and progesterone receptor (PR) positive. They are distinguished by the

expression of the proliferation marker Ki67 [244]. Luminal A tumors express only low levels of Ki67 and are slow growing, whereas luminal B tumors express high Ki67 levels. Luminal tumors are the most common form of breast cancer and show the best prognosis [245]. However, luminal B tumors due to the higher proliferation rate are more prone to early recurrence and worse prognosis [244, 245]. Basal like, which is also known as triple negative breast cancer, and HER2+ breast tumors are negative for the hormone receptors ER and PR and are distinguished by HER2 expression, with basal like being negative for HER2 [244]. These two subtypes have significantly worse survival rates and poorer outcomes than the luminal subtypes [245].

Improvements in screening and detection methods such as mammography and ultrasonography allow early detection of the primary tumor. This, in combination with more efficient surgery and follow-up treatment, has improved the outcome of the disease [246]. Nevertheless, the metastatic stage of the disease is still incurable in many cases [246]. In breast cancer mainly four metastatic sites are observed: lung, liver, bone and brain (**Table 1**). The organ tropism as well as the latency time for metastatic progression is dependent on the molecular subtype [246, 247]. Luminal tumors show significant longer relapse-free and overall survival than basal like tumors. While patients with a luminal tumor are at risk for metastatic progression even decades after removal of the primary tumor, the majority of patients with basal disease, will develop metastasis already between one and two years after diagnosis [245-247]. Over all, bone metastases are most common in breast cancer patients with different preferences in the individual subtypes. In basal like tumors, the lung is a common metastatic location, whereas bone metastases are most common in hormone receptor positive tumors. Her2+ tumors, in comparison tend to metastasize to brain and liver [247, 248]. However, no metastatic site is exclusive to one subtype and there is no predictor for the organ tropism of an individual patient.

Dependent on the stage of the disease and the molecular subtype, different treatment strategies can be applied in the clinic. Surgery of the primary tumor, which is performed as mastectomy or breast-conserving surgery dependent on size and invasiveness of the tumor, is the most common treatment option [246]. In case of an invasive disease at diagnosis, surgery is often combined with neoadjuvant therapy, which means systemic treatment prior to surgery to eliminate locally spread cells, shrink the tumor and test the efficiency of the treatment [249]. Adjuvant therapy is given to patients with micrometastases at the time of surgery to reduce the relapse risk. Commonly applied therapy options are radiotherapy especially in the neoadjuvant setting [246], endocrine therapy, blocking hormone receptor signaling for luminal tumors [250], targeted therapy using trastuzumab for HER2 positive tumors [250-252] and chemotherapy, which is the only treatment option for triple negative tumors [253]. The efficiency of the treatment is highly dependent on the molecular subtype with luminal and Her2+ tumors responding well to targeted therapy and basal-like tumors being difficult to treat due to lack of specific treatments. Furthermore, there is only limited knowledge on the factors influencing metastatic growth leading to risk of overtreatment or underestimation of disseminated cancer cells.

Once metastases are established, a palliative treatment is employed with the focus on reducing pain rather than cure of the disease [246]. A better understanding of metastatic progression of breast cancer is needed to develop therapies targeting the early steps of breast cancer metastasis.

5 Aim of the study

Metastasis is the main cause for cancer related deaths in most tumor entities and can occur months or even years after the removal of the primary tumor. This indicates that a subpopulation of cancer cells can survive in secondary organs for extended time and maintain the ability to trigger metastatic growth. Despite the essential role of such long-lived, non-proliferative cancer cells for patient outcome, only little is known about the molecular mechanisms that regulate survival, dormancy and reactivation of latent cancer cells. A deeper understanding of the mechanisms driving these processes is essential to being able to target these cells and optimize anti-metastatic therapy.

Since breast cancer is among the cancer entities, in which clinical dormancy plays an important role in disease progression, the aim of this work was the comprehensive analysis of latent breast cancer cells. The first step of the project was to identify locations of latent disseminated cancer cell (DCC) pools independent of the metastatic organ tropism in *in vivo* models. A potential pathological role of these DCCs was investigated through functional characterization with a focus on the tissue location of DCCs within the organs, their survival capacity, growth potential as well as therapy response *in vivo*. Having identified latent DCC pools with possible clinical significance, transcriptomic profiling was performed with the aim to identify molecular mechanisms that characterize DCCs compared to the mammary tumor and lung metastasis. Based on these results, the goal was to identify molecular mediators of survival and therapy responses of DCCs. Such targets and mechanisms may potentially be useful for the development of therapeutic strategies targeting latent DCCs in breast cancer.

6 Materials and methods

6.1 Cell culture

All media used for cultivation of cell lines as well as patient samples are summarized in **Table 2**.

Table 2 composition of all used cell culture media

medium	cells	ingredients	supplier	final concentration
D10f				
	HEK293T	DMEM GlutaMAX™	Thermo	
	4175	Fetal calf serum (FCS)	Thermo	10% (vol/vol)
	MDA-MB-231	Penicillin/Streptomycin (10000 U/ml)	Thermo	1% (vol/vol)
	MCF7	Amphotericin B	US biologics	1% (vol/vol)
DMEM/F12				
	SUM159	DMEM/F12 GlutaMAXTM	Thermo	
		Fetal calf serum (FCS)	Thermo	5% (vol/vol)
		Penicillin/Streptomycin (10000 U/ml)	Thermo	1% (vol/vol)
		human Insulin	Sigma	5 µg/ml
M199				
	BA	Medium 199	Thermo	
	BPE	Fetal calf serum (FCS)	Thermo	2.5% (vol/vol)
		Penicillin/Streptomycin (10000 U/ml)	Thermo	1% (vol/vol)
		L-Glutamin	Thermo	200 mM
		human Insulin	Sigma	10 µg/ml
		Hydrocortisone	Sigma	0.5 µg/ml
		epidermal growth factor (EGF)	Sigma	20 ng/ml
		Cholera toxin	Sigma	100 ng/ml
M87				
	BA	DMEM/F12 GlutaMAXTM	Thermo	
	BPE	Fetal calf serum (FCS)	Thermo	2% (vol/vol)
	DCCs	Insulin-Transferrin-Selenium	Life technologies	0.7x
		Penicillin/Streptomycin (10000 U/ml)	Thermo	1% (vol/vol)
		epidermal growth factor (EGF)	Sigma	5 ng/ml
		Hydrocortisone	Sigma	0.3 µg/ml
		Cholera toxin	Sigma	0.5 ng/ml
		Triiodo-L-Thyronine	Sigma	5 nM
		Estadiol	Sigma	0.5 nM
		Isopreterenol	Sigma	5 µM
		Ethanolamine	Sigma	50 nM
		Phosphoryl ethanolamine	Sigma	50 nM

medium	cells	ingredients	supplier	concentration
Onco2				
	SUM159	HUMEC Basal serum-free medium	Thermo	
	MDA-MB-231	Penicillin/Streptomycin (10000 U/ml)	Thermo	0.25% (vol/vol)
		human Insulin	Sigma	5 µg/ml
		epidermal growth factor (EGF)	Sigma	20 ng/ml
		human fibroblast growth factor- basic (bFGF)	Invitrogen	10 ng/ml
		B27 supplement	Thermo	2% (vol/vol)

6.1.1 Cultivation of cell lines

MDA-MB-231, 4175 [16] (both provided by Joan Massague) and MCF7 cells (ATCC) were cultured in D10f medium (**Table 2**). SUM159 cells (Asterand Bioscience) were cultured in DMEM/F12 complete medium (**Table 2**). All cell lines were maintained in adherent monolayers at 37°C and 5% CO_2. All cell lines were previously transduced with a TGL-reporter construct [254], which encode for a green fluorescence protein (GFP), the firefly luciferase (luc+) and the herpes simplex virus Thymidine kinase 1.

For passaging, the cells were incubated with 0.25% Trypsin-EDTA (Thermo) at 37°C for 5-10 min. Once detached, trypsin was deactivated by addition of fetal calf serum (FCS) containing medium. Cells were passaged in 1:5 – 1:12 ratios depending on the cell type and experiment. For cryopreservation, cells were pelleted at 330 g for 5 min and re-suspended in their appropriate cell culture medium, with addition of 10% Dimethyl Sulfoxide (DMSO) Hybri-Max® (Sigma). Cells were progressively frozen to a temperature of -80 °C in a freezing container filled with Isopropanol (Sigma). After at least 6h, the cells were transferred to liquid nitrogen for long-time storage. Cells were reconstituted by thawing in a 37°C water bath and re-suspended in complete culture medium. 18-24h after thawing, the medium was exchanged to fresh medium. Cell numbers were assessed using a Vi-CELL XR (Beckman Coulter).

6.1.2 Pleural effusion and ascites cultures

Pleural effusion and ascites samples were collected from breast cancer patients at the Department of Gynecology, University Clinic Mannheim, University of Heidelberg, and the National Center for Tumor Diseases Heidelberg (NCT). The studies were approved by the ethical committee of the University of Mannheim (case number 2011-380N-MA) and the University of Heidelberg (case number S-295/2009) and were performed according to the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report. Written informed consents are available from all patients.

Upon obtainment, effusion samples were centrifuged at 300g for 5 min and the cell pellets were re-suspended in ACK red blood cell lysis buffer. The lysis reaction was incubated for 5 min at RT and stopped by addition of 30ml PBS. Cells were cultured in adherent monolayers in a 1:1 mix of M199 and M87 [255] medium (**Table 2**) at 37°C and 5% CO_2.

6.1.3 Preparation of cell suspensions from mouse organs

Cell suspensions from mouse lung, liver, pancreas, spleen, kidney and mammary tumors were prepared enzymatically. The adrenal glands were removed from the kidneys and the gallbladder was removed from the liver. The organs were cut in small pieces in Phosphate-buffered saline (PBS) (Sigma) containing 1% (w/vol) Dispase II (Life technology), 0.5% Collagenase III (PAN Biotech) and 1µg/ml DNAse (Sigma). The enzymatic digestion of the tissue was carried out at 37°C for 1h for lung, kidney, spleen and mammary tumors. Pancreas and liver were incubated at room temperature (RT) for 45 min followed by incubation at 37 °C for further 15 min. After complete digestion, the reaction was stopped by addition of PBS and the suspension was filtered through 70 µm EASYstrainer^{TM} (Gibco). The cells were centrifuged at 330 g for 5 min and re-suspended in 1x ACK Lysis buffer (Lonza) to lyse red blood cells. The mix was incubated for 5 min at RT and the reaction was stopped by addition of 15-20 ml PBS. For the pancreas 1mg/ml Trypsin inhibitor from *Glycine max* (Sigma) was added to each step. Bone marrow cells were obtained by crushing the bone in PBS. Red blood cells were lysed by incubation in ACK buffer for 5 min. The obtained cell suspensions were subjected to further analysis.

6.1.4 Cultivation of DCCs

To enrich for DCCs in culture, tissue cell suspensions were prepared as described above. Cells were cultured in M87 medium [255] (**Table 2**). 2 µg/ml puromycin (Sigma) was added to the medium to select for tumor cells that carry a puromycin resistance gene in addition to the TGL-vector. Wild-type NSG organs were used as selection control. Puromycin treatment was stopped when no living cells remained on the control plates. The selection for cancer cells was confirmed by detection of GFP using a Nikon Eclipse Ti-E microscope.

6.1.5 Oncosphere cultures

For oncosphere cultures, MDA-MB-231 were detached from the culture plates, re-suspended in PBS and counted using a Vi-CELL XR (Beckman Coulter). Cells were plated in a cell density of 25,000 cells/ml in Onco2 medium (**Table 2**) in ultra-low adhesion culture flasks (Corning). For quantification and size measurement of spheres 10,000 cells/ml were plated in ultra-low adhesion 96-well plates (Corning). Spheres were analyzed 5-7 days after seeding.

Quantitative analysis of sphere number and sizes was performed using the ImageJ 2.0.0 software by counting the number of spheres and measuring sphere area to determine the size.

6.1.6 Label retaining experiment

Label retaining experiments were performed in MDA-MB-231 that were cultured in oncosphere conditions. Cells were labeled using the PKH26 Fluorescent cell linker kit (Sigma). Therefore, adherent growing cells were trypsinized, counted using the ViCell XR and spin down at 400 g for 5 min. The cell pellet was resuspended at $1x10^7$ cells/ml in Diluent C and the same volume of Dye-solution ($4x10^{-6}$ M) was added and mixed well. The solution was incubated for 5 min at RT with occasional mixing. The reaction was stopped by addition of the same volume FCS and incubation for 1 min. The cells were pelleted at 400 g for 5 min, were washed two times in 10 ml FCS containing culture medium (**Table 2**) and once in 10 ml Onco2 medium. PKH26 labelled cells were seeded in oncosphere conditions as described above.

6.2 Generation of stable knockdown cell lines

MDA-MB-231 or SUM159 cells were genetically modified to stably knockdown the tetraspanin genes TSPAN1 and TSPAN8 or YAP1 using the miRE lentiviral vector system [256]. The miRE technology is based on the miR30 backbone, which has been modified by Fellman *et al.* to increase mature shRNA levels and improve knockdown efficiency [256]. The original SGEP vectors were generously provided by Johannes Zuber (IMP- Research institute for molecular pathology, Vienna). In our laboratory, the GFP expression sequence was replaced by a tagBFP or tdTomato cassette as indicated. In addition the vectors carry an antibiotic selection marker. In this study the BFP or GFP cassette was used in combination with a puromycin resistance and the tdTomato with a zeocin resistance casette. Vector maps are depicted in **Figure 7.**

shRNAs targeting TSPAN8, TSPAN1 or YAP1 were designed using the shERWOOD shRNA library [257]. The sequences are listed in **Table 3**.

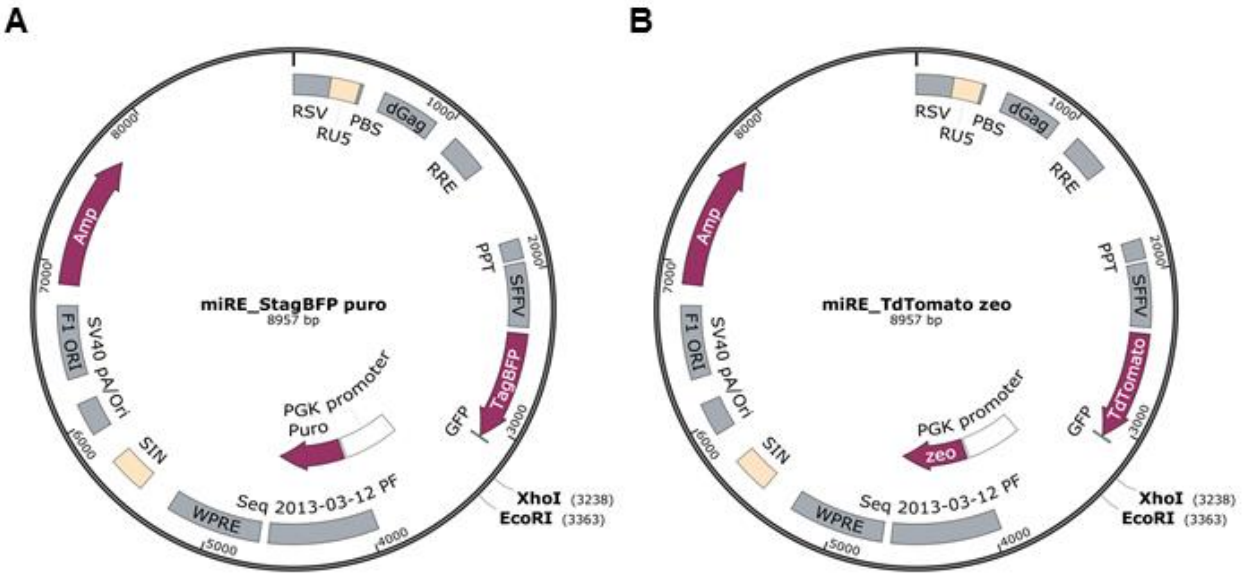


Figure 7 miRE lentiviral vector backbones

The lentiviral miRE vector system [256] was used for the generation of stable knock-down cell. The original SGEP vector was obtained from Johannes Zuber (IMP). Two different backbones were generated in our laboratory by exchanging the GFP sequence by a tagBFP cassette (**A**) or a tdTomato (**B**). In addition, the vectors express an antibiotics resistance gene under the control of the PGK promoter. The tagBFP was combined with a puromycin (puro) resistance (**A**), the tdTomato with a zeocin (zeo) resistance gene (**B**). The restriction sites for the enzymes XhoI and EcoRI that were used to clone in the knockdown sequences are highlighted.

Table 3 small hairpin RNA (shRNA) sequences used to generate knockdown cell lines

name	vector backbone	oligosequence (5`-3`)
NonSil	miRE StagTomato-zeo miRE StagBFP-puro miRE StagGFP-puro	TGCTGTTGACAGTGAGCGTAGATAAGCATTATAATTCCTTAGTGAAGCCAC AGATGTAAGGAATTATAATGCTTATCTACTGCCTCGGA
TSPAN1 2-1	miRE StagTomato-zeo	TGCTGTTGACAGTGAGCGCGCTGTGGCTTCACCAACTATATAGTGAAGCC ACAGATGTATATAGTTGGTGAAGCCACAGCATGCCTACTGCCTCGGA
TSPAN1 6-4	miRE StagTomato-zeo	TGCTGTTGACAGTGAGCGAAGGAAGACTTCACTCAAGTGATAGTGAAGC CACAGATGTATCACTTGAGTGAAGTCTTCCTGTGCCTACTGCCTCGGA
TSPAN8 3-3	miRE StagBFP-puro	TGCTGTTGACAGTGAGCGCGAGACCTGTATTTCTTTCATATAGTGAAGCCA CAGATGTATATGAAAGAAATACAGGTCTCTTGCCTACTGCCTCGGA
TSPAN8 4-4	miRE StagBFP-puro	TGCTGTTGACAGTGAGCGCAGGAGCTGTTTTCAAATCTAATAGTGAAGCC ACAGATGTATTAGATTTGAAAACAGCTCCTATGCCTACTGCCTCGGA
YAP1 3-2	miRE StagGFP-puro	TGCTGTTGACAGTGAGCGCCACATCGATCAGACAACAACATAGTGAAGC CACAGATGTATGTTGTTGTCTGATCGATGTGATGCCTACTGCCTCGGA
YAP1 4-3	miRE StagGFP-puro	TGCTGTTGACAGTGAGCGACATGTTCGAGCTCATTCCTCATAGTGAAGCC ACAGATGTATGAGGAATGAGCTCGAACATGCTGCCTACTGCCTCGGA

6.2.1 Molecular cloning

In the first step, the hairpins (**Table 3**) were PCR amplified using primers that had restriction site extensions. The forward primer (TGAACTCGAGAAGGTATATTGCTGTTGACAGTGAGCG) contained a XhoI restriction site, the reverse primer (TCTCGAATTCTAGCCCCTTGAAGTCCGAGGCAGTAGGC) and EcoR1 restriction site. In the PCR reaction 1 ng of the desalted hairpins (Sigma) was amplified with 1x HF buffer, 0.2 mM dNTPs, 0.5 mM of each primer and 1 µl Phusion High-fidelity Flex Polymerase (all components from NEB) in a total volume of 50 µl. The reaction was carried out using the following program: 30 sec at 98 °C, 33 rounds of amplification at 98°C for 10 sec, 58 °C for 20 sec and 72 °C for 5 sec followed by a final elongation for 600 sec at 72 °C. The PCR product was resolved on a 2% agarose gel and the 150bp band containing the hairpin was purified using a gel purification kit (Qiagen) following the manufactures instructions. The PCR product was eluted in 30 µl dH_2O. In the next step, the Hairpin was cut with the restriction enzymes XhoI and EcoRI. Therefore, 1x Cutsmart buffer (NEB), 2µl EcoRI- high fidelity (HF) (NEB) and 2 µl XhoI (NEB) were added to 30 µl of eluate and were incubated in a total reaction volume of 60 µl for 1 h at 37 °C. The product was purified using PCR purification kit (Qiagen) and was eluted in 30 µl of dH_2O. The concentration was measures using a NanoDrop 1000 spectrophotometer (Peqlab).
10 µg of the miRE lentiviral vectors (**Figure 7**) were cut as described above with EcoRI-HF (NEB) and XhoI (NEB) in 1x Cutsmart buffer (NEB). The reaction was carried out for 1 h at 37 °C. The cut vectors were dephosphorylated by addition of 1x Arctic phosphatase buffer (NEB) and 1 µl arctic phosphatase (NEB). The dephosphorylation reaction was incubated for 15 min at 37 °C followed by enzyme inactivation at 65 °C for 5 min. The cut vectors were purified on a 2% agarose gel. The vectors were extracted from the 8000 bp band using a gel purification kit (Qiagen). The plasmids were eluted in 30 µl dH_2O and the concentrations were measured using a NanoDrop 1000 spectrophotometer.

For the ligation reaction 60 fmol of the vector backbone and 180 fmol of the hairpin were incubated in 1x ligation buffer and 1 µl T4-ligase (Thermo) for 1 h at RT. 1 µl of the ligated vector was transformed into 19 µl electro competent electroMax™ Stabl4™ bacteria (Thermo) using a MicroPulser® electroporator (BioRad). The bacteria were diluted in 480 µl S.O.C. medium (Thermo) and incubated shaking at 37 °C for 1 h. 200 µl of the bacteria suspension were plated on Luria-Bertani (LB) Agar-plates containing 0.1 mg/ml ampicillin (Sigma) and were incubated at 37 °C. 24 h after platting clones were picked and expanded in 5 ml LB-cultures for further 24 h at 37 °C. DNA was extracted using a DNA MiniPrep kit (Qiagen) and positive clones were identified by sequencing (performed at GATC). Correct clones were further expanded in 200 ml of LB-medium with 0.1 mg/ml ampicillin for further 18-24 h. Plasmid DNA was isolated using Quick MaxiPrep Kit (Qiagen) and stored at -20 °C for further usage.

6.2.2 Lentiviral production

Lentiviral particles containing knockdown plasmids were produced in HEK293T cells. The cells were grown to a confluency of 80% in D10f medium (**Table 2**) in a 10 cm dish and transfected with 7.2 µg of the vector of interest, 2.4 µg envelope plasmid (pMD2.G) and 6.4 µg packaging plasmid (psPAX2). The plasmid mix was diluted in a total volume of 1 ml Opti-MEM Reduced Serum Medium (Thermo). In a separate tube 48 µl of Lipofectamin2000 Transfection Reagent (Thermo) was added to 950 µl of Opti-MEM medium and incubated for 10 min at RT. The plasmid mix was added to the Lipofectamin2000 trop-wise and incubated for 30 min at RT. HEK293T cells were washed with PBS and 8 ml fresh DMEM GlutaMAX^{TM} (Thermo) containing 10% (vol/vol) FCS (Thermo) was added. The plasmid-Lipofectamin mix was added to the HEK293T cells. 18-24 h after transfection the medium was replaced by fresh medium D10f medium. Virus-containing supernatant was harvested 24 h and 48 h later and was filtered through 0.2 µm Acrodisc Supor Membranes (Pall Corporation) before being aliquoted and frozen at -80 °C.

6.2.3 Lentiviral infection

Cell lines were transduced at a confluency of 60-70%. To obtain that, 600,000 MDA-MB-231 or 450,000 SUM159 cells were seeded in 6 cm culture plates 24 h prior to infection. For the infection, the culture medium was replaced by 1.2 ml of fresh medium, containing 16 µg/ml Polybrene (Sigma). 1.2 ml of non-concentrated virus particles was added to the cells. 24 h after infection the medium was replaced by fresh culture medium (**Table 2**). Further 24h later, the cells were passaged and infected clones were selected by addition of 2 µg/ml Puromycin (Sigma) or 500 µg/ml Zeocin (Sigma), respectively depending on the selection marker present on the transduced plasmid (**Figure 7**). One culture plate containing control cells that were not infected, was always used as selection control. The antibiotic selection was stopped, when no living cells remained on the control plate.

6.3 Protein analysis

6.3.1 Fluorescence activated cell sorting

Fluorescence activated cell sorting (FACS) was performed on cultured cells as well as on cells isolated from mouse organs. Cultured cells were trypsinized as described above, washed in PBS and $1*10^6$ cells were used for FACS analysis. For surface marker staining the cells were incubated with the appropriate concentrations of antibodies (**Table 4**) diluted in FACS buffer for 30 min on ice in the dark. FACS buffer was composed of PBS supplemented with 2% FCS and 2 mM EDTA (Sigma). Following labelling, the cells were washed with 4 ml FACS buffer and centrifuged at 330 g for 5 min. For analysis the cells were re-suspended in 500 µl of FACS buffer containing 3µg/mL DAPI (Biolegend).

Cell suspensions prepared from mouse organs as described on page35 were incubated in 1x mouse FC-Receptor Block (BD) diluted in FACS buffer for 10 min on ice to minimize background staining. The same volume of FACS buffer containing the antibody cocktail (**Table 4**) was added to the cells and incubated for further 30 min protected from light on ice. For analysis 3µg/mL DAPI (Biolegend) was added and the cells were re-suspended in 500 µl of FACS buffer.

Table 4 Fluorophore coupled antibodies used for flow cytometry

gene	fluorophore conjugate	source	company	clone	dilution
mCD31	eFlour450	rat, IgG2a, kappa	eBioscience	390	1:3000
mCD11b	eFlour450	rat, IgG2b, kappa	eBioscience	M1/70	1:2000
mCD45	pacific blue	rat, IgG2b, kappa	Biolegend	30-F11	1:500
mH2kD	pacific blue	mouse, IgG2a kappa	Biolegend	SF1-1.1	1:30
hCD298	PE	mouse, IgG1 kappa	Biolegend	LNH-94	1:150
hTSPAN8	Alexa647	rat, IgG2b, kappa	R&D	45881	1:50
hTSPAN1	Alexa750	mouse, IgG3	R&D	819202	1:15
hHLA-ABC	APC	mouse, IgG1 kappa	BD	G46-2.6	1:15
hHLA-DR	PE-Cy7	mouse, IgG2a kappa	Thermo	L243	1:15

AnnexinV staining was performed by diluting AnnexinV-Alexa647 (Biolegend) 1:20 in 1x AnnexinV binding buffer (Biolegend). The reaction was carried out for 15 min at RT in the dark.

Samples were analyzed on LSR Fortessa (BD) and LSRII (BD) cytometers. Cell sorting was performed on Aria I (BD), Aria II (BD) or Fusion (BD) cytometer of the Cytometry core facility of DKFZ. The data was analyzed using the FlowJo Version 10.1r7 software (FlowJo LLC).

6.3.2 Immunofluorescence staining

Mouse organs were resected and fixed in 10% formalin (Sigma) for 7-8 h at 4 °C rotating, followed by dehydration in 30% (w/vol) Sucrose (Sigma) in PBS overnight at 4 °C. The organs were imbedded in Tissue Tek OCT Standard (Sakura) in Tissue TeK CryoMold-cassettes (Sakura) and frozen on dry ice. Frozen organs were stored at -80°C for further usage.

For immunofluorescence (IF) staining the OCT blocks were cut in 5-8 µm slices on Superfrost Plus™ glass slides (Thermo) using a Microtome MH525 cyrotome (Thermo). The cuts were air-dried for 1-2 h at RT and the slides were washed in PBS twice for 5 min each. The slides were blocked using TNB buffer containing 0.1 M Tris-HCl (pH7.5), 0.15 M sodium chloride (NaCl) and 0.5% (w/vol) Blocking Reagent (Perkinelmer, FP1020) for 1 h at RT. After blocking, the primary antibodies (**Table 5**) diluted in TNB buffer were pipetted on the cuts and incubated overnight at 4°C. The slides were washed three time in PBS + 0.05% (vol/vol) Tween20 (Sigma) for 5 min and were incubated with the secondary antibodies (**Table 5**) diluted in TNB buffer for 1 h at RT. Following the antibody incubation, slides were again washed three times in PBS + 0.05% (vol/vol) Tween20 for 5 min. The cuts were imbedded in FluoromountG (Southern Biology) and covered using coverslips (MetzerGlas). The coverslips were sealed with nail polish for long time storage.

Table 5 Primary and secondary antibodies used for immunofluorescence staining

primary antibodies	**source**	**supplier**	**clone**	**dilution**
GFP	chicken IgY	Abcam	ab13970	1:1000
CD31	rat IgG2 kappa	BD	MEC13.3	1:100
phospho-Histone H3 (Ser10)	rabbit IgG	Thermo	9HCLC	1:50
cleaved Caspase3 (Asp179)	rabbit IgG	Cell signaling	D3E9	1:100
secondary antibodies	**source**	**supplier**	**fluorophore**	**dilution**
anti-chicken	goat IgG	Themo	GFP	1:500
anti-rabbit	goat IgG	Themo	Cy5	1:500
anit-rat	goat IgG	Themo	Cy3	1:500

The stained slides were analyzed using a Zeiss LSM710 ConfoCor 3 confocal microscope (Zeiss). Pictures were taken with the magnifications indicated in the figures. Staining was quantified using the Zen 2.3 software (Zeiss).

6.4 Hematoxylin and Eosin staining

Hematoxylin and Eosin (H&E) staining was performed on formalin fixed and OCT embedded lungs, generated as described in the paragraph above. The staining steps are summarized in **Table 6.** After the staining, the slides were air dried, imbedded in Cytoseal XYL (Thermo Fischer) and sealed with coverslips. The slides were dried overnight, before they were analyzed using a Zeiss Observer.Z1 microscope.

Table 6: H&E staining protocol

reagent	supplier	time
Xylol	VWR	3 min
Xylol		3 min
100% EtOH	VWR	3 min
95% EtOH		3 min
80% EtOH		3 min
70% EtOH		3 min
50% EtOH		3 min
dH2O water		1-2 min
Hematoxylin solution according to Mayer	Sigma	6 min
dH2O water		5-10 min
Tap water		5 min
70% EtOH + 0.25% HCl		30-60 s
Tap water		5 min
dH2O water		5 min
Eosin Y solution	Sigma	1 min
95% EtOH		1 min
100% EtOH		1 min
Xylol	VWR	15 min

6.5 Gene expression analysis

6.5.1 RNA extraction

RNA extraction of cells maintained in cell culture was performed by directly lysing the cells on the culture plate in RLT buffer (RNeasy Kit Qiagen) supplemented with 1% (vol/vol) β-Mercaptoethanol (Thermo). The lysate was passed through RNAshredder columns (Qiagen) by centrifugation at 13,000 rpm for 3 min. The same volume of 70 % Ethanol was added to the cell lysate and the RNA was extracted using the RNeasy® Mini kit (Qiagen) following the manufacturer's instructions. The DNase step was only performed for samples that were subjected to gene expression analysis. The RNA was eluted in 30-50 µl of RNAse free ddH2O and the RNA concentration was measured using a NanoDrop 1000 spectrophotometer (Peqlab).

RNA of sorted cancer cells, isolated from mouse organs, was extracted using Arcturus pico pure RNA isolation kit (Thermo) following the manufacturer's instructions. Up to 50,000 cells were directly sorted into 150 µl extraction buffer and frozen at -80°C until extraction. RNA purification columns were preconditioned with 250 µl Conditioning buffer and incubated for 5 min at RT. The columns were centrifuged at 16,000 g for 1 min. The RNA extracts were supplemented with an equal volume of 70 % Ethanol and added on the pre-conditioned columns. RNA was bound by centrifugation at 100 g for 2 min, followed by centrifugation for 30 sec at 16,000 g. The columns were washed by addition of 100 µl Wash buffer1 and centrifuged at 8,000 g for 1 min. For DNase digestion, 5 µl DNase in 35 µl RRD buffer (Quiagen) was added on the membranes and incubated for 15 min at RT. After incubation, 40 µl of Wash Buffer1 were added on the membranes and the columns were centrifuged for 30 sec at 8,000 g. Another washing step was performed with 100 µl of Wash buffer2, which was followed by centrifugation at 16,000 g for 1 min. RNA was eluted in 10 µl of Elution buffer, which was incubated for 10 min at RT. RNA concentration was measured using a NanoDrop 1000 spectrophotometer.

Whole organ RNA isolation was performed in multiple steps, starting with the mechanical dissociation of the tissue using a GentleMACS dissociator (Miltenyi). Therefore, frozen organs were placed in GentleMACS TM M tubes (Miltenyi) in 600 µl RLT-buffer (RNeasy Kit; Qiagen) that was supplemented with 1% (vol/vol) β-Mercaptoethanol, and dissociated using the RNA2.1 program on the dissociator. The cell suspension was passed through RNAshredder columns (Qiagen) by centrifugation at 13,000 rpm for 3 min. The eluate was supplemented with 600 µl of 70% Ethanol and RNA was extracted using RNease RNA extraction kit (Qiagen) following the manufacturer's instructions.

6.5.2 Reverse transcription

RNA samples below a concentration of 50 ng/µl were reverse transcribed and amplified using the Ovation PicoSL WTA system V2 (Nugen). 5 µl of total RNA samples with a maximum of 50 ng RNA were used as input. The amplification was performed following the manufacture´s protocol. cDNA was purified using PCR-purification kit (Qiagen) and was eluted in 30 µl RNAse-free ddH_2O. cDNA concentrations were measured using a NanoDrop 1000 spectrophotometer. cDNA samples were diluted to a concentration on 10 ng/µl and stored at -20 °C.

RNA samples above a concentration of 50 ng/µl were reverse transcribed using High-Capacity cDNA Reverse Transcription Kit (Thermo). 500-2,000 ng of total RNA were reverse transcribed following the manufacturer's instructions. cDNA concentration was measured on a NanoDrop 1000 spectrophotometer and was diluted to a concentration of 10 ng/µl.

6.5.3 qRT-PCR

Quantitative real-time PCR (qRT-PCR) was performed using the SYBR Green protocol for detection of double-stranded PCR products. Per reaction 10-80 ng of cDNA, 7.5 µl SYBR® Green PCR Master Mix (Thermo) and 5 µM of each forward and reverse primer (**Table 7**) were used. The qRT-PCR reaction was performed on the ViiA ™ 7 Real-Time PCR System (Applied Biosystems) according to the following program: 50 °C for 2 min, 95 °C for 10 min (enzyme activation phase), 95 °C for 15 sec, 60 °C for 1 min for 40 cycles (amplification phase), 95 °C for 15 sec, 60 °C for 1 min, 95 °C for 15 sec (enzyme inaction and dissociation phase). The ViiA 7 ™ Software (Applied Biosystems) or the DataAssist Software were used for data analysis using the $2^{-\Delta\Delta Ct}$ method. Human TATA-box binding protein (hTBP), luciferase+ (luc+) or mouse Beta-2-Microglobulin (mB2M) were used as housekeeping genes (refer to **Table 7** for primer sequences) for normalization.

Primer specificity was confirmed by analysis of melting curves. Human specificity of the primers was tested using cDNA of un-injected organs of NSG mice. Mouse specificity was tested using cDNA from human breast cancer cell lines.

Table 7 Primer sequences used for qRT-PCR

symbol	gene name	forward primer (5`-3`)	reverse primer (5`-3`)
hALDOC	Aldolase C, fructose-bisphosphate	CAAATTGGGGTGGAAAACAC	TAGAGGCACCACACCCTTGT
hCTGF	Connective tissue growth factor	CCAATGACAACGCCTCCTG	TGGTGCAGCCAGAAAGCTC
hCYR61	Cellular Communication Network Factor 1	GTGACGAGGATAGTATCAAGGACC	ATTTCTGGCCTTGTAAAGGGTTG
hGLUT1	Glucose transporter type 1	ACGCTCTGATCCCTCTCAGT	GCAGTACACACCGATGATGAAG
hHLA-B	Major Histocompatibility Complex, Class I, B	CTAGCAGTTGTGGTCATCGGAG	ACAGCTGTCTCAGGCTTTTCA
hHLA-DMB	Major Histocompatibility Complex, Class II, DM beta 1	TAACCCCCTCTTACGGGGAC	AGAGGAGTGTAACTAGAGTGGC
hHLA-DPB	Major Histocompatibility Complex, Class II, DP beta 1	GCCTGGATAGTCCTGTCACC	TCCTGGAGCCAGATGCTAAC
hHLA-DQB	Major Histocompatibility Complex, Class II, DQ beta 1	TGGAGTGGCGGGCTCA	TTCTGGGCAGGCATAAGCAG
hHLA-DRA	Major Histocompatibility Complex, Class II, DR Alpha	ACAGAGCGCCCAAGAAGAAA	CTCGCCTGATTGGTCAGGAT
hNANOG	Nanog Homeobox	CACCTATGCCTGTGATTTGTG	AAGTGGGTTGTTTGCCTTTG
hOCT4	POU Class 5 Hemeobox 1	GATGGCGTACTGTGGGCCC	CAAAACCCGGAGGAGTCCCA
hPGK	Phosphoglycerate kinase	AAGGGAAGCGGGTCGTTAT	CCCAGCAGAGATTTGAGTTCTA
hSNAI2	Snail family transcriptional repressor 2	TGTTGCAGTGAGGGCAAGAA	GACCCTGGTTGCTTCAAGGA
hSOX2	SRY-Box 2	TGTCAAGGCAGAGAAGAGAGTG	GCCGCCGATGATTGTTATTA
hTBP	TATA-box binding protein	AGACCATTGCACTTCGTGCC	AAAATCAGTGCCGTGGTTCG
hTSPAN1	Tetraspanin 1	CACCATGAAAGGGCTCAAGT	ACCCACGGTGACTGCATTAG
		CCCTCGTGACGTTCTTCTTC	TCATGGTGGTGTTCCACACT
hTSPAN8	Tetraspanin 8	GCAGGTGTGAGTGCCTGTAT	TCGTACCCATATTGCTAATGC
		ACTTGCCTGGAGATAGCCTTT	AGGATCAAGATACCACATAGCCAG
hVIM	Vimentin	CTTCAGAGAGAGGAAGCCGA	ATTCCACTTTGCGTTCAAGG
hYAP1	Yes associated protein 1	CCTTCTTCAAGCCGCCGGAG	CAGTGTCCCAGGAGAAACAGC
luc+	Luciferase+	TGTCGCTCTGCCTCATAGAA	ACCGTGATGGAATGGAACAA
mB2m	Beta-2-Microglobulin	CCTGGTCTTTCTGGTGCTTG	CCGTTCTTCAGCATTTGGAT

6.5.4 Standard curve for qRT-PCR

To associate the measured CT-values with the number of cancer cells present in the organs, a standard curve was generated by spiking different numbers of cancer cells into NSG organs. Therefore, cell suspensions of NSG lungs and liver were generated as described on page 35. MDA-MB-231 cells were trypsinized and counted. 25, 50, 100, 250, 500, 1,000, 2,500, 5,000 or 10,000 cancer cells were added to 600,000 mouse cells. Each condition was prepared in five replicates. RNA was isolated and reverse transcribed to cDNA. qPCR was performed using hTBP and mB2m primers. The measured ΔCT of hTBP/mB2m was plotted against the logarithmic number of cancer cells spiked into the organs.

6.5.5 Gene expression analysis on Affymetrix human Genome U133 plu2.2 arrays

Gene expression analysis was performed to compare the gene expression profiles of MDA-MB-231 cancer cells, isolated from kidney, pancreas, lung and mammary tumor by FACS sorting. The cells were either directly sorted into RNA extraction buffer or were cultured for 48 h in M87 medium (**Table 2**). In a second experiment, control or TSPAN8 and TSPAN1 double knockdown cancer cells isolated from pancreas were gene expression profiled. Whole RNA of sorted DCCs was isolated with the Arcturus Pico Pure RNA isolation kit as described above (page 43). RNA quality and quantity was measured using Bioanalyser pico chip. As the RNA concentration was below 50 ng/µl, the RNA was pre-amplified using the non-exponential T7 *in vitro* transcription technology. The amplified samples were labeled and hybridized to Affymetrix Human Genome U133 plus2.0 arrays. Quality control, amplification, labeling and hybridization were performed by the Genomics and Proteomics core facility of DKFZ.

6.6 Bioinformatics analysis of gene expression data

The gene expression data obtained on Affymetrix arrays was analyzed using Chipster (version 3.11.X) [258]. Raw data was normalized using the Robust Multichip Average (RMA) normalization method with variance stabilization. The normalized data was filtered by standard deviation (SD) with a cut-off of 50%. Principal component analysis was performed on the filtered normalized data. Differentially expressed genes were assessed by empirical Bayes test using a Bonferroni and Hochberg (BH) adjusted p-value of 0.05. The DCC gene signatures were generated by comparison of pancreas DCCs or kidney DCCs with cells from the mammary tumor, respectively. All genes that were down- or upregulated with a fold change of at least 1.5-fold and a BH corrected p-value below 0.05 were used for the signatures. Gene Ontology (GO) analysis was performed in Chipster with the up- or downregulated signatures, respectively. A p-value cut-off of 0.05 was used. Gene set enrichment analysis (GSEA) was performed with the unfiltered RMA normalized data as described previously [259]. For the analysis, all genes are ranked regarding their fold change within the comparison between samples (DCCs) and controls (mammary tumor cells). Genes from publically available gene signatures are then aligned within the ranked gene list and enrichment score are calculated. An enrichment score of 1 indicates that all genes of a signature are randomly distributed within the ranked gene list. A positive enrichment score indicates that the genes of the signature are primarily found in the top of the ranked gene list and are therefore overexpressed in the DCC samples. A negative enrichment score indicates that the majority of genes from the signature are located in the bottom of the ranked list and therefore upregulated in the mammary tumor controls. Nominal p-values were determined on basis of random permutations of the gene set. P-values smaller than 0.05 or false discovery rates (FDR) below 0.25 were considered as statistically significant.

6.7 Mouse studies

All animal care and procedures were performed consistent with German laws and were reviewed and approved by Regierunspraesidium Karlsruhe for the state of Baden-Wuerttemberg. Experiments were performed under the authorization numbers G65/17 and G82/16. Mice were held and bred in the central animal laboratory of the DKFZ in individually ventilated cages under temperature control and with enrichment. Non-obese diabetic-severe (NOD) scid gamma null (NSG) female mice 6-10 weeks of age were used for all *in vivo* experiments.

6.7.1 Mammary fad pad injections

Orthotopic tumor experiments were performed by injection of cancer cell lines into the 4th mammary fat pad on both sites of NSG mice. The mice were injected subcutaneous with 5 mg/kg body-weight Carprofen (Rimadyl 50mg/ml injection solution) and were narcotized with isoflurane (Eduphar). 500,000 MDA-MB-231 cells were injected in a total volume of 50 µl per site in 1:1 PBS and growth factor reduced Matrigel (BD). Tumor growth was followed by bioluminescence imaging every 5-7 days. Therefore, the mice were injected intraperitoneal (IP) with 150 mg/kg body-weight D-Luciferin (Biothyn) and incubated for 10 min. Bioluminescence was measured using an IVIS Spectrum In Vivo Xenogen Imaging System (PerkinElmer). The experiment was ended after 5 weeks or when the tumor diameter reached 1 cm. At the endpoint, mice were again injected IP with D-Luciferin and narcotized with isoflurane. 10 min after injection the mice were sacrificed, lung, liver, pancreas, spleen, kidney and bones were resected and *ex vivo* bioluminescence of the organs was measured in the IVIS device. Organs were snap-frozen in liquid nitrogen for extraction of RNA or fixed in formalin for immunofluorescence staining.

6.7.2 Transplantation of patient derived Xenografts

Patient-derived xenograft (PDX) models, which have only been passaged *in vivo* and are therefore considered maintaining patient characteristics, have generously been provided by Prof Alana Welm. PDXs [255] were implanted into the 4th mammary fat pad of NSG mice. Tumor pieces of 3-6 mm diameter, which were frozen in FCS with 10% DMSO, were thawed at 37°C in a water bath and washed twice in M199 medium (Error! Reference source not found.). Tumor ieces were kept in fresh medium on ice until the injection. Mice were injected with 5 mg/kg body-weight Carprofen and narcotized with isoflurane. Small incisions were made into both mammary fat pads. The tumor pieces were placed into the cuts and the incisions were loosely closed. Tumor growth was followed by caliper measurements and the experiment was ended when the tumors reached 1 cm in diameter.

6.7.3 Tumor resection

To follow tumor cell dissemination for a longer time period, mammary tumors were resected when they reached 5-6 mm of diameter. Therefore, the mice were injected subcutaneous with 5 mg/kg body-weight Carprofen and were narcotized with isoflurane. The skin was cut open and the mammary tumor was removed using an electro cauterizer. The wound was closed and the mice were kept for further 6-8 weeks. Thereafter, the animal were sacrificed, organs were resected and used for further analysis.

6.7.4 Perfusion

When indicated, the blood of the animals was systemically replaced by PBS via perfusion [260] at the endpoint of an orthotopic *in vivo* experiment,. Therefore, the mice were narcotized by IP injection of 120 mg/kg body weight Ketamin und 14 mg/kg body weight Xylazin in PBS. Once the narcosis was very deep, which was secured by absence of reflexes, the mice were taped on to an inclined, to 37°C pre-warmed table. The thorax was cut open using scissors. As long as the heart was still pumping, a butterfly needle, which is connected to a syringe with PBS, was inserted into the left heart ventricle. A small cut was placed in the right atrium to allow the blood to flow out. Immediately, an automatic pump was started, pumping 30 ml PBS at a constant flow rate of 1.1 ml/min through the mouse body. The success of the perfusion was monitored by color changes in the organs and the tail. Organs were harvested after successful perfusion and analyzed by flow cytometry.

6.7.5 Tail-vein injection experiment

Mammary tumor, kidney, spleen and pancreas were harvested 5 weeks after orthotopic injection and cancer cells were isolated from organ cell suspensions by FACS. The cells were sorted into PBS. For each organ, a separate control was used containing the exact same number of cancer cells sorted from the mammary tumor. All samples were centrifuged at 350 g for 5 min and the cells were re-suspended in 100 µl PBS. The cells were immediately injected intravenously (IV) into the tail vein of NSG mice. Tumor growth was follow by bioluminescence measurement every 5-7 days. D-luciferin (150 mg/kg body-weight). was injected IP and luminescence was measured 10 min later. The experiment was terminated when the luminescence signal of control (mice injected with cancer cells from mammary tumor) or experimental animals (mice injected with tumor cells from kidney, pancreas or spleen) reached $8x10^7$ - $2x10^8$ p/s/cm^2/sr. At the endpoint, the mice were injected with luciferin and incubated for 10 min. Thereafter, the mice were sacrificed, organs resected and bioluminescence of the organs was measured ex vivo.

6.7.6 Chemotherapy treatment *in vivo*

Eight days after orthotopic injection of MDA-MB-231 breast cancer cells into NSG mice, the mice were injected IP with a combination of 50 mg/kg body weight Cyclophosphamid (Sigma) and 1.5 mg/kg body weight Doxorubicin (LC labs) or 20 mg/kg body weight Paclitaxel (Absource). The injection was repeated every fifth day for four treatment rounds. At each treatment day, the mice were weighted and mammary tumors were measured using a caliper. The mammary tumor volume was calculated using the formula:

Volume = 3/4*Π*length*width*height

The experiment was ended at day 28 or when the mice lost more than 20% of weight compared to the starting point. At the endpoint, the mice were sacrificed and organs were harvested and snap-frozen for RNA extraction.

7 Results

7.1 Identification of breast cancer cells in kidney, pancreas and spleen

In the first part of this study it is demonstrated that breast cancer cell dissemination in orthotopic *in vivo* models is not restricted to organs that are prone for the development of breast cancer metastasis, namely lung, liver, bone and brain. Disseminated breast cancer cells of cell lines as well as PDX models were detected in organs that do not develop metastatic growth, pancreas, spleen and kidney using various techniques (**Figure 8, Figure 9** and **Figure 10**).

Our data further indicates that the majority of disseminated cancer cells in kidney, spleen and pancreas extravasated from blood vessels but remained in close proximity to vessels (**Figure 12**).

7.1.1 Breast cancer cells disseminate to organs that do not support metastatic growth

In many tumor types, metastases only grow in selective organs, an observation, which is referred to as organ tropism of metastatic growth [173]. In breast cancer lung, liver, brain and bone are almost exclusively affected by metastasis. This specificity is caused by complex regulations during different steps of the metastatic cascade. On one hand, vessel structure and body architecture can dictate which organs are reached by cancer cells. On the other hand, tumor intrinsic signals as well as reactive cues from the microenvironment influence the growth potential of metastatic cells at the secondary site [223]. With this foundation, it was determined in breast cancer models, whether dissemination is restricted to lung, liver and bone, organs that are prone for metastasis development. To test this, the triple negative and highly metastatic breast cancer cell line MDA-MB-231, which has been transduced with a TGL reporter construct encoding thymidine kinase, green fluorescence protein (GFP) and luciferase [254], was injected orthotopically into immune deficient NSG mice (**Figure 8 A**). Five weeks after injection, lung, liver and bone, as well as organs not supporting metastatic growth of breast cancer cells, namely spleen, kidney and pancreas, were analyzed for the presence of cancer cells. To do so we made use of the luciferase enzyme expressed by the cancer cells, which photo-converts its substrate luciferin to oxyluciferin while producing light. Injection of luciferin into mice harboring tumor cells, allowed us to specifically detect cancer cells and follow their growth *in vivo* and *ex vivo* by measuring the intensity of the emitted light. In line with previous reports [215], MDA-MB-231 cells were detected in lung, bone and liver. In addition, luminescence signal was also detected in pancreas, spleen and kidney (**Figure 8 B**) but was about 10^3-times lower compared to the signal in the lung, where MDA-MB-231 cells form growing metastasis [16].

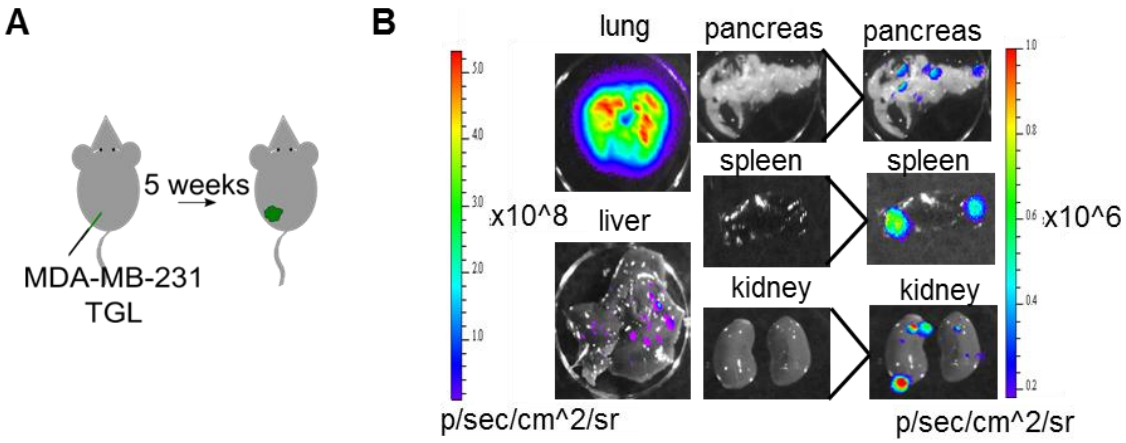


Figure 8 Breast cancer cells can be detected in several different organs

A Experimental set up used for all *in vivo* experiments if not stated otherwise: the triple negative breast cancer cells, MDA-MB 231, labeled with a TGL reporter construct expressing thymidine kinase, a GFP and a luciferase gene [254] were injected orthotopically into the fourth mammary fat pad of NSG mice. Five weeks after injections the mice were injected with luciferin, sacrificed 10 min after injection and *ex vivo* luminescence of the organs was measured using an IVIS Spectrum In Vivo Imaging System**.** Organs were snap frozen for RNA extraction, used freshly for FACS analysis or fixed and imbedded in Tissue Tek OCT for IF staining.

B Representative *ex vivo* images of lung, liver, pancreas, spleen and kidney of a NSG mouse five weeks after injection with MDA-MB-231 TGL cells.

Next, it was investigated whether the detected luminescence was triggered by the presence of cancer cells in the organs or if particles shaded from tumor cells, containing the luciferase enzyme may be the cause for the luminescence signal. To answer this question, I set up a fluorescence activated cell sorting (FACS) scheme to specifically isolate cancer cells (**Figure 9**). For FACS, cell homogenates of whole organs were prepared enzymatically and cells were stained with FACS antibodies targeting mouse and humane specific proteins. The gating strategy is shown in **Figure 9 A**. In the first step of the gating strategy, doublets were excluded based on their forward scatter (FSC) and sideward scatter (SSC) signal (**1**). Next, we excluded dead cells by excluding cells stained with DAPI (**2**), which intercalates into DNA of cells with a disrupted cell membrane. Next, we excluded mouse cells using an antibody cocktail targeting endothelial cells (mCD31+ cells) and immune cells (mCD45+, mCD11b+). In addition, we excluded all cells expressing the murine MCH class I protein (mH2kD+) (**3**). Having excluded these murine cell types, we positively selected cancer cells using the tumor cell endogenously expressed GFP as well as an antibody targeting the human β3 Na+/K+ ATPase subunit (CD298) (**4**) (**Figure 9 A**). The CD298 protein has previously been used as selective marker for human breast cancer cells in various mouse models [261]. The specificity and efficiency of CD298 for our models was confirmed using cell homogenates from organs of NSG mice as negative control and *in vitro* cultured breast cancer cell lines as positive controls. Representative results are shown in **Figure 9 B** for a pancreas homogenate and MDA-MB-231 cells. 100% of cancer cells expressed CD298 on the surface, while almost no background staining was detected in NSG organs. Similar results were obtained with different breast cancer cell lines, PDX models and NSG organs.

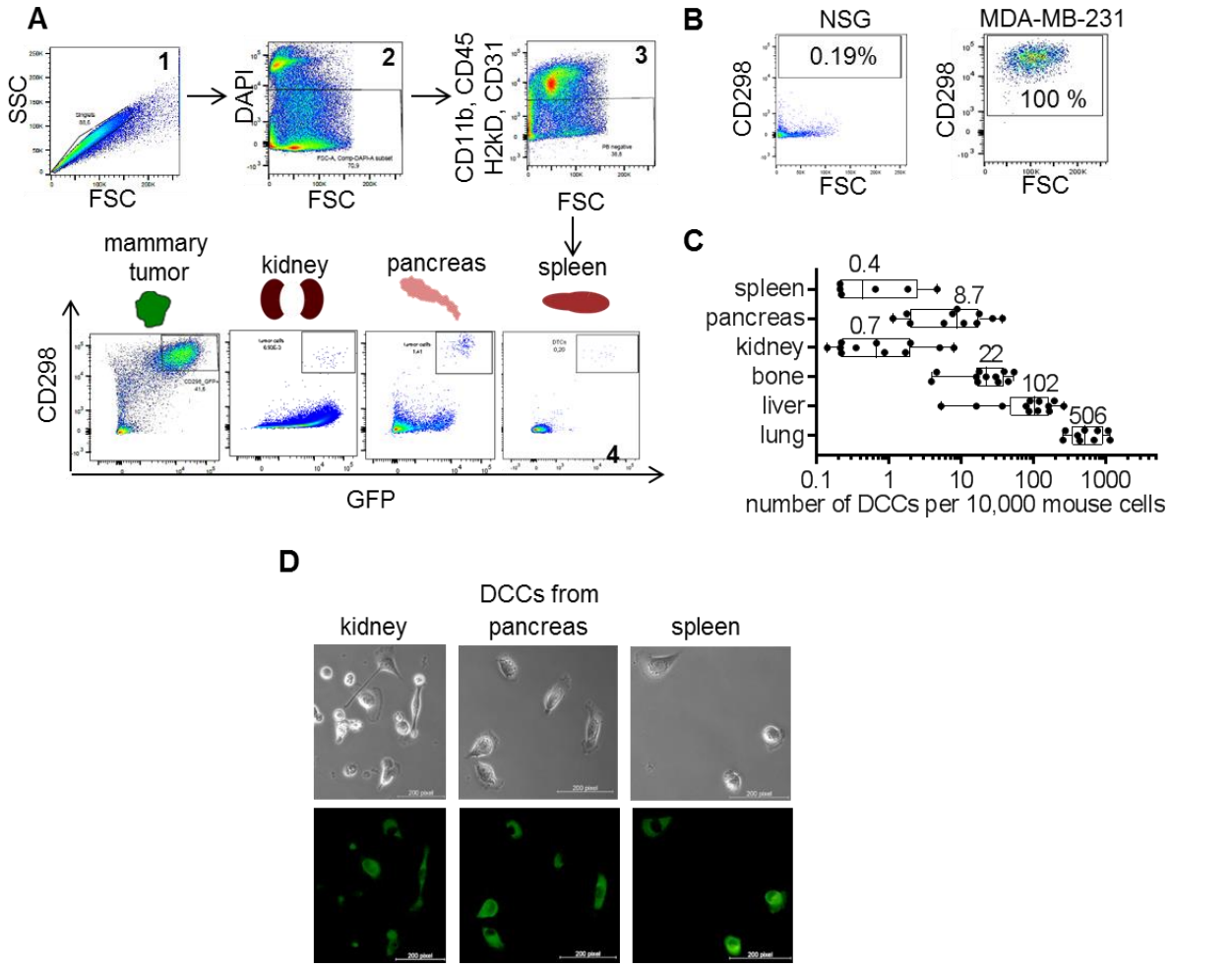


Figure 9 Breast cancer cells disseminate to several different organs including pancreas, spleen and kidney

A FACS strategy for isolation of cancer cells from the mammary tumor and different organs: whole organ cell homogenates were generated enzymatically using dispase and collagenase. Reb blood cells were lysed using ACK buffer and the following gating strategy was used to identify cancer cells. **1**) Singlets were selected based on their forward scatter (FSC) and sideward scatter (SSC) distribution. **2**) Dead cells were excluded with DAPI. **3**) Mouse cells of endothelial and immune origin were excluding using an antibody cocktail with mCd11b, mCd45, mH2kD and mCd31. **4**) Cancer cells were positively selected using CD298 expression [261] and the cell intrinsic GFP.

B Target specificity of the used CD298 antibody was confirmed with NSG organs (representative plot of a pancreas is shown) as negative controls. The efficiency of the antibody was tested for MDA-MB-231, SUM159 cells and the used PDX models. A presentative image obtained with MDA-MB-231 cells is shown.

C The number of MDA-MB-231 cancer cells in lung, liver, bone, pancreas, kidney and spleen of NSG mice five weeks after tumor cell injection was determined by FACS analysis using the staining strategy described in **A**. Cancer cell numbers are shown per 10,000 live mouse cells. The mean of 6-9 biological replicates is shown for each organ. Each data point represents one biological replicate.

D MDA-MB-231 TGL tumor cells harboring a puromycin resistance marker were isolated from kidney, pancreas and spleen of NSG mice five weeks after tumor cell injection. The cells were selected with puromycin (20 µg/ml) containing medium for 72h. Pictures were taken on a Nikon Eclipse Ti-E microscope with a 20x magnification. Upper panel: bright field; lower panel: GFP channel.

Using this sorting strategy, we were able to isolate and quantify tumor cells in different organs (**Figure 9 C**). As expected, the highest number of cancer cells was detected in the lung, with about 500 cancer cells in 10,000 analyzed live cells. Liver and bone, organs that can support metastatic growth of breast cancer cells, harbored about 100 and 20 cancer cells per 10,000 cells, respectively. In pancreas, spleen and kidney significantly less cancer cells were detected with 9, 0.4 and 0.7 cells per 10,000 mouse cells, respectively.

The presence of viable cancer cells in these organs was further confirmed by isolation and *in vitro* cultivation of cancer cells from kidney, pancreas and spleen (**Figure 9 D**). Therefore, MDA-MB-231 cells were transduced with a lentiviral vector carrying a puromycin resistance gene. Cancer cells were selected in cultures of organ suspensions by addition of puromycin (20 µg/ml) to the culture medium. GFP expression was detected under a microscope and confirmed that the selected cells are cancer cells (**Figure 9 D**).

Taken together it was demonstrated that viable MDA-MB-231 cells disseminated beyond their metastatic organ tropism to kidney, pancreas and spleen.

7.1.2 Widespread dissemination is not restricted to aggressive breast cancer cell lines

The until now used cell line MDA-MB-231 belongs to the triple-negative breast cancer subtype, which has poor prognosis and develops metastasis relatively early (1-5 years after removal the primary tumor). The hormone receptor positive luminal A and B subtypes are generally less aggressive, have better prognosis and form metastasis often after long relapse periods [244].

To determine whether less aggressive luminal subtypes are also able to disseminate to multiple organs, organ dissemination patterns of breast cancer cell lines as well as patient derived xenografts (PDX) belonging to different subtypes were analyzed (**Figure 10**). PDX models have only been passaged *in vivo* and are thought to maintain characteristics of tumors in patients [255]. Cancer cells in different organs were detected using qRT- PCR. As PDX models did not express any reporter construct, the presence of tumor cells in the organs was assessed using primers targeting human specific housekeeping genes. Cancer cells from cell lines carrying a TGL construct were detected using luciferase amplifying primers. The specificity of the primers was tested using organs of NSG mice.

Figure 10 A shows a representative qRT-PCR amplification plot, illustrating the normalized magnitude of fluorescence signal generated at each time point (dRN) with luciferase primers in lung, pancreas, spleen and kidney samples harboring MDA-MB-231 cells. The lower the cycle number at which fluorescence signal (dRN) is detected, the higher is the luciferase expression in the organs. As we know from the flow cytometry experiments that the TGL-reporter expression of tumor cells does not vary in different organs, fluorescence signal intensity can be correlated with the number of tumor cells present in the organs. Similar to the results obtained with flow cytometry and *ex vivo* luminescence imaging, significantly less tumor cells were detected in pancreas, spleen and kidney as compared to the lung with an amplification difference of about 10 cycles (**Figure 10 A**). The amplification plot further demonstrates that reliable signal could be

obtained also from organs with low tumor burden, making qRT-PCR a suitable method for detection of DCCs.

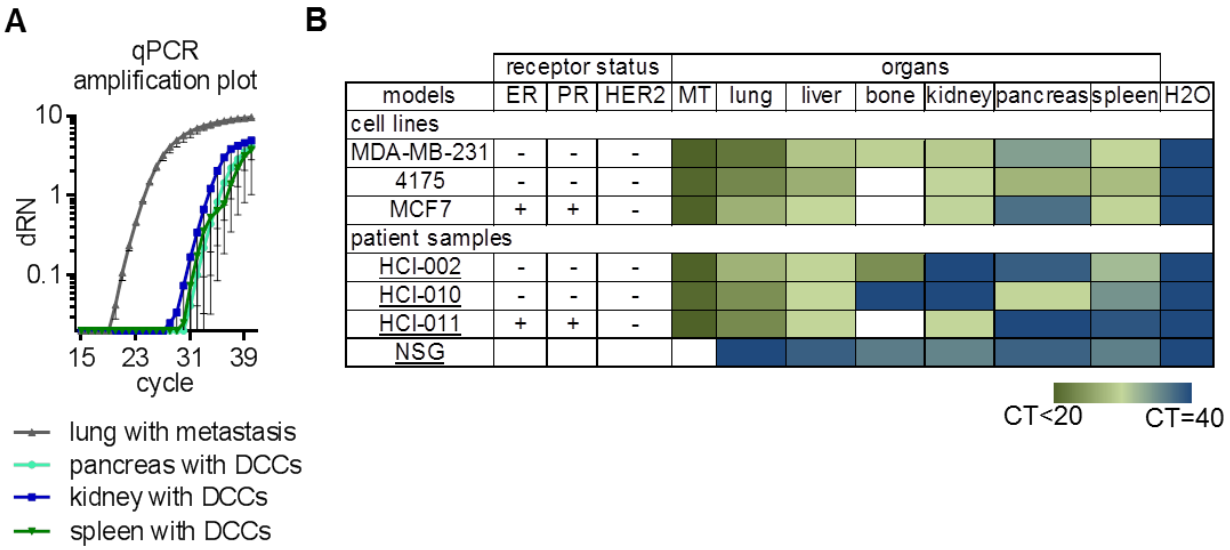


Figure 10 Validation of breast cancer dissemination to many organs in different cell line and PDX models

A qRT-PCR amplification curves of MDA-MB-231 TGL cancer cells in whole organ homogenates of lung, kidney, pancreas or spleen using specific primers against luciferase. The fluorescence signal generated at each time point (dRN) is plotted against the number of amplification cycles. Means of 6-9 replicates with standard deviation are shown.

B Heat map of amplification cycles (CT values) determined for breast cancer cell lines and PDX models of different breast cancer subtypes. The cells were injected orthotopically into NSG mice. Organs were harvested when the mammary tumors reached a diameter of 1 cm. Cancer cells were detected by qRT-PCR using luciferase specific primers for the cell lines and human TATA-box binding protein (hTBP) primers for the patient derived samples. The CT values of the organs of NSG mice without tumor cells amplified with hTBP primers and a H_2O control are shown as controls. The CT values were calculated as mean of 3-6 biological replicates. A CT-value of 40 was considered as not expressed. ER: Estrogen receptor; PR: progesterone receptor; MT: mammary tumor

The comparison of different cell lines and PDX models indicated that widespread dissemination to kidney, pancreas and spleen was not restricted to highly metastatic samples. Also less aggressive models such as the luminal cell line MCF7 that is poorly-metastatic [262] as well as luminal and basal PDXs that resemble more closely tumors in patients and did not form macroscopic metastasis in our hands, disseminated to several organs (**Figure 10 B**). The tumor burden in the organs differed from sample to sample, probably reflecting interpatient variations

.

7.1.3 The majority of DCCs has extravasated

The metastatic process can be divided into two main steps: (1) dissemination and (2) colonization at the secondary site (**Figure 3**) [2]. While thousands of cancer cells leave the primary tumor, survival and metastatic growth at the distant site are thought to be the rate limiting steps in many malignancies [37]. Therefore, it was investigated whether the tumor cells, detected in kidney, pancreas and spleen, were extravasated and located in the parenchyma of the organ.

A perfusion experiment was performed, in which the blood of tumor bearing mice was systemically replaced by PBS via PBS injections into the left ventricle of the heart prior to analysis of the organs [260]. Tumor cells located in blood vessels would be removed and only cells present in the parenchyma remain. Perfusion efficiency was confirmed by a color change in the organs caused by blood clearance. Representative images of brain, kidney and pancreas from perfused and non-perfused mice are depicted in **Figure 11 A**.

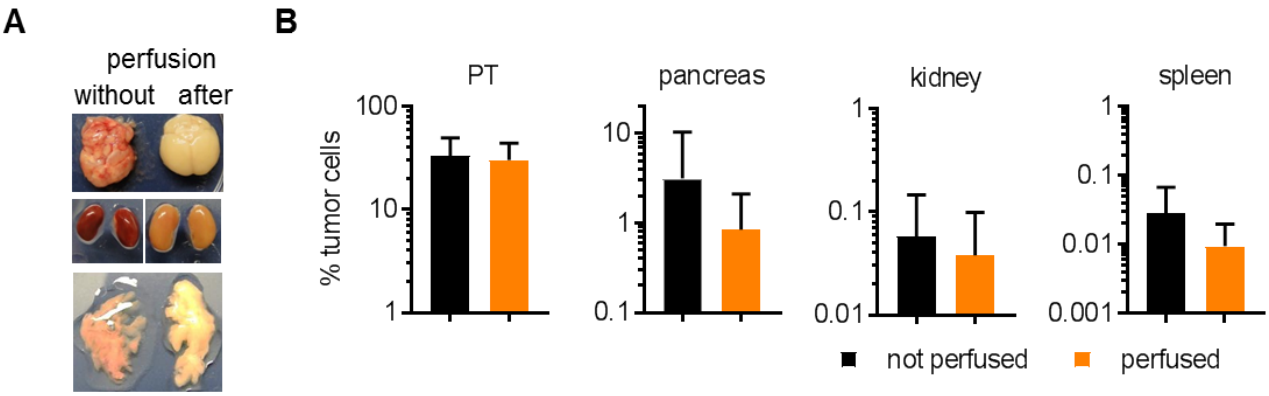


Figure 11 The majority of disseminated breast cancer cells in kidney, pancreas and spleen have extravasated from vessels

To assess if the cancer cells in kidney, pancreas and spleen were extravasated, a perfusion experiment was performed. 30 ml of PBS were injected prior to resection of the organs into the left ventricle of the heart using a constant flow rate of 1.1 ml/min [260].

A Representative images of brains (upper panel), kidneys (middle panel) and pancreas (lower panel) without and after perfusion are shown.

B The number of cancer cells in kidney, pancreas and spleen without (black) and after (orange) perfusion of mice was quantified by flow cytometry. cancer cell numbers are shown as percentage of the cell population staining negative for mCD11b, mCD45, H2KD and mCD31. Means with standard deviations are shown for six biological replicates performed in two independent experiments. The statistical significance of the difference in tumor cells in perfused vs. not perfused organs was tested with two-tailed unpaired t-test. Only p-values <0.05 are shown.

Perfusion resulted in a non-significant reduction of cancer cell numbers in pancreas and spleen and no change in kidney DCC numbers as detected by flow cytometry (**Figure 11 B**). This indicates that the majority of cancer cells in kidney, pancreas and spleen had extravasated from blood vessels.

To confirm these results and to gain more insights into the location of cancer cells within the organ, immunofluorescence (IF) staining of kidneys of mice injected orthotopically with MDA-MB-231 cells was performed five weeks after the injection. Tumor cells were stained with an antibody targeting GFP, blood vessels with an anti-CD31 antibody and nuclei with DAPI. We detected DCCs primarily as single cells and small clusters in the kidney (**Figure 12 A**). Cancer cells were analyzed according to their location relative to CD31 positive blood vessels and the extravasated cells were quantified. Extravasation of cancer cells was quantified in two biological replicates with 30-35 cancer cells per mouse. 88% of cancer cells were clearly located in the parenchyma of the kidney and only 12% may reside inside of CD31-positive vessels (**Figure 12 B**). These results confirm that the majority of cancer cells in the kidney had extravasated and were lodged in the tissue. Interestingly, about 80% of extravasated cancer cells remained in direct contact with blood vessels (**Figure 12 C**).

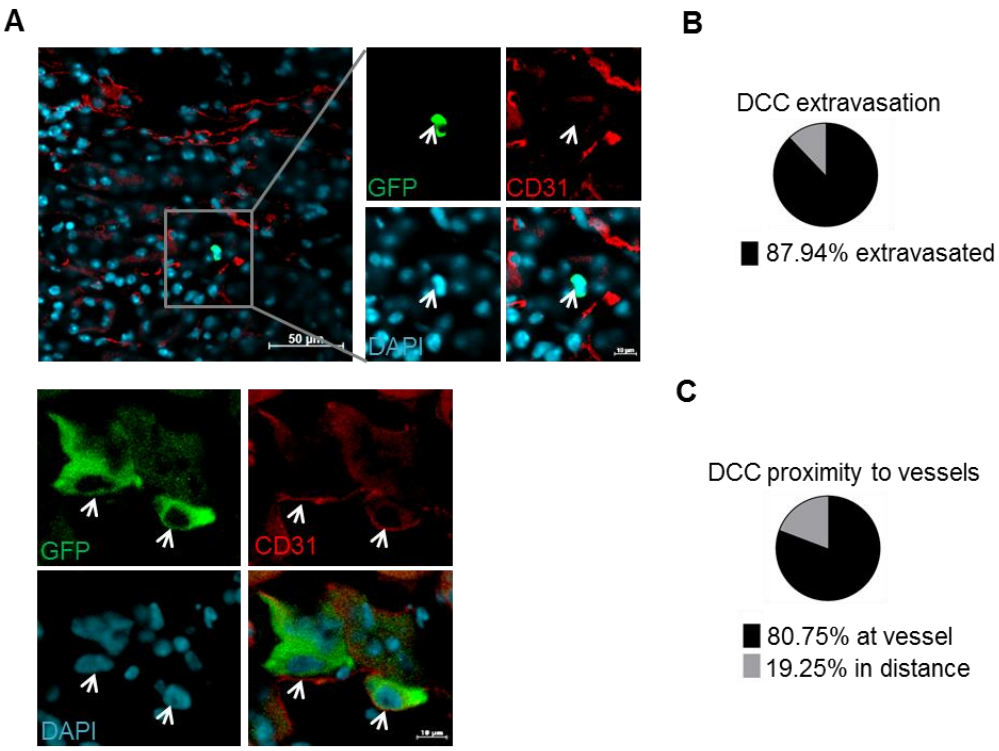


Figure 12 Disseminated breast cancer cells in the kidney are primarily located in close proximity to vessels

A Representative immunofluorescence (IF) images of MDA-MB-231 cancer cells in the kidney of NSG mice five weeks after orthotopic injection. Tumor cells were stained with an anti-GFP antibody (green), blood vessels with an anti-CD31 antibody (red) and nuclei with DAPI (blue). Pictures were taken with a Zeiss LSM710 ConfoCor 3 confocal microscope at a 40x magnification. The scale bare represents 10 µm with exception of the first picture (50 µm).

B Quantification of the percentage of cancer cells located in the parenchyma of the kidney using IF images. 30-35 tumor cells were analyzed in two biological replicates each.

C Quantification of the proximity of tumor cells to vessels in the parenchyma of the kidney. 25-30 extravasated cancer cells in two biological replicates were analyzed.

These data demonstrate that the majority of disseminated cancer cells in kidney, pancreas and spleen had extravasated from blood vessels in our breast cancer model. In the kidney, extravasated DCCs furthermore remained in close proximity to vessels.

7.2 Functional characterization of DCCs in kidney, pancreas and spleen

Following up on the identification of disseminated breast cancer cells in the parenchyma of kidney, pancreas and spleen, these cancer cells were functionally characterized. In this chapter it is demonstrated that disseminated breast cancer cells in kidney, pancreas and spleen that do not proliferate (**Figure 13**), survive an extended time period after mammary tumor resection (**Figure 14**) and maintain their growth potential (**Figure 15**).

7.2.1 Disseminated cancer cells in the kidney do not proliferate

As DCCs did not initiate macroscopic metastatic growth in kidney, pancreas and spleen, the question was asked whether these cells remained dormant. Therefore, IF staining of MDA-MB-231 cells in the kidney was performed using the mitotic marker phospho-Histone H3. Histone H3 gets phosphorylated at Serin10 and 28 during chromosome condensation in early prophase of mitosis. The antibody, we used, detects phosphorylated Ser10 of Histone H3, thus staining for cells in mitosis [263].

IF staining was performed in three biological replicates per organ with more than 1000 analyzed cells per mammary tumor, 500 cancer cells per lung and 27-34 cancer cells per kidney (90 cells total). Indeed, breast cancer cells in the kidney did not express phospho-Histone H3, whereas about 2% of cancer cells in the mammary tumor and 4% of metastatic cells in the lung expressed the mitosis marker (**Figure 13**).

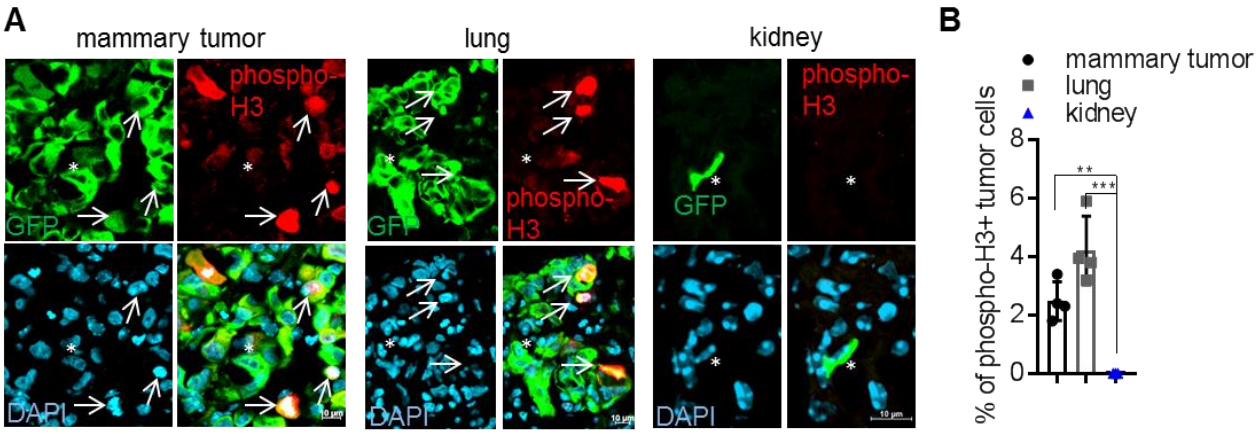


Figure 13 Disseminated cancer cells in the kidney are not proliferating

A Representative images of MDA-MB-231 cancer cells in the mammary tumor, the lung and the kidney of NSG mice five weeks after orthotopic cancer cell injection. Cancer cells were identified with GFP (green) and stained for proliferation using an anti-phospho Histone H3 antibody (red). This antibody detects Histone H3 phosphorylated at Ser10 and thus stains cells in mitosis [263]. Arrows mark cancer cells in mitosis (phospho-H3 positive) and * mark phospho-H3 negative cancer cells.

B The percentage of phospho-Histone H3 positive tumor cells was calculated as the median of three biological replicated. More than 1000 cells per replicate were analyzed in the mammary tumor and more than 500 in the lung. 90 cells were analyzed in the kidney. In the kidney no phospho-Histone H3 positive tumor cell was detected in all replicates. The median of the biological replicates is shown with interquartile range. p-values comparing kidney to lung or tumor, respectively were calculated using Dunnett`s multiple comparison test. ** p< 0.01; *** p< 0.001.

These results indicate that cancer cells in the parenchyma of the kidney are not proliferating

7.2.2 DCCs in kidney, pancreas and spleen survive for extended time

Patients may develop metastatic disease years or even decades after removal of the primary tumor, indicating that cancer cells that can reinitiate metastatic growth must survive in the body for that time [19]. We investigated whether the disseminated cells detected in pancreas, spleen and kidney may be long-lived and be able to contribute to disease progression.

It was analyzed whether DCCs in different organs were able to survive even after resection of the mammary tumor. Two PDX models, HCI009 and HCI010 [255], were used to determine the survival capability of these cells. These models did not develop macroscopic metastasis in the timespan of analysis ensuring that disseminated cells were not fueled from a secondary source. Mammary tumors were resected when the tumors reached a diameter of about 0.8 cm and mice were kept for further six weeks. The number of DCCs in the organs was analyzed by FACS (**Figure 14 A**).

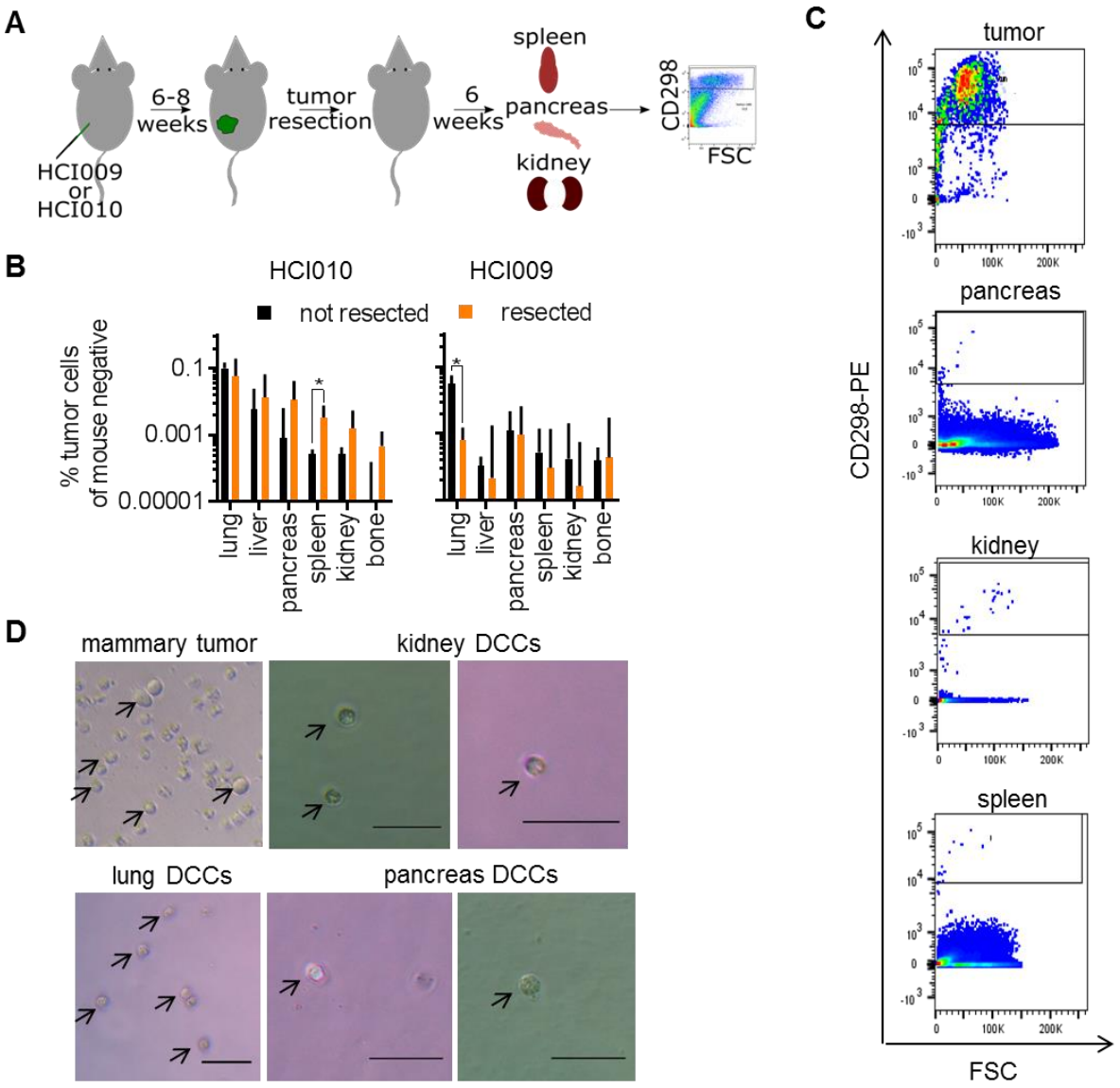


Figure 14 DCCs from PDX models survive without growth in kidney, pancreas and spleen after resection of the mammary tumor

A Experimental scheme: tumor pieces of HCI009 or HCI010 PDX models [255] were implanted into the 4th mammary fat pad of NSG mice. The tumor was resected when it reached a diameter of 0.8 cm. Six weeks later the number of cancer cells in the organs was accessed by FACS using a human specific CD298 antibody. At the time point of resection, some mice were sacrificed and the numbers of DCCs in the organs were measured as non-resected controls.

B Percentage of HCI010 (left) or HCI009 (right) cancer cells in different organs at tumor resection (black) or 6 weeks after resection (orange), respectively. Percentage of CD298 positives cancer cells in the cell population staining negative for CD11b, H2kD, CD45 and CD31 are shown as means of 3 to 6 replicates. Statistical testing was performed comparing cancer cell numbers at resection and six weeks later for each organ individually using t-test. Only significant comparisons are marked. * $p < 0.05$

C Representative FACS plots of the mCd45, mCd31, mCd11b and mH2kD negative fraction in pancreas, kidney and spleen of mice injected with HCI010 cells, analyzed 6 weeks after resection and a mammary tumor at the time point of resection. CD298 expression is plotted against the forward scatter (FSC)

D Representative images of HCI010 cancer cells from mammary tumor, lung, kidneys and pancreas that were isolated by FACS 6 weeks after the mammary tumor was resected. The cells were cultured in DMEM medium with 10% FCS and images were taken 48 h after isolation of the cells with 10x magnification. Arrows highlight the cells.

For the detection of cancer cells different murine stromal cell types were excluded as shown in **Figure 9** and cancer cells were selected by staining for CD298. Using flow cytometry, the number of cancer cells in lung, liver, kidney, pancreas, spleen and bone, detected at the time point of resection was compared to the number measured six weeks after resection. The percentage of HCI009 cancer cells in the cell population staining negative for murine stromal markers did not significantly change within 6 weeks after removal of the mammary tumor in most of the organs. Only the number of HCI009 cells in the lung was significantly reduced 6 weeks after resection (**Figure 14 B** right panel). In the second PDX model (HCI010), a significant increase of the number of cancer cells was detected in the spleen. However, even after 6 weeks the percentage of cancer cells in the spleen was only about 0.002%, which corresponds to a mean of 25 cancer cells in the whole organ. The number of tumor cells in the other organs was similar to the time point of resection (**Figure 14 B** left panel). Representative flow cytometry plots, depicting CD298 staining against the forward scatter are shown in **Figure 14 C**. These images demonstrate that cancer cells detected in kidney, pancreas and spleen had a similar size as cells from the mammary tumor suggesting that the cells in the organs were intact. The integrity of cancer cells in pancreas, spleen and kidney was further confirmed by sorting out the cells and culturing them. Microscope pictures were taken 48h after seeding. The cells did not adhere on the cell culture plates but survived for several days. The shape of the cancer cells in kidney, pancreas and spleen was comparable to sorted cells from the mammary tumor as well as lung (**Figure 14 D**).

The data presented in **Figure 14** suggest that disseminated cancer cells in kidney, pancreas and spleen survived prolonged time periods after resection of the mammary tumor.

7.2.3 DCCs in kidney, pancreas and spleen maintain their growth potential

In the next step of the study it was investigated whether non-proliferating disseminated cancer cells may be able to initiate metastasis in another location. Therefore, MDA-MB-231 cells carrying the TGL reporter construct were injected orthotopically into the mammary fat pad of NSG mice. Five weeks after injection, cancer cells from kidney, pancreas, spleen and the mammary tumor were isolated using FACS. The highest possible number of cancer cells was sorted from each organ. At the same time, the exact same number of cancer cells from the mammary tumor was isolated and used as control. The cancer cells were immediately reinjected intravenously (IV) into the lateral tail vein of individual NSG mice. Intravenous injection of tumor cells is often used to model lung colonization [16, 52]. Metastatic growth in the lung was determined by bioluminescence detection. Mice were sacrificed and luminescence of the organs was measured *ex vivo* when the lung signal reached about 10^8 p/sec (**Figure 15 A**).

Cancer cells isolated from spleen, pancreas and kidney were able to initiate growth in the lung after tail vein injection at a comparable rate to cancer cells from the mammary tumor. The *ex vivo* luminescence in lungs injected with DCCs from the kidney (3 out of 3) or pancreas (2 out of 3) were similar to lungs injected with cells from the mammary tumor. The luminescence signal was between 100- and 1000-fold higher than the background signal (**Figure 15 B**). The growth potential in the lung of DCCs isolated from kidney, pancreas or spleen was further analyzed.

Growth in the lung was assessed positive when the *ex vivo* luminescence signal was above the background signal in healthy NSG lungs that have been injected with luciferin. In total six biological replicates from two independent experiments were analyzed. Metastatic growth was detected in all six lungs injected with kidney DCCs. Pancreas DCCs grew in four of the six biological replicates and spleen DCCs in five (**Figure 15 C**). Representative *ex vivo* luminescence pictures of lungs injected with cancer cells from kidney, pancreas and spleen are shown in **Figure 15 D**. Development of metastatic nodules was confirmed by Hematoxylin and Eosin staining of lungs. A representative picture of a lung injected with DCCs isolated from the spleen is shown in **Figure 15 E**.

Bioluminescence was additionally measured in spleen, pancreas and kidney, the organs where the disseminated cancer cells were previously isolated, to determine dissemination and possible growth of aggressive clones. The luminescence signal in spleen and pancreas, detected after IV injection of cancer cells isolated from those organs, was similar to the signal detected after IV injection of cells from the mammary tumor. Luminescence signal in the kidney after injection with disseminated cancer cells from the kidney was non-significantly higher than the signal resulting from injection of the same number of mammary tumor cells (**Figure 15 E**). This indicates that we do not select for a more aggressive subpopulation of cancer cells.

Taken together these results indicate that disseminated cancer cells from kidney, pancreas and spleen maintain their growth potential and can initiate metastases in the lung, an environment that supports metastatic growth of MDA-MB-231 cells.

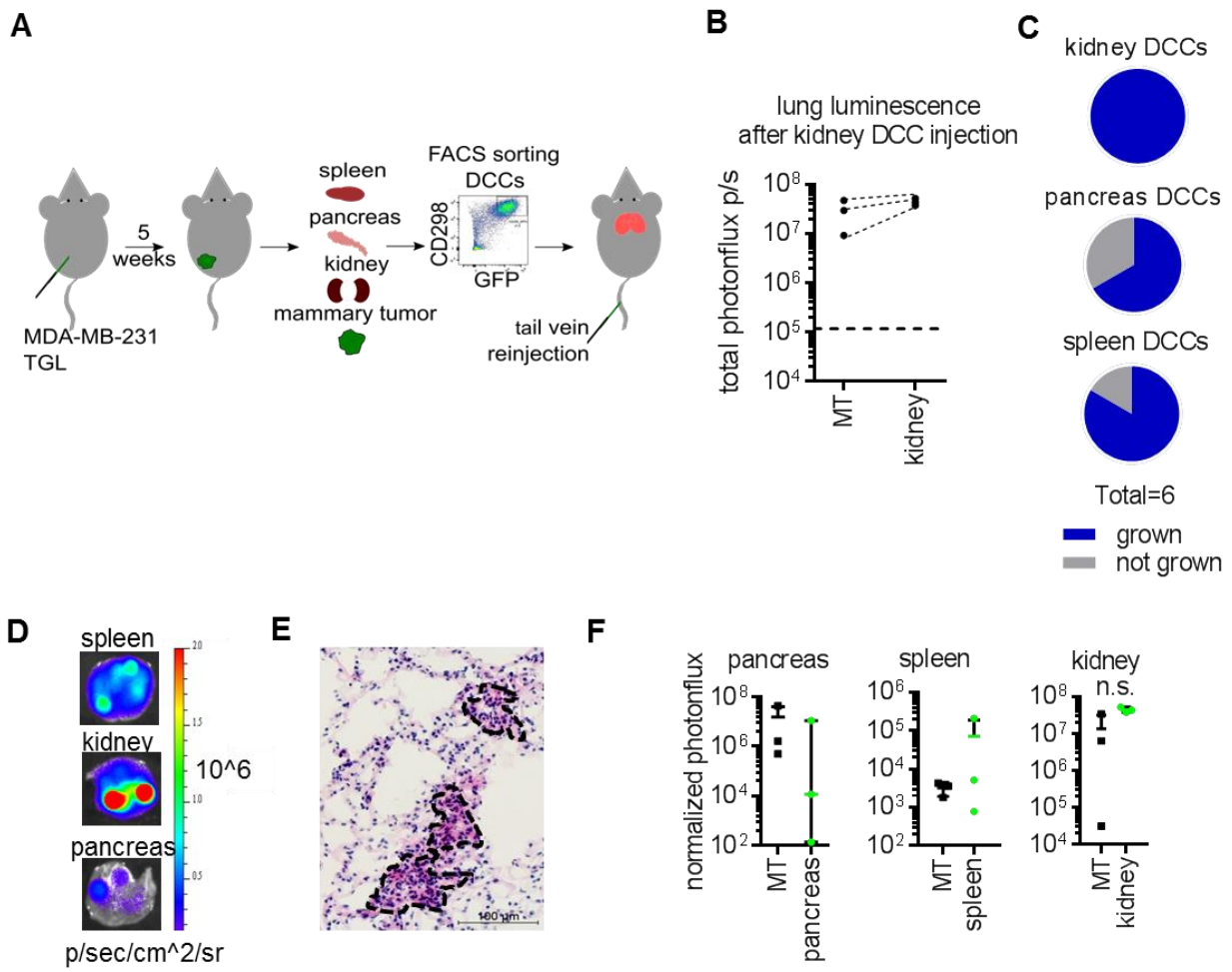


Figure 15 Cancer cells in kidney, pancreas and spleen maintain their growth potential

A Experimental scheme: MDA-MB-231 TGL cells were injected orthotopically into the 4th mammary gland of NSG mice. Five weeks after injection cancer cells from spleen, kidney, pancreas and the mammary tumor were isolated via FACS. The sorted cells were directly reinjected into the tail vein of healthy NSG mice. From pancreas, spleen and kidney as many cells as possible were reinjected. The number of mammary tumor cells was matched to the number of cells sorted from the organs, respectively. Metastatic growth in the lung was followed by luminescence imaging.

B *Ex vivo* luminescence signal of the lung injected with tumor cells from the mammary tumor, pancreas, spleen or kidney was measured. Representative luminescence signal of lungs injected with kidney DCCs or cancer cells from the mammary tumor (MT) are shown. The background signal of lungs without cancer cells is shown as horizontal line. Each data point represents one biological replicate.

C Lungs injected with DCCs from kidney, pancreas and spleen were analyzed for metastatic growth by *ex vivo* luminescence imaging. Metastasis was counted as growing if the luminescence signal was above the baseline signal in non-injected lungs as shown in **B.** six biological replicates performed in two independent experiments were analyzed for growth per organ.

D Representative luminescence images of lungs injected with cancer cells freshly isolated from kidney, spleen or pancreas, respectively

E Representative Hematoxylin and Eosin (H&E) staining of a lung injected with DCCs isolated from a spleen. Pictures were taken at 10x magnification. The bar represents 100 µm. Tumor regions are framed with black margins.

F *Ex vivo* luminescence of pancreas (left), spleen (middle) or kidney (right) of mice injected intravenously with DCCs isolated from the respective organ (green) or the same number of cells from the mammary tumor (MT) (black). n=3, Mean with standard deviations are shown. Statistical significance was tested using two-tailed unpaired t-test.

7.3 Gene expression analysis of DCCs in kidney and pancreas

Given the extended survival capability and the maintained growth potential of DCCs in kidney, pancreas and spleen, the subpopulation of cancer cells present in kidney and pancreas was further characterized by gene expression profiling compared to the mammary tumor to identify molecular differences that may be used therapeutically.

While the transcriptional profile separates DCCs from mammary tumor cells (**Figure 16**) and lung metastasis (**Figure 20**), similarities were observed between DCCs from kidney and pancreas (**Figure 17**). The identified characteristics of DCCs from kidney and pancreas will be discussed in more detail in this chapter and are listed below:

- Downregulation of cell death mechanisms (**Figure 21**)
- Repression of EMT regulators and mesenchymal genes(**Figure 22**)
- Metabolic switch from glycolysis to oxidative phosphorylation (**Figure 24** and **Figure 25**)
- Repression of MHC II mediated antigen presentation (**Figure 28**)
- Upregulation of genes involved in DNA repair and cell cycle checkpoints (**Figure 30**)
- Intrinsic chemotherapy resistance (**Figure 31**)

7.3.1 DCCs in kidney and pancreas repress most of the regulated genes

Cancer cells from kidney, pancreas and the mammary tumor were isolated from organ homogenates by FACS as described above (**Figure 9**). RNA was isolated from sorted cells and expression of 47,000 transcripts was analyzed on Affymetrix Human genome U133.2plus arrays in triplicates (**Figure 16 A**). The purity of the sorted samples was confirmed by qRT-PCR using cancer cell (luc+) and mouse specific (mB2m) primers. High expression of a gene is reflected by low CT values, whereas low mRNA levels are reflected by high CT values. Cultured MDA-MB-231 cells were used as positive control and healthy NSG organs as negative background control. MDA-MB-231 cells as wells as the sorted samples showed heterogeneous mB2m expression. However, the mB2m CT values measured for all samples were above the CT values of NSG organs and the CT values of the sorted samples were in range of the CT values of cultured cells. In addition, the luc+ expression of the sorted samples was homogeneously on the level of cultured cells whereas NSG organs did not show luc+ expression reflected by high CT values, confirming the purity of the samples (**Figure 16 B**).

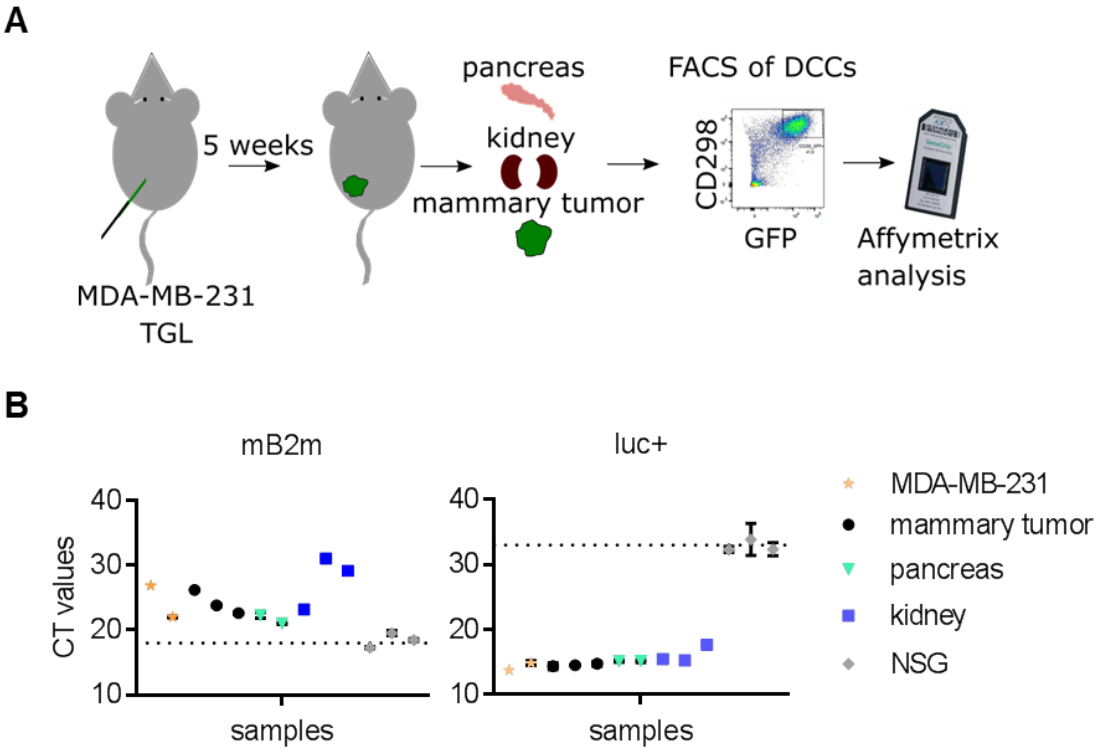


Figure 16 Gene expression profiling of DCCs from kidney and pancreas compared to the mammary tumor

A Experimental set-up: MDA-MB-231 TGL cells were injected orthotopically into NSG mice. Five weeks after injection pancreas, kidney and the mammary tumor were harvested, digested and cancer cells were FACS isolated using the antibody panel described in **Figure 9.** The cancer cells were analyzed by gene expression profiling on Affymetrix Human genome U133.2plus arrays. (MT, n=4; pancreas, kidney, n=3 each)

B The sorted cancer cell samples used for gene expression profiling were checked for purity by qRT-PCR using mouse (mB2M) or cancer cell (luc+) specific primers. Cultured MDA-MB-231 cells were used as positive and non-injected NSG organs as negative controls. Each data point represents one biological replicate that was analyzed in three technical replicates. For one mammary tumor and one pancreas insufficient material was left to perform the analysis. Amplification cycles (CT) are plotted for each biological replicate.

The array raw data was RMA normalized using Chipster version 3.11.X and a SD filter of 50% was applied. Principal component analysis (PCA) performed on the RMA normalized data of samples from kidney, pancreas and the mammary tumor showed that disseminated cancer cells from kidney and pancreas cluster separately from cells of the mammary tumor (**Figure 17 A**). The difference between cancer cells from the mammary tumor and DCCs accounted with 30% for the major differences observed in the array. This was confirmed by unsupervised hierarchical clustering using the spearman correlation method. Cancer cells from kidney and pancreas, in contrast were not clearly separate from each other by their gene expression (**Figure 17 B**). This suggests that disseminated cancer cells, despite been located in different organs, have significantly overlapping properties

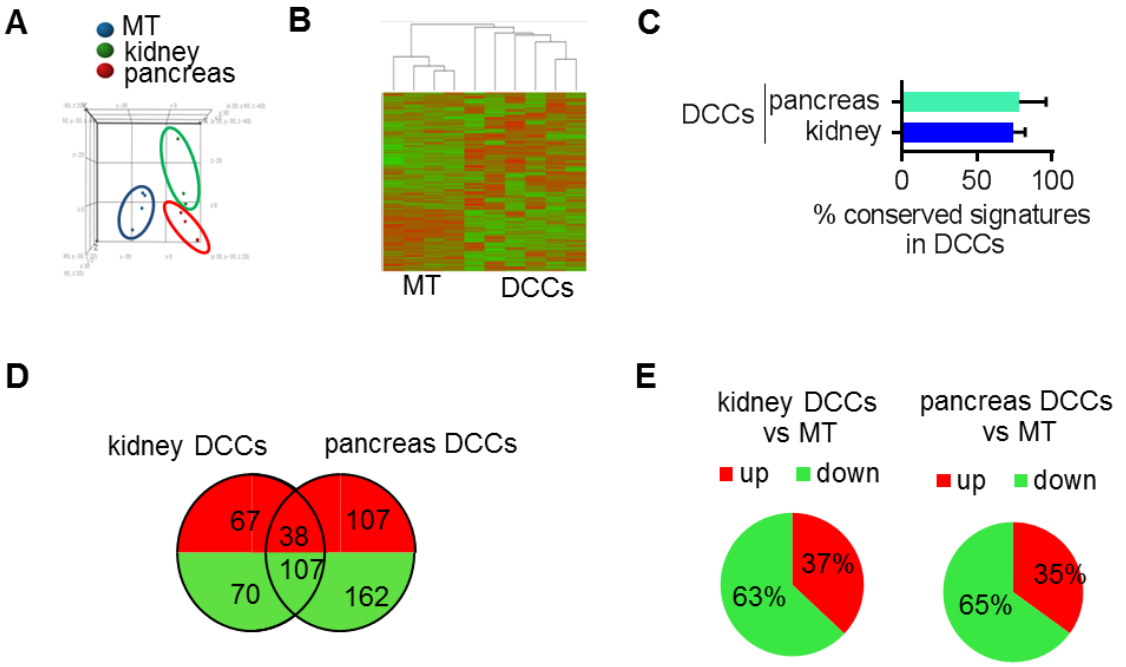


Figure 17 The gene expression of DCCs in kidney and pancreas overlaps significantly

A Principal component analysis (PCA) of the gene expression data generated from cancer cells from kidney, pancreas and mammary tumor (MT) after RMA normalization with variant stabilization. (PC1: 29%, PC2: 15%, PC3:13%)

B Unsupervised hierarchical clustering of the normalized data of mammary tumor cells and DCCs from kidney and pancreas was performed using Spearman Correlation method.

C Gene set enrichment analysis (GSEA) was performed comparing kidney DCCs or pancreas DCCs, respectively with the mammary tumor. Gene signatures of the HALLMARK, the GO and the KEGG databases were used. The significantly ($p<0.05$, FDR<0.25) enriched and repressed gene signatures in kidney DCCs relative to the mammary tumor were compared to the terms significantly enriched and repressed to in pancreas DCCs and vice versa. The percentages of conserved gene signatures in pancreas and kidney DCCs are shown as the mean of the three used databases.

D Empirical Bayes test was performed comparing cancer cells from kidney or pancreas to cells from the mammary tumor. Benjamini-Hochberg (BH) multiple testing correction was applied and a corrected p-value of 0.05 or lower was considered significant. The number of significantly upregulated (red) and downregulated (green) genes are shown for both comparisons. The numbers of genes that were found in kidney (left) as well as pancreas DCCs (right) compared to the mammary tumor and those genes only found in one comparison are indicated.

E Quantification of the number of significantly regulated genes ($p<0.05$) in kidney (left) or pancreas DCCs (right) compared to cells from the mammary tumor using empirical Bayes test. The percentages of up- and downregulated genes are shown, respectively.

The similarities of DCCs from kidney and pancreas were further explored by Gene Set Enrichment Analysis (GSEA). GSEA was performed using the HALLMARK, KEGG and GO signature collections available from the Broad Institute. The significantly ($p<0.05$ or FDR<0.25) enriched and repressed signatures in kidney or pancreas DCCs, respectively were compared.

Using the HALLMARK collection, 96% of the gene signatures enriched in kidney DCCs were also significantly enriched in cells from pancreas and 74% of the enriched signatures in pancreas DCCS were also found in kidney DCCs. Of the GO signature collection, 78% of the enriched terms in kidney DCCs were conserved in pancreas DCCs and 82% of enriched signatures in pancreas tumor cells in the kidney. Using the KEGG collection, 58% of enriched signatures in kidney DCCs were conserved in tumor cells from the pancreas and 59% vice versa. The results are summarized in **Figure 17 C** by presentation of the mean of the three datasets.

To further explore the datasets, kidney and pancreas DCCs signatures were generated by performing empirical Bayes tests comparing kidney DCCs with mammary tumor cells or pancreas DCCs with the mammary tumor. Multiple testing correction was performed using the Benjamini-Hochberg (BH) method and the p-value cutoff was set to 0.05. Similar to the results obtained with GSEA, kidney and pancreas DCC signatures showed a significant overlap especially in the downregulated genes. Compared to the mammary tumor, the expression of 282 genes was significantly changed in kidney DCCs. 177 genes were downregulated in kidney DCCs and 106 up. In pancreas DCCs, the expression of 414 genes was significantly changed compared to the mammary tumor. 269 of these genes were downregulated in pancreas DCCs and 145 up. Of those genes, 145 overlapped in kidney and pancreas DCCs – 38 were up and 107 downregulated (**Figure 17 D**). Summing up, 51% of the kidney DCC signature was conserved in pancreas DCCs and 35% of the pancreas DCC signature in disseminated cancer cells in the kidney. More than 63 and 65% of the significantly (BH-p< 0.05) regulated genes were downregulated in disseminated tumor cells from kidney and pancreas, respectively (**Figure 17 D and E**).

Taken together, a general analysis of the gene expression profile of disseminated cancer cells from kidney and pancreas compared to cancer cells from the mammary tumor indicated that disseminated cells were distinct from the bulk mammary tumor. Moreover, disseminated cells from different organs shared about 50% of their gene expression profile when comparing regulated genes and even 75 % when analyzing enriched gene signatures. Furthermore, the majority of differentially expressed genes compared to the mammary tumor were downregulated in pancreas and kidney DCCs.

7.3.2 Numerous transcriptomic changes in DCCs are regulated intrinsically

The observed similarities of disseminated cancer cells in kidney and pancreas raise the question whether these changes are caused by common cues in the microenvironment of both organs or are regulated cancer cell intrinsically. To investigate the impact of cell intrinsic versus microenvironmental gene expression regulation on the transcriptome of DCCs, a second gene expression analysis was performed. Cancer cells from kidney, pancreas, lung and the mammary tumor were isolated by FACS as described before. Sorted cells were cultured for 48h *in vitro* to remove the immediate impact of the microenvironment. Thereafter, RNA was isolated and gene expression profiling was performed on Affymetrix Human genome U133.2plus arrays (**Figure 18 A**).

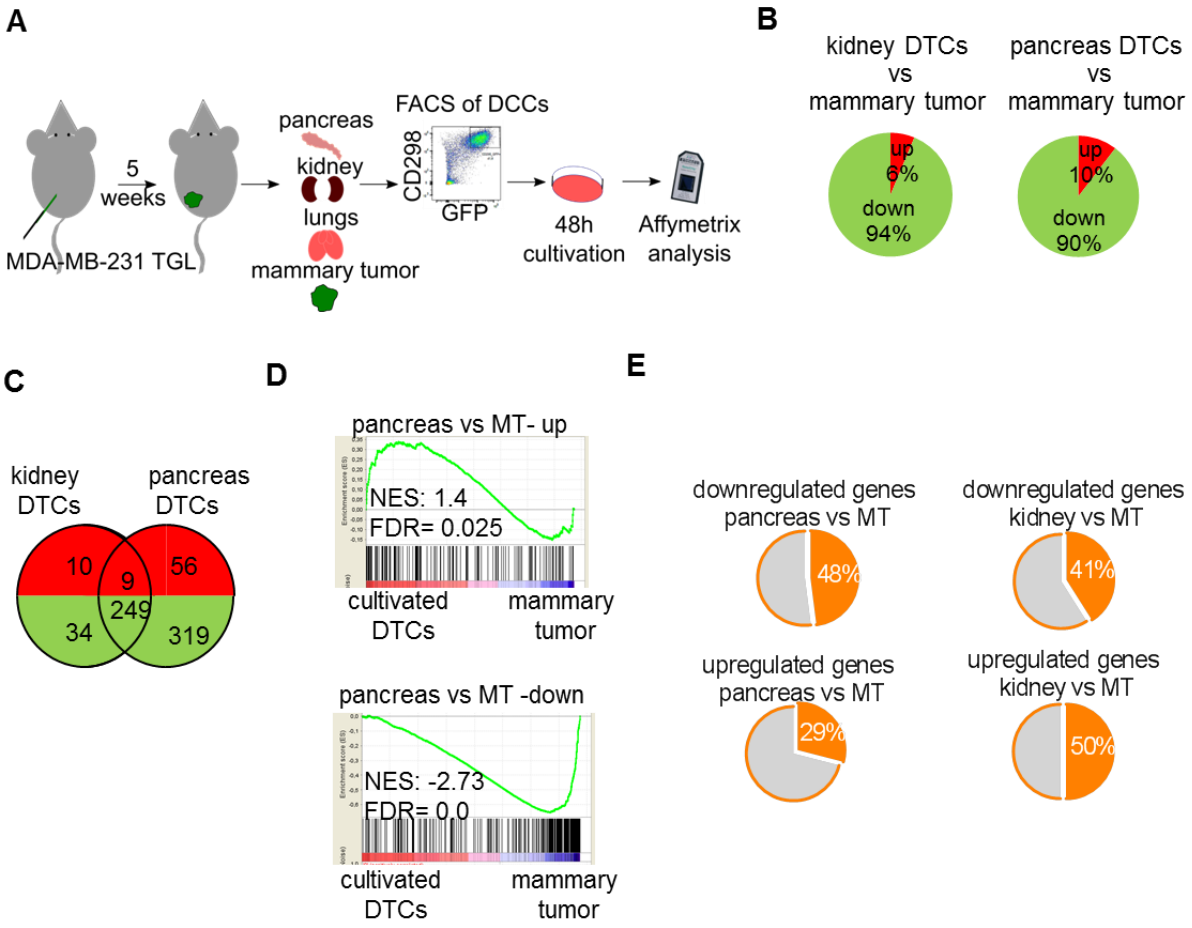


Figure 18 A significant percentage of the gene signatures of DCCs from kidney and pancreas is stable in short term culture

A Experimental set-up: To investigate whether the gene signatures of DCCs from kidney and pancreas compared to the mammary tumor are intrinsically regulated independent of the microenvironment, MDA-MB-231 TGL cells were injected orthotopically into NSG mice. Five weeks after injection, pancreas, kidney, lung and mammary tumors were harvested and cancer cells were isolated via FACS. The isolated cells were cultivated for 48 h under low serum conditions and RNA was harvested for gene expression analysis on Affymetrix Human genome U133.2plus arrays. (n=2)

B Empirical Bayes test was performed after RMA normalization of the raw data comparing cultivated kidney or pancreas DCCs with the mammary tumor (BH-p<0.05). Of the genes with at least 1.5 fold changed expression compared to the mammary tumor, 94% and 90% were downregulated in kidney and pancreas DCCs, respectively.

C Analysis of the significantly enriched genes identified by empirical Bayes test that were shared between cultivated DCCs in kidney and pancreas or that were unique to pancreas or kidney DCCs, respectively.

D GSEA plots of cultivated DCCs compared to the mammary tumor using the signatures of immediately analyzed pancreas DCCs vs mammary tumor (MT), generated as shown in **Figure 17**. Cultivates DCCs from kidney and pancreas were treated as one group in this analysis. NES: normalized enrichment score. FDR: false discovery rate.

E Percentage of significantly up- or downregulated genes from the immediately analyzed DCC signatures generated in **Figure 17** that were also core enriched in DCCs after cultivation. The core enriched genes were obtained from the GSEA analysis shown in **D**.

The raw data was RMA normalized and SD filtered. Empirical Bayes test was performed comparing cultured kidney and pancreas DCCs to the mammary tumor, respectively. The Benjamini-Hochberg corrected p-value threshold was set to 0.05 and the fold change cutoff to 1.5.

As the aim of this experiment was the comparison between immediately analyzed and cultured DCCs, the gene expression profile of cultured DCCs was also analyzed for the percentage of downregulated genes and the similarities of DCCs from kidney and pancreas. In cultured DCCs the tendency towards downregulation of gene expression was even more pronounced than in immediately analyzed DCCs (**Figure 17**). Compared to the mammary tumor 94% and 90% of genes with changed expression in kidney DCCs and pancreas DCCs, respectively are repressed (**Figure 18 B**). Comparing the gene signatures of cultivated pancreas DCCs with those of cultured kidney DCCs, a significant number of genes was shared between the disseminated cells from both organs. 88% of downregulated genes in kidney DCCs were also downregulated in pancreas DCCs and 44% of downregulated genes in pancreas DCCs were also detected in kidney DCCs. 47% of upregulated genes in kidney DCCs and 13.8% in pancreas DCCs were common to both DCC pools (**Figure 18 C**).

To further compare immediately analyzed versus cultivated DCCs, GSEA was performed on the normalized data of the cultivated DCCs using the immediately analyzed DCC signatures shown in **Figure 17**. The signatures with upregulated or downregulated genes in pancreas and kidney DCCs were significantly enriched in the cultured disseminated cells (**Figure 18 D**). The genes upregulated in pancreas DCCs were significantly enriched in the cultivated DCCs with a normalized enrichment score (NES) of 1.4 and a false discovery rate (FDR) of 0.025. The genes with downregulated expression in pancreas DCCs were also underrepresented after cultivation with a NES of -2.73 and FDR of 0.0 (**Figure 18 D**). The effect of cultivation of DCCs was further investigated by analyzing the overlap between the immediately analyzed DCC signature and the genes that are still core enriched in cultured DCCs. In DCCs from the pancreas, 48% of downregulated and 29% of upregulated genes were still core-enriched in the DCCs after 48h cultivation. In kidney DCCs, the overlap between the immediate signatures and the core-enriched genes after cultivation was 41% of the genes with repressed and 50% with increased expression (**Figure 18 E**).

The overlap between the signatures generated with freshly analyzed or cultivated DCCs, led us to further explore the impact of intrinsic regulation on gene expression by functional analysis. Genes that were significantly repressed in kidney or pancreas DCCs compared to the mammary tumor, as identified by empirical Bayes test, were subjected to Gene ontology (GO) analysis. The analysis was performed on the signatures generated from DCCs that were analyzed immediately after isolation as well as from cultured DCCs independently (**Figure 19**). To be able to compare the results, GO terms belonging to the same molecular functions were grouped. All GO terms with a p-value of at least 0.01 and five detected genes in the gene signatures were taken into account. Molecular functions, for which at least three GO-terms were detected in one of the gene sets, are plotted in **Figure 19**.

As already observed for the significantly regulated genes, the overlap between the cellular functions regulated in immediately analyzed kidney as well as pancreas DCCs was significant. All reoccurring categories identified in pancreas DCCs were also detected in kidney DCCs with similar numbers of GO-terms. Especially strong was the regulation of metabolic GO-terms. In pancreas DCCs, 20 GO-terms associated with glucose metabolism, 16 with nucleotide

metabolism and 2 with energy metabolism were enriched in the genes with repressed expression. In kidney DCCs, 37 glucose metabolism, 28 nucleotide and 6 energy metabolism terms were enriched. In addition, cell death related GO-terms, GO-terms associated with oxygen and hypoxia response, stress and stimulus response and reactive oxygen species (ROS) production and scavenging were significantly enriched in DCCs from kidney and pancreas (**Figure 19 left panels**). The GO-terms enriched in DCCs after cultivation were very similar. Only the number of GO-terms associated with metabolism was reduced after 48h *in vitro* culture to 8 glucose and nucleotide metabolism related terms each and two energy metabolic terms in pancreas and 10 nucleotide and 2 energy metabolic terms in kidney DCCs (**Figure 19 right panels**). Furthermore, in kidney and pancreas DCCs GO-terms associated with cell motility were enriched in the repressed genes after *in vitro* cultivation. Cell motility terms have not been detected in freshly analyzed samples and may therefore be induced by the culture conditions.

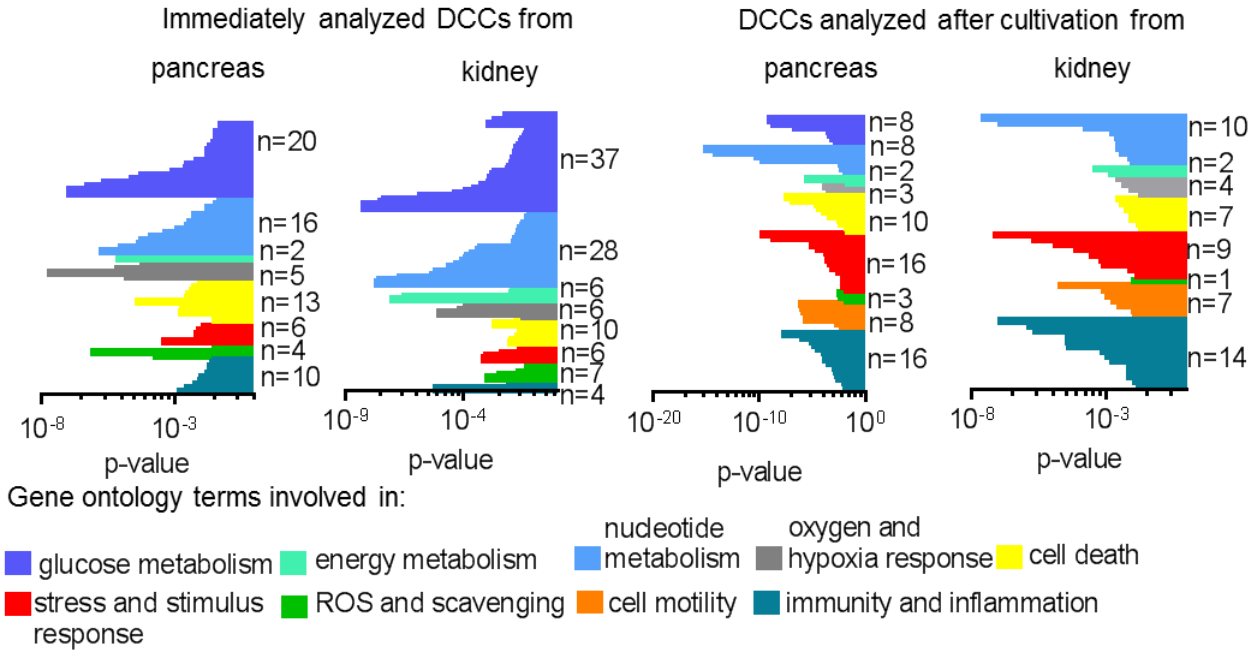


Figure 19 Many of the cellular functions changed in disseminated MDA-MB-231 cells in kidney and pancreas are regulated intrinsically

Gene Ontology (GO) analysis was performed using the genes that were significantly downregulated with a fold change of at least 1.5 fold and a BH-corrected p-value of 0.05 comparing kidney or pancreas DCCs with cells from the mammary tumor, respectively. The differentially expressed genes were calculated using empirical Bayes test. Gene ontology analysis was performed for kidney and pancreas DCCs analyzed immediately after isolation or after 48h *in vitro* cultivation, separately. Enriched GO terms were manually grouped intro reappearing categories. All GO terms with a p-value of at least 0.01 and 5 observed genes in the analyzed gene signatures were taken into account. The following functional terms were assigned (top to bottom): Glucose metabolism, nucleotide metabolism, energy metabolism, hypoxia and oxygen response, cell death, stress and stimulus response, ROS and scavenging, cell motility and immunity and inflammation. Each row represents one GO-term, which is plotted against its logarithmic p-value. The numbers (n) on the right site of the graphs indicate the number of GO-term describing a particular cellular function.

Taken together, the comparison of immediately analyzed disseminated MDA-MB-231 cells in kidney and pancreas with DCCs that were cultured for 48h suggests that a significant percentage of genes with changed expression in pancreas and kidney DCCs is regulated intrinsically and not by the direct influence of the microenvironment.

7.3.3 DCCs in kidney and pancreas are different from metastatic cells in the lung

To better characterize disseminated cancer cells in pancreas and kidney, I compared them to metastasis forming cells from lung tissue. Therefore, MDA-MB-231 cells isolated from lung were added to the analysis performed earlier (**Figure 18)**. The presence of proliferative metastasis in the lung in our set-up was already confirmed by phospho-Histone H3 staining. Quantification and a representative image of a lung metastasis are shown in **Figure 13**. Cancer cells isolated from lung were cultivated for 48h before analysis to compare cell intrinsic features of metastatic and disseminated cells. The experimental setup is depicted in **Figure 18 A**.

Principal component analysis (PCA) on RMA normalized data was performed to compare the similarities of the study groups. Disseminated cells from kidney and pancreas that were cultured *in vitro* for 48h clustered closely together and were clearly separated from cells from the mammary tumor (**Figure 20 A**). This resembles the results obtained with the samples that were analyzed immediately after isolation as shown in **Figure 17 A**. Cultured cancer cells isolated from lungs also clustered apart from disseminated cells from kidney and pancreas and showed even more pronounced differences to DCCs (**Figure 20 A**).

Genes with significantly different expression comparing kidney and pancreas DCCs to metastatic cells from the lung were calculated using empirical Bayes test with a BH corrected p-value of 0.05. For this analysis kidney and pancreas DCCs were summarized in one group and jointly compared to metastatic cells isolated from lung. The expression of 618 genes was significantly changed in kidney and pancreas DCCs by at least 1.5-fold. Similar to the comparison of disseminated cells with cells from the mammary tumor (**Figure 17** and **Figure 18**), the majority of significantly regulated genes was downregulated comparing kidney and pancreas DCCs to cancer cells isolated from lungs. The expression of 85% of the affected genes was repressed in DCCs (**Figure 20 B**). To further compare metastatic to disseminated cancer cells, GO analysis was performed using the genes with repressed expression in kidney and pancreas DCCs compared to lung metastasis. GO-terms belonging to the same cellular function were grouped together manually. GO-terms with a p-value of 0.05 and at least 5 observed genes in the signature were taken into account. Functional groups with at least 3 significantly enriched GO-terms are depicted in **Figure 20 C**.

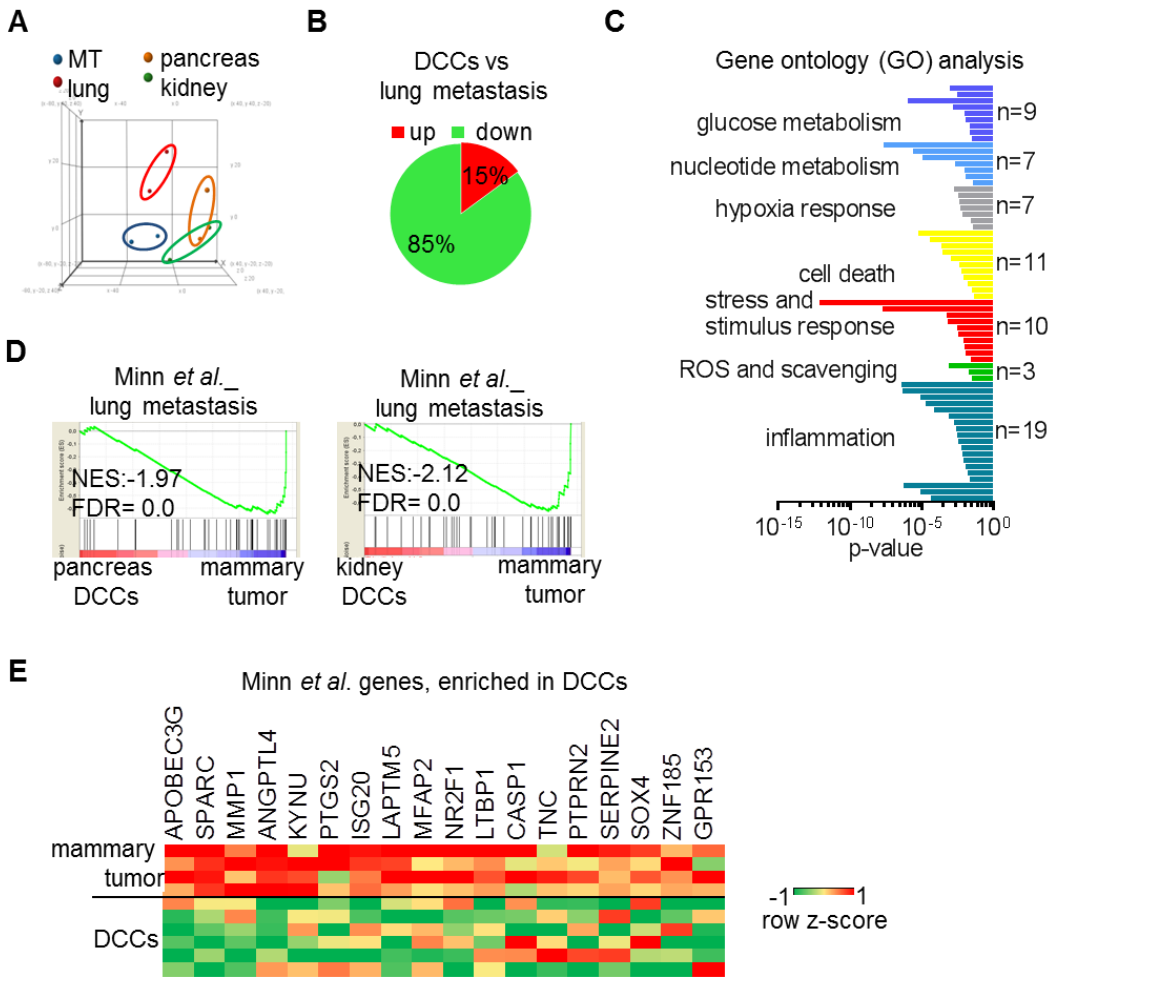


Figure 20 DCCs in kidney and pancreas are distinct from metastatic cells in the lung

A The experimental set-up described in **Figure 18** was used to compare the cell intrinsic gene expression profile of DCCs from kidney and pancreas to metastatic cells from the lung. Principal component analysis of the RMA normalized data is shown. Two biological replicates per study group were analyzed. Blue: cells from the mammary tumor (MT), red: cancer cells from lung, green: cancer cells isolated from kidney, orange: MDA-MB-231 cells isolated from pancreas. PC1 40,7%, PC2: 15.9%, PC3: 13%

B Significantly regulated genes (BH-p<0.05; FC>1.5) between DCCs and lung metastasis were calculated using empirical Bayes test. Samples isolated from kidney and pancreas were group together for this analysis and were compared jointly against samples isolated from lungs. 85% of the 618 genes with changed expression in disseminated cells were repressed.

C Gene ontology (GO) analysis was performed with the genes downregulated in DCCs vs lung metastasis (BH-p<0.05; FC>1.5). Reappearing terms were grouped into categories. The numbers (n) indicate the numbers of GO-terms that were enriched in the repressed gene signature with a p-value of at least 0.05 and five observed genes. Each row represents one GO-term, which was plotted against the logarithmic p-value.

D GSEA was performed comparing immediately analyzed pancreas and kidney DCCs, respectively to cells from the mammary tumor (experimental set-up **Figure 17**) using the gene signature generated by Minn *et al.*. The Minn *et al.* signatures covers 48 genes overexpressed in a lung metastatic derivative of MDA-MB-231 cells compared to the parental cells [16]. NES: normalized enrichment score; FDR: false discovery rate

E Heat map of genes of the Minn *et al.* lung metastasis gene signature that were core enriched in DCCs in the analysis performed in **D**. Row z-scores are shown.

The functional groups regulated in disseminated cells compared to cells from lung metastasis were very similar to those obtained in the comparison of DCCs with mammary tumor cells (**Figure 19**). Also in comparison to metastatic cells from the lung, GO terms enriched in the repressed genes of DCCs in kidney and pancreas were involved in glucose and nucleotide metabolism, hypoxia response, cell death, stress and stimulus response and ROS and scavenging. GO-terms involved in inflammation were also identified (**Figure 20 C**).

To show that freshly analyzed DCCs, still carrying some signaling cues of the microenvironment, are distinct from metastatic cells in the lung, the metastatic gene signature published by the laboratory of Joan Massagué [16] was used. The investigators established a lung metastatic derivative of the MDA-MB-231 breast cancer cells by sequential intravenous injection of cancer cells isolated from the lung [16]. They generated a gene set associated with lung metastasis and comprising 48 overexpressed genes by comparing the second generation of MDA-MB-231 cells from the lung to cells from the mammary tumor [16]. GSEA was performed on the normalized data of kidney DCCs, pancreas DCCs and mammary tumor cells that were analyzed directly after isolation using the Minn *et al.* 48-gene set. The Minn *et al.* signature was upregulated in the mammary tumor compared to DCCs from kidney or pancreas, respectively (**Figure 20 D**). These results confirm the results obtained with the cultured samples. In the PCA as well as the GSEA, metastatic cells from the lung cluster closer to the mammary tumor than to DCCs from kidney and pancreas. This finding was further strengthened by examining the expression of genes of the Minn *et al*. signature that were repressed in DCCs (**Figure 20 E**). The expression of well-established mediators of lung metastasis, such as Tenascin C (TNC) [52], SPARC [264] and MMP1 [87] that are enriched in lung metastatic cells compared to the mammary tumor, were down regulated in disseminated cells compared to mammary tumor cells (**Figure 20 E**).

Taken together the data indicate that disseminated cancer cells from kidney and pancreas do not only show a differential gene expression profile compared to cancer cells from the mammary tumor but are also distinctly different from metastatic cells in the lung.

7.3.4 Cellular functions that are regulated in DCCs from kidney and pancreas

Building on the differences between DCCs and mammary tumor cells, the gene expression changes in DCCs were characterized in more detail using the transcriptomic profiles that were generated immediately after isolation of cancer cells. The following paragraphs focuses on the molecular functions that were differentially regulated in disseminated cancer cells isolated from kidney as well as pancreas compared to cancer cells from the mammary tumor. A list summarizing all used gene signatures and GO-terms is provided in the appendix in **Table S 1**.

7.3.4.1 Apoptotic responses are repressed in DCCs from kidney and pancreas

A major challenge of disseminated cancer cells in unfavorable microenvironments is survival. As we were able to show that cancer cells of PDX models, located in kidney, pancreas and spleen, survived for several weeks after tumor resection (**Figure 14**), it was analyzed whether cell death gene responses were changed in DCCs. Gene ontology analysis revealed that GO-terms associated with cell death and apoptosis were repressed in disseminated cells from kidney and pancreas (**Figure 21 A**). In total 10 GO-terms associated with cell death were enriched in the genes with repressed expression in kidney DCCs and 12 terms in pancreas DCCs compared to cancer cells from the mammary tumor (**Figure 21 A**). The genes, involved in the GO-term "GO:0008219- cell death" that were downregulated in kidney DCCs compared to cancer cells from the mammary tumor with a p-value of 0.019 (**Figure 21 B**) were analyzed in more detail. These genes included several pro-apoptotic genes (8 in total), such as Caspase4 (CASP4), which gets cleaved during the executive phase of apoptosis and Egl9-family hypoxia inducible factor 3 (EGLN3), which inhibits the anti-apoptotic effects of BCL2. However, components of the GO-term 0008219 with repressed expression in disseminated cancer cells from kidney included also anti-apoptotic genes (4 in total). For examples, the expression of TNF Receptor Superfamily F Decoy (TNFRSF10D) was reduced in DCCs, which is a TRAIL-receptor with a truncated death domain and is therefore not able to induce TRAIL mediated apoptosis but inhibits it (**Figure 21 B**). A reduction in apoptosis related genes in DCCs compared to metastatic cancer cells from the lung was also confirmed by GSEA using different gene signatures related to apoptosis. A representative enrichment plot of the HALLMARK_apoptosis signature, which was underrepresented in DCCs with a normalized enrichment score of 1.96 and FDR of 0.0, is shown in **Figure 21 C**.

To clarify if the cell death associated genes repressed in kidney DCCs lead to increased or repressed apoptosis in disseminated cancer cells, immunofluorescence staining of cleaved caspase 3 was performed (**Figure 21 D**) on tissue slides of MDA-MB-231 mammary tumors, lungs and kidneys five weeks after orthotopic injection of tumor cells. Cleaved caspase 3 is a commonly used marker for apoptosis [265]. Upon apoptosis activation caspase 3 is cleaved leading to DNA fragmentation and morphological changes and cleavage is induced by several stimuli in different cell types. In this study, cleaved caspase 3 was detected in about 1% of cancer cells in mammary tumors and in about 4% of cancer cells in the. Out of 220 analyzed DCCs in the kidney, not a single cancer cell was positive for cleaved caspase 3 (**Figure 21 D** and **E**).

This confirms the initial findings from the gene expression data indicating that apoptosis is reduced in DCCs from the kidney compared to cancer cells from the mammary tumor as well as metastatic cells from the lung.

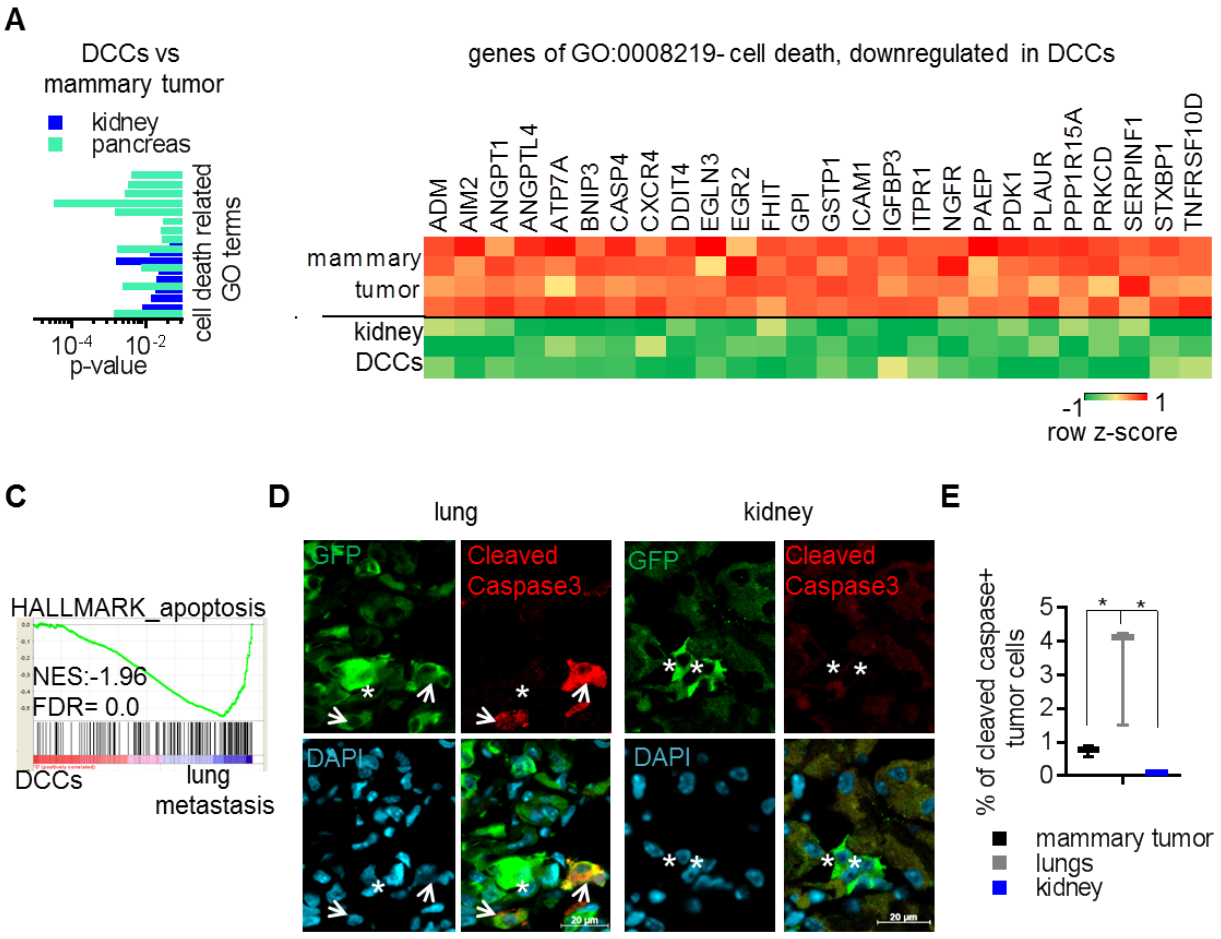


Figure 21 Cell death is reduced in DCCs in kidney and pancreas

A A collection of gene ontology signatures (see **Table S 1** for the complete list) associated with cell death and apoptosis was significantly enriched in the genes downregulated in DCCs from kidney and pancreas compared to cells from the mammary tumor (experimental setup**: Figure 17**). Each row represents one GO-term, which is plotted against the logarithmic p-value. Dark blue: kidney DCCs, light blue: pancreas DCCs.

B Heat map of the genes belonging to the Gene ontology term GO:0008219- "cell death", which are significantly regulated between kidney DCCs and cells from the mammary tumor. Row z-scores are shown.

C GSEA-plot of the HALLMARK_"apoptosis" gene signature comparing DCCs from kidney and pancreas to metastatic cells from the lung in the experimental setup described in **Figure 18.** Normalized enrichment score (NES) and false discovery rate (FDR) are shown.

D IF staining for cleaved caspase 3 in lungs with MDA-MB-231 metastasis and kidneys with DCCs five weeks after orthotopic tumor cell injection to confirm downregulation of apoptotic activity. Cancer cells are stained with GFP (green), cleaved-caspase 3 staining is shown in red and nuclei are stained with DAPI (blue). Arrows indicate cancer cells staining positive for cleaved caspase, * mark cancer cells that are negative for cleaved caspase 3. The scale bar shows 20 µm.

E.Quantification of cleaved caspase 3 positive MDA-MB-231 cells in the mammary tumor, lung and kidney identified by IF staining. Three biological replicates were analyzed. In each replicate more than 1000 cancer cells from the mammary tumor and more than 500 cancer cells in the lung were analyzed. In the kidney, 223 cancer cells were analyzed in total.

7.3.4.2 DCCs in kidney and pancreas repress EMT

During the invasive stage of cancer metastasis, cancer cells obtain a mesenchymal phenotype via epithelial to mesenchymal transition (EMT) allowing them to migrate and degrade the ECM [36]. Growth at the secondary site in contrast requires reversion to a proliferative epithelial state via MET [80]. However, direct clinical evidence for EMT in patients is difficult to obtain [65]. In context of this discussion, the epithelial and mesenchymal characteristics of disseminated cancer cells in kidney and pancreas are of interest as these cells may represent a subpopulation at an intermediate state of the metastatic cascade during a latency period.

GSEA was performed on the DCC dataset using signatures associated with EMT or MET (**Table S 1**). Among others, the EMT signature from the HALLMARK database, mesenchymal and epithelial signatures in colon cancer cells published by Rokavec *et al.* [266], a signature from Jechlinger *et al.* with genes that are upregulated in mesenchymal vs epithelial EpRas cells [267] and a signature of genes negatively regulated by the EMT transcription factor ZEB1 in MDA-MB-231 cells published by Aigner *et al.* [268] were applied (**Figure 22 A** and **B**). A heat map depicting the NES of the signatures that were significantly enriched in DCCs versus cells from the mammary tumor with a p-value of at least 0.05 or a FDR of 0.25 or lower is shown in **Figure 22 A**. Signatures associated with a mesenchymal phenotype were underrepresented in disseminated cells from kidney and pancreas, respectively, whereas signatures of epithelial cells were positively enriched in DCCs (**Figure 22 A**).

The enrichment plots of the Rokavec *et al.* signatures [266] are shown in **Figure 22 B**. Rokavec *et al.* induced EMT in colon cancer cells by IL6 treatment and selected stable mesenchymal clones by sphere and anchorage independent growth assays [266]. By comparing these two populations they generated a signature of genes upregulated by EMT. In the next step, they intravenously injected the mesenchymal cells into mice. *In vivo,* the cells reacquired an epithelial phenotype by MET [266]. The comparison between the reacquired epithelial cells and the injected mesenchymal cells yielded a MET signature. GSEA on the DCCs dataset was performed using the EMT (Rokavec_mesenchymal vs epithelial) as well as the MET signature (Rokvec_epithelial vs mesenchymal). The EMT gene signature was significantly underrepresented in kidney DCCs compared to cells from the mammary tumor with a NES of -1.8 (FDR=0.04). The MET gene signature was positively enriched in kidney DCCs with a NES of 1.4 and a FDR of 0.2 (**Figure 22 B**).

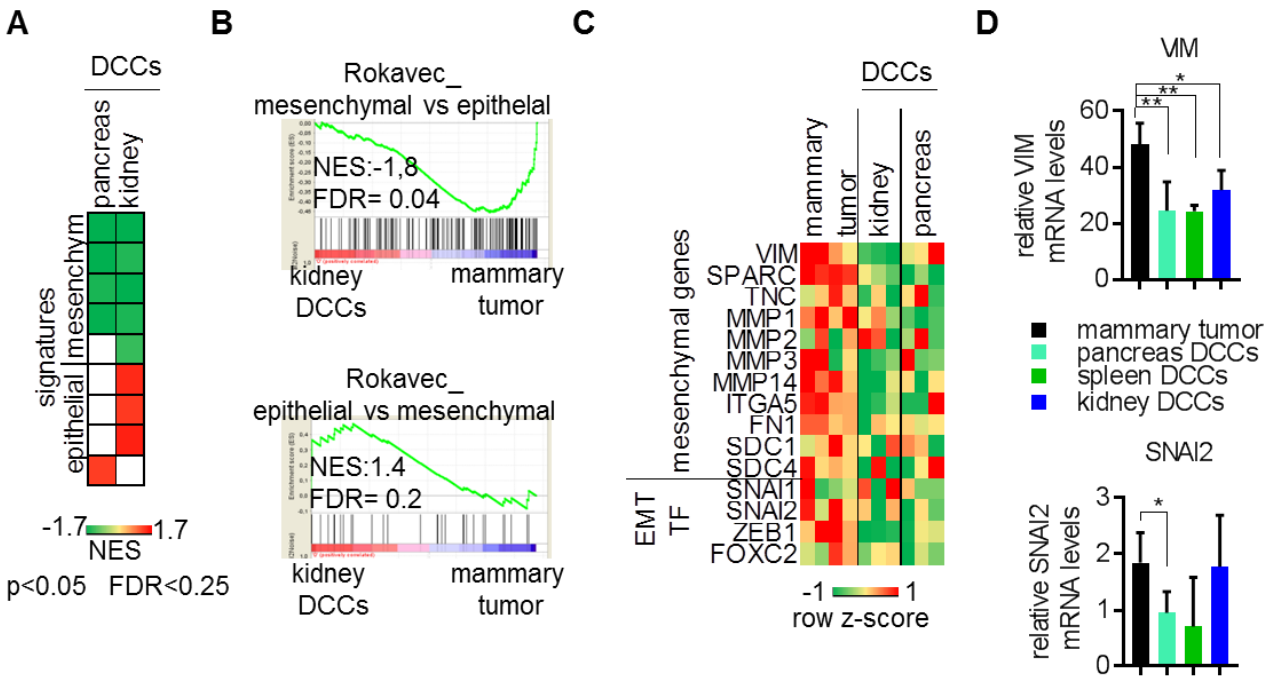


Figure 22 DCCs in kidney and pancreas repress a mesenchymal phenotype

A Collection of gene signatures related to epithelial to mesenchymal transition (EMT) (see **Table S 1** for a summary of all signatures), which were tested by GSEA in the dataset from DCCs in kidney or pancreas compared to the mammary tumor. The normalized enrichment scores (NES) are shown of signatures significantly enriched (FDR<0.25 or p<0.05)

B Enrichment plots of the gene signatures generated by Rokavec *et al*. [266] in the kidney DCC dataset compared to the mammary tumor. Upper panel: gene signature enriched in mesenchymal colon cancer cells after EMT compared to epithelial cells. Lower panel: genes enriched in epithelial colon cancer cells after MET compared to mesenchymal.

C Heat map of selected mesenchymal marker genes and transcription factors (TF) involved in epithelial to mesenchymal transition (EMT). Row z-scores are shown for the expression values of these genes in the Affymetrix array performed on MDA-MB-231 cells isolated from mammary tumor and DCCs from kidney or pancreas.

D Validation of downregulation of SNAIL2 and VIM in disseminated cancer cells from kidney, pancreas and spleen using qRT-PCR. qRT-PCR was performed on a second set of samples with 3-6 biological replicates. Expression values were normalized to human TATA-box binding protein (hTBP) expression. P-values comparing DCCs individually to the mammary tumor were calculated using one-way Anova. Only comparisons with a p<0.05 are marked. * p<0.05; ** p<0.001

The GSEA results suggest that the expression of genes involved in EMT and a mesenchymal phenotype is downregulated in disseminated cells in kidney and pancreas compared to cancer cells from the mammary tumor. To strengthen these results the expression of known mesenchymal and epithelial markers as well as EMT mediators was analyzed in DCCs compared to mammary tumor cells. As expected from the GSEA results, the expression of several established mesenchymal markers was repressed in disseminated cancer cells from kidney and pancreas compared to cells from the mammary tumor (**Figure 22 C**). Repressed

mesenchymal genes included Vimentin (VIM), SPARC, Tenascin C (TNC), Matrix metallopeptidases (MMP) 1, 2, 3 and 14, Integrin alpha 5 (ITGA5), Fibronectin 1 (FN1) and Syndecan (SDC) 1 and 4. In addition, the expression of SNAIL 1 and 2, FOXC1 and ZEB1, well established transcription factors regulating EMT, was repressed in disseminated cells (**Figure 22 C**). Reduced expression of VIM and SNAI2 in DCCs was additionally confirmed by qRT-PCR using an independent sample set of sorted DCCs and mammary tumor cells (**Figure 22 D**).

These data suggest that disseminated cancer cells in kidney and pancreas repress mesenchymal features and downregulate EMT mechanisms compared to cancer cells from the mammary tumor

7.3.4.3 DCCs undergo a metabolic switch from glycolysis to oxidative phosphorylation

The gene ontology analysis, depicted in **Figure 19**, was performed to gain a first overview over the cellular processes that were regulated in DCCs from kidney and pancreas compared to mammary tumor cells. Especially in the samples that were analyzed immediately after isolation, many of the enriched GO-terms were associated with metabolism. 36 metabolic GO-terms were enriched in the downregulated genes in pancreas DCCs and even 71 in kidney DCCs. Based on this data, a more thorough analysis of the metabolic phenotype of disseminated cancer cells was performed. GSEA comparing disseminated cells from kidney and pancreas to cancer cells from the mammary tumor was run using gene signatures associated with metabolism (**Figure 23 A**). Three types of metabolism were distinguished: carbohydrate metabolism, all biochemical reactions involved in formation, degradation and conversion of carbohydrates, energy metabolism, which summarizes all reactions involved in the production of ATP and NADPH and nucleotide metabolism, the process of nucleotide production and degradation. Gene signatures belonging to all three categories were significantly downregulated in disseminated cancer cells from kidney and pancreas (**Figure 23**). Representative enrichment plots of one gene signature of each of these categories are shown in **Figure 23 B**. As example for a carbohydrate metabolism signature, the Glycolysis signature of the HALLMARK dataset was negatively enriched with a NES of -2.03 (FDR 0.0) in pancreas DCCs. The GO-term “ATP production from ADP” is shown representative for an energy metabolism signature with a NES of -2.14 and a FDR of 0.0 and the GO-term “regulation of cyclic nucleotide metabolism” with a NES of -1.56 (FDR 0.28) is shown as nucleotide metabolic signature. A list with all used signatures is provided in the appendix in **Table S 1**.

This analysis shows that genes involved in carbohydrate metabolism, energy production as well as nucleotide homeostasis are repressed in disseminated cancer cells in kidney and pancreas.

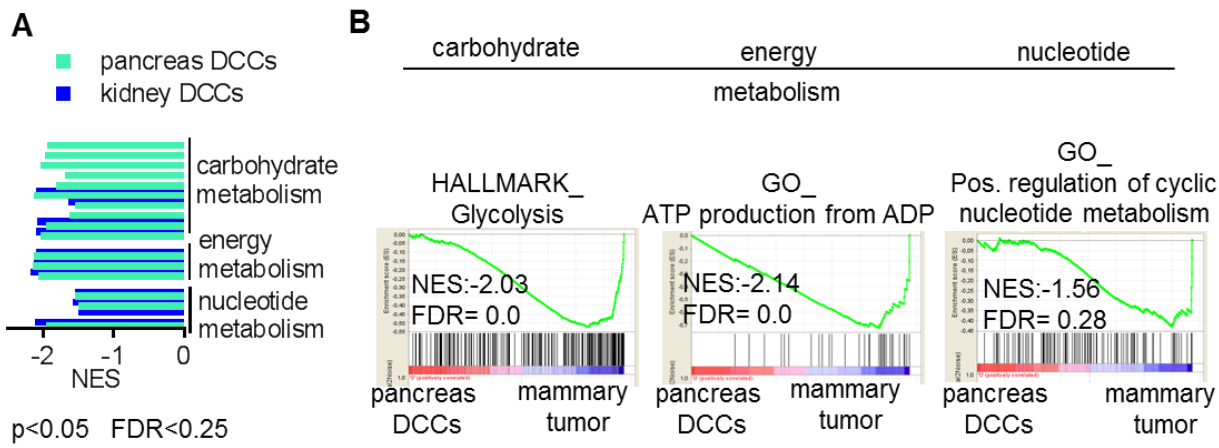


Figure 23 Metabolic processes are downregulated in DCCs in kidney and pancreas

A GSEA was performed on DCCs from kidney and pancreas, respectively compared to cells from the mammary tumor using gene signatures associated with metabolism. A collection of gene signatures related to carbohydrate, energy or nucleotide metabolism that were significantly (FDR<0.25 or p-value<0.05) downregulated in DCCs from kidney or pancreas compared to the mammary tumor are shown. The following gene signatures were included: “glycolysis” signatures from the Hallmark, Reactome and Biocarta platform respectively, GO-terms “glucose catabolic metabolism”, “cellular response to carbohydrates”, “response to carbohydrates” and “pyruvate metabolic processes”, KEGG “Galactose” and “Fructose and Mannose metabolism” signatures and Reactome “glucose metabolism”. Each bar represents one signature, which is plotted against its normalized enrichment score (NES).

B Representative enrichment plots of gene signatures belonging to nucleotide, energy or carbohydrate metabolism, respectively selected from (**A**). Plots are shown for the comparison of pancreas DCCs with cancer cells from the mammary tumor. NES: normalized enrichment score; FDR: false discovery rate.

The regulation of genes involved in carbohydrate metabolism was especially prominent in the GO-term analysis as well as in GSEA. Cells have two main pathways for energy production: aerobic oxidative phosphorylation, when enough oxygen is available or anaerobic glycolysis and production of lactate. During oxidative phosphorylation a cell gains 36 molecules of ATP per molecule glucose, whereas the energy output of one glucose molecule after glycolysis is only 2 ATP molecules. Nevertheless, cancer cells often use the glycolytic pathway even if oxygen is available, which is referred to as Warburg effect [269]. Following up on the results that disseminated cancer cells in kidney and pancreas seem to repress carbohydrate metabolism, the genes involved in these pathways were analyzed more closely. The following gene signatures were implemented in the analysis (**Table S 1**): the “glycolysis” signatures from the Hallmark, Reactome and Biocarta platform respectively, the “glucose catabolic metabolism”, “cellular response to carbohydrates”, “response to carbohydrates” and “pyruvate metabolic processes” signatures from the GO platform, the “Galactose, Fructose and Mannose metabolism” signature from KEGG as well as the Reactome “glucose metabolism” signature.

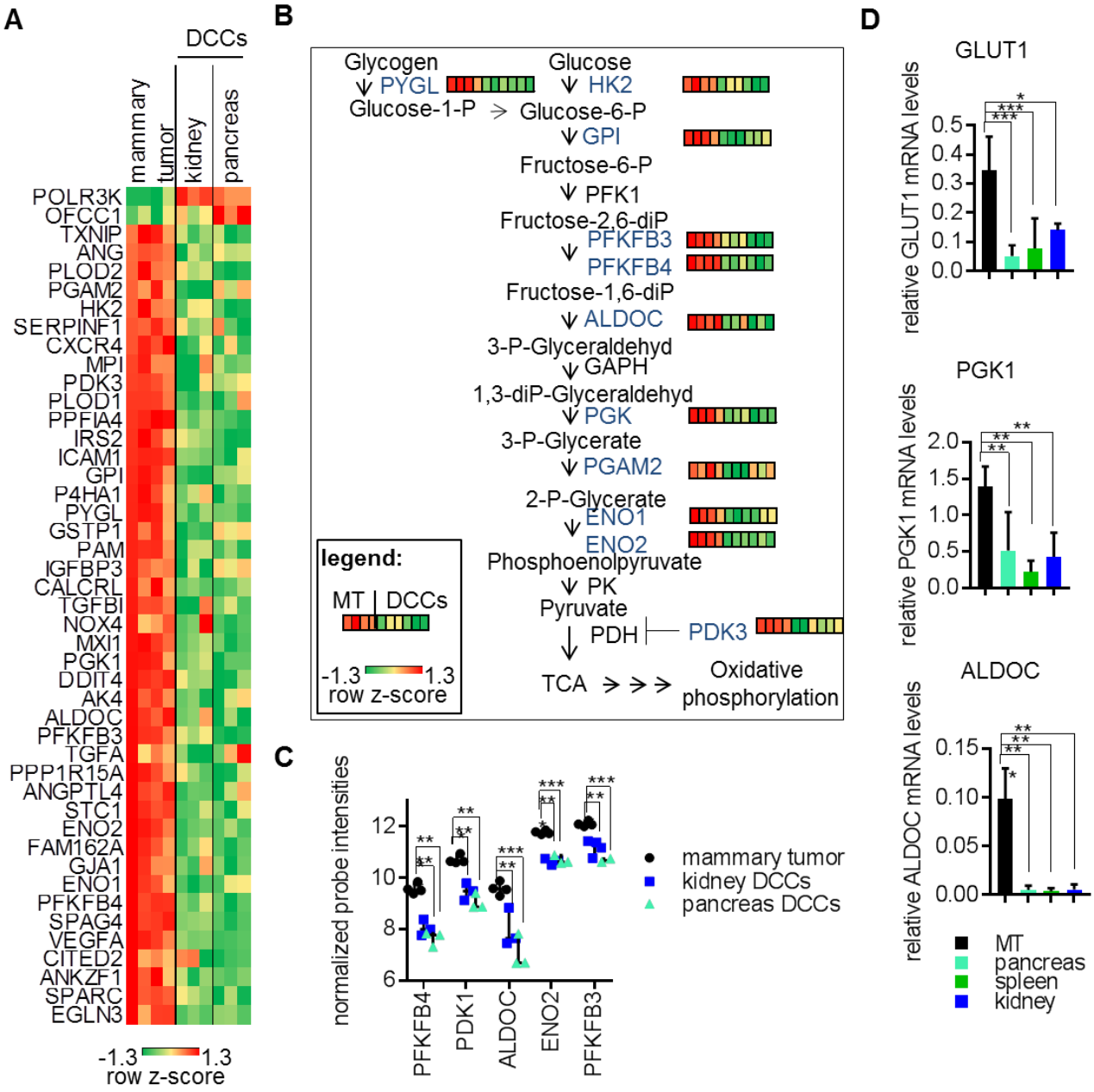


Figure 24 Glucose metabolism is repressed in DCCs in kidney and pancreas

A Heat map of all genes belonging to the carbohydrate metabolism gene signatures used in **Figure 23 A** that were significantly changed (FC>1.5; p<0.05) in kidney or pancreas DCCs compared to the mammary tumor. Row z-scores are shown for all genes.

B Graphical summary of the significantly downregulated genes (p < 0.05; FC > 1.5) in DCCs that encode enzymes of the glycolysis pathway. All genes with significantly altered expression in kidney and pancreas DCCs are written in blue. Heat map of the row z-scores is depicted for each gene. MT: mammary tumor

C Affymetrix normalized probe intensities of some of the glycolytic enzymes marked in (**B**), that were significantly repressed in DCCs from kidney and pancreas compared to the PT. ** p<0.01; *** p<0.005

D qRT-PCR validation of the transcriptional repression of genes, involved in glycolysis, in DCCs from kidney, pancreas and spleen was performed in an independent sample set. Selective primers for GLUT1, phosphoglycerate kinase 1 **(**PGK1) and aldolase C (ALDOC) were used. Expression was normalized to human TATA-box binding protein (hTBP). p-values were calculated via one-way Anova. *p<0.05; **p<0.001 ***p<0.0001. MT: mammary tumor

A heat map including all genes from those metabolism signatures with significantly changed expression in kidney or pancreas DCCs ($p < 0.05$ and $FC > 1.5$) is shown in **Figure 24 A**. Each gene is shown only once even if it was including in more than one signature. Going through this gene list, it was apparent that several of the genes are directly involved in the glycolysis pathway. Glycolysis refers to the pathway converting glucose to pyruvate. Genes transcribing enzymes of almost each step of glycolysis were downregulated in DCCs from kidney and pancreas (**Figure 24 B**). Examples are fructose- 2,6 biphosphatase 3 and 4 (PFKFB3 and 4), catalyzing the reaction from fructose 2,6- di-Phosphate to fructose 1,6- di-Phosphate, aldolase C (ALDOC), which converts fructose 1,6- di-Phosphate further to 3-Phospho Glyceraldehyde and Enolase 1 and 2 (ENO1 and 2) that catalyze the reaction from 2-Phospho Glycerate to Phosphoenolpyruvate further downstream of the pathway. The normalized probe intensities of these genes, as detected on the Affymetrix arrays are plotted in **Figure 24 C**. Based on these data we assumed that also glucose uptake into the cells, which is primarily mediated by GLUT1 glucose transporter, may be downregulated in disseminated cancer cells from kidney and pancreas. To test this, qRT-PCR was performed in a second independent sample set using primers selectively targeting GLUT1. Indeed, a downregulation of GLUT1 transcription was observed in DCCs (**Figure 24 D**). In addition, ALDOC and phosphoglycerate kinase 1 (PGK1) downregulation was confirmed by qRT-PCR (**Figure 24 D**).

The expression of pyruvate dehydrogenase kinases (PDK) 1 and 3 was also significantly downregulated in disseminated cancer cells in kidney and pancreas compared to cells from the mammary tumor (**Figure 24**). These two enzymes are not directly involved in glycolysis but inhibit pyruvate dehydrogenase (PDH), the enzyme that converts pyruvate to Acetyl-CoA, which is further processed in the oxidative phosphorylation, taking place in mitochondria [270]. Therefore, it was assumed that the Warburg effect may be reversed in DCCs in kidney and pancreas and a switch from anaerobe energy production to oxidative phosphorylation may have taken place. This was tested by GSEA using gene signatures related to oxidative phosphorylation. Six of seven tested signatures were significantly enriched in DCCs isolated from kidney, whereas only the oxidative phosphorylation signature, published by Mootha *et al*. [271] was enriched with a p-value below 0.05 in pancreas DCCs (**Figure 25 A**). The enrichment plots of the Mootha signature in kidney and pancreas DCCs are shown in **Figure 25 B**. In kidney DCCs the signature was enriched with a NES of 1.82 and a FDR of 0.017 and in pancreas DCCs with a score of 1.38 at a FDR of 0.29. The enrichment plots indicated that the enrichment of oxidative phosphorylation is weaker in disseminated cells from the pancreas compared to DCCs from the kidney, because the involved genes were less strongly upregulated in pancreas DCCs than in kidney DCCs (**Figure 25 B**).

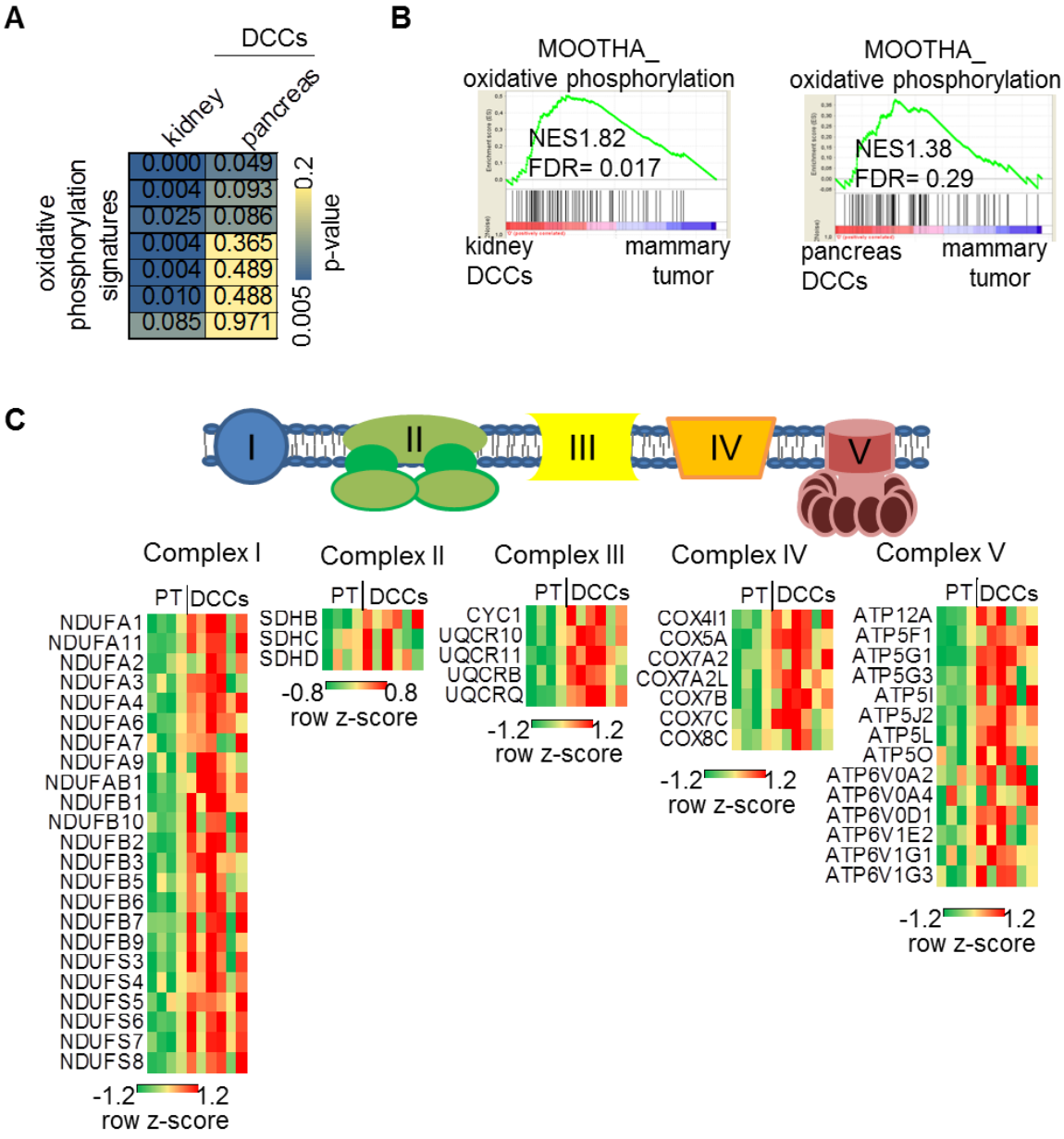


Figure 25 Genes involved in oxidative phosphorylation are upregulated in DCCs in kidney and pancreas

A Different gene signatures regarding oxidative phosphorylation and the mitochondrial electron chain (see **Table S 1** for the complete list of signatures) were enriched in kidney and pancreas DCCs compared to the mammary tumor. The graph shows the p-values of the tested signatures. The following signatures were used: VOXPHOS, HUMAN MITODP_6_2002 and MITOCHONDRIA signatures published by Mootha *et al*. [271], Respiratory chain and PROTON_TRANSPORTING_ATP_SYNTHASE_COMPLEX from GO, KEGG "oxidative phosphorylation" and MODULE_152 from the Broad Institute.

B Representative enrichment plot of the Mootha *et al*. gene signature for oxidative phosphorylation [271], which is also included in the table in (**A**) in the first row. The left graph shows the plot for the comparison of kidney DCCs vs mammary tumor, the right for pancreas DCCs vs mammary tumor.

C Graphical summary of all genes encoding components of the mitochondrial electron chain, which are covered by the gene signature „oxidative phosphorylation" from the KEGG database and were core enriched in kidney DCCs compared to cells from the mammary tumor. The genes are grouped according to the electron chain complexes they belong to. Complex I: NADPH:Ubiquinone Oxidoreductase; complex II: Succinate Dehydrogenase; complex III: Ubiquinol-Cytochrome C Reductase; complex IV: Cytochrome C oxidase and complex V: ATP synthase. Row z-scores are depicted in heatmaps.

The electron chain in mitochondria is composed of five complexes: NADPH:Ubiquinone Oxidoreductase (complex I), Succinate Dehydrogenase (complex II), Ubiquinol-Cytochrome C Reductase (complex III), Cytochrome C oxidase (complex IV) and ATP synthase (complex V) [272] (**Figure 25 C**). A possible enrichment of genes, coding for these complexes, in DCCs was tested using the KEGG "oxidative phosphorylation" signature. The KEGG signature was significantly enriched in disseminated cancer cells in the kidney with a p-value of 0.004. In pancreas DCCs the signature was only slightly enriched with p-value of 0.365. Several genes, encoding subunits of the electron chain complexes, were enriched in kidney DCCs compared to cancer cells from the mammary tumor (**Figure 25 C**). For a better overview, enriched genes in kidney DCCs were grouped into the five mitochondrial complexes. Genes encoding members of all five complexes were significantly enriched in disseminated cancer cells in kidney compared to cancer cells from mammary tumor with an emphasis on complex I and V with 23 and 14 core enriched genes, respectively. In addition, three genes belonging to complex II, five genes of complex III and seven genes encoding components of complex IV were core enriched in kidney DCCs. Although the enrichment of the KEGG signature in disseminated cancer cells from the pancreas was not significant, the majority of the core-enriched components of the mitochondrial electron chain, in kidney DCCs were also upregulated in pancreas DCCs with the exception of genes encoding for complex III (**Figure 25 C**).

The analysis of the metabolic changes occurring in disseminated cancer cells located in kidney and pancreas compared to cancer cells isolated from the mammary tumor suggests the DCCs switched their energy metabolisms from anaerobic production of lactate via glycolysis to oxidative phosphorylation. The expression of genes, encoding enzymes involved in the glycolytic pathway, was repressed in DCCs. The energy loss resulting from reduced glycolytic activity seems to be compensated by an increase in oxidative phosphorylation. The expression of genes encoding components of all electron chain complexes in mitochondria was increased in disseminated cancer cells in kidney and pancreas compared to cells from the mammary tumor. Furthermore, the enzymatic reaction catalyzing the conversion of Pyruvate into Acetyl-co, which initiates oxidative phosphorylation, is likely increased as the expression of the inhibitors PDK1 and 3 was reduced in DCCs.

7.3.4.4 DCCs in kidney and pancreas downregulate hypoxia response

In the context of glycolytic repression that was observed in DCCs in kidney and pancreas as compared to cells from the mammary tumor (**Figure 24**), downregulation of hypoxia response signatures in kidney and pancreas DCCs, respectively as observed by GSEA was of interest (**Figure 26 A**). Hypoxia, which refers to reduced oxygen availability that leads to molecular changes in cells, is a key metabolic regulator. Under hypoxic conditions the glycolytic pathway is upregulated as several molecules involved in glycolysis are direct targets of the transcription factor HIF1 (Hypoxia induced transcription factor 1), which is the master transcriptional regulator of cellular responses to hypoxia [273]. HIF1 targets include GLUT1, PGK1, PFKFB3, HK2, ENO1 and ALDOC, genes with significantly downregulated expression in disseminated cancer cells in kidney and pancreas (**Figure 24**). Hypoxia related signatures that were significantly underrepresented in DCCs from kidney and pancreas compared to cancer cells from the

mammary tumor by GSEA included the signatures published by Elvidge *et al.* [274], Fardin *et al.* [275] and Semenza [276]. Elvidge *et al.* performed gene expression microarray analysis of the breast cancer cell line MCF7 under hypoxic vs normoxic conditions as well as upon knockdown of HIF1 and 2 compared to control cells [274]. All signatures compiling genes that are upregulated under hypoxic conditions were repressed in disseminated cancer cells. Whereas signatures composed of genes downregulated under hypoxia were positively enriched in disseminated cancer cells in kidney and pancreas. Representative enrichment plots are shown in **Figure 26 B**.

In 2004, Schofield and Ratcliffe compiled all direct transcriptional targets of HIF1, for which experimental evidence was available at that time [277]. Based on this study, the expression of these direct HIF target genes was investigated in our dataset. The means of the row z-scores of the HIF target genes in kidney DCCs, pancreas DCCs or mammary tumor were calculated separately. Genes were analyzed for their expression change in disseminated cancer cells in kidney and pancreas compared to cancer cells from the mammary tumor. Genes were considered as downregulated if the mean z-scores in kidney as well as pancreas DCCs were lower than the mean row z-scores in the mammary tumor samples (colored green in **Figure 26 C**). Genes were counted as upregulated if the row z-scores in both, kidney and pancreas DCCs were higher than the score of the mammary tumor samples (colored red in **Figure 26 C**). Genes, for which the direction of expression change compared to the mammary tumor was different in kidney and pancreas DCCs or which were not changed at all, were not considered (colored black in **Figure 26** C). In summary, the majority of HIF target genes were downregulated in DCCs compared to cells from the mammary tumor with 26 repressed genes and 8 upregulated of 42 analyzed genes. In their review, Schofield and Ratcliffe categorized the HIF targets into functional groups [277]. In the DCC dataset, the downregulated genes were distributed over all functional groups with a special emphasis on genes involved in energy metabolism (**Figure 26 C**). With exception of GAPDH all 10 genes belonging to this category were repressed in kidney and pancreas DCCs compared to the mammary tumor. All other functional groups contained also genes with increased expression in DCCs, making it difficult to draw a conclusion of the functional output in disseminated cells in kidney and pancreas.

The data obtained by GSEA suggest that disseminated cancer cells in kidney and pancreas have a lower hypoxia response signature than cancer cells isolated from the mammary tumor. Analysis of the gene expression changes of direct HIF target genes [277] in DCCs from kidney and pancreas confirmed the repression of hypoxia associated genes in DCCs.

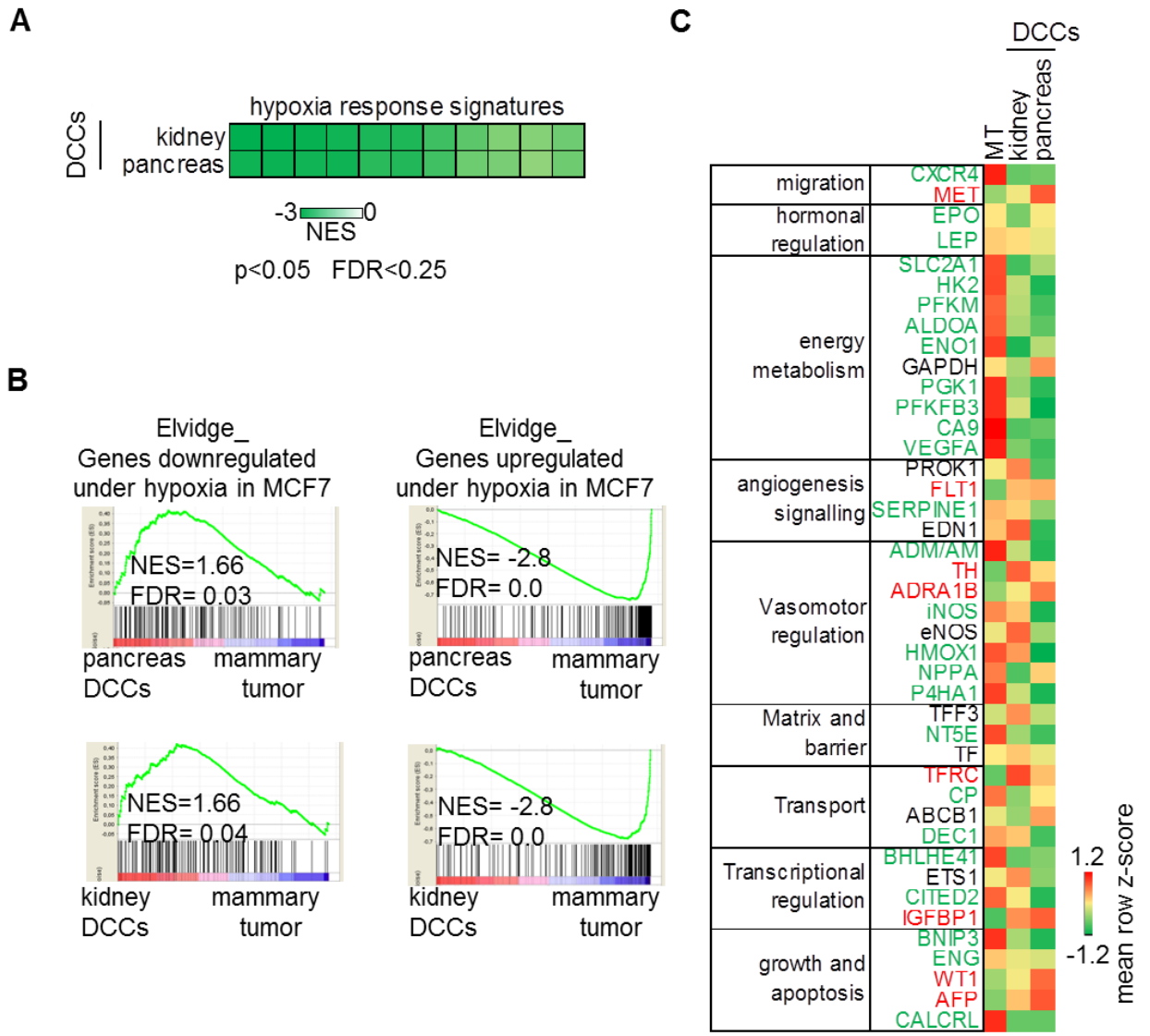


Figure 26 The expression of hypoxia response genes is repressed in DCCs from kidney and pancreas

A Hypoxia response signatures were significantly downregulated in DCCs from kidney and pancreas compared to the mammary tumor. Normalized enrichment scores (NES) of the signatures are depicted for kidney and pancreas DCCs respectively compared to the mammary tumor as calculated by GSEA. A list with all used signatures is provided in the appendix **Table S 1**

B Representative GSEA enrichment plots of the gene signatures generated by Elvidge *et al.* in DCCs from pancreas (upper panel) and kidney (lower panel), respectively. Elvidge *et al*. profiled MCF7 breast cancer cells under hypoxic and normoxic conditions [274]. NES and FDR of all comparisons are shown.

C Schofield and Ratcliffe summarized genes, which have been reported as direct transcriptional targets of hypoxia inducible factor (HIF) [277]. The heat map shows the expression of those HIF target genes in the DCCs data set. Expression is plotted as mean row z-scores in the experimental groups (n (mammary tumor):4, n (DCCs): 3 each). Genes were grouped according to the functional groups assigned by Shofield and Ratcliffe. The genes are colored according to their expression changes in DCCs compared to the mammary tumor (MT): genes that were downregulated in DCCs from both organs are written in green, genes that were upregulated in both DCCs compared to the mammary tumor are written in red and genes of which the expression was not changed at all or not consistent in pancreas and kidney DCCs are shown in black.

7.3.4.5 Antigen presentation is repressed in DCCs in kidney and pancreas

During the metastatic process, cancer cells are subjected to negative pressure by cells of the immune system. To be able to form metastasis, cancer cells need to avoid clearance by immune cells [149].

Unsupervised GSEA using the GO and Reactome datasets resulted in the negative enrichment of several gene signatures associated with antigen presentation (**Figure 27 A**). Antigen presentation is mediated by major histocompatibility complexes (MHC), receptors that bind and present antigens at the surface of cells and are encoded by the human leucocyte antigen (HLA) genes [278]. In the DCC dataset, the GO-terms “MHC protein complex”, “MHC protein complex binding”, “MHC class II protein complex binding”, “peptide antigen binding” and “antigen binding” were underrepresented in DCCs isolated from kidney and pancreas compared to cells from the mammary tumor. In addition, the Reactome signature “MHC class II antigen presentation” was significantly repressed in pancreas DCCs (**Figure 27 A**). Representative GSEA enrichment plots of the GO-term “antigen binding” are shown in **Figure 27 B** for kidney and pancreas DCCs, respectively. In kidney DCCs, the signature was enriched with a NES of -1.52 at a p-value of 0.015 and a FDR of 0.3 and in disseminated cancer cells from the pancreas with a NES of -1.94 at a FDR of 0.037. Among the core enriched genes of the “antigen binding” signature in DCCs were almost exclusively HLA genes (except for CD74). Six of these core-enriched HLA molecules belonged to class I and eleven to class II (**Figure 27 C**).

MHC complexes, expressed on the surface of cells, can be recognized by immune cells to identify and clear foreign or malfunctioning cells. In this context, the surface expression of the class I HLA molecules A,B and C as well as of the class II molecule HLA-DR on cancer cells was analyzed using flow cytometry. MDA-MB-231 cancer cells from the mammary tumor, the lung, liver, pancreas, kidney, spleen and bone were analyzed five week after orthotopic injection of cancer cells. The number of cancer cells in the pancreas was too low for reliable analysis and was therefore excluded. About 65% of cancer cells in the mammary tumor expressed HLA A,B and C. This percentage increased in cancer cells from lung and liver to 90% and in bone to 75%. The HLA class I expression of cancer cells isolated from kidney and spleen did not significantly change with a mean expression of 80 and 88% of cancer cells, respectively (**Figure 27 D** upper panel). The MHC class II protein HLA-DR was expressed on 22 % of cancer cells in the mammary tumor. Again, this percentage was slightly increased to 26% and 30% in cancer cells isolated from lung and liver and stayed in the bone with 21.5% of HLA-DR expressing cancer cells within the range of the mammary tumor. In contrast to these findings, significantly less DCCs located in kidney and spleen expressed the MHC II molecule HLA-DR with 7% positive cells in each organ (**Figure 27 D** lower panel). Thus, the decreased gene expression of HLA class II genes was also detected on protein level with a 3-fold reduction of cancer cells expressing HLA-DR on their surface in kidney and spleen compared to cancer cells in the mammary tumor.

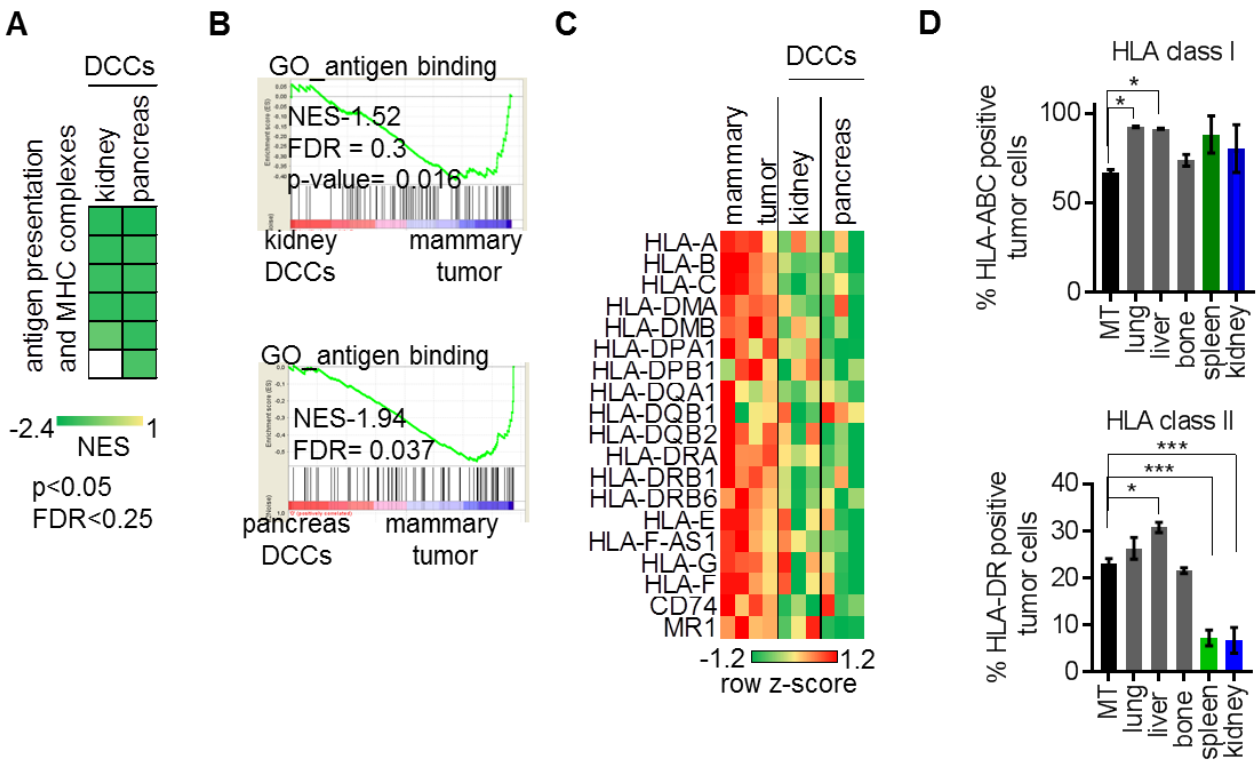


Figure 27 Antigen presentation is repressed in DCCs in kidney and pancreas by downregulation of HLA class II molecules

A Gene set enrichment analysis was performed with gene signatures of the GO and Reactome datasets associated with antigen presentation and major histocompatibility complex (MHC) assembly (**Table S 1**). A heat map with the normalized enrichment scores (NES) of the gene signatures that were significantly negatively enriched in cancer cells in kidney and pancreas compared to cells from the mammary tumor are shown ($p<0.05$ or $FDR<0.25$).

B Example of the gene set enrichment plots for the GO gene signature „antigen binding“, included in **A**, that was significantly enriched in cells from the mammary tumor compared to DCCs**.**

C Genes of the „antigen binding“ gene signature of the GO database (GSEA plot in **B**) that were core-enriched in the mammary tumor compared to DCCs from kidney and pancreas. Row z-scores of the core enriched genes are shown.

D Confirmation of the downregulation of the MHC class II molecule HLA-DR on MDA-MB-231 DCCs from kidney and spleen compared to cancer cells in lung, bone, liver and the mammary tumor (MT) on protein level by flow cytometry. The surface protein expressions of the MHC class I molecules HLA-A, B and C were not significantly changed on cancer cells from kidney and spleen compared to cells from other organs. (n= 2) The cells were analyzed five weeks after orthotopic injection of MDA-MB-231 cells in NSG mice. Cancer cells were identified using the FACS strategy described in **Figure 9**. P-values comparing all organs individually to the mammary tumor were calculated using one-way ANOVA. Only significant p-values are shown. * $p<0.05$; *** $p<0.001$

Following up on the repressed expression of genes as well as surface proteins of the HLA class II family, molecular regulation of the repression was further investigated. In this study immune compromised NSG-mice were used for all *in vivo* experiments. These mice are severely immune deficient and lack any mature adaptive immune cells as well as natural killer cells and express only defective macrophages and dendritic cells. Only innate immune response is partially active

in NSG mice [279]. On this background, extrinsic immune stimuli were excluded as trigger for the downregulation of HLA-DR due to the immune compromised nature of the NSG model. Therefore, cell intrinsic mechanisms leading to downregulation of HLA class II molecules were considered.

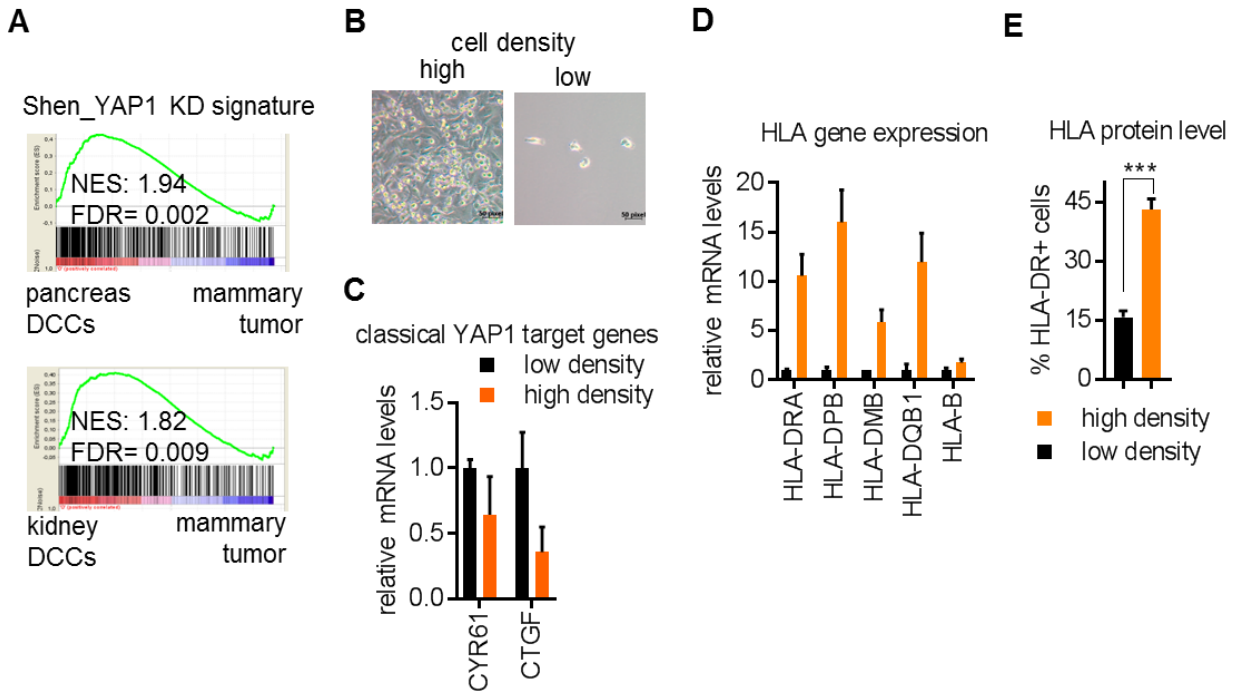


Figure 28 Downregulation of HLA class II genes may be regulated by the HIPPO pathway

A Gene set enrichment plots of the YAP1 signature published by Shen *et al.* [280] in DCCs from pancreas (upper plot) and kidney (lower plot) compared to mammary tumor cells, respectively. Normalized enrichment scores (NES) and false discovery rates (FDR) are shown.

B The impact of the HIPPO pathway on surface expression of HLA molecules was tested functionally using cell density experiments. MDA-MB-231 cells were cultivated under high density conditions with $5x10^6$ cells in a 6 cm culture dish or under low density conditions with $0.1x10^6$ cells seeded in a 15 cm culture dish. Representative images of both conditions are shown at a 4x magnification.

C Regulation of the gene expression of the classical YAP1 target genes CYR61 and CTGF by cell density was confirmed by qRT-PCR analysis of MDA-MB-231 cells, plated under the conditions shown in **B**. Expression of CYR61 and CTGF was normalized against expression of hTBP. Relative expression values compared to the low density samples are shown.

D mRNA levels of the HLA class II genes HLA-DRA, HLA-DPB, HLA-DMB and HLA-DQB as well as the HLA class I gene HLA-B were measured at different cell densities (**B**) by qRT-PCR. The expression values were normalized against expression of hTBP. mRNA levels relative to the low density samples are shown.

E Upregulation of the cell surface expression of the HLA-DR protein at high cell density was confirmed by flow cytometry analysis of MDA-MB-231 cells seeded in the conditions depicted in **B**. HLA surface expression was analyzed by flow cytometry 48h after seeding. (n=3) p-value was calculated using unpaired two-tailed t-test. ** $p<0.01$

Recent literature suggests that the HIPPO pathway, regulating YAP1 activity, is a cancer cell intrinsic regulatory pathway for immune escape [281-283]. The regulation of YAP1 targets in disseminated MDA-MB-231 cancer cells was tested, performing GSEA with a YAP1 knockdown signature published by Shen and Stanger [280]. They performed gene expression profiling of YAP1 knockdown HUMEC cells compared to control cells on the Affymetrix human exon gene chip 1.0 [280]. This YAP1 signature was significantly enriched in DCCs isolated from kidney and pancreas compared to cancer cells from the mammary tumor, respectively using GSEA (**Figure 28 A**). In pancreas DCCs the Shen signature was enriched with a NES of 1.94 at a FDR of 0.002 and in kidney DCCs with a NES of 1.82 and a FDR of 0.009.

Cell density is one of the most important triggers regulating YAP1 activity via the HIPPO pathway. At low cell density YAP1 target genes are active, whereas YAP1 signaling is inactive at high cells density [284, 285]. The impact of YAP1 activity on HLA gene expression was tested by measuring HLA mRNA levels dependent on cell density. Therefore, $0.1x10^6$ MDA-MB-231 cells were seeded on a 15 cm culture dish to establish low cell density conditions and $5x10^6$ cells were seeded on a 6 cm dish for high cell density. Representative microscope pictures illustrating the cell densities are shown in **Figure 28 B**. 48h after seeding, the regulation of YAP1 activity by the selected cell densities was confirmed by qRT-PCR amplifying the classical YAP1 target genes CYR61 and CTGF. The mRNA levels of both genes were downregulated under high cell density as compared to low density (**Figure 28 C**). Therefore, the expression of the HLA class I molecule HLA-B and the HLA class II molecules HLA-DRA, -DPB, -DMB and -DQB was tested by qRT-PCR under the same conditions. While HLA-B mRNA levels were not affected by changes in cell density, all tested HLA class II genes were upregulated under high density conditions (**Figure 28 D**). The upregulation of HLA-DR was also confirmed on protein level. The cell surface expression on HLA-DR was significantly increased upon high cell density as seen by flow cytometry analysis. Under high cell density, which results in low YAP1 activity, cell surface HLA-DR protein expression was detected on about 40% of cancer cells. At low cell density with high YAP1 signaling, the percentage of HLA-DR expressing cells was reduced to about 15% (**Figure 28 E**). The experiment was performed in three independent replicates. Also by FACS no difference in the percentage of HLA –A, B and C expressing cells was detected (data not shown).

The impact of YAP1 on HLA class II expression was further tested using YAP1 knockdown constructs. YAP1 knockdowns were generated in MDA-MB-231 or SUM159 cells using the miRE system [256]. Two different shRNAs targeting YAP1, obtained from the Sherwood database [257], were transduced into the triple negative breast cancer cell lines. A non-targeting vector was used as control (NonSil). YAP1 knockdown in both cell lines was confirmed by qRT-PCR. In MDA-MB-231 cells the knockdown results in a 70-80% reduction of YAP1 mRNA levels compared to control cells. In SUM159 cells the knockdown efficiency was about 50% (**Figure 29 A**). The mRNA levels of the HLA class I and II genes were measured by qRT-PCR in the YAP1 knockdown cells compared to the control cells. Also knockdown of YAP1 did not result in significant changes of HLA-B expression in both cell lines. However, mRNA levels of HLA-DRA, -DPB and DMB were increased in the knockdown cells in MDA-MB-231 and less strong also in SUM159 cells (**Figure 29 B**). In addition, HLA-DR protein expression was detected on knockdown and control MDA-MB-231 cells by FACS. HLA-DR cell surface expression was detected on 20% of the control cells. The percentage of MDA-MB-231 cells expressing HLA-DR was increased to about 50% by both YAP1 knockdown constructs (**Figure 28 D**).

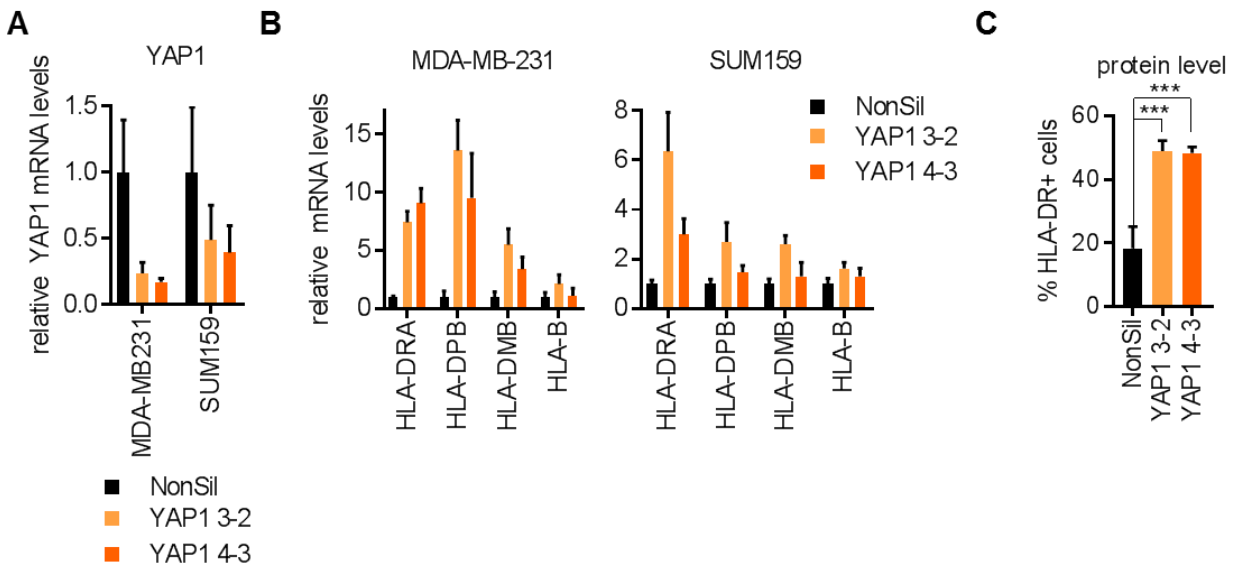


Figure 29 YAP1 expression regulates expression of HLA class II molecules

A Regulation of HLA class II molecules by YAP1 was tested, using YAP1 knockdown MDA-MB-231 or SUM159 cells. shRNA knockdown of YAP1 was generated using the miRE System. Two shRNAs targeting YAP1 (YAP1 3-2 and 4-3) were selected. Knockdown efficiency was confirmed by qRT-PCR analysis of knockdown versus control cells (NonSil) seeded at low density ($0.1x10^6$ cells in a 15 cm culture dish). mRNA levels were normalized against expression of hTBP and are shown relative to the control cells.

B mRNA levels of the HLA class II genes HLA-DRA, HLA-DPB and HLA-DMB as well as the HLA class I gene HLA-B were measured in YAP1 knockdown versus control cells by qRT-PCR. The expression values were normalized against expression of hTBP. mRNA levels are shown relative to the NonSil control cells.

C The increase in HLA-DR expression upon YAP1 knockdown was confirmed on protein level by flow cytometry analysis. HLA-DR cell surface expression was measured by flow cytometry on knockdown and control MDA-MB-231 cells seeded at low cell density 48h after seeding. (n=3) p-values were calculated using One-way ANOVA.

Taken together, antigen presentation mediated by MHC class II molecules seems to be repressed by DCCs from kidney and pancreas compared to cancer cells from the mammary tumor. YAP1 mediated functionw may be responsible for the repression of HLA class II molecules observed in disseminated cancer cells in kidney and pancreas as YAP1 signatures were enriched in DCCs compared to cells from the mammary tumor. Furthermore, increased expression of YAP1 downstream targets at a low cell density was associated with reduced HLA-DR protein expression, whereas reduced YAP1 target gene expression obtained through high cell density or YAP1 knockdown led to increased HLA-DR protein expression.

7.3.4.6 DNA repair mechanisms are upregulated in DCCs in kidney and pancreas

Using immunofluorescence staining of phospho-histone H3 it was confirmed that DCCs located in the kidney of NSG mice were not actively proliferating five weeks after orthotopic tumor injections (**Figure 13**). Paradoxically, the “E2F target” and “G2M checkpoint” gene signatures of the Hallmark dataset were among the strongest enriched signatures in kidney as well as pancreas DCCs when compared to cancer cells from the mammary tumor via GSEA. The “E2F targets” signature was enriched with a NES of 2.05 (FDR=0.0008) and 2.2 (FDR=0.0) in kidney and pancreas DCCs, respectively (**Figure 30 A**). To better characterize the genes responsible for the unexpected enrichment of the “E2F target” signature, GO-term analysis was performed using the core enriched genes in kidney and pancreas DCCs. 95 of the GO-terms that were enriched in the core enriched genes ($p< 0.05$) were associated with negative regulation of cell cycle, cell cycle checkpoint or DNA repair. A list of GO-terms associated with these three categories and the corresponding FDR of their enrichment in DCCs is illustrated in **Figure 30 B**. In comparison, only 17 GO-terms were associated with positive regulation of cell cycle. The complete list with all significantly enriched GO terms is presented in the appendix **Table S 2**. The results suggest that the enrichment of the “E2F targets” signature in DCCs can be explained by the increased expression of genes involved in DNA repair and cell cycle checkpoints.

The transcriptional upregulation of genes involved in DNA repair and cell cycle checkpoints was further investigated in our dataset comparing DCCs from kidney and pancreas with cancer cells from the mammary tumor. In DCCs from kidney and pancreas the HALLMARK signatures “G2M checkpoint” and the “cell cycle checkpoints” signature from Reactome were enriched ($p<0.05$ or FDR 0.25). In addition, the “DNA repair” signatures of the Reactome and GO database, the “DNA synthesis involved in DNA repair” signature from GO as well as several signatures of DNA repair pathways such as “homologous recombination” (KEGG), “Fanconi pathway” (PID and Reactome) and “double strand break repair” (GO) were enriched in DCCs in kidney and pancreas (**Figure 30 C**).

As these results suggest that DCCs, despite their lower proliferation rate, have more active DNA repair machinery than cancer cells from the mammary tumor, the genes responsible for the increase were further analyzed. Genes of the Reactome signature “DNA repair”, core enriched in DCCs were categorized into the four DNA repair pathways homology directed repair (HDR), miss match and nucleotide excision repair (MMR/NER), non-homologous end-joining (NHEJ) and base excision repair (BER). Genes belonging to all four pathways were core enriched in DCCs compared to the mammary tumor. Most of the genes (13) were associated with the HDR pathway. Nine genes were categorized to the MME/NER pathway and further four and two to BER and NHEJ, respectively (**Figure 30 D**).

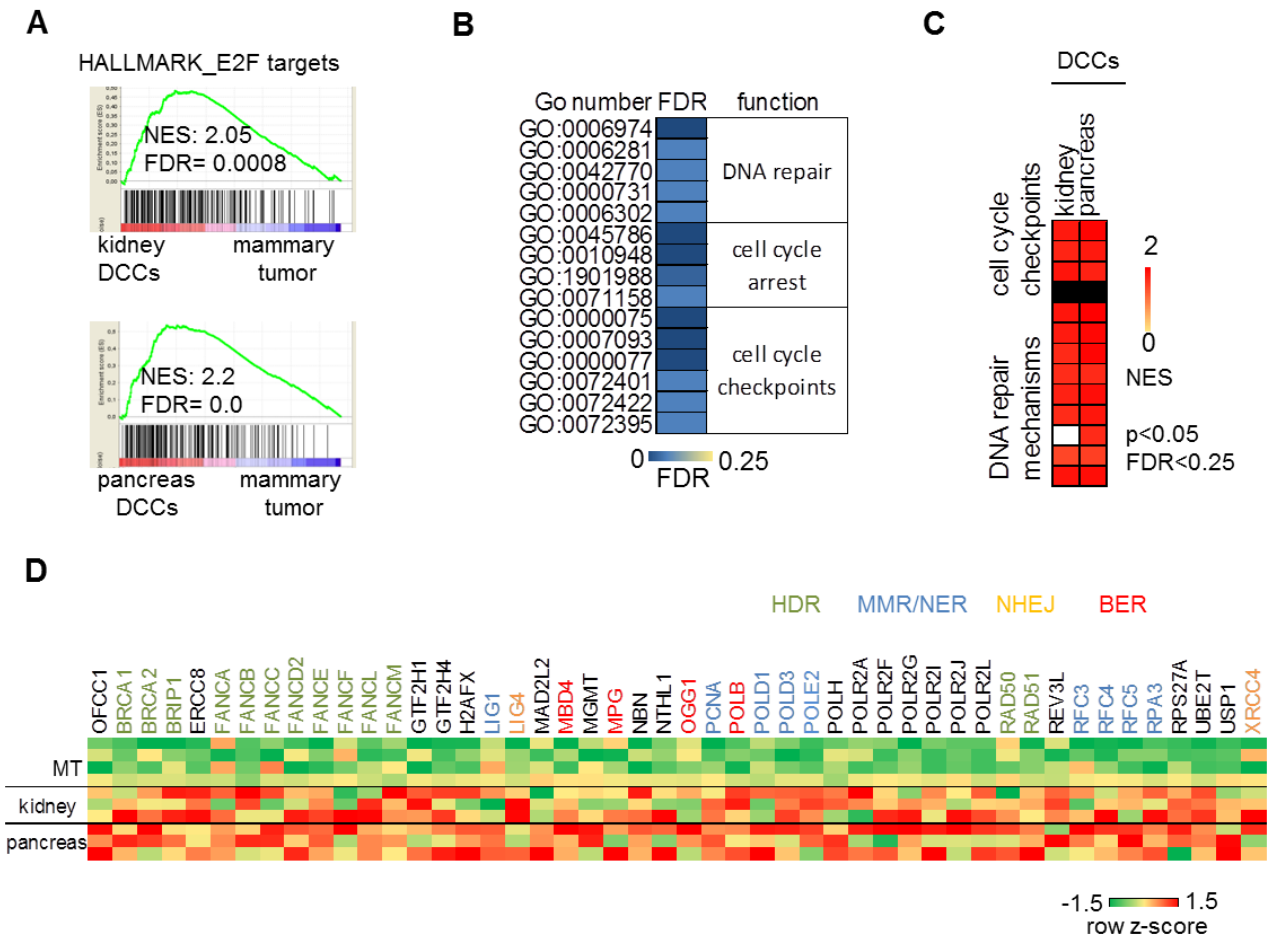


Figure 30 DNA repair mechanisms and cell cycle checkpoints are enriched in DCCs in kidney and pancreas

A Gene set enrichment plot of the HALLMARK "E2F targets" signature in DCCs from kidney and pancreas compared to cells from the mammary tumor. Normalized enrichment scores (NES) and false discovery rates (FDR) are given.

B Gene ontology analysis was performed using the genes of the "E2F targets" signature that were enriched in DCCs (**A)**. A heat map with the FDRs of selected GO-terms associated with DNA repair, cell cycle arrest and cell cycle checkpoint activity are shown. The GO identification numbers of the GO-terms are shown. The list of all GO-terms (p<0.05) is provided in the appendix in **Table S 2**

C Cell cycle checkpoint and DNA damage response signatures are significantly (FDR< 0.25 or p-value<0.05) enriched in DCCs from kidney and pancreas compared to cells from the mammary tumor by GSEA. NES of the enriched signatures are shown in the heat map.

D Heat map of genes from the signature „DNA repair" of the Reactome database that are enriched in DCCs from kidney and pancreas. The genes are color coded according to their function in different DNA repair pathways: green: homology directed repair (HDR), blue: miss match mediated repair (MMR) and nucleotide excision repair (NER), yellow: non-homologous end-joining (NHEJ), red: base excision repair (BER). Black colored genes are not direct members of DNA repair pathways. The row z-scores of the genes are depicted.

In summary, genes associated with DNA damage response and cell cycle checkpoints were increased in disseminated cancer cells in kidney and pancreas compared to the mammary

tumor. The involved genes indicated that especially homology directed repair and nucleotide excision repair may be induced in DCCs.

7.4 Therapy resistance of disseminated cancer cells in kidney and pancreas

In patients, metastasis may not only occur after long latency periods but even after systemic chemotherapy treatment suggesting that a subpopulation of tumor cells can resist therapy. Therefore, we aimed to test whether disseminated cancer cells in kidney and pancreas may resemble such a subpopulation of cancer cells being resistant to chemotherapy. The comprehensive analysis of the gene expression profiles obtained from disseminated MDA-MB-231 cells isolated from kidney and pancreas compared to mammary tumor cells, presented in section **7.3.4,** indicated that several molecular functions possibly involved in therapy response of cancer cells were changed. These functions include repression of apoptosis (**Figure 21**), reduced proliferation (**Figure 13**) and upregulation of DNA repair mechanisms (**Figure 30**) in DCCs from kidney and pancreas.

To further investigate therapy response of DCCs the gene expression profile of DCCs in kidney and pancreas was analyzed for a signature associated with response to Doxorubicin that has been published by Kang *et al.* [230]. They established a gene signature distinguishing doxorubicin resistant gastric cancer cell lines from sensitive ones [230]. In our dataset, this signature was significantly enriched in pancreas DCCs (NES= 1.94; FDR= 0.004) as well as in kidney DCCs (NES= 1.94; FDR= 0.004) (**Figure 31 A**). This suggests that disseminated breast cancer cells located in kidney and pancreas partially have the gene expression profile, acquired by cells able to resist doxorubicin treatment. It was further tested whether signatures associated with therapy response in patients were enriched in the analyzed DCCs. Therefore, we made use of the Balko *et al.* signature, which was associated with relapse and resistance to chemotherapy in a cohort of breast cancer patients that received neoadjuvant chemotherapy [286]. This signature was significantly enriched in DCCs from kidney and pancreas compared to the mammary tumor. In kidney DCCs, the Balko signature was enriched with a NES of 1.1 and an FDR or 0.18 and in pancreas DCCs with a NES of 1.35 and a FDR or 0.009 (**Figure 31 B**).

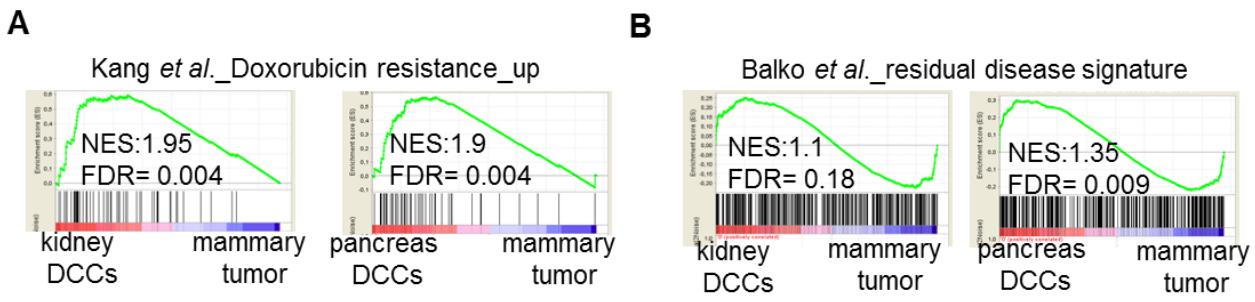


Figure 31 DCCs in kidney and pancreas express genes associated with resistance to chemotherapy

A GSEA was performed testing the kidney and pancreas DCCs datasets for enrichment of the Doxorubicin resistance signature published by Kang *et al.* in 2004 [230]. Enrichment plots are shown including normalized enrichment scores (NES) and false discovery rate (FDR) for kidney and pancreas DCCs, respectively compared to cells from the mammary tumor.

B Enrichment plots showing the enrichment of the residual disease signature published by Balko *et al.* [286] in the DCC datasets. The signature was generated in breast cancer patients that had residual disease after neoadjuvant chemotherapy treatment [286]. NES and FDR for the enrichment of the signature in kidney DCCs and pancreas DCCs compared to the mammary tumor are shown.

The enrichment of gene signatures associated with therapy resistance in experimental models as well as breast cancer patients, suggests that DCCs in kidney and pancreas may represent a subpopulation of intrinsically therapy resistant cancer cells. Therefore, the potential chemotherapy resistance of DCCs from kidney, pancreas and spleen was further tested using functional *in vivo* experiments. MDA-MB-231 cells were injected orthotopically into NSG mice. Eight days after tumor cell injection, mice were injected either with Paclitaxel (20 mg/kg) or a combination of Doxorubicin (1.5 mg/kg) and Cyclophosphamide (50 mg/kg) intraperitoneally. The drug treatment was repeated every fifth day for four treatment cycles. At day 27, the mice were sacrificed and the cancer cell burden in the organs was analyzed by qRT-PCR (**Figure 32 A**).

We followed tumor growth and therapy efficiency *in vivo* by measuring mammary tumor sizes using a caliper. Upon Doxorubicin and Cyclophosphamide (Doxo + Cylco) treatment, the tumor volume was reduced by about 3-fold from a mean of 770 mm^3 in the control group to 280 mm^3 in the treatment group at day 23 after injection. Also the linear growth observed with the controls over three weeks was stopped by treatment with Doxo + Cylco. Treatment with Paclitaxel resulted in a decrease in tumor volume at day 23 after injection to a mean of 520 mm^3, although the growth of the mammary tumors could not be stopped completely (**Figure 32 B**). To analyze the presence of cancer cells in different organs, RNA of whole lung, kidney, pancreas and spleen was isolated, reverse transcribed and qRT-PCR was performed with human (hTBP) and mouse specific (mB2m) primers. To be able to correlate the expression values with the number of tumor cells present in the organs, a spike-in experiment was performed. Nine cell numbers reaching from 25 to 10,000 cancer cells were spiked into organ suspensions of lungs and livers of healthy NSG mice (600,000 mouse cells per sample). qRT-PCR was performed using hTBP and mB2m primers. A standard curve associating the cancer cell numbers to the achieved ΔCT-value of hTBP and mB2M was drawn (**Figure 32 C**).

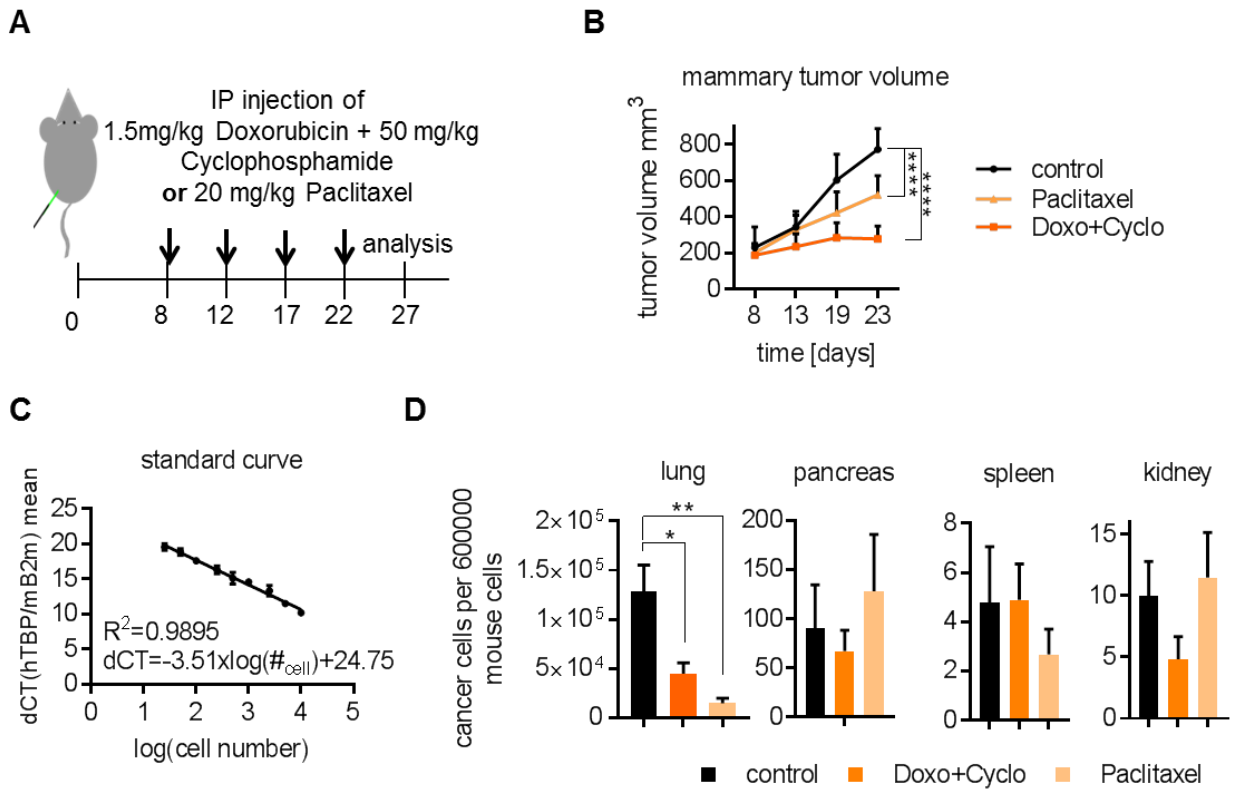


Figure 32 Disseminated cancer cells in different organs are resistant to chemotherapy

A Experimental set-up: MDA-MB-231 cells were injected orthotopically into the 4th mammary fat pad of NSG mice. Eight days after the injection, mice were treated with 1.5 mg/kg Doxorubicin and 50 mg/kg Cyclophosphamide or 20 mg/kg Paclitaxel via intraperitoneal injections. The treatment was repeated every fifth day for four rounds. On day 27 the mice were sacrificed and the organs were analyzed for tumor cell burden.

B *In vivo* growth of mammary tumors was followed over time by measuring tumor sizes every fifth day using a caliper. The volume was calculated using the formula: 3/4*Π*length*width*height. Means are shown with SD, n=6.

C A spike-in experiment was performed to associate the measured ΔCT(hTBP/mB2m) values with the number of cancer cells present in the organs. Therefore, 9 different cell numbers reaching from 25-10,000 cancer cells were spiked into 600,000 mouse cells isolated from liver and lung. qRT-PCR was performed with these samples amplifying hTBP and mB2m. A standard curve was calculated by plotting ΔCT(hTBP/mB2m) values against the logarithmic cell numbers. This standard curve was used to extrapolate the number of tumor cells present in the organs of control or chemo treated mice, which are shown in **D**.

D Comparative assessment of tumor cell burden in lung, kidney, spleen and pancreas of mice treated either with Doxorubicin and Cyclophosphamide (dark orange), Paclitaxel (light orange) or with DMSO as control (black) using qRT-PCR. Human specific primers targeting the TATA-box binding protein (hTBP) were used to detect cancer cells in the organs. The standard curve shown in **C** was used to calculate the number of cancer cells in the organs based on the ΔCT(hTBP/mB2m) values measured by qRT-PCR. Representative results of one of two independent experiments for Paclitacel and one of four experiments for Doxorubicin + Cyclophosphamide with 5 biological replicates each are shown. Means are shown with standard error of the mean. p-values were calculated for all organs comparing control versus treatment groups using t-test and only significant comparisons are marked. * p<0.05; ** p<0.01

Using this standard curve, the number of cancer cells detected in the organs with or without chemotherapy treatment was calculated. Doxo + Cylco treatment led to an almost 3-fold reduction of cancer cells in the lung from 130,000 cells in the control group to 45,000 cells in the treatment group. Metastatic burden was even reduced more efficiently using Paclitaxel, resulting in 15,000 cancer cells in the lung of treated mice (**Figure 32 D**). The reduced tumor growth in combination with the reduction of metastasis burden in the lung confirmed the efficiency of Paclitaxel and Doxo + Cylco treatment. However, despite the successful inhibition of mammary tumor growth and reduction of metastatic cells in the lung, the number of cancer cells in kidney, pancreas and spleen did not significantly change upon chemotherapy. 90 cancer cells per 600,000 mouse cells were detected in the pancreases of control mice. In Doxo + Cyclo treated mice 70 cancer cells were measured and in Paclitaxel treated mice 130. In the spleens of control mice four cancer cells were detected in 600,000 mouse cells. This number was not changed by treatment with Doxo + Cyclo and was slightly reduced to three cells in Paclitaxel treated spleens. Kidneys of control mice harbored ten cancer cells per 600,000 mouse cells. This number was reduced to five cancer cells by Doxo + Cyclo treatment but was not affected by Paclitaxel treatment with 11 cancer cells per 600,000 cells of treated mice (**Figure 32 D**).

These i*n vivo* experiments in combination with gene expression analysis of DCCs show that disseminated MDA-MB-231 cancer cells in kidney, pancreas and spleen are less sensitive to chemotherapy treatment than cells from the mammary tumor as well as metastatic cells from the lung.

7.5 The role of TSPAN8 and TSPAN1 for survival and chemotherapy resistance of DCCs

Based on the findings that disseminated breast cancer cells in kidney, pancreas and spleen are resistant to chemotherapy in animal models, we set out to identify potential molecular mediators of therapy resistance and survival. In this chapter, the results indicating that TSPAN8 and TSPAN1, two members of the tetraspanin (TSPAN) gene family that were upregulated in DCCs (**Figure 33** and **Figure 34**) and expressed in breast cancer patients' samples (**Figure 35**), may contribute to chemo resistance and survival of DCCs, are described. Knockdown of TSPAN8 and TSPAN1 sensitized DCCs to chemotherapy and significantly reduced the number of cancer cells in kidney, pancreas and spleen *in vivo* (**Figure 40**).

7.5.1 TSPAN8 and TSPAN1 are upregulated in DCCs and may play a role for therapy resistance

The molecular cause for the resistance of DCC to chemotherapy treatment was further investigated by comparing the Balko *et al.* gene signature [286] with the genes of at least 1.5-fold increased expression in kidney and pancreas DCCs relative to the mammary tumor. Only nine genes were shared between the Balko *et al.* signature and the DCC signatures. Three of those genes (SCUBE2, NAT1 and AGR2) were part of the Balko *et al.* signature and were specifically increased in kidney DCCs and four additional genes (P2RY2, DPYSL3, SERPINB5, SDR16C5) were shared between the Balko *et al.* signature and the specifically upregulated genes in pancreas DCCs. Only two genes, TSPAN1 and TSPAN8, were common to pancreas and kidney DCCs as well as part of the Balko *et al.* signature (**Figure 33 A**). The TSPAN1 and TSPAN8 genes belong to the Tetraspanin superfamily of transmembrane receptors [287]. Among the upregulated genes in disseminated cancer cells in kidney and pancreas, TSPAN8 was the gene with the highest increase of 11-fold in kidney DCCs and 12-fold in pancreas DCCs compared to the mammary tumor. TSPAN1 expression was 2.5-fold higher in kidney DCCs and 2.1-fold in pancreas DCCs when compared to cancer cells of the mammary tumor (**Figure 33 B**).

The transcriptional upregulation of TSPAN1 and TSPAN8 in kidney and pancreas DCCs was confirmed by qRT-PCR in a second sample set (**Figure 33 C**). DCCs isolated from kidney showed a 5.5-fold upregulation of TSPAN8 and a 7.3-fold of TSPAN1 compared to mammary tumor cells. In pancreas DCCs, TSPAN8 expression was 60-fold upregulated and TSPAN1 1.6-fold. In addition to the organs included in the array analysis, cancer cells isolated from lung, bone and spleen were also integrated into the qRT-PCR experiment. TSPAN8 expression was significantly increased in metastatic cells in the lung (7-fold), while TSPAN1 expression was 1.6-fold increased. Furthermore, the expression of TSPAN8 as well as TSPAN1 in spleen DCCs was increased compared to cancer cells from the mammary tumor 14.4 and 60-fold, respectively. Elevated expression of TSPAN1 and TSPAN8 was not detected in bone DCCs (**Figure 33 C**).

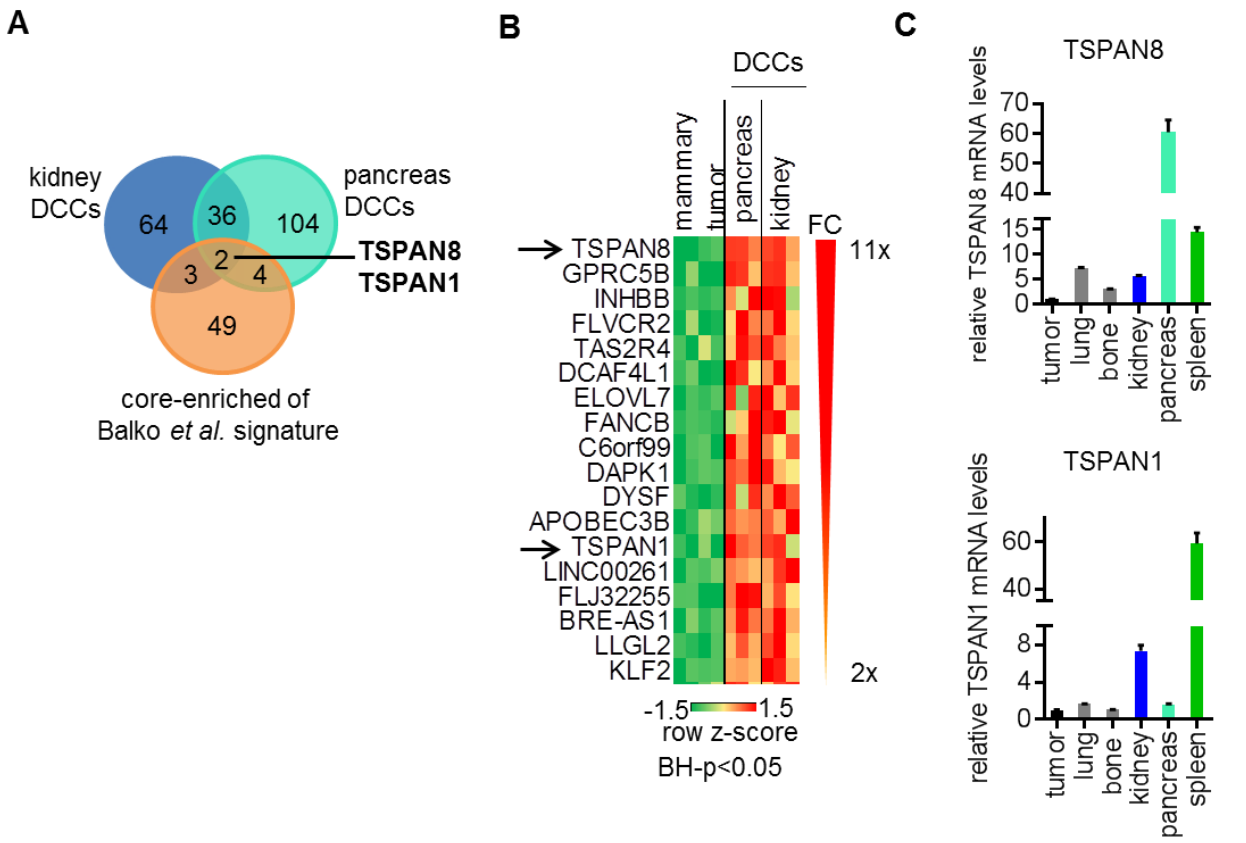


Figure 33 TSPAN8 and TSPAN1 are upregulated in DCCs in kidney and pancreas

A The genes with significant upregulation (FC>1.5; adj. p<0.05) in DCCs in kidney and pancreas, respectively were compared to the genes of the Balko *et al.* signature [286] that were core enriched in DCCs compared to the mammary tumor. The number of overlapping genes among these three groups is shown. Only two genes, TSPAN1 and TSPAN8, are shared between all three gene sets.

B List of significantly upregulated genes in DCCs from kidney and pancreas compared to the mammary tumor with a fold increase of at least 2-fold (BH p-value< 0.05). Genes are ranked according to their fold change (FC). Heat map of the row z-scores is shown. Members of the Tetraspanin gene family (TSPAN) are marked with arrows.

D Validation of the upregulation of TSPAN8 and TSPAN1 in independent sample sets of sorted and immediately analyzed cancer cells from different organs by qRT-PCR. Gene expression was normalized to the expression of human TATA-box binding protein (hTBP). Expression values are shown relative to the expression in the mammary tumor (MT), n=6

As tetraspanins are membrane bound proteins, it was of interest to study whether the transcriptional increase in TSPAN1 and TSPAN8, detected in DCCs in kidney and pancreas compared to cancer cells from the mammary tumor, was also translated into increased protein level at the plasma membrane. Five weeks after cancer cell injection into NSG mice, organ homogenates were prepared and cancer cells in the mammary tumor, lung, bone, kidney, pancreas and spleen were detected by flow cytometry using the gating strategy shown in **Figure 9**. The protein level of TSPAN1 and TSPAN8 at the membrane of cancer cells was measured in six biological replicates using specific antibodies. Due to very low tumor burden in spleens, three biological replicates contained less than 100 cancer cells and were therefore excluded from

analysis. Representative FACS plots of CD298 and GFP+ cancer cells isolated from mammary tumor or pancreas and stained for TSPAN1 (y-axis) and TSPAN8 (x-axis) are shown in **Figure 34 A**. Depending on the TSPAN protein expression, cancer cells were categorized into four subgroups: cancer cells that were double negative for both TSPAN1 and TSPAN8, cancer cells positive for either TSPAN1 (T1) or TSPAN8 (T8) and cancer cells that expressed both TSPAN1 and TSPAN8 (T1/ 8) (**Figure 34 A**). Background was subtracted using an unlabeled control. The specificity of the antibodies was tested using TSPAN8 and TSPAN1 knockdown cells.

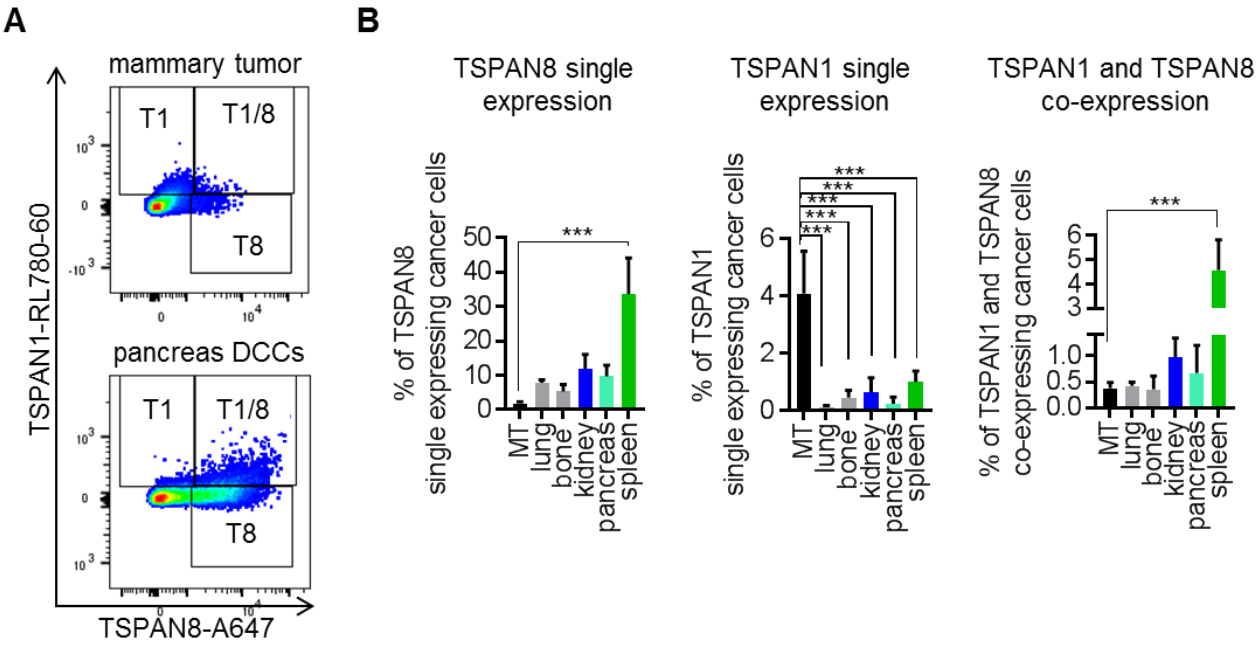


Figure 34 The proportion of TSPAN8 and TSPAN1 double positive cancer cells is increased in DCCs

The upregulation of TSPAN8 and TSPAN1 on MDA-MB-231 cancer cells in kidney and pancreas relative to cancer cells from the mammary tumor was confirmed on protein level via flow cytometry. Cancer cells were detected in organ homogenates five weeks after orthotopic injections based on the gating strategy shown in **Figure 9**. TSPAN1 and TSPAN8 expression on cancer cells was measured using specific antibodies.

A Representative FACS plots of TSPAN staining on mammary tumor cells (upper panel) and cancer cells from pancreas (lower panel). TSPAN8 was detected using a specific antibody coupled to Alexa647 (A647) and TSPAN1 using a specific antibody coupled to Alexa750 (RL780-60). Three cancer cell populations were quantified: cells expressing only TSPAN1 (T1) or TSPAN8 (T8) and cells expressing both TSPAN proteins (T1/ 8).

B Percentages of the three cancer cell populations defined in **A** (n (spleen)=3; n(other organs)=6). Means and standard deviations are shown. P-values were calculated using One-way Anova. Only significant comparisons are marked. * p<0.05; *** p<0.0001; MT: mammary tumor

In the mammary tumor only a subpopulation of cancer cells expressed TSPAN1 (4%) and TSPAN8 (2%) with very few cells (0.34%) expressing both proteins. The percentage of TSPAN8 single expressing cancer cells was increased to 7.8% in the lung, 5% in the bone, 9% in the

pancreas, 12% in the kidney and 33% in the spleen (**Figure 34 B** left panel). Surprisingly, the percentage of cancer cells expressing TSPAN1 alone dropped from 4% in the mammary tumor to 0.1% in the lung, 0.45% in the bone, 0.24% in the pancreas, 0.65% in the kidney and 1% in the spleen (**Figure 34 B** middle panel). The transcriptional increase of TSPAN1 shown in **Figure 33** was translated into an increase in TSPAN8 and TSPAN1 double expressing DCCs in kidney, pancreas and spleen. The mean percentage of TSPAN8 and TSPAN1 expressing cells was doubled in pancreas DCCs, three-fold higher in kidney DCCs and even 12-fold higher in spleen DCCs compared to cells from the mammary tumor. No difference in the percentage of TSPAN8 and TSPAN1 double expressing cancer cells was measured in cancer cells in lung and bone compared to the mammary tumor (**Figure 34 B** right panel). These results suggest that a higher percentage of cancer cells in kidney, pancreas and spleen expressed TSPAN8 and TSPAN1 than cancer cells in the mammary tumor or the lung. Furthermore, individual DCCs also seemed to express higher levels of TSPAN8 and TSPAN1 as the label intensity on DCCs was higher than on cell from the mammary tumor.

To address a potential clinical association, TSPAN1 and TSPAN8 gene and protein expression was measured in pleural effusion and ascites samples of breast cancer patients. During the course of the disease, fluids can accumulate between the pleural membrane and the lung (pleural effusion) or in the abdomen (ascites) of breast cancer patients [288]. These fluids contain cancer cells that can be isolated via *in vitro* cultivation. The expression of TSPAN1 and TSPAN8 in cancer cells isolated from ascites and pleural effusion samples was determined via qRT-PCR using gene specific primers. In total five patient samples, two pleural effusion (BPE) and three ascites (BA) samples were analyzed (**Figure 35 A**). Expression of TSPAN8 and TSPAN1 was normalized to the expression of human TATA-box binding protein (hTBP). The cancer cell lines MDA-MB-231 and SUM159 were used as reference. TSPAN8 expression in MDA-MB-231 cells was detected at a CT-value of 28 and in SUM159 at a CT-value of 25.5. The TSPAN8 expression in patient samples was distributed between CT-values of 23.6 and 27.4 (**Figure 35 A** right panel). TSPAN1 expression in MDA-MB-231 cells was detected at a CT-value of 26 and in SUM159 at cycle 25. In patient samples, expression of TSPAN1 was lower and was detected between CT-values of 27.8 and 29.5 (**Figure 35 A** left panel). The qRT-PCR results showed that patient derived cancer cells expressed TSPAN1 and TSPAN8 mRNA. The level of TSPAN8 expression was within the range of cell lines, whereas TSPAN1 expression was detected about two cycles later by qRT-PCR.

Protein levels of TSPAN1 and TSPAN8 were also measured in four patient samples using flow cytometry (**Figure 35 B**). The gating strategy introduced in **Figure 34** was applied to identify TSPAN8 and TSPAN1 single expressing cell populations as well as TSPAN8 and TSPAN1 co-expressing cells. The cell lines MDA-MB-231 and SUM159 were again used as reference. 60.5% MDA-MB-231 cells were single positive for TSPAN8 and 2.57% positive for TSPAN1 and TSPAN8. The percentage of TSPAN1 single expressing cells was below the background with 0.17%. In SUM159 cells, 26.3% of the cells were staining positive for TSPAN8 and 3% for both, TSPAN8 and TSPAN1. In these samples, no TSPAN1 single expressing cells were detected. All four tested patient samples showed protein expression of TSPAN8, which varied between 16.7% in BPE30 to 48% in BPE28. Patient derived samples contained almost no TSPAN1 single positive cells. In the samples BPE28, BA18-2 and BA19-2, TSPAN8 and TSPAN1 co-expressing cells were detected with 2.1%, 2.025% and 0.9% probability, respectively. Only BPE30 did not contain double positive cells (**Figure 35 B**). These data suggest that TSPAN8 and TSPAN1 expressing cancer cells were present in some patients with metastatic breast cancer.

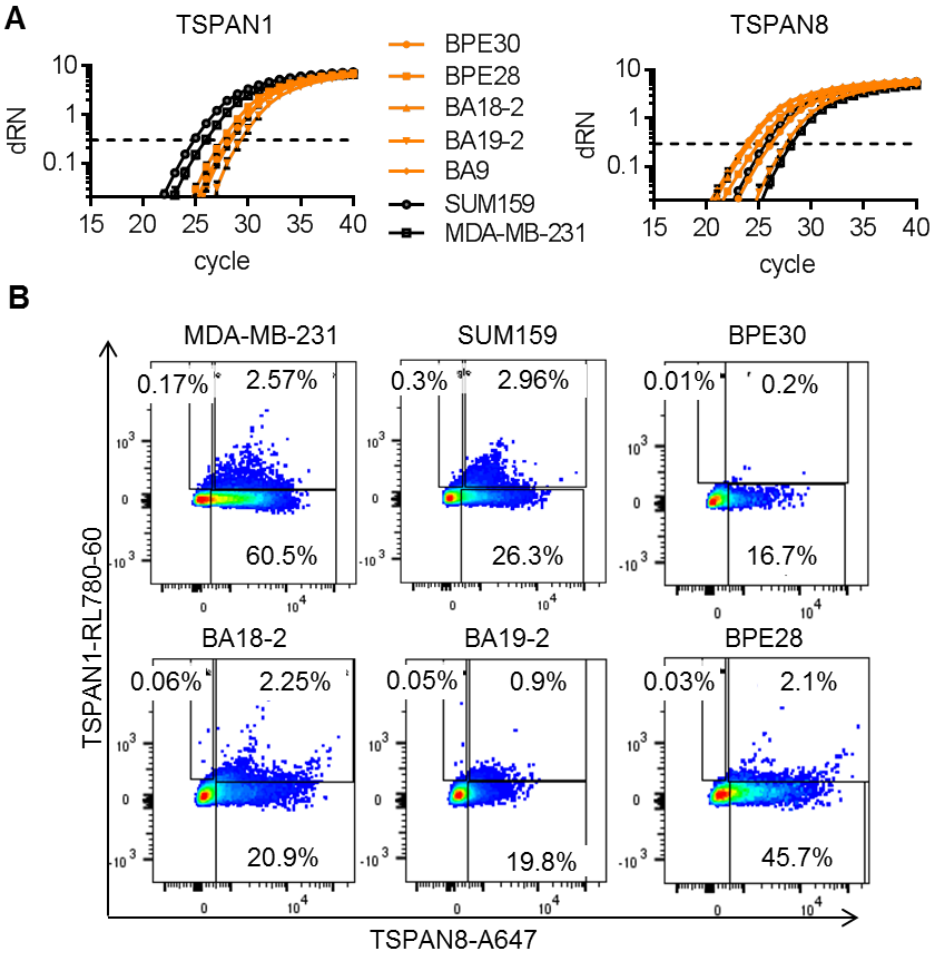


Figure 35 TSPAN8 and TSPAN1 are expressed in cancer cells of metastatic breast cancer patients

A TSPAN1 and TSPAN8 gene expression was measured in cancer cells isolated from pleural effusions (BPE) and ascites (BA) samples from breast cancer patients. TSPAN1 and TSPAN8 gene expression in patient samples was detected by qRT-PCR using gene specific human primers. Five different patient samples were analyzed and MDA-MB-231 and SUM159 cells were used as references. dRN, the normalized and background subtracted reporter signal, was plotted against the amplification cycle. The dotted line represents the threshold, at which fluorescence signals was above background. The patient derived cells are shown in orange, the cell lines in black.

B TSPAN1 and TSPAN8 protein expression was measured on BPE and BA samples of breast cancer patients. Flow cytometry was performed on four of the five patient samples shown in **A** using specific antibodies targeting TSPAN8 and TSPAN1. Cancer cells were analyzed for single or co-expression of TSPAN1 and TSPAN8 after dead cell exclusion using DAPI. Percentages of TSPAN8 and TSPAN1 single and double expressing cells of the DAPI negative cell population are shown.

The results demonstrated in this chapter show that disseminated cancer cells located in kidney, spleen and pancreas upregulated TSPAN1 and TSPAN8 at the mRNA as well as the protein level. Especially the subpopulations of TSPAN8 single and TSPAN8 and TSPAN1 double expressing cancer cells were increased in DCCs, whereas TSPAN1 single expressing cells were decreased. The upregulation of TSPAN8 and TSPAN1 double expressing cells was restricted to DCCs in kidney, pancreas and spleen, while TSPAN8 single expressing cells were also increased in lung and bone compared to the mammary tumor. Furthermore, we confirmed

expression of TSPAN8 and TSPAN1 in primary breast cancer cells using pleural effusion and ascites samples from breast cancer patients.

7.5.2 TSPAN8 and TSPAN1 play a role in maintaining DCCs in pancreas and spleen

The functional importance of the Tetraspanin genes for disseminated cancer cells was investigated using TSPAN1 and TSPAN8 knockdown cells. The knockdowns were generated in MDA-MB-231 breast cancer cells using the miRE shRNA system. Two independent shRNAs targeting TSPAN1 (TSPAN1 2-1 and TSPAN1 6-4) or TSPAN8 (TSPAN8 3-3 and TSPAN8 4-4) were selected from the shERWOOD database for each gene to account for off-target effects [257]. Double knockdown cells were transduced with TSPAN1 2-1 and TSPAN8 3-3 knockdown constructs or TSPAN1 6-4 and TSPAN8 4-4 knockdown constructs, respectively. shRNAs targeting the Renilla luciferase (Ren713) gene were used as non-silencing control (NonSil). Knockdown efficiency was measured by qRT-PCR on transcription level and by flow cytometry on protein level (**Figure 36** and appendix **Figure S 1** and **2**). For the qRT-PCR two independent primer pairs were used for both genes and expression was normalized against expression of the TATA-box binding protein. The Tetraspanin expression in uninfected MDA-MB-231 cells was measured to ensure that the NonSil control construct had no influence on TSPAN gene expression. In MDA-MB-231 cells, the mRNA level of TSPAN8 was decreased by 97% with both shRNAs compared to the control construct (appendix **Figure S 1**). The TSPAN1 expression level was reduced by 83% using the TSPAN1 2-1 shRNA and by 73% using TSPAN1 6-4. The results for both primer pairs per gene were well comparable and the NonSil construct did not significantly change TSPAN1 or TSPAN8 expression (appendix **Figure S 1** and **2**). Therefore, the TSPAN1 and TSPAN8 expression in the double knockdown cells was only evaluated using one primer pair per gene. In the MDA-MB-231 cells transduced with TSPAN8 3-3 and TSPAN1 2-1 constructs the TSPAN8 expression was reduced by 96% and the TSPAN1 expression by 85% compared to control cells transduced twice with the NonSil control construct. The TSPAN8 4-4 and TSPAN1 6-4 constructs led to a reduction of TSPAN8 transcription of 93% and of TSPAN1 of 70% (**Figure 36 A**). The reduction of TSPAN1 and TSPAN8 protein levels was measured by flow cytometry using specific antibodies targeting either TSPAN1 or TSPAN8. Doublets were excluded based on their forward scatter properties, dead cells via DAPI staining and the percentage of TSPAN8 and TSPAN1 positive cells in control and knockdown MDA-MB-231 cells was quantified. Representative FACS plots indicating the proportion of TSPAN8 and TSPAN1 positive cells in control and knockdown cells are shown in **Figure 36 B**. TSPAN8 protein expression was strongly reduced by 97% using the TSPAN8 4-4 shRNA and by 61% using TSPAN8 3-3 shRNA. Also TSPAN1 protein expression was decreased by 94% using TSPAN1 6-4 shRNA. Thus, analysis of mRNA and protein level confirmed the significant downregulation of TSPAN1 and TSPAN8 genes and proteins in MDA-MB-231 cells using the miRE shRNA system.

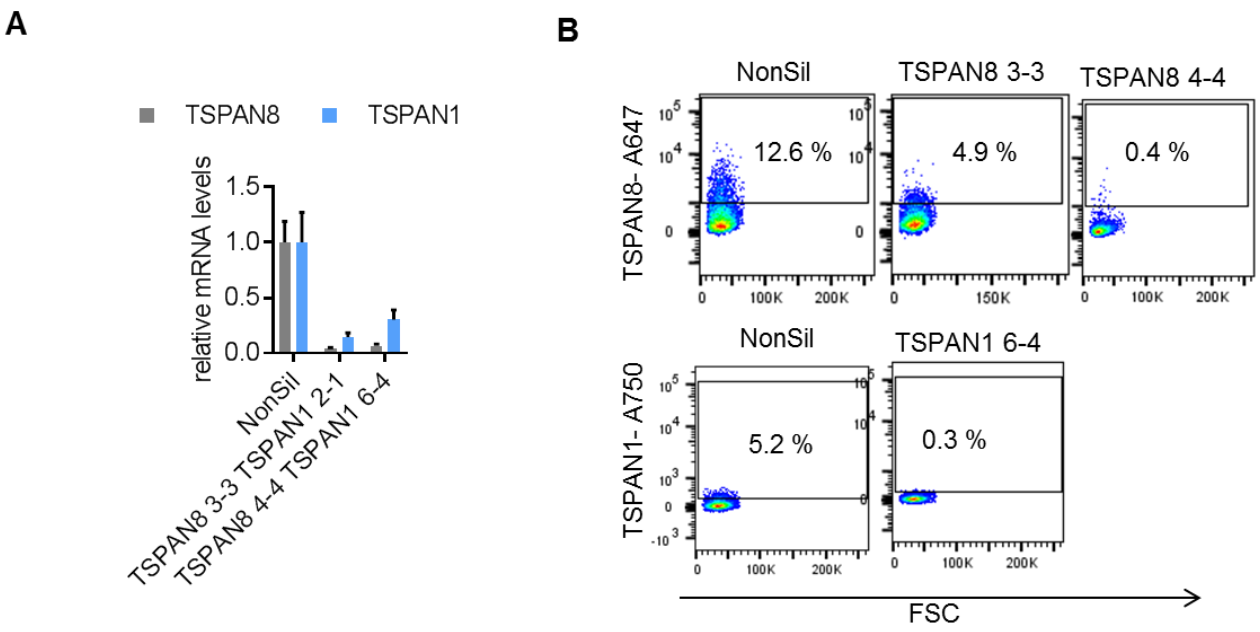


Figure 36 Confirmation of TSPAN8 and TSPAN1 knockdown in MDA-MB-231 breast cancer cells

Knockdowns of TSPAN1 and TSPAN8 in MDA-MB-231 breast cancer cells were generated using the miRE shRNA backbone. Two shRNAs targeting either TSPAN8 (TSPAN8 3-3 and TSPAN8 4-4) or TSPAN1 (TSPAN1 2-1 and TSPAN1 6-4) were selected from the shERWOOD database [256] and were cloned in lentiviral vectors carrying a fluorescent protein and an antibiotic resistance gene. A vector carrying a shRNA targeting the Ren713 gene was used as non-silencing control (NonSil).

A Knockdown efficiency of TSPAN8 and TSPAN1 was measured on mRNA level by qRT-PCR. Expression was normalized against expression of the TATA-box binding protein. TSPAN1 and TSPAN8 mRNA levels are shown relative to the mRNA levels measured in control infected MDA-MB-231 cells.

B Knockdown of TSPAN1 and TSPAN8 was in addition validated on protein level by flow cytometry using specific antibodies targeting TSPAN8 and TSPAN1. Percentages of TSPAN8 and TSPAN1 positive cells were calculated after doublet and dead cell exclusion. Representative flow cytometry plots are shown for TSPAN1 and TSPAN8 knockdown and control MDA-MB-231 cells.

The knockdown cells were used to assess the functional importance of TSPAN1 and TSPAN8 for mammary tumor growth, lung metastasis and DCCs of MDA-MB-231 TGL cells *in vivo*. Control cells or knockdown cells were injected orthotopically into the 4th mammary fat pad of NSG mice. Five weeks after injection, mammary tumor volume was measured *ex vivo* using a caliper and tumor load of the organs was determined by *ex vivo* bioluminescence measurement. Six mice were analyzed per group. Neither TSPAN8 nor TSPAN1 single knockdown did impact mammary tumor growth (**appendix Figure S 1** and **2**). The tumor burden of the organs was determined by ex vivo bioluminescence measurements. The luminescence signal detected in organs harboring TSPAN8 or TSPAN1 knockdown cells was not significantly different from the signal obtained in organs with control cells (**appendix Figure S 1** and **2**). These results suggest that neither TSPAN8 nor TSPAN1 knockdown alone do influence mammary tumor growth, lung metastasis or maintenance of DCCs in the used *in vivo* model.

Following up on the observed increase in TSPAN1 and TSPAN8 co-expressing cancer cells in kidney, pancreas and spleen compared to the mammary tumor (**Figure 34**), we also tested the

in vivo effect of combined knockdown of TSPAN1 and TSPAN8 in MDA-MB-231 cells. Double knockdown or control cells were injected orthotopically into NSG mice. Two different combinations of shRNAs targeting TSPAN1 and TSPAN8 were used and six biological replicates were analyzed per group. Five weeks after injections, the mice were sacrificed and tumor load in the organs was determined by *ex vivo* luminescence. Mammary tumor growth was assessed by tumor weight. Combined knockdown of TSPAN8 and TSPAN1 in breast cancer cells led to a minor reduction of mammary tumor growth. The tumor weight was slightly reduced from a mean of 0.59 g in the control group to 0.3 g and 0.44 g in the knockdown groups, respectively (**Figure 37 A**). Knockdown of TSPAN8 and TSPAN1 did not affect lung metastasis as the bioluminescence signal in lungs from animals injected with knockdown or control cancer cells was comparable. However, the luminescence signal measured in spleen and pancreas of animals injected with TSPAN8 3-3 TSPAN1 2-1 knockdown cells was significantly reduced compared to control organs and showed a trend using the TSPAN8 4-4 TSPAN1 6-4 constructs (**Figure 37 B**). The luminescence signal measured in kidneys of animals injected with the knockdown constructs was not changed compared to control cells (**Figure 37 B**).

Taken together, our results suggest that TSPAN1 or TSPAN8 alone do not influence metastatic growth or dissemination of MDA-MB-231 cells in orthotypic *in vivo* models. Depletion of both molecules however resulted in reduced cancer cell numbers in pancreas and spleen.

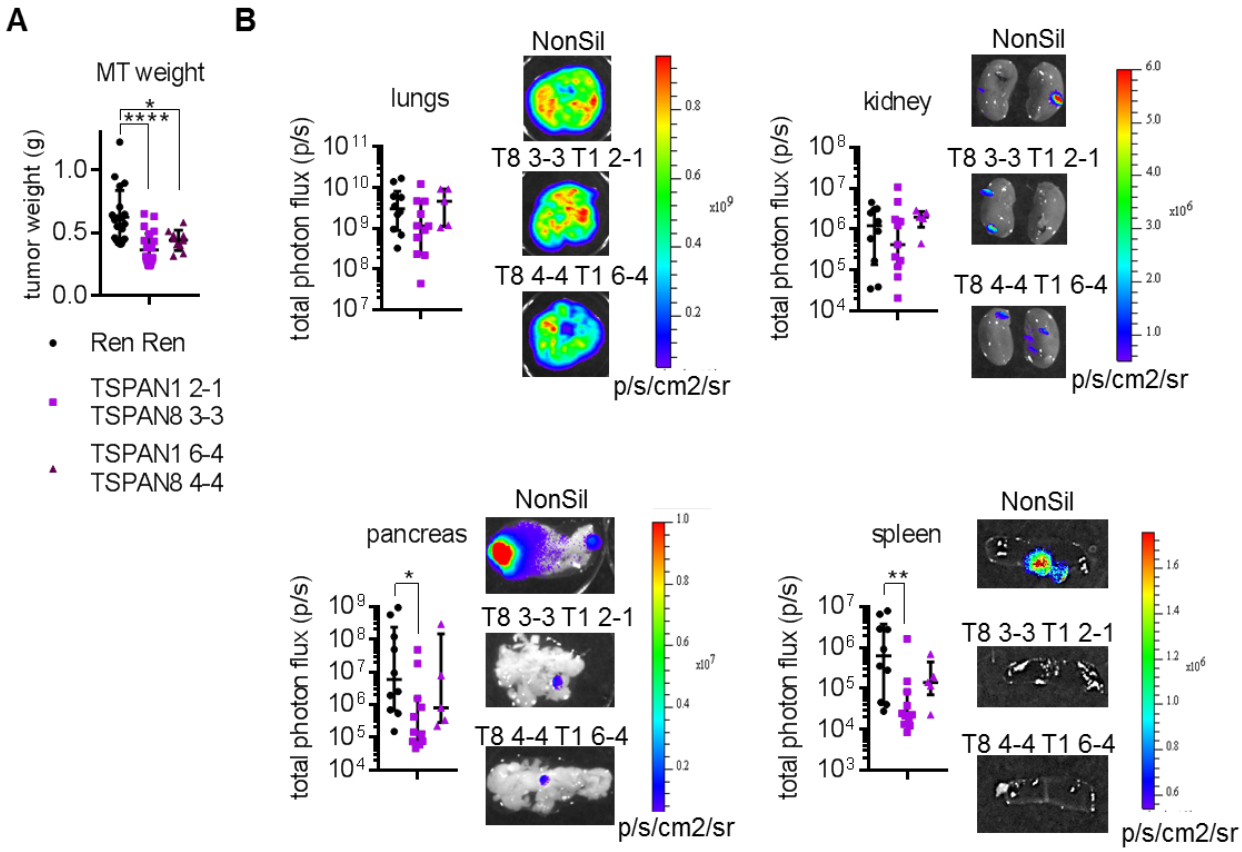


Figure 37 Combined knockdown of TSPAN8 and TSPAN1 affects DCCs in spleen and pancreas

MDA-MB-231 TGL cells expressing a control construct (NonSil) or one of two combinations of TSPAN8 and TSPAN1 targeting shRNAs, TSPAN8 3-3 with TSPAN1 2-1 or TSPAN8 4-4 and TSPAN1 6-4, were injected into the 4th mammary fat pad of NSG mice. Five weeks after injection, mammary tumor weight was determined *ex vivo* and tumor burden in lung, kidney, spleen and pancreas was measured via bioluminescence *ex vivo*. The results for control and TSPAN8 3-3 TSPAN1 2-1 group were generated in two independent experiments with five biological replicates each. Five biological replicates were analyzed for TSPAN8 4-4 TSPAN1 6-4.

A tumor weight in gram [g] per mammary tumors (MT) is shown. All replicates and the median values are shown with interquartile range. P-values were calculated using Kruskal Wallis test. ** p<0.005

B Cancer cell burden in lung, pancreas, kidney and spleen was analyzed by *ex vivo* bioluminescence imaging of the organs. Mice were injected with luciferin and incubated for 10 min. Mice were sacrificed, organs were resected and bioluminescence was measured using an IVIS live imaging device. The total photon flux [p/s] for all replicates is shown with the median per group and the interquartile range of all replicates. p-values were calculated comparing each knockdown construct to control cells using One-way ANOVA and are only shown for significant comparisons. * p-values< 0.05; ** p-value< 0.01. Representative bioluminescence pictures of the organs injected with cancer cells expressing control (NonSil) or TSPAN knockdown constructs are shown.

7.5.3 TSPAN8 and TSPAN1 are associated with quiescence and survival in breast cancer cells

TSPAN8 as well as TSPAN1 have been reported in the context of normal stem cell maintenance and functions using different systems [51, 289]. TSPAN8 has been linked to stemness in the mammary gland where it marks a population of quiescent stem cells [51] and TSPAN1 has been found to be expressed on adult pluripotent stem cells in planarians [289]. With these findings in mind, it was tested whether TSPAN8 and TSPAN1 may also be involved in regulation of stemness in disseminated cancer cells. To answer this question we made use of sphere cultures, an *in vitro* culture method at low adhesion and without serum that has been shown to enrich for stem cell properties [290] (**Figure 38 A**).

MDA-MB-231 cells were cultured under sphere conditions for 7 days. Spheres were harvested and TSPAN8 and TSPAN1 expression was measured by qRT-PCR. TSPAN8 and TSPAN1 mRNA levels were 2.5- and 2-fold higher, respectively in spheres compared to adherent MDA-MB-231 cells (**Figure 38 B**). To further analyze the role of TSPANs in stemness we performed a label retaining experiment. MDA-MB-231 cells were labeled with the membrane dye PKH26 and were seeded under sphere culture conditions. 100% PKH26 expression immediately after labelling was confirmed by flow cytometry analysis. Cells with high PKH26 expression (label retaining) did not divide during culture, resembling a quiescent population, whereas cells that lost PKH26 expression (label non-retaining) did proliferate. Spheres were harvested 5 days after labeling and were stained for TSPAN8 and TSPAN1 expression. The gating strategy is shown in **Figure 38 C.** First, doublets were excluded based on their FSC and SSC gated and live cells were detected by DAPI staining. Next, live cells were either gated for TSPAN8 and TSPAN1 expression (from now on referred to as TSPAN bulk) or label retaining and non-retaining populations based on their PKH26 labeling. The PKH26+ and PKH26- populations were further analyzed based on their TSPAN expression (referred to as TSPAN of PKH26+ or PKH26-, respectively (**Figure 38 C**).The percentage of TSPAN1/8 bulk cells that were also PKH26+ or PKH26- were then calculated. As the PKH26+ population did not account for exactly 50% of live cells, the TSPAN1/8 of PKH26- population was corrected for the difference between the PKH26+ and PKH26- populations. The same calculation was performed for the TSPAN8 single population. It was observed that the majority of TSPAN8 and TSPAN1 co-expressing cells were also positive for PKH26 (**Figure 38 D**). In the mean about 70% of TSPAN1 and TSPAN8 double positive cells were within the label retaining compartment. Only about 20% of TSPAN double positive cells did proliferate in sphere cultures. However, TSPAN8 single positive cells were equally distributed within the PKH26+ and PKH26- fractions (**Figure 38 D**). These results suggest that TSPAN8 and TSPAN1 co-expression marks a quiescent cell population.

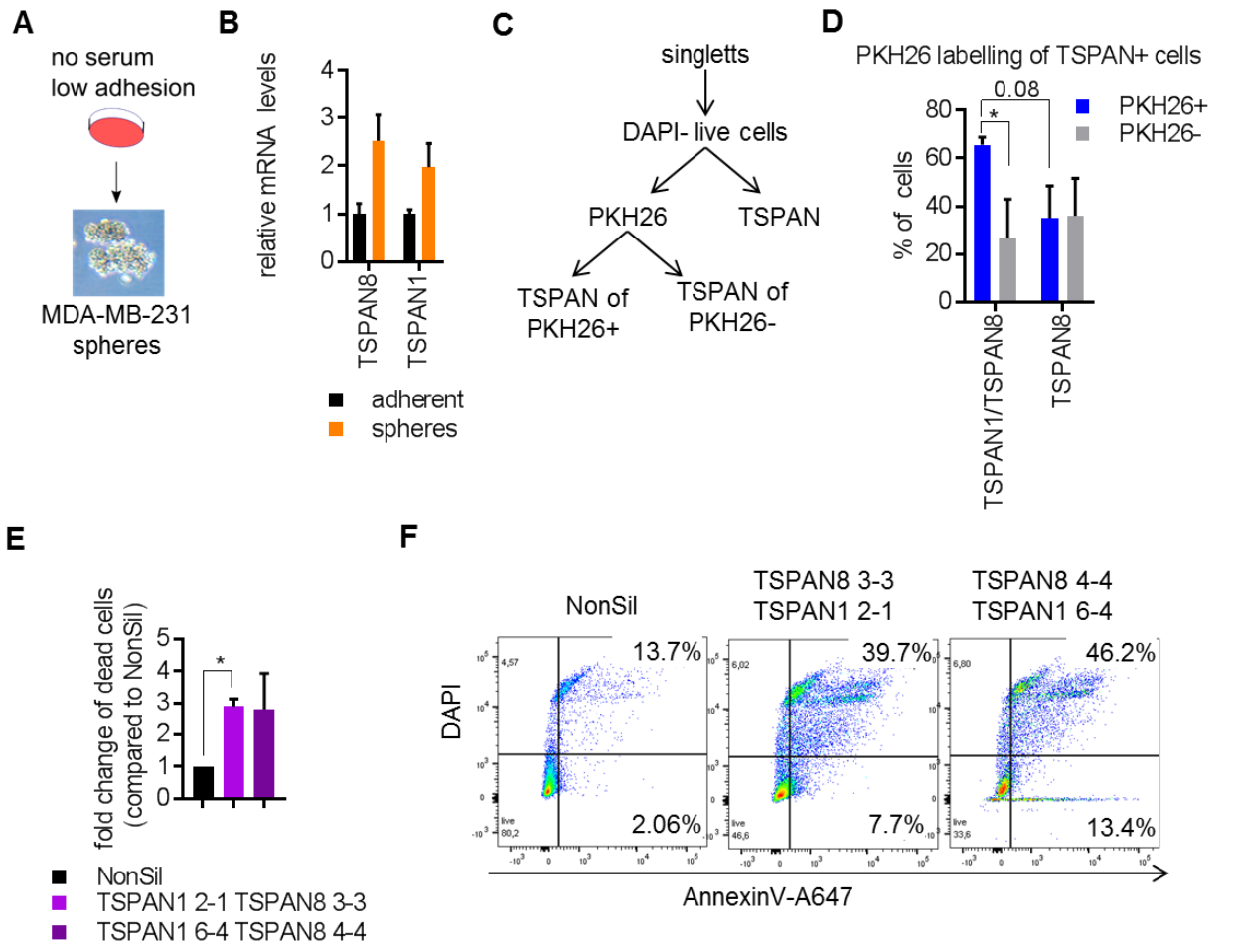


Figure 38 TSPAN8 and TSPAN1 mark a label retaining population in sphere and are essential for survival

A MDA-MB-231 cells were cultured at low adhesion, without serum and with addition of growth factors for 7 days to enrich for stem cell properties, as it has previously been established [290]. A representative image of spheres taken at 10x magnification is shown.

B TSPAN8 and TSPAN1 gene expression was measured in 7 day spheres compared to MDA-MB-231 cells cultured under adherent conditions for 48h. mRNA levels were measured via qRT-PCR using human TATA-box binding protein as housekeeping gene. The results are presented for one of two independent experiments with similar results. The mean for three technical replicates is shown with standard deviation.

C MDA-MB-231 cells were labeled with the membrane dye PKH26 immediately before being seeded under sphere culture conditions. Five days after seeding, spheres were harvested and stained for TSPAN8 and TSPAN1 by FACS. The gating strategy that has been used to determine the percentage of TSPAN8/1 or TSPAN8 cells that are either PKH26+ or PKH26-, is shown.

D Quantification of the PKH26+ (label retaining) and PKH26- (label non-retaining) fractions of TSPAN8/1 double or TPAN8 single expressing cells. As the PKH26+ and PKH26- fractions were not exactly 50% of live cells, the percentages of “TSPAN of PKH26- cells” were corrected for this difference. Means and standard deviations of three independent experiments are shown. P-values were calculated using unpaired, two-sited t-test. * $p<0.05$

E+F Cell death rates in control and TSPAN8 and TSPAN1 knockdown MDA-MB-231 spheres were measured by Annexin V and DAPI staining by flow cytometry. Spheres were analyzed 7 days after seeding. **C** shows the quantification of two independent experiments. The means with standard deviation are shown. P-values were calculated using One-way Anova comparing each of the knockdown construct individually to the control cells. Results are only shown for significant comparisons. * $p<0.05$. **D** shows representative FACS plots of control and TSPAN knockdown MDA-MB-231 cells gated for Annexin V and DAPI staining.

Furthermore, it was tested whether TSPAN knockdown affected survival of cancer cells under sphere conditions. Therefore, TSPAN knockdown and control MDA-MB-231 were cultured under sphere culture conditions and stained for Annexin V and DAPI to assess the cell death rate. Knockdown of TSPAN8 and TSPAN1 resulted in significantly more cell death under sphere conditions. The cell death rate was 3-fold higher in knockdown MDA-MB-231 spheres compared to control cells (**Figure 38 D+E**).

Taken together our data suggest that TSPAN8 and TSPAN1 mark a quiescent cell population in spheres. Upon knockdown of the two molecules, cell death was increased under sphere conditions.

7.5.4 TSPAN8 and TSPAN1 regulate therapy resistance in DCCs

Following up on the *in vitro* results suggesting an association of TSPAN8 and TSPAN1 with stem cell properties and survival under stress, it was further investigated whether this mechanism may also be in place in DCCs in kidney and pancreas. Thus, control or TSPAN8 and TSPAN1 knockdown cells were injected into the mammary fat pad of NSG mice. Five weeks after injection, mice were sacrificed and MDA-MB-231 cancer cells, located in the pancreas, were isolated from organ homogenates using FACS. RNA of the sorted control and knockdown cancer cells was prepared and gene expression profiling was performed using Affymetrix Human Genome U133 plus2.0 arrays (**Figure 39 A**). Raw data were normalized by RMA normalization and was used for GSEA comparing both knockdown constructs (TSPAN8 3-3 TSPAN1 2-1 and TSPAN8 4-4 TSPAN1 6-4) to control cells.

First, we investigated if the knockdown of TSPAN8 and TSPAN1 affected therapy resistance of DCCs. The Balko *et al.* residual disease signature, which was associated with residual disease and relapse of breast cancer patients treated with neoadjuvant therapy [286] was used. This signature was enriched in DCCs from kidney and pancreas compared to cancer cells from the mammary tumor and its overlap with the DCC signatures led to the discovery of TSPAN8 and TSPAN 1 (**Figure 31** and **Figure 33**). Knockdown of TSPAN8 and TSPAN1 resulted in significant repression of the Balko *et al.* signature (**Figure 39 B**). This data suggests that TSPAN8 and TSPAN1 may essentially contribute to therapy resistance of DCCs, which can be impaired by TSPAN knockdown.

Based on the *in vitro* results demonstrating that TSPAN8 and TSPAN1 are enriched in the label retaining fraction of spheres and are involved in survival of cancer cells under sphere conditions, we tested the enrichment of stem cell signatures in TSPAN knockdown DCCs. Several published stem cell signatures were significantly repressed in TSPAN8 and TSPAN1 knockdown DCCs compared to control cells (**Figure 39 C**). These signatures included a signature of murine mammary stem cells compared to progenitor cells [291] as well as a human pan adult stem cell signature that has been compiled by Smith *et al.* by ranking several stem cells signatures from different epithelial origins. This signature was used to predict outcome in epithelial cancers [292]. In addition, the transcriptional profile of normal quiescent human mammary stem cells published by Pece *et al.* [293] was tested. Enrichment plots of these signatures in the TSPAN knockdown

data set are shown in **Figure 39 D**. The repression of stem cell signatures in TSPAN knockdown cells are in line with the results obtained in sphere cultures (**Figure 38**).

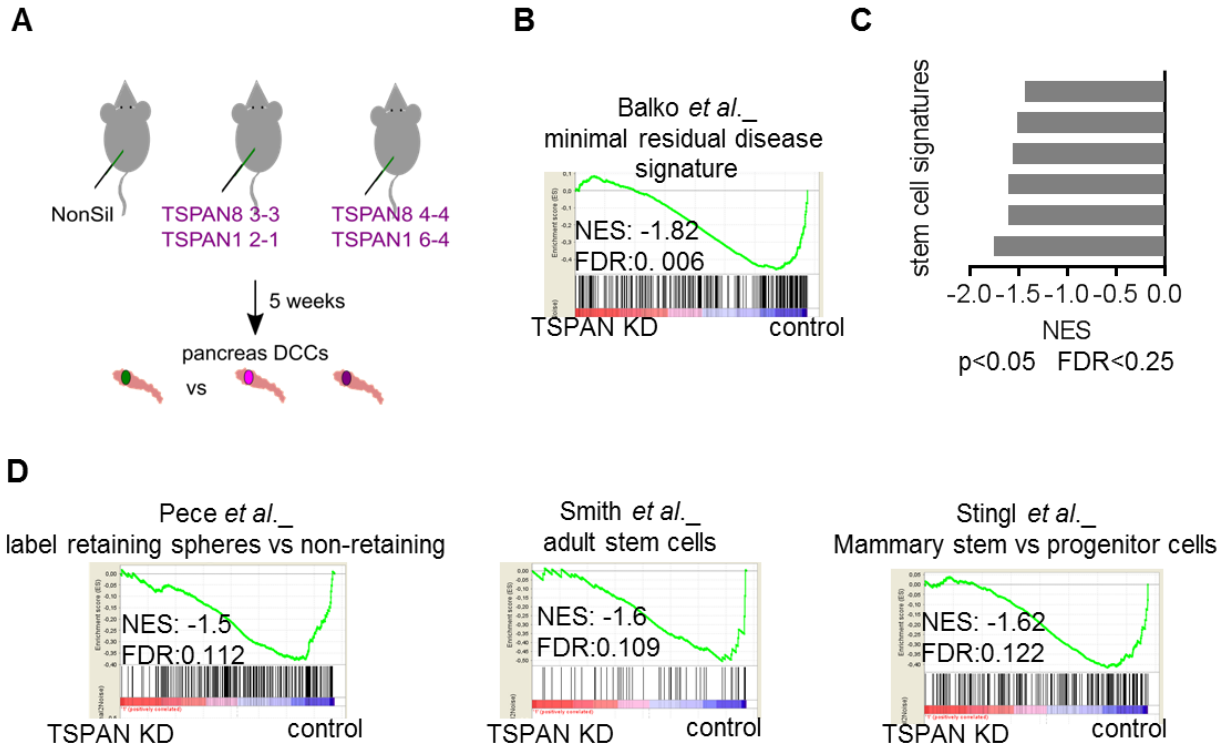


Figure 39 TSPAN8 and TSPAN1 regulate therapy resistance and stemness in DCCs

A Experimental set-up: MDA-MB-231 cells carrying a control construct or one of two TSPAN8 and TSPAN1 knockdown constructs (TSPAN8 3-3 TSPAN1 2-1 or TSPAN8 4-4 TSPAN1 6-4) were injected orthotopically into NSG mice. 5 weeks after injection, the mice were sacrificed, and the pancreas was resected. Cancer cells were isolated from the pancreas by organ homogenization and FACS. RNA was prepared from sorted cells and the gene expression profile of knockdown versus control cells was analyzed on Affymetrix human genome U133 plus2.0 arrays.

B The RMA normalized raw data generated as described in **A** was analyzed using Gene Set Enrichment Analysis (GSEA). Both knockdown sequences were jointly compared to control cells. The enrichment plot of the Balko *et al.* residual disease signature, generated in breast cancer patients treated with neoadjuvant chemotherapy [286], is shown in the TSPAN knockdown dataset. Normalized Enrichment Score (NES) and false discovery rate (FDR) are given.

C Several stem cells gene signatures, were underrepresented in TSPAN knockdown cells (p<0.05 and FDR<0.25). A list with all used signatures is presented in **Table S 1** of the appendix.

D Representative enrichment plots of three of the stem cell signatures used in **C**. left: Stingl *et al.* compared the transcription profile of murine mammary stem cells with progenitor cells [291]; middle: Smith *et al.* compiled a human " pan adult stem cell signature" by combining several stem cell signatures from different epithelial origins in a ranked manner [292]; right: Pece *et al.* generated transcription profiles of human mammary stem cells by comparing label retaining cells in mammospheres to label non-retaining cells [293]. Normalized enrichment scores (NES) and false discovery rates (FDR) are shown for all comparisons.

7.5.5 TSPAN8 and TSPAN1 mediate chemotherapy resistance in DCCs *in vivo*

The results thus far showed that combined knockdown of TSPAN1 and TSPAN8 reduces mammary tumor size as well as cancer cell burden in pancreas and spleen (**Figure 37**). Furthermore, TSPAN1 and TSPAN8 may play a role in chemotherapy resistance of disseminated cancer cells based on gene expression analysis of TSPAN8 and TSPAN1 knockdown DCCs (**Figure 39**). Therefore, the effect of combined knockdown of both tetraspanin molecules on chemotherapy sensitivity was tested *in vivo*.

TSPAN8 and TSPAN1 knockdown or control MDA-MB-231 cells were injected orthotopically into the mammary fat pad of NSG mice. Eight days after cancer cells injection, mice were injected IP with a combination of Doxorubicin (1.5 mg/kg) and Cyclophosphamide (50 mg/kg). As shown in **Figure 32**, the chemotherapeutic agents were injected every fifth day for four cycles. Three groups were included into the experiment: MDA-MB-231 cells expressing the control construct with DMSO control treatment (group 1), control cells treated with chemotherapy (group 2) and TSPAN8 3-3 TSPAN1 2-1 expressing knockdown cancer cells treated with Doxorubicin and Cyclophosphamide (group 3) (**Figure 40 A**). Mammary tumor size was measured with a caliper and organs were harvested for RNA extraction. Results from two independent experiments with nine and six mice in groups 1 and 2, respectively and 6 mice in group 3, each were pooled. Efficacy of Doxorubicin and Cyclophosphamide treatment was confirmed by reduced mammary tumor growth as well as lung metastasis of control cells treated with chemotherapy compared to DMSO treatment. The tumor volume in mice that have been treated with Doxorubicin and Cyclophosphamide was reduced to 42% of the volume measured in DMSO control treated mice (**Figure 40 B**). The observed reduction in tumor size was comparable to the results shown in **Figure 32**. Combination of TSPAN1 and TSPAN8 knockdown with chemotherapy resulted in a slight additional reduction of tumor growth compared to chemotherapy treatment alone. The lung metastatic burden was measured by qRT-PCR in whole organ mRNA using luciferase specific primers to ensure the effectiveness of the chemotherapy treatment on metastatic cells. Doxo + Cyclo treatment reduced the metastatic burden in the lung significantly by about 70%, confirming the efficiency of the chemotherapy treatment in this experiment. (**Figure 40 C**). Combination of TSPAN knockdown and chemotherapy resulted in a moderate increase of tumor cells in the lungs compared to chemotherapy alone. However, compared to the control group, the increase was only minor and the tumor burden in the lung was significantly reduced by the combination therapy.

The effect of combined TSPAN1 and TSPAN8 knockdown on disseminated cancer cells in pancreas, spleen and kidney was analyzed via qRT-PCR using mouse beta-2- microglobulin as reference gene and primers detecting the luciferase gene, expressed in the TGL vector (luc+), for the identification of cancer cells. The impact of tetraspanin deficiency on chemo sensitivity of cancer cells was investigated comparing luc+ levels between mice injected with knockdown cells and treated with chemotherapy and mice only treated with chemotherapy. Luc+ levels were analyzed for each organ individually and normalized against the group with control cells and chemotherapy treatment. **Figure 40 D** shows the mean fold changes of luc+ mRNA levels with the combination treatment compared to chemotherapy alone in two independent experiments with 5-9 biological replicates each. Combined knockdown of TSPAN8 and TSPAN1 in cancer cells led to a significant sensitization of cancer cells in kidney, pancreas and spleen compared to chemotherapy treatment with doxorubicin and cyclophosphamide. Combination of chemotherapy with the knockdown of TSPAN8 and TSPAN1 resulted in a 90%, 60% and 66% reduction in the

tumor burden in kidney, pancreas and spleen, respectively compared to chemotherapy alone (**Figure 40 D**).

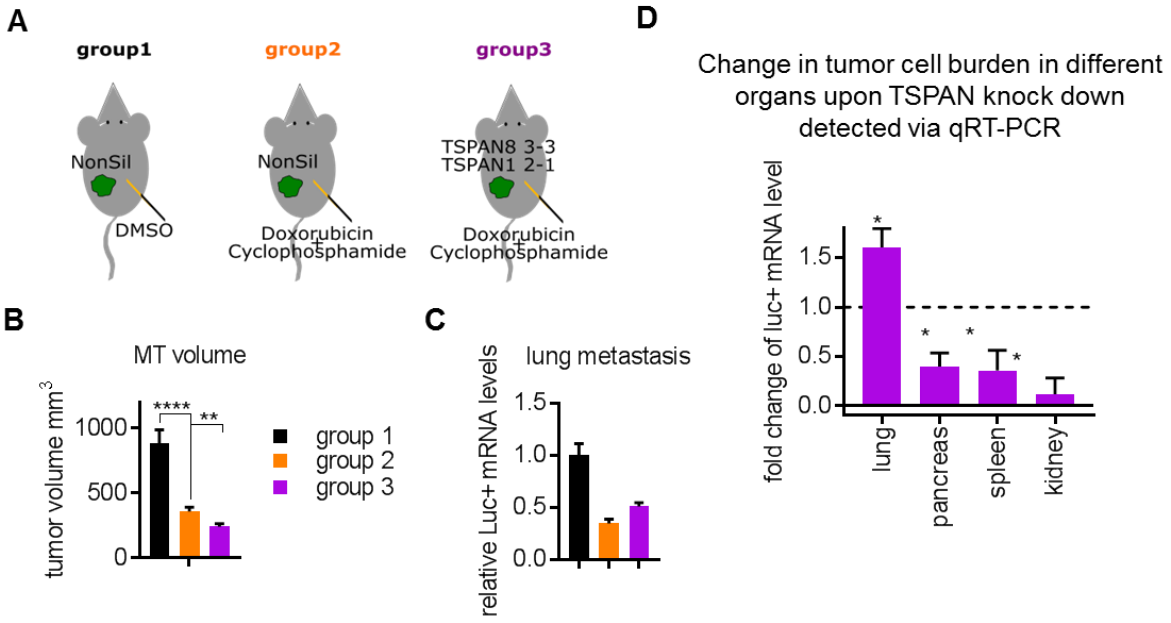


Figure 40 Combined knockdown of TSPAN8 and TSPAN1 sensitizes DCCs to chemotherapy treatment *in vivo*

MDA-MB-231 cells expressing either control constructs (NonSil) or the shRNAs TSPAN8 3-3 and TSPAN1 2-1, which target TSPAN8 and TSPAN1 gene expression, respectively, where injected into the mammary fat pad of NSG mice. Eight days after cancer cells injection, treatment with 1.5 mg/kg Doxorubicin and 50 mg/kg Cyclophosphamide was started. Chemotherapy injections were performed IP every five days for four cycles. The experimental scheme is shown in **Figure 32**. Results are shown for two independent experiments with a total of 15 biological replicates in the control groups and 10 in the TSPAN8 3-3 TSPAN1 2-1 knockdown group.

A Three groups were used for the experiment. Group 1: MDA-MB-231 cells expressing the NonSil control vector were injected and mice were treated with DMSO as control. Group2: NonSil expressing MDA-MB-231 control cells were injected and mice were treated with Doxorubicin and Cyclophosphamide (Doxo+Cyclo). Group 3: MDA-MB-231 cells expressing TSPAN8 3-3 and TSPAN1 2-1 shRNA constructs were injected and mice were treated with Doxo+Cyclo

B Tumor volumes were measured *ex vivo* using a caliper. Means of all replicates are shown with standard error of the mean (SEM). p-value was calculated using t-test. **** p-value<0.0005.

C Tumor burden in the lung was determined by qRT-PCR using luciferase (luc+) specific primers and was normalized against mouse specific beta-2-microglobulin (mB2m) expression. Luc+ level relative to group1 of control cells with DMSO treatment are shown.

D Cancer cell burden in lung, pancreas, spleen and kidney was measured by qRT-PCR using specific primers targeting the luc+ gene. Luc+ expression was normalized to expression of mB2m. Luc+ mRNA levels were normalized for each organ individually to group 2, control cells that have been treated with chemotherapy. Results are shown as mean of the two independent experiments with standard deviations. P-values were calculated comparing for each organ group 3 versus group 2 using unpaired two-sited t-test. * p-value< 0.05

These results indicate that TSPAN1 and TSPAN8 are jointly important for chemotherapy resistance of DCCs in different organs. Depletion of TSPAN1 and TSPAN8 significantly sensitized DCCs to chemotherapy *in vivo*. Metastatic cells in the lung, which were sensitive to chemotherapy in the first place were not affected by TSPAN8 and TSPAN1 knockdown.

8 Discussion

8.1 Dissemination of breast cancer cells

This work demonstrates that breast cancer cells disseminate to multiple different organs in orthotopic *in vivo* models (**Figure 8, Figure 9, Figure 10**). Using orthotopic breast cancer models, disseminated cancer cells were detected in organs that frequently harbor metastasis in human patients [197] – namely lung, liver and bone. In addition, cancer cells were also detected in pancreas, spleen and kidney, organs that do not support metastatic growth of breast cancer cells. In 2006, Suzuki *et al.* analyzed the dissemination of several MDA-MB435 clones with different metastatic potential. Using a highly metastatic clone, they frequently detected disseminated cancer cells in lung, lymph node, spleen, liver and kidney [294]. With the lowly metastatic cells, DCCs were only detected in lung and lymph node [294]. However, in this study wide spread dissemination was not restricted to highly aggressive cancer cell lines but was also confirmed in less aggressive cells as well as in patient derived xenografts (**Figure 10**). The lack of detection of rare DCCs by Suzuki *et al.* may be due to less sensitive approaches used. In their study, Suzuki *et al.* primarily used microscopy for detection of cancer cells in different organs and *in vitro* antibiotic selection for isolation [294]. In addition to those methods, highly sensitive qRT-PCR and flow cytometry were used for detection of cancer cells in the presented work. These methods allow detection of even single cells in organ homogenates. The low percentage of cancer cells detected e.g. in spleen and pancreas (**Figure 9**) may be missed with less sensitive detection methods.

Next to sensitivity issues, cell intrinsic characteristics of the used cell lines may affect the results. Breast cancer is not a single disease but can be further divided into subtypes depending mainly on the hormone receptor status of the cells. As it has been shown that the molecular subtype influences organ tropism of breast cancer metastasis [246-248] it cannot be excluded that the subtype also affects dissemination to and survival in pancreas, spleen and kidney. In this study, primarily triple negative breast cancer cell lines and PDX models were primarily used as this subtype is associated with the worst prognosis due to lack of targeted therapies [245, 253]. In addition, luminal cell lines and PDX models were used as less aggressive counterparts. Luminal breast cancers are less aggressive and have often long latency periods [244, 245]. Detection of DCCs even in this subtype emphasizes the generality of the process in breast cancer as well as its importance. Suzuki *et al*. performed their experiments in MDA-MB435, which is a Her2+ positive cell line. As Her2+ cancer cells were not tested in this study, it is also possible that the intrinsic breast cancer subtype influences the dissemination pattern.

Besides the molecular subtype, technical aspects such as prolonged culturing may change the gene expression profile of DCCs, thus influencing functional readout. To eliminate this potential influence, freshly isolated DCCs were analyzed in this study. To our knowledge it was shown for the first time in the present work that freshly isolated DCCs from organs that are non-proliferative at the time point of isolation maintain their growth potential and are capable of initiating metastasis. In previous studies, reaching the same conclusion [294, 295], DCCs were cultured for an extended period of time to generate stable cell lines before assessing growth potential. Culturing on plastic and in ideal conditions with serum and growth factors, may lead to gene expression changes affecting DCC biology. In the present study, DCCs isolated via FACS from kidney, pancreas and spleen were immediately injected intravenously into healthy NSG mice.

This approach was chosen as gene expression changes such as TSPAN8 downregulation were observed upon prolonged culture (data not shown). Furthermore, comparison of the gene expression profile of immediately analyzed cancer cells with cells that have been cultured for only 48h confirmed the impact of cultivation on the gene expression profile of DCCs (**Figure 18** and **Figure 19**). In this comparison, a 50-60% overlap of the expression profile was observed; confirming that cultivation of DCCs may affect their phenotypic landscape. Certain biological characteristics of DCCs are stable in short term culture. Nevertheless, some features such as the metabolic rewiring is weakened upon *in vitro* cultivation. This reduction of metabolism related GO term may be due to the culture conditions including 2% serum and growth factor supplements. Thus, only injection of freshly isolated cells allows unbiased analysis of their growth potential.

Due to the novelty of the approach, the gained insights in the biology of DCCs in kidney, pancreas and spleen as well as a possible impact on clinical practice will be put in context and discussed in the next paragraphs.

8.1.1 DCCs´ close contact to blood vessels

As an early step of the study it was important to investigate the location of DCCs within the organs as colonialization is the rate limiting step of the metastatic cascade and only a subpopulation of disseminated cells is able to reach the organ parenchyma and survive there. Thus the location of DCCs provides valuable information on the biology of the cells. Perfusion experiments and Immunofluorescence imaging of DCCs in the kidney confirmed extravasation and localization within the kidney parenchyma for the majority of cancer cells (**Figure 11** and **Figure 12**). Interestingly, the majority of extravasated cancer cells were still located in close proximity to blood vessels (**Figure 12**). These results are in line with findings by Ghaja *et al.* Malladi *et al*.. Ghaja *et al.* have demonstrated that breast cancer DCCs in brain, lung and bone marrow reside on micro-vessels in these microenvironments [193]. They revealed that stable perivasculature promotes DCC dormancy via secretion of thrombospondin 1 [193]. Malladi *et al* have demonstrated that latency competent cells (LCCs) in breast and lung cancer reside in distant organs in close proximity to blood vessels [296]. The authors identified a subpopulation of LCCs that have stem cell characteristics and enter a quiescent state that enables them to evade immune responses and to survive extended periods [296]. The presented data suggest that DCCs in kidney, pancreas and spleen are also dormant as they do not proliferate (**Figure 13, Figure 30**) but can exit this reversible state when being relocated to the lung (**Figure 15**).

Endothelial cells, lining vessel walls, have not only been shown to regulate dormancy and proliferation, it has also been demonstrated that endothelial cells play a role in chemotherapy resistance. It has been shown that direct contact of cancer cells to endothelial cells in the bone marrow via integrin mediated adhesion promotes chemotherapy resistance of breast cancer cells [234]. Due to the close proximity of kidney DCCs to blood vessels, it cannot be excluded that the microenvironment also plays a role in therapy resistance of DCCs in kidney, pancreas and spleen. Further investigations are necessary to analyze if and how endothelial cells in kidney and pancreas influence DCC survival, growth inhibition and therapy resistance.

In addition to the supportive roles of endothelial cells for cancer cell dormancy and therapy resistance, the close proximity to vessels may be beneficial for nutrient and oxygen supply of DCCs as starvation has been linked to apoptosis [175-178]. In line with the assumption that contact to blood vessels gives DCCs access to oxygen, the hypoxia response, which is the cellular response to low oxygen levels is reduced in DCCs from kidney and pancreas compared to cells from the mammary tumor (**Figure 26**). Only recently Kumar *et al.* have reported that the distance to blood vessels influences the phenotype of glioma cells. They have demonstrated that cancer cells in direct contact with blood vessels grow under normoxic conditions leading to increased oxidative phosphorylation and chemotherapy resistance [297]. As these features match some of the characteristics of DCCs, it is possible that the close proximity to blood vessels may partially drive the gene expression changes observed in DCCs and may be partially responsible for therapy resistance of DCCs.

8.1.2 DCCs in kidney and pancreas are in a state of dormancy

Following the result that DCCs in kidney, pancreas and spleen had efficiently extravasated, we wanted to learn more about their phenotype.

The transcription profiles of DCCs from kidney and pancreas were compared to either lung metastasis or cancer cells from the mammary tumor. These comparisons showed that DCCs were clearly distinct from cancer cells from the mammary tumor (**Figure 17**) as well as from actively growing metastatic cells (**Figure 20**). Analysis of the GO term enriched in the downregulated genes in DCCs in both comparisons was performed to gain deeper insight into the changed molecular functions. Interestingly, we observed an overlap between the pathways downregulated in DCCs compared to the mammary tumor as well as downregulated in DCC in comparison to lung metastasis (**Figure 19** and **Figure 20**). In one of the gene expression profiles (**Figure 20**) lung metastasis and mammary tumor cells were analyzed in the same experiment, allowing direct comparison of the similarities of metastasis and mammary tumor compared to DCCs. Principal component analysis showed that the expression profile of lung metastases was even more distinct from DCCs than the primary tumor. A possible explanation for the similarity of the mammary tumor and lung metastasis in the PCA and the functional overlap is that lung metastases as well as the mammary tumor were proliferative at the time point of analysis whereas disseminated tumor cells in kidney and pancreas were not. This assumption is strengthened by the experimental setup used. The *in vivo* experiments were performed five weeks after cancer cell injection, which represents a late stage of the disease with tumors of about 0.7-0.9 cm of diameter. As tumor growth often exhibits an exponential growth phase followed by a plateau, it was hypothesized that the tumors in our analysis already reached this plateau, resulting in reduced proliferation [298]. This assumption is supported by phospho-Histone H3 (**Figure 13**) and cleaved Caspase 3 staining (**Figure 21**), detecting less proliferative and apoptotic cells in the mammary tumor than in lung metastasis. The slower growth rate of the mammary tumor may explain the less pronounced transcriptomic differences in DCCs compared to the mammary tumor than compared to the lung.

Several results, presented in this work, support the view of DCCs as subpopulation during a metastatic latency period. These results include that DCCs were exravasated (**Figure 11, Figure 12**), excluding that they represent cells in circulation, and that they did not proliferate (**Figure 13, Figure 30**), excluding that they represent fully metastatic cells. The last assumption was also confirmed by the transcriptional difference between DCCs and lung metastasis, elaborated above. Furthermore, the capacity for regrowth (**Figure 15**) and the resistance to chemotherapy (**Figure 32**), which has been demonstrated, are well established characteristics of dormant, latent cancer cells [173]. Regarding the regrowth potential, previous studies have shown that dormant bone marrow DCCs can be reawakened by transplantation into irradiated bones, which resemble a permissive microenvironment [14]. In our study, translocation of dormant DCCs from kidney, pancreas or spleen into the lung, which permits growth of MDA-MB-231 cells [15, 215], also resulted in metastatic growth of DCCs. This result stresses the context dependency of growth arrest in kidney, pancreas and spleen and the impact of the microenvironment. The resistance of chemotherapy, which is known for dormant cells [173] and has been demonstrated for DCCs is most likely linked to the lack of proliferation of DCCs [232] in combination with molecular characteristics of DCCs such as reduced cell death (**Figure 21**) and upregulation of the Tetraspanin molecules (**Figure 33**).

Molecularly, metastatic latency can be caused by cues from the microenvironment resulting in dormancy at a colony level [62, 149] as well as intrinsic signals of cancer cells themselves, restricting proliferation at a cellular level [174, 181]. Dormancy on a colony level implies that cell extrinsic mechanisms such as lack of angiogenesis or the immune system restrict the metastatic growth [175]. Cancer cells, kept dormant by these means, would proliferate at a pace comparable to metastatic cells and would undergo apoptosis due to the microenvironmental restrictions. The reduced proliferation rate (**Figure 13**) as well as apoptosis rate (**Figure 21**) of DCCs compared to lung metastasis render this possibility less likely and suggests that dormancy of DCCs is mediated by cellular quiescence However, further investigations on DCCs in pancreas and spleen are needed to address the potential contribution of a balance between proliferation and apoptosis.

In conclusion, DCCs in kidney, pancreas and spleen represent a non-proliferative, latent population of cancer cells in an intermediate state of the metastatic cascade. The collected data suggest that a large majority of DCCs are kept dormant on a cellular level rather than by a balance of proliferation and apoptosis.

8.1.3 Clinical evidence for widespread cancer dissemination

In breast cancer, it is well known that the disease may undergo latency [19, 171]. The results of our study showing that DCCs in kidney, pancreas and spleen may resemble latent cancer cells raised the question of the clinical translatability of the findings. Direct proof for the existence of DCCs in kidney, pancreas and spleen of breast cancer patients is difficult to achieve as no tissue material of theses organs is collected from breast cancer patients. Nevertheless, indirect evidence may support the notion that widespread dissemination can occur in breast cancer patients.

Multiple case report studies have been published of organ donations from cured cancer patients to healthy recipients. Although there were no signs of cancer in the donor, recipient nevertheless developed metastases later in life [299-304]. In most of these cases donors were successfully treated and were counted as cured for many years [303]. In most cases the transplanted organ belonged to the organ tropism of the cancer of origin. However, there are examples of patients that carried cancer cells in organs that are not prone to harbor metastasis [302, 304] demonstrating that wide spread dissemination can also occur in patients. These cases are mostly reported for kidneys as this organ is frequently transplanted. Most case studies report development of metastasis in the donated organ itself [302]. This metastatic growth is most likely related to the immune suppression of the recipients as interruption of immune repression led to regression of the cancer in several cases [303]. Some recipients developed additional metastases in secondary organs [302, 304]. These examples suggest that DCCs can also start migrating and initiate growth at distant sites. However, from these reports it is not clear if migration is initiated after induction of growth or if dormant DCCs can also translocate to other more permissive organs without growth in the transplanted organ. This question can be answered in mouse models by organ transplantations of kidney, pancreas or spleen with DCCs into healthy mice. If the DCCs in the organ of origin remain quiescent but lung, liver or bone macro-metastasis are established in the recipients, dormant DCCs can migrate through the body.

Alternatively, photo-convertible reporter constructs can be used [305]. In our set-up, photo-conversion would be activated locally in the kidney by laser application leading to a change of reporter fluorescence emission from green to red. Detection of red cancer cells in blood, lung or liver would indicate that dormant DCCs are able to translocate through the body. The applicability of this system has recently been demonstrated by Pereira *et al.*. In their study photo-conversion was performed in lymph nodes and cancer cells were detected in the lung [116]. Although we successfully demonstrated that DCCs, translocated to the lung, can induce metastasis (**Figure 15**), these experiments are indispensable to demonstrate that dormant DCCs in pancreas, spleen and kidney can be sources for breast cancer lung, liver or bone metastasis.

Taken together, the transplantation reports confirm the presence of breast cancer DCCs in many organs including the kidney that does not support metastatic growth. Metastatic growth in organ recipients of cells originating from the donor shows that these cells can survive many years after successful therapy and maintain their growth potential. Further research is needed to demonstrate that these DCCs can initiate distant metastasis in particular in an immunocompetent setting.

8.1.4 Comparison of different disseminated cancer cell populations

Talking about clinical translatability it is also of interest to reflect on our results in the context of the current clinical state of the art. In the clinic, metastatic spread of breast cancer is shown by the presence of CTCs in patients` blood [134, 306] or DCCs in the bone marrow [307, 308]. Therefore, the comparison and relationship of the analyzed DCCs from kidney and pancreas to the clinically detected CTCs and bone marrow DCCs would be of high interest.

In this study, CTCs were not analyzed in comparison to DCCs from kidney, pancreas and spleen. As CTCs are of clinical relevance and can be analyzed non-invasively [309] it will be of interest to also compare CTCs and DCCs. It is well known that dissemination is occurring frequently and only a small proportion of CTCs is able to survive at the distant site [14], implying further selection during the process. Thus, it is likely that DCCs resemble a rather small subpopulation of CTCs and comparison of CTCs and DCCs may allow further discrimination between factors that are essential for dissemination and survival of cancer cells in secondary organs.

Another interesting question arising from the presented data is the similarity of bone marrow DCCs and DCCs in kidney, pancreas and spleen. On one hand, the high similarity of DCCs in different organs would suggest that analysis of bone marrow DCCs, which are more accessible, is sufficient to understand and estimate DCC biology and presence in the whole body. On the other hand, the bone belongs to the organ tropism of breast cancer metastasis and cancer cells present in this organ may be rather similar to cells from the lungs, which are clearly distinct from kidney and pancreas DCCs as discussed earlier. Direct comparison of DCCs from kidney, pancreas and spleen to bone marrow DCCs regarding functions and transcription profiles is needed to address differences and similarities of the DCC pools.

8.1.5 Determinants of organ tropism in metastasis

An important characteristic of metastases of many cancer types is organ tropism, which refers to preferred metastatic growth of cancer cells in particular organs. Organ tropism is highly dependent on the cancer type with breast cancer showing tropism to lung, liver, bone and brain [62]. Due to the high complexity of the metastatic process, the molecular mechanisms regulating organ tropism are still not fully understood [310]. It is known today that a complex interplay between physiological characteristics of the organs, e.g. vessel diameter, and signaling cues from the microenvironment as well as intrinsic from cancer cells determine this tropism [62, 222]. If organ structures would predominantly be responsible for organ tropism by restricting dissemination, breast cancer cells would only be able reach the parenchyma of lung, liver, bone and brain. The here presented results indicate that physiological characteristics of organs alone cannot explain organ tropism as hundreds to thousands of cancer cells were detected and survived in kidney, pancreas and spleen. Thus, it was hypothesized by the author that the microenvironment selects for a subpopulation of cancer cells with transcriptional characteristics enabling survival in unfavorable microenvironments.

8.2 Transcriptomic changes in disseminated breast cancer cells

Transcriptional profiling of disseminated cancer cells from kidney and pancreas compared to cancer cells from mammary tumors (**Figure 17**) was performed to characterize the molecular and functional changes necessary for survival at the distant site. Interestingly, the changes occurring in DCCs in kidney and pancreas affect similar molecular functions. Especially the repressed molecular functions in DCCs from kidney and pancreas were very similar with about 80% of the repressed gene signatures being shared between kidney and pancreas DCCs (**Figure 17**). Thus, it was hypothesized that dormant DCCs in different organs apply similar survival mechanisms in the hostile microenvironments. This assumption is supported by results that the majority of molecular changes identified in DCCs in kidney and pancreas could also be confirmed in spleen DCCs. These changes include downregulation of genes involved in glycolysis (**Figure 24**), repressed EMT mechanisms (**Figure 22**), downregulation of MHC class II molecules (**Figure 27**) and upregulation of TSPAN8 and TSPAN1 (**Figure 33** and **Figure 34**). Therefore, it was concluded that these changes occur intrinsically in DCCs and are not directly regulated by the microenvironment. A recent study published by Echeverria *et al.*, demonstrating that lung, liver, bone and brain metastases of triple negative breast cancer PDX carried the same high-abundant genetic clone [311]. In line with the present work, this study suggests that sub-clones exist in the primary tumor that are able to seed independently of the secondary organ [311]. In my work it was demonstrated that DCCs present in different organs share transcriptomic features that may be beneficial independent of the target organ. Analysis of organ specific alteration may allow new insights into adaptation to the microenvironment and cross talk with e.g. stromal cells. Nevertheless, the presented analysis was focused on mechanisms common to DCCs from kidney and pancreas, as molecular similarities may be of interest for therapeutic targeting of DCCs.

8.2.1 Metabolic changes

One of the most pronounced changes that was observed in DCCs from kidney and pancreas was a metabolic switch. Compared to the mammary tumor, DCCs repressed gene signatures associated with glucose, energy and nucleotide metabolism (**Figure 23**). A general downregulation of metabolic pathways may occur in context of cellular dormancy as the cells need less energy and building blocks. However, a closer look at the involved genes revealed that glycolysis was especially repressed in DCCs compared to cells from the mammary tumor (**Figure 24**). In contrast, the gene expression of components of the mitochondrial electron chain responsible for oxidative phosphorylation was increased in DCCs (**Figure 25**). These data suggest that DCCs switch from anaerobe glycolysis to oxidative phosphorylation. It is well established that cancer cells preferably use anaerobic glycolysis for energy production in a process referred to as Warburg effect. Nevertheless cancer cells have functional mitochondria and are able to switch when appropriate [312]. The use of oxidative phosphorylation has been reported in slow cycling cancer cells such as cancer stem cells, which rely on oxidative phosphorylation [312-315]. Considering the characteristics of cancer stem cells such as reduced proliferation, therapy resistance and tumor initiating capacity [315], DCCs in kidney and

pancreas may resemble a population of CSCs or metastasis initiating cells. Recently, it has been shown by Basnet *et al.* that pulmonary micrometastases from breast cancer upregulate mitochondrial complex I [316]. In the present work these findings were expanded to DCCs in kidney and spleen, suggesting a more general role of oxidative phosphorylation during the early steps of breast cancer metastases.

Metabolic pathways are tightly regulated by cell intrinsic signaling as well as by cues from the microenvironment, as demonstrated by the close link between hypoxia and glycolysis [273]. Hypoxia, the lack of oxygen, which is often found in the core of primary tumors induces glycolysis as multiple glycolytic enzymes as well as glucose transporters (GLUT) are direct transcriptional targets of HIF1 [312, 317]. In line with the downregulation of glycolysis signatures in DCCs (**Figure 24**), hypoxia response signatures were repressed in DCCs in kidney and pancreas. The downregulated HIF target genes was not restricted to metabolic genes (**Figure 26**), indicating that the observed downregulation of hypoxia response signatures is not passive due to the metabolic shift. The downregulation of hypoxia response signaling in DCCs may be due to the close proximity of kidney DCCs to blood vessels, allowing diffusion of oxygen.

Also epithelial to mesenchymal transition (EMT) has been linked to metabolic regulation. Dong *et al.* revealed that the EMT-TF SNAIL induces glycolysis via methylation of the promotor of the fructose-1,6 biphosphatase (FBP1) gene [318]. The loss of FBP1 led to reduced oxygen consumption and suppression of complex I of the mitochondrial electron chain [318]. Thus, this study links EMT to increased glycolytic activity and is in line with the presented findings that EMT mechanisms, including SNAIL, as well as glycolysis are repressed, and oxidative phosphorylation is increased in DCCs. The provided evidence for the metabolic switch of DCCs in kidney and pancreas is based on gene expression analysis only. Further proof for a switch from glycolysis to oxidative phosphorylation can be achieved by the analysis of metabolites and proteins. However, the low number of cancer cells present in kidney and pancreas (500-10000) limits metabolomics and proteomic analysis.

8.2.2 Epithelial expression profile of DCCs

In the last decade evidence accumulated suggesting that cancer cells undergo epithelial to mesenchymal transition (EMT) to acquire a migratory phenotype [80, 81]. The induction of metastatic growth at the distant site however requires regain of the epithelial proliferative phenotype via MET [84, 85]. To the knowledge of the author, the EMT-MET state of cancer cells during metastatic latency periods is not known so far. The EMT profile of CTCs from breast cancer patients was comprehensively analyzed by Yu *et al.* They screened CTCs from 17 breast cancer patients for the expression of epithelial and mesenchymal related RNA. While the exact distribution of epithelial and mesenchymal cells was dependent on the breast cancer subtypes, almost all patients harbored epithelial as well as mesenchymal cells in their blood [138], suggesting that both phenotypes may be able to reach the organs. Nevertheless, they found a significantly correlation with disease progression and therapy resistance only for mesenchymal CTCs [138]. The present work suggests that DCCs in kidney and pancreas, although not growing at the time of analysis, express an epithelial genotype with downregulation of EMT-TFs

and mesenchymal markers (**Figure 22**) compared to cancer cells from mammary tumor. In *in vitro* culture, MDA-MB-231 cells are mesenchymal with stable expression of vimentin [319] and may thus already have undergone EMT. However, several studies show that MDA-MB-231 cells maintain their plasticity [320, 321] and can be used for studying EMT.

As EMT has been shown to induce stem cell properties it may seem contradictory that dormant DCCs show an epithelial gene expression profile. However, Schmidt *et al.* demonstrated that maintenance of the stem like properties obtained by Twist1 activation requires Twist1 inactivation [93]. Thus, the epithelial phenotype obtained upon MET is not identical to the one before EMT and some EMT characteristics can be maintained also during MET [93]. Further evidence suggests that EMT is a continuous process with intermediate stages [94]. Thereby, cancer cells can be partially EMT or MET, expressing both epithelial and mesenchymal characteristics [94] In addition, the existence of epithelial like CSCs has been shown in squamous cell carcinoma [322], suggesting that stem cell characteristics can be acquired by EMT independent mechanisms. A further indication that uncoupling of stem cell properties and EMT can be possible is the high plasticity of cancer cells in regulating signaling networks. One example for a pathway that is flexibly regulated in dormant cancer cells is the AKT pathway that results in activation of mTOR signaling [323]. Several studies have shown that downregulation of AKT signaling is involved in proliferation arrest of dormant cancer cells [189]. Schewe *et al.* demonstrated that dormant cells can uncouple AKT signaling from mTOR activation and can maintain mTOR signaling independent of AKT [191]. The cited studies thus suggest that stemness can be regulated independently of EMT.

8.2.3 Primary chemotherapy resistance of disseminated cancer cells

Interestingly, the here presented data suggest that the chemotherapy resistance observed in disseminated cancer cells in kidney, pancreas and spleen (**Figure 32**) is imprinted in DCCs and not acquired. This hypothesis is based on the enrichment of chemotherapy resistance signatures in untreated DCCs compared to the mammary tumor (**Figure 31**) as well as the upregulation of TSPANs (**Figure 33**), which were linked to therapy resistance (**Figure 40**). In addition, several other characteristics of DCCs such as downregulation of apoptotic processes (**Figure 21**), lack of proliferation (**Figure 13**) and upregulation of DNA repair genes (**Figure 30**) may contribute to chemotherapy resistance of DCCs in kidney, pancreas and spleen as these mechanisms have previously been linked to resistance [173]. Furthermore, environmental factors have been shown to contribute to the therapy resistant phenotype of cancer cells [234]. If and how the microenvironment may be involved in therapy resistance of DCCs in kidney, pancreas and spleen requires further investigations

8.2.4 Immune regulation of DCCs

All *in vivo* experiments performed in this study were conducted in immune compromised NSG mice. Therefore, it was surprising that gene signatures associated with antigen presentation were downregulated in DCCs from kidney and pancreas compared to cancer cells from the primary tumor (**Figure 27**). Antigen presentation is part of the innate immune response and facilitates presentation of pathogenic, tumor derived or misfolded proteins to T-lymphocytes, leading to activation of the adaptive immune response [278]. In humans, antigen presentation is mediated by MHC molecules, encoded by human leucocyte antigen (HLA) genes. Two types of MHC molecules exist: MHC class I, expressed by all cell types and MHC class II expressed only by antigen presenting cells (APCs) such as dendritic cells, macrophages and B-lymphocytes [278, 324]. In the case of MHC class I mediated antigen presentation, cytosolic antigens get processed mainly by the proteasome, bind to MHC molecules in the endoplasmic reticulum and get transported to the cell surface, where they are presented to and recognized by CD8+ T-cells [278, 325]. Antigens presented by MHC class II molecules are derived from extracellular sources that get internalized, are primarily bound in the phagosome and presented to CD4+ T-cells [278]. In DCCs, especially MHC class II mediated antigen presentation was repressed.

As MHC class I response triggers cytotoxic CD8+ T-cell response it is suggested that cancer cells downregulate MHC I genes to avoid T-cell cytotoxicity [326, 327]. Downregulation of MHC class I genes has indeed been reported on bone marrow DCCs of breast cancer patients and was associated with poorly differentiated tumors and poor survival [328, 329]. In the present study, about 65% of cancer cells in the mammary tumor expressed the MHC class I proteins HLA-A, B and C on their cell surface. This expression was not changed in DCCs in kidney, spleen (**Figure 27**). That no change was observed in the expression of MHC class I molecules on cancer cells, may be due to the used model system, lacking the adaptive immune compartment.

While MHC class I molecules are ubiquitously expressed by all cell types including cancer cells, MHC class II expression can be gained by cancer cells of several tumor entities, including breast cancer [330, 331]. In this work, about 20% of MDA-MB-231 cells in the mammary tumor and in lung, liver or bone expressed the MHC class II gene HLA-DR (**Figure 27**). Although still poorly understood, an inflammatory environment via interferon γ and interleukin 10 is known to induce MHC class II expression [330]. The role of MHC class II expression on cancer cells is still unclear as a correlation with clinical outcome is so far missing and also the role of CD4+ T-cells in tumor biology is not completely understood [330]. However, there are studies suggesting that the recruitment of CD4+ T-cells without co-stimulatory molecules induces immune tolerance to tumors [330, 332]. An important role for MHC class II mediated CD4+ T-cell activation in anti-tumor immunity is suggested by the study of Pyke *et al.,* who demonstrate an evolutionary selection pressure in tumors towards mutations that are not bound by MHC class II molecules [333]. They conclude that this selection leads to recruitment of only CD4+ T-cells that do not have anti-tumor activity [333]. In this context, downregulation of MHC class II molecules on DCCs in kidney and pancreas (**Figure 27**) may be part of the dormancy program to avoid recruitment of active immune cells. However, further analyses including *in vitro* experiments for T-cell activation and *in vivo* experiments in the immune competent system are needed to clarify which role MHC class II molecules play in the mammary tumor and at the distant site for DCC survival. The downregulation of antigen presentation by DCCs may open new treatment

opportunities as it has been shown that repression of antigen presentation can limit response to immune therapy [334, 335]. Therefore, it will be interesting to investigate if gain of MHC class II molecules renders DCCs sensitive for immunotherapy in the immune competent system.

In the used experimental *in vivo* setting no cellular triggers for immune evasion, such as mature B-cells, macrophages and T-cells were present. Thus, it was hypothesized that the downregulation of MHC molecules observed in DCCs in NSG mice is regulated by cancer cell intrinsic mechanisms of transcriptional regulation of HLA genes. Recently, several groups have demonstrated that intrinsic signaling mechanisms in cancer cells are able to manipulate immune responses [336]. Miao *et al.* even demonstrate that in squamous cell carcinoma in particular CSCs have the ability to evade immune responses. They further show that the evasion mechanism is linked to the stem cell properties, namely TGFβ response [337]. Several other studies report an involvement of the HIPPO pathway in cancer cell immune evasion [281-283, 338, 339]. The HIPPO pathway is an evolutionary conserved kinase cascade that leads to degradation of the transcription factor YAP1 and thereby to repression of proliferation, self-renewal and activation of apoptosis [283]. In the context of immune regulation pro- and anti-tumor effects have been reported for the HIPPO pathway. Two studies have shown that activation of YAP via knockdown of LATS1/2 leads to recruitment of M2 macrophages via cytokine secretion, thus establishing a pro-tumorigenic microenvironment [338, 339] and promoting survival of tumor-initiating cells [338]. An immune suppressive pro-tumorigenic role of YAP1 activation independent of direct immune cell recruitment was elucidated by Janse van Rensburg *et al. [281]*. They showed that YAP activation induces PD-L1 expression on breast cancer cells, which leads to inhibition of T-cell functions [281]. Contradictory, Moroishi *et al.* demonstrated an anti-tumorigenic role of Lats1 knockdown (YAP1 activation) in breast cancer models [282]. They showed that Lats1 deficient cancer cells secreted exosomes that triggered interferon type 1 signaling leading to increased antigen presentation on dendritic cells and thus induced anti-tumor immune reactions [282]. The here present data based on cell density experiments and YAP knockdown cells suggesting that YAP signaling, which was induced in DCCs compared to cells from the mammary tumor, is involved in downregulation of MHC class II molecules on MDA-MB-231 cells (**Figure 28**). This suggests that YAP activation may have different downstream effects in our study as compared to Moroishi *et al.* One explanation for the different results may be the species used in the studies: whereas I used human cell lines, Moroishi *et al.* worked with murine breast cancer cells. The group of Xialong Yang has shown that YAP1 regulates different target genes in mouse than in human cells due to different promoter structures [281, 340].

8.3 The role of Tetraspanins for DCC survival and therapy resistance

In this study the tetraspanin gene family (TSPAN) members TSPAN8 and TSPAN1 were identified as potential targets on dormant disseminated cancer cells mediating survival and chemo resistance (**Figure 40**). Tetraspanins are a heterogeneous family of membrane receptors characterized by four transmembrane (TM) regions, a conserved CCG motive and conserved residues in the extracellular loop [287] (**Figure 41 A**). Although they are expressed by nearly all cell types and tissues and are conserved from sponges to mammals, their functions are not well understood and only few members such as CD81 and CD9 have been studied in more detail [341, 342]. TSPANs do not have conventional receptor or enzymatic functions, instead they interact with tetraspanins as well as further transmembrane receptors and cytosolic proteins to build functional tetraspanin enriched microdomains so called tetraspanin webs [343] (**Figure 41 B**). TSPAN-TSPAN interactions primarily stabilize the microdomains, while lateral interactions between TSPANs and other proteins mediate downstream functions by influencing stability and functionality of their binding partners [343]. Many interactions with other proteins are mediated by the large extracellular domain that has been shown to be heterogeneous among different TSPANs [344], explaining that binding partners differ greatly between TSPAN molecules [343]. However, integrins are the most prominent interaction partners [342]. In addition to membrane bound interaction partners bound by the extracellular domain, TSPANs can also bind intracellular signaling molecules with their cytosolic tails [343], thereby influence intracellular signaling pathways [343, 345-347]. A direct involvement in signaling by being phosphorylated itself was so far only demonstrated for CD37 [348]. Upon drug treatment CD37 gets phosphorylated at the cytosolic tails leading to recruitment of additional phosphorylated proteins, ultimately resulting in AKT inhibition and regulation of apoptosis [348]. Furthermore, several TSPANs harbor an internalization signaling sequence that leads to endocytotic packaging of the membrane domain [287] (**Figure 41 A**), thus influencing membrane expression of TSPANs and their binding partners [343]. Additionally, as proteins on and in exosomes maintain their functionally, they may also influence the target cells. However, the impact of TSPANs in exosomes and their target cells is only poorly understood [342].

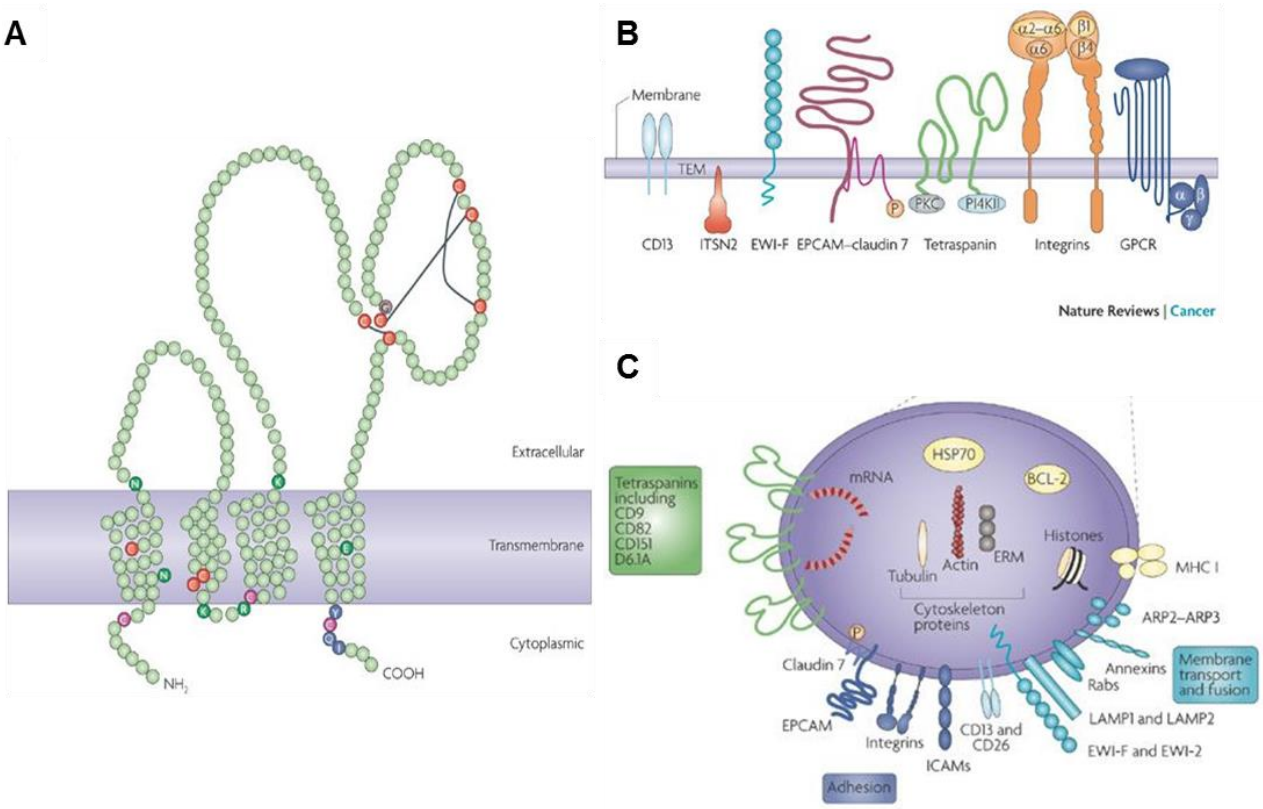


Figure 41 Structure and binding partners of tetraspanin molecules

A Tetraspanins (TSPAN) are transmembrane receptors that are characterized by four transmembrane domains and conserved cysteine residues (red). They harbor two extracellular loops, of which the bigger one is highly dynamic and is responsible for the heterogeneity of different TSPANs regarding binding partners. In addition, TSPANs have two intercellular tails with the carboxyl-tail of many TSPANs harboring an internalization signal (blue). The transmembrane domains contain polar amino acids (green) and the cytosolic tails have palmitoylation sites (pink) [342]. Recent studies have shown that the four transmembrane domains are divided in two pairs by a cholesterol molecule [349].

B Tetraspanins are not conventional signaling receptors but interact with a variety of binding partners to build so called TSPAN webs. Within these webs, TSPAN can interact with other family members as well as with different membrane bound and cytosolic partners. The interaction partners of the rat TSPAN8 ortholog, D6.1A, are shown. D6.1A binds to several integrins, epithelial adhesion molecule (EPCAM), G-protein coupled receptors (GPCR), CD13, intersectin 2 (ITSN2), EW1-F, protein kinase C (PKC), type II phosphatidylinositol-4-kinase (PI-4K).

C Via their internalization signal (presented in blue in **A**), TSPANs (including D6.1A/TSPAN8) are often packed into exosomes. Although the exact composition is dependent on the cell type, TSPAN positive exosomes were also found to contain MHC class I molecules, proteins involved in membrane transport and fusion (light blue), adhesion (dark blue) and cytoskeleton proteins. Only direct interaction partners of TSPANs are depicted. Signal transduction molecules, involved in TSPAN mediated exosomal signaling include in addition GTP-binding proteins, Src proteins, ERK proteins, SH2 phosphatases and catenins. Adapted from [342]

TSPANs have been shown to participate in cell adhesion, migration, invasion and survival processes of cells – all mechanisms that are essential during cancer development and progression [341, 342]. Nevertheless, the functional output is highly dependent on the TSPAN molecule: several TSPANs including CD151 and TSPAN8 have been shown to promote tumorigenesis and metastasis, whereas others like CD82 and CD9 function as metastasis

suppressors [341]. In the following paragraphs the role of TSPAN1 and TSPAN8 in homeostasis and during cancer progression will be discussed in more detail.

8.3.1 TSPAN8 and TSPAN1 during homeostasis and development

In this work, we showed that TSPAN8 and TSPAN1 are upregulated in breast cancer DCCs in kidney, pancreas and spleen and are functionally importance for survival and chemotherapy resistance of those cells. Thus, the following paragraph will discuss the current knowledge on the role of TSPAN8 and TSPAN1 in homeostasis and tumorigenesis.

In line with a general involvement of TSPAN molecules in several essential cellular processes such as migration, invasion and survival, TSPAN8 has been linked to tumor development and progression. The binding partners of TSPAN8 are depicted in **Figure 41 B** and include integrins as well as EPCAM, G-protein coupled receptors and cytosolic proteins (PKC, PI-4 K). In the tumor context, TSPAN8 was shown to be involved in angiogenesis [350], induction of migration [351] and inhibition of apoptosis [352, 353]. TSPAN8 was found to be overexpressed in several cancer entities [342] and was associated with poor differentiation and intrahepatic metastasis in hepatocellular carcinoma [354]. In line with the association with poor differentiation, TSPAN8 has also been linked to stemness in the healthy mammary gland [51]. Fu *et al.* revealed that the stem cell hierarchy in the mammary gland is determined by a mammary epithelial stem cell population that expresses both, TSPAN8 and LGR5, and is quiescent [51].

Since TSPAN8 contains an internalization sequence (**Figure 41 A**), it is a frequent component of exosomes and thus has been suggested to contribute to pre-metastatic niche formation [342]. The group of Margot Zöller has demonstrated over the past years that TSPAN8 containing exosomes contribute to cancer cell invasion via modulation of the microenvironment [355-357]. They demonstrated that TSPAN8+ exosomes mediate interaction with endothelial cells via GPCRs and receptor tyrosine kinases (RTK) thus stimulating angiogenesis [356]. Furthermore, TSPAN8+ exosomes regulate matrix degradation and stromal activation and promote metastasis of non-metastatic pancreatic adenocarcinoma cells [355]. Besides metastasis initiation, TSPAN8 also plays a role in therapy resistance. In multiple drug resistant gastric cancer cells TSPAN8 expression was elevated compared to sensitive controls and was shown to function via promotion of WNT signaling [358]. Based on these and other studies, demonstrating the importance of TSPAN8 for metastatic progression, therapeutic antibodies against TSPAN8 have been developed by several groups [349, 359, 360].

In a healthy organism, the role of TSPAN1 is only poorly understood. Recently, Zheng *et al.* characterized the stem cell population in planarias, one of few organisms harboring pluripotent stem cells in the adult animal [289]. They unraveled that these pluripotent stem cells, able to regrow the entire organism, expressed TSPAN1, suggesting a role of TSPAN1 in stemness [289]. However, no mechanistic insights into the function of TSPAN1 in this context are available to date. Similar to TSPAN8, TSPAN1, is upregulated in several tumor entities and has been associated with disease progression [361-365]. In prostate cancer TSPAN1 is controlled by androgen receptor signaling, mediates migration and promotes a mesenchymal phenotype [364]. Also in cholangiocarcinoma, TSPAN1 expression is associated with poor prognosis and

was shown to regulate Snail expression via PI3K signaling [365]. In line with the results for TSPAN8, TSPAN1 was also shown to mediate cancer invasion [366], survival and migration [367]. Copy number alteration analyses of matched lymph node metastases and primary breast tumors as wells as metastatic and non-metastatic breast tumors in contrast, suggest that TSPAN1 expression in this cancer entity is lost during cancer progression and metastasis [368]. Thus, depending on the context TSPAN1 may act as a tumor suppressor or promoter.

In the last years, TSPANs have become understood as crucial players during tumor development and progression. While the pro-tumorigenic role of TSPAN8 is well characterized, the functions of TSPAN1 are less well understood. In addition, the high flexibility and context dependency of TSPAN interaction with its partner molecules makes predictions of its functions in a specific situation difficult. A recent study published by Voglstaetter *et al*. emphasizes the context dependency of the functions of TSPANs. They have characterized the function of TSPAN8 in breast cancer cells using TSPAN8 overexpressing cells in different matrices. They have observed that the impact of TSPAN8 on breast cancer cells proliferation varies on the ECM composition of the matrix. TSPAN8 overexpression inhibited growth of cancer cells on a collagen I containing matrix, while it induced growth on basal membrane extract, containing laminin, collagen IV and proteoglycans [369]. Further investigations are necessary to elucidate the role and mode of action of TSPAN8 and TSPAN1 in disseminated cancer cells.

8.3.2 The role of TSPAN8 and TSPAN1 in DCCs

In this study, TSPAN8 and TSPAN1 were identified to be amongst the highest upregulated genes in DCCs from kidney and pancreas compared to cancer cells from the mammary tumor. Furthermore, both molecules are also part of the Balko *et al.* signature associated with minimal residual disease in breast cancer patients treated with neoadjuvant therapy [286] (**Figure 33**). Functional *in vivo* analyses confirmed the critical involvement of TSPAN8 and TSPAN1 in mediating chemotherapy resistance of DCCs in kidney, pancreas and spleen. Double knockdown of TSPAN8 and TSPAN1 sensitized DCCs in kidney, pancreas and spleen to chemotherapy (**Figure 40**). This is in line with results from Li *et al.* deciphering a critical role of TSPAN8 for multi drug resistance of gastric cancer cells via regulation of WNT signaling [358]. Further investigations are required to elucidate if WNT regulation also plays a role in TSPAN mediated therapy resistance of dormant DCCs. As TSPANs mediate their functions by regulation and modification of their binding partners, the downstream effect is highly dependent on the availability of the partners [369]. While it is well established that TSPAN molecules can interact with other family members [342], a direct binding of TSPAN8 and TSPAN1 has not been documented. The lack of functional impact of TSPAN8 or TSPAN1 single knockdown on DCC therapy resistance in this study suggests a collaborative effect of both molecules.

The aforementioned study of Voglstaetter *et al.* recently investigated the functional role of TSPAN8 in breast cancer metastasis. They demonstrated *in vivo* that overexpression of TSPAN8 increases liver and spleen metastasis [369]. Furthermore, they showed that TSPAN8 induces MET and leads to a more epithelial phenotype [369]. This data is in line with our observations that DCCs repressed EMT mediators and mesenchymal gene expression (**Figure**

22). Whether the epithelial phenotype in DCCs is linked to the upregulation of TSPAN8 and TSPAN1 needs further investigations.

Despite the strong effect of TSPAN8 and TSPAN1 knockdown on chemotherapy resistance the population of TSPAN8 and TSPAN1 double-positive cells is rather small with about 1% to 5% in DCCs from kidney, pancreas and spleen (**Figure 34**). Thus, it was hypothesized that the effect of TSPAN8 and TSPAN1 is mediated not only on the TSPAN expressing cells themselves but rather on the microenvironment. This is in line with previous studies demonstrating that the effect of TSPANs is only present in the *in vivo* situation and is not manifested *in vitro [349]*. In addition, many TSPANs, including TSPAN8 contain an internalization sequence leading to packaging of the web into exosomes. Secretion of these exosomes can then alter the microenvironment [357]. Previous results have demonstrated that exosomal derived TSPAN8 is involved in exosome targeting them to endothelial cells [370]. In addition, endothelial cells can contribute to cancer cells therapy resistance [234]. Together with the data that DCCs remain in close proximity to blood vessels in the kidney (**Figure 12**), this suggests a possible link between TSPAN8 expression, endothelial cell interaction and chemotherapy resistance. However, it is not known, whether TSPAN1 is processed into exosomes and further experimental proofs confirming the secretion of TSPANs in exosomes in the context of cancer cell dormancy and an involvement of exosomes in therapy resistance will be needed.

In vitro analysis using sphere culture assays suggested that TSPAN8 and TSPAN1 double expressing cells were associated with quiescence and survival since TSPAN8 and TSPAN1 expression was enriched in the label retaining population and knockdown of both molecules increased apoptosis under sphere conditions (**Figure 38**). Gene expression analysis of DCCs harboring a knockdown construct for TSPAN8 and TSPAN1 was performed compared to a control construct to investigate the role TSPAN8 and TSPAN1 in DCCs *in vivo*. Gene expression analysis of knockdown DCCs revealed that the Balko *et al.* therapy resistance signature as well as stem cells signatures was lost in the knockdown cells (**Figure 39**). This confirms the critical role of the TSPANs on therapy resistance as well as the effect of TSPAN8 and TSPAN1 on stem cell characteristics *in vivo*. These findings are in line with studies describing TSPAN8 and TSPAN1 as markers of normal stem cells [51, 289]. Moreover, this raises the question whether TSPAN8 and TSPAN1 positive DCCs in kidney and pancreas represent a population of metastasis initiating stem like cancer cells. Although the presented data suggests that these cancer cells have stem like properties, sorting of TSPAN8 and TSPAN1 positive and negative cells and *in vivo* injections are necessary to confirm the tumor initiating capacity of double positive cells.

8.4 Clinical implications of the results

Currently, metastatic spread of breast cancer can be detected by the presence of CTCs in patients` blood [134, 306] or DCCs in the bone marrow [307, 308]. The presented results suggest that chemotherapy resistant DCCs, which maintain their metastatic potential, are lodged in numerous organs of the body that do not normally support metastatic growth. While detection of DCCs in some of these organs, such as kidney, pancreas and spleen is difficult due to limited accessibility, it is important to keep in mind that not detecting cancer cells in the bone marrow and blood does not exclude the presence of DCCs in other organs.

The need to pay attention to the presence of DCCs in kidney, pancreas and spleen is underscored by several results presented in this study. The data suggest that DCCs in kidney and pancreas and the molecular mechanisms that were identified may play a role in breast cancer patients and disease relapse. This assumption is based on the enrichment of the Balko *et al*. signature in DCCs as well as the detection of TSPAN8 and TSPAN1 expression in primary patient samples. The expression of TSPAN8 and TSPAN1 in breast cancer patients was further confirmed in pleural effusion and ascites samples (**Figure 35**).

The signature generated by Balko *et al.* was associated with minimal residual disease and relapse after neoadjuvant chemotherapy in breast cancer patients [286]. The enrichment of this signature in DCCs from kidney and pancreas (**Figure 31**) suggests that these DCCs exhibit an overlapping transcriptional pattern compared to a therapy resistant cancer cell population present in breast cancer patients. As the Balko *et al.* signature uses primary tumor material [286], a subpopulation of cancer cells similar to DCCs may already be present in the primary tumor. Thus, a minority of cancer cells with the DCCs transcriptome may be selected from the bulk primary tumor due to survival advantages rather than being acquired de novo during the metastatic process. The existence of a small subpopulation of TSPAN8 and TSPAN1 expressing cells in the mammary tumor (**Figure 34**) and the transcriptomic similarity of kidney and pancreas DCCs (**Figure 17**) supports this assumption. However, comprehensive analysis on a single cell level would be required to answer this question. The finding that the Balko *et al.* signature is lost upon TSPAN8 and TSPAN1 double knockdown in DCCs (**Figure 39**) indicates a significant association of these two molecules to therapy resistance in human patients.

Furthermore, this study does not only impact potentially novel diagnostic means, but also the therapeutic strategy applied, as significant transcriptional changes between growing lung metastasis and DCs were identified (**Figure 20**). Several pro-metastatic genes and pathways that are potential therapeutic targets for lung metastasis, are downregulated in DCCs. Examples include MMPs [371], SPARC [264] and the niche component TNC [372], which is activated via the JNK pathway [373]. Thus, therapies targeting these molecules may leave DCCs in kidney, pancreas and spleen as minimal residual disease. The finding that these changes translate into functional differences, such as reduced apoptosis and proliferation in DCCs, suggests that different therapeutic strategies may be necessary for lung metastasis and DCCs. This study emphasizes the need for identification of different states of the disease prior to therapy as a dormant disease may need a different interventional strategy than a micro- or macro-metastatic disease [374, 375].

Another important question regarding the therapeutic approach is whether a dormant disease is beneficial for the patient or not. In theory three therapeutic approaches could be considered

against dormant DCCs: (1) maintaining dormancy of DCCs, (2) triggering reawakening of DCCs followed by chemotherapy targeting proliferating cells or (3) targeting survival of DCCs to eliminate them while being dormant [191, 374, 376]. Maintaining cancer cell dormancy would not cure cancer but render it a chronic disease. A similar strategy has been applied for HIV infections [376]. However, our results (**Figure 15**) as well as recent studies [199] demonstrated that dormant DCCs can be reawakened leading to metastatic relapse. For as long as the molecular mechanisms and environmental cues involved in silencing and reawakening are not fully understood, it might also be a risky strategy to maintain dormant cells in patients. An alternative strategy is therapeutic reawakening of dormant cells, which can then be targeted by chemotherapy. Lastly, the survival of dormant disseminated cancer cells could be targeted directly, avoiding reawakening and proliferation [191]. This work suggests that TSPAN1 and TSPAN8 may be potential target molecules of dormant breast cancer DCCs (**Figure 33, Figure 35**) as they are enriched in a non-proliferative subpopulation, are involved in survival and most importantly in chemotherapy resistance of DCCs. Targeting of TSPAN8 and TSPAN1 may lead to combinatory reactivation of dormant DCCs and immediate killing of dormant cells via apoptosis induction (**Figure 38**). Thus, combining inhibition of TSPANs with chemotherapy may sensitize cancer cells for therapy and yield the best results (**Figure 40**).

Taken together, this study provides novel insights in the biology of dormant disseminated cancer cells that may lead to development of therapeutic strategies to target residual breast cancer. The results further underline the need for precisely distinguishing the different steps of the metastatic cascade for ideal therapeutic intervention. Finally, TSPAN1 and TSPAN8 were identified as promising therapeutic targets regulating survival and chemotherapy resistance of dormant DCCs. Depletion of TSPAN1 and TSPAN8 rendered DCCs sensitive to chemotherapy and reduced the number of DCCs in kidney, pancreas and spleen.

9 Conclusion and outlook

In the present study we identified and characterized disseminated breast cancer cells in organs that normally do not support metastatic growth in patients, namely kidney, pancreas and spleen. We show that these cells, despite being non proliferative, can play a role in metastatic progression as they survive extended time spans, maintain their growth potential and are resistant to chemotherapy. These cells represent an early stage of the metastatic process that is so far only poorly understood. Our data demonstrate that disseminated cancer cells can behave differently to therapy than primary tumor and metastasis, stressing the need for precise diagnosis of the cancer stage for optimal treatment outcome. Thus, we performed transcriptomic profiling of these disseminated cancer cells and compared them to the bulk mammary tumor to identify DCC specific mechanisms. We elucidated that disseminated cancer cells in kidney, pancreas and spleen have an epithelial phenotype, repress cell death, rewire the metabolism from glycolysis to oxidative phosphorylation, upregulate DNA repair mechanisms and repress antigen presentation. This study sheds light on the intermediate stage of latent breast cancer and may help to identify potential therapeutic targets to interfere with metastasis in an early state.

Following up on the expression profile, we identified TSPAN8 and TSPAN1 as crucial mediators for survival and therapy resistance of disseminated cancer cells in kidney, pancreas and spleen. Both molecules were upregulated in DCCs, were important for stemness *in vitro* and *in vivo* and mediated therapy resistance of DCCs *in vivo*. The analysis should be expanded to additional breast cancer cell lines and further investigations are necessary to identify the relationship between TSPAN8 and TSPAN1 as well as the molecular mechanisms mediating therapy resistance and survival. Our findings suggest that TSPAN8 and TSPAN1 may be promising therapeutic targets for dormant breast cancer. A combination of anti-TSPAN treatment with chemotherapy may enable the elimination of dormant cancer cells that are currently spared by conventional treatment.

10 Appendix

10.1 Supplemental figures

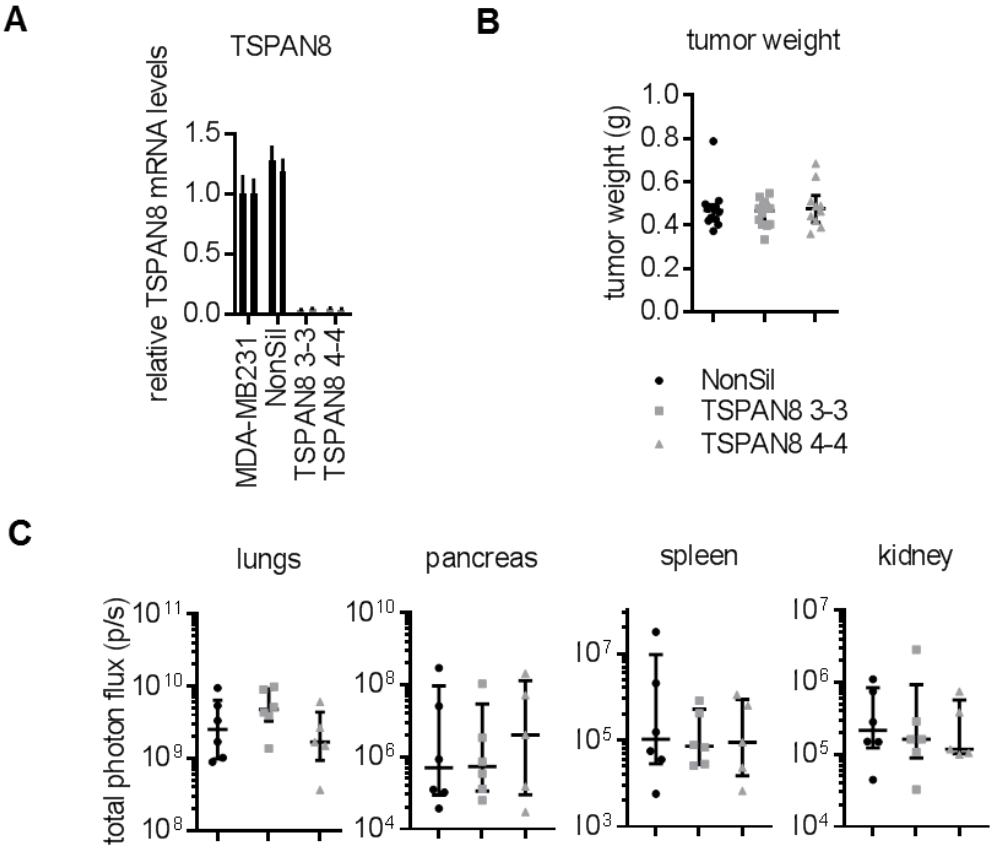


Figure S 1 TSPAN8 single knockdown does not affect lung metastases and DCCs in kidney, pancreas and spleen

A Knockdown of TSPAN8 was generated in MDA-MB-231 cells using two different shRNA hairpins (TSPAN8 3-3 and TSPAN8 4-4). Efficiency of TSPAN8 knockdown in MDA-MB-231 cells was measured by qRT-PCR. MDA-MB-231 carrying a control construct were used as control (NonSil). Uninfected MDA-MB-231 cells were included into the analysis to confirm that the control construct does not affect TSPAN8 expression. Two different primer pairs amplifying TSPAN8 were used and expression was normalized to hTBP expression. Expression values relative to the uninfected MDA-MB-231 cells are shown.

B and **C** NonSil Control or TSPAN8 knockdown cells were injected orthotopically into NSG mice. Five weeks after injection, mammary tumor weight was determined *ex vivo* (**B**) and tumor burden in lung, kidney, spleen and pancreas were measured via bioluminescence *ex vivo* (**C**). Mammary tumor weight is shown in gram in **B**. Six biological replicates were analyzed per group. All replicates and the median values are shown with interquartile range. The tumor cell burden in lung, pancreas, kidney and spleen was analyzed by *ex vivo* bioluminescence imaging of the organs. Mice were injected with luciferin and incubated for 10 min. Mice were sacrificed, organs were resected and bioluminescence was measured using an IVIS live imaging device. Six biological replicates were analyzed per group. The measurements for all replicates are shown in **C** with the median and the interquartile range of all replicates.

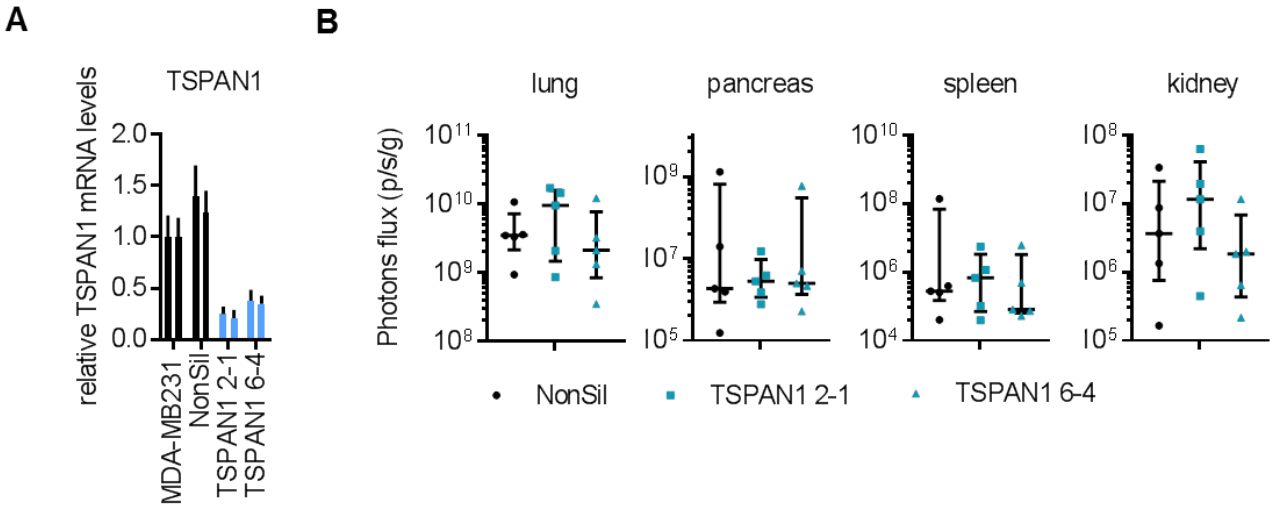


Figure S 2 TSPAN1 knockdown does not affect DCCs in kindey, pancreas and spleen

A Knockdown of TSPAN1 was generated in MDA-MB-231 cells using two different shRNA hairpins (TSPAN1 2-1 and TSPAN1 6-4). Efficiency of TSPAN1 knockdown in MDA-MB-231 cells was measured by qRT-PCR. MDA-MB-231 carrying a control construct were used as control (NonSil). Uninfected MDA-MB-231 cells were included into the analysis to confirm that the control construct does not affect TSPAN1 expression. Two different primer pairs amplifying TSPAN1 were used and expression was normalized to hTBP expression. Expression values relative to the uninfected MDA-MB-231 cells are shown.

B NonSil Control or TSPAN1 knockdown cells were injected orthotopically into NSG mice. Five weeks after injection tumor burdens in lung, kidney, spleen and pancreas were measured via bioluminescence *ex vivo.* Mice were injected with luciferin and incubated for 10 min. Mice were sacrificed, organs were resected and bioluminescence was measured using an IVIS live imaging device. Five biological replicates were analyzed per group. The measurements for all replicates are shown with the median and the interquartile range of all replicates.

10.2 Supplemental tables

Table S 1 collection of all used signatures and GO terms listing the source and the figure they are displaying in

topic	source	name	description	figure
whole database	GSEA	HALLMARK collection	all signatures collected in the HALLMARK database	18
	GSEA	GO collection	all signatures collected in the GeneOntology database	18
	GSEA	KEGG collection	all signatures collected in the KEGG database	18
DTCs profiling	self generated	pancreas vs MT_up	genes upregulated in MDA-MB-231 cells isolated from pancreas vs cancer cells from MT; Analysis immediately after isolation	19
	self generated	pancreas vs MT_down	genes, downregulated in MDA-MB-231 cells isolated from pancreas vs cancer cells from MT; Analysis immediately after isolation	19
lung metastasis	[16]	Minn et al_lung metastasis	48 genes upregulated in lung metastatic 4175 cells compared to the parental MDA-MB-231 cells	21
cell death	GSEA	KEGG_apoptosis	apoptosis	22
	Gene Ontology	GO:0097285	obsolete cell-type specific apoptotic process	22
	Gene Ontology	GO:0043524	negative regulation of neuron apoptotic process	22
	Gene Ontology	GO:1904035	regulation of epithelial cell apoptotic process	22
	Gene Ontology	GO:0070059	intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stres	22
	Gene Ontology	GO:0008219	cell death	22
	Gene Ontology	GO:0043069	negative regulation of programmed cell death	22
	Gene Ontology	GO:2000352	negative regulation of endothelial cell apoptotic process	22
	Gene Ontology	GO:0097190	apoptotic signaling pathway	22
	Gene Ontology	GO:0043281	regulation of cysteine-type endopeptidase activity involved in apoptotic process	22
	Gene Ontology	GO:0043067	regulation of programmed cell death	22
	Gene Ontology	GO:1902236	negative regulation of endoplasmic reticulum stress-induced intrinsic apoptotic signaling pathway	22
	Gene Ontology	GO:0012501	programmed cell death	22
	Gene Ontology	GO:0016265	obsolete death	22
	Gene Ontology	GO:0010942	positive regulation of cell death	22
	Gene Ontology	GO:0043065	positive regulation of apoptotic process	22
	Gene Ontology	GO:1902042	negative regulation of extrinsic apoptotic signaling pathway via death domain receptors	22
EMT	GSEA	HALLMARK_Epithelial_mesenchymal_Transition	Genes defining EMT, as in wound healing, fibrosis and metastasis.	23

topic	source	name	description	figure
EMT	[148]	PADUA_TGFB response signature	genes, upregulated in Breast cancer cells after 3h TGFb treatment	23
	[284]	ROKAVEC_genes up in mesenchymal vs epithelial	genes, upregulated in EMT induced DLD1 cells by incubation with IL6	23
	GSEA	LEF1_UP.V1_UP	genes up upon LEF1 OE	23
	[285]	JECHLINGER_ genes up in EMT	genes, upregulated in EpRas cells after induction of EMT with TGFb	23
	[284]	ROKAVEC_ genes up in epithelial vs mesenchymal	genes, upregulated in DLD1 cells that have undergone MET *in vivo*	23
	Gene Ontology	GO:0010719	negative regulation of epithelial to mesenchymal transition	23
	GSEA	LEF1_UP.V1_DN	genes down upon LEF1 OE	23
	[286]	AIGNER_genes negatively regulated by ZEB1	genes downregulated by ZEB1 in MDA-MB-231 cells	23
glucose metabolism	GSEA	HALLMARK_ GLYCOLYSIS	Genes encoding proteins involved in glycolysis and gluconeogenesis.	24
	GSEA	REACTOME_ GLYCOLYSIS	Genes involved in Glycolysis	24
	GSEA	BIOCARTA_ GLYCOLYSIS_ PATHWAY	Glycolysis Pathway	24
	GSEA	KEGG_GALACTOSE_ METABOLISM	Galactose metabolism	24
	Gene Ontology	GO:0006007	glucose catabolic process	24
	GSEA	REACTOME_ GLUCOSE_ METABOLISM	Genes involved in Glucose metabolism	24
	GSEA	KEGG_FRUCTOSE_ AND_MANNOSE_ METABOLISM	Fructose and mannose metabolism	24
	Gene Ontology	GO:0071322	cellular response to carbohydrate stimulus	24
	Gene Ontology	GO:0009743	response to carbohydrate	24
	Gene Ontology	GO:0006090	The chemical reactions and pathways involving pyruvate, 2-oxopropanoate	24
energy metabolism	Gene Ontology	GO:0046031	The chemical reactions and pathways involving ADP, adenosine 5'-diphosphate	24
	Gene Ontology	GO:0006757	The process of introducing a phosphate group into ADP to produce ATP	24
	Gene Ontology	GO:0006734	The chemical reactions and pathways involving reduced NADH, a coenzyme present in most living cells and derived from the B vitamin nicotinic acid	24
nucleotide metabolism	Gene Ontology	GO:0009185	The chemical reactions and pathways involving a ribonucleosite diphosphate, a compound consisting of a nucleobase linked to a ribose sugar esterified with diphosphate on the sugar.	24
	GSEA	KEGG_AMINO_SUGAR _AND_NUCLEOTIDE_ SUGAR_ METABOLISM	Amino sugar and nucleotide sugar metabolism	24

topic	source	name	description	figure
nucleotide metabolism	Gene Ontology	GO:0006140	Any process that modulates the frequency, rate or extent of the chemical reactions and pathways involving nucleotides	24
	Gene Ontology	GO:0030801	Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways involving cyclic nucleotides	24
Hypoxia	GSEA	ELVIDGE_HYPOXIA_BY_DMOG_UP	Genes up-regulated in MCF7 cells (breast cancer) treated with hypoxia mimetic DMOG	27
	GSEA	ELVIDGE_HIF1A_AND_HIF2A_TARGETS_DN	Genes down-regulated in MCF7 cells after knockdown of both HIF1A and HIF2A by RNAi.	27
	GSEA	ELVIDGE_HYPOXIA_UP	Genes up-regulated in MCF7 cells (breast cancer) under hypoxia conditions.	27
	GSEA	ELVIDGE_HIF1A_TARGETS_DN	Genes down-regulated in MCF7 cells (breast cancer) after knockdown of HIF1A by RNAi.	27
	GSEA	HALLMARK_HYPOXIA	Genes up-regulated in response to low oxygen levels (hypoxia).	27
	GSEA	FARDIN_HYPOXIA_11	Genes in the hypoxia signature, based on analysis of 11 neuroblastoma cell lines in hypoxia and normal oxygen conditions.	27
	GSEA	SEMENZA_HIF1_TARGETS	Genes that are transcriptionally regulated by HIF1A	27
	GSEA	REACTOME_REGULATION_OF_HYPOXIA_INDUCIBLE_FACTOR_HIF_BY_OXYGEN	Genes involved in Regulation of Hypoxia-inducible Factor (HIF) by Oxygen	27
	GSEA	GROSS_HIF1A_TARGETS_DN	Genes down-regulated in SEND cells (skin endothelium) at normal oxygen (normoxia) conditions after knockdown of HIF1A by RNAi.	27
	GSEA	RESPONSE_TO_HYPOXIA	Genes annotated by the GO term GO:0001666.	27
ox. phosphorylation	GSEA	MOOTHA_VOXPHOS	Genes involved in oxidative phosphorylation; based on literature and sequence annotation resources and converted to Affymetrix HG-U133A probe sets.	26
	GSEA	GO:0098803	respiratory chain complex	26
	Gene Ontology	GO:0045259	A proton-transporting two-sector ATPase complex that catalyzes the phosphorylation of ADP to ATP during oxidative phosphorylation	26
	GSEA	KEGG_OXIDATIVE_PHOSPHORYLATION	Oxidative phosphorylation	26
	GSEA	MOOTHA_HUMAN_MITODB_6_2002	Mitochondrial genes; based on literature and sequence annotation resources and converted to Affymetrix HG-U133A probe sets.	26
	GSEA	MOOTHA_MITOCHONDRIA	Mitochondrial genes	26
	GSEA	MODULE_152	Oxidative phosphorylation and ATP synthesis.	26
antigen presentation	Gene Ontology	GO:0042611	A transmembrane protein complex composed of an MHC alpha chain and, in most cases, either an MHC class II beta chain or an invariant beta2-microglobin chain.	28
	Gene Ontology	GO:0023023	Interacting selectively and non-covalently with the major histocompatibility complex.	28
	Gene Ontology	GO:0023026	Interacting selectively and non-covalently with the class II major histocompatibility complex.	28

topic	source	name	description	figure
antigen presentation	Gene Ontology	GO:0042605	Interacting selectively and non-covalently with an antigen peptide	28
	Gene Ontology	GO:0003823	Interacting selectively and non-covalently with an antigen, any substance which can induce a specific immune response and react with the products of that response, the antibody orT-lymphocytes.	28
	GSEA	REACTOME_MHC_ CLASS_II_ANTIGEN_ PRESENTATION	Genes involved in MHC class II antigen presentation	28
YAP1 targets	[297]	Shen_YAP1 KD signature	genes, downregulated in HUVEC cells upon siRNA KD of YAP1	29
proliferation	GSEA	HALLMARK_E2F_ Targets	Genes encoding cell cycle related targets of E2F transcription factors.	31
cell cycle checkpoints	GSEA	HALLMARK_G2M_ CHECKPOINT	Genes involved in the G2/M checkpoint, as in progression through the cell division cycle.	31
	GSEA	REACTOME_G2_M_ CHECKPOINTS	Genes involved in G2/M Checkpoints	31
	GSEA	REACTOME_CELL_ CYCLE_CHECKPOINT S	Genes involved in Cell Cycle Checkpoints	31
DNA repair	GSEA	PID_FANCONI_ PATHWAY	Fanconi anemia pathway	31
	GSEA	REACTOME_DNA_REP AIR	Genes involved in DNA Repair	31
	GSEA	KEGG_HOMOLOGOUS _RECOMBINATION	homologous recombination	31
	GSEA	REACTOME_DOUBLE_ STRAND_BREAK_ REPAIR	Genes involved in Double-Strand Break Repair	31
	Gene Ontology	GO:0000725	A DNA repair process that involves the exchange, reciprocal or nonreciprocal, of genetic material between the broken DNA molecule and a homologous region of DNA.	31
	Gene Ontology	GO:0000731	Synthesis of DNA that proceeds from the broken 3' single-strand DNA end and uses the homologous intact duplex as the template.	31
	Gene Ontology	GO:0006302	The repair of double-strand breaks in DNA via homologous and nonhomologous mechanisms to reform a continuous DNA helix.	31
	Gene Ontology	GO:0006281	The process of restoring DNA after damage.	31
	GSEA	REACTOME_FANCONI _ANEMIA_PATHWAY	Genes involved in Fanconi Anemia pathway	31
chemotherapy resistance	[246]	Kang et al._Doxorubicine resistance_up	genes upregulated in Doxorubicin resistant gastric cancer cells	32
	[303]	Balko et al._residual disease signature	genes associated with minimal residual disease and relapse of breast cancer patient treated with neoadjuvant chemotherapy	32
stem cell signatures	[377]	LIM_GENES UP IN MASC MOUSE AND HUMAN	genes common to human and murin mammary gland stem cells	40
	[308]	STINGL_genes up in mammary stem cells vs progenitor MA-CFC	genes upregulated in mammary stem cells compared to progenitors	40

topic	source	name	description	figure
stem cell signatures	[309]	SMITH ET AL_ADULT STEM CELL	human pan adult stem cell signature obtained by ranking of existing gene signatures	40
	[378]	SMITH_GENES UP IN CD49HI POPULTATION PROSTATE CANCER AND BENIGN	genes upregulated in the CD49hi population of prostate cancer cells compared to CD49lo cells	40
	[310]	PECE_GENES UP IN LABEL RETAINING CELLS COMPARED TO NOT_ALL POOLS	genes upregulated in label-retaining mammary spheres isolated from the healthy mammary gland	40
	[308]	STINGL_GENES DOWN IN MAMMARY STEM CELLS MRU VS TOTAL NEG FRACTION MYO	genes downregulated in mammary stem cells compared to all negative fractions	40

Table S 2 all GO terms that were significantly (p<0.05) enriched in the genes of the HALLMARK_E2F_Targets, core-enriched in kidney and pancreas DCCs

GO:0072528	GO:0045814	GO:0044772	GO:0018130	GO:0048522	GO:0051252
GO:0045786	GO:1902679	GO:0010564	GO:0019438	GO:0006950	GO:0006271
GO:0000075	GO:0051172	GO:0051301	GO:0008152	GO:0000281	GO:0042127
GO:0007093	GO:0051784	GO:0000819	GO:0000226	GO:0085020	GO:0034502
GO:0010948	GO:0010605	GO:0090304	GO:0007051	GO:0048478	GO:0045935
GO:0000077	GO:0010629	GO:0007346	GO:0007052	GO:0007569	GO:0010965
GO:0031570	GO:1902750	GO:0007059	GO:0000732	GO:0080090	GO:0006312
GO:1901988	GO:0044818	GO:0044260	GO:0008283	GO:0006323	GO:0071236
GO:0045930	GO:0031324	GO:0098813	GO:0000086	GO:0031323	GO:0007127
GO:0006281	GO:1901796	GO:0000280	GO:0044839	GO:0050789	GO:0006913
GO:1901991	GO:0006298	GO:0046483	GO:0006275	GO:0051128	GO:0033365
GO:0042770	GO:0009892	GO:0006139	GO:1902850	GO:0071103	GO:0009149
GO:0000731	GO:0045839	GO:0006725	GO:0051983	GO:0009157	GO:1905818
GO:0030330	GO:0007050	GO:1901360	GO:0007017	GO:0019222	GO:0061982
GO:0044773	GO:0010972	GO:1901987	GO:1902749	GO:0022414	GO:0051169
GO:0044774	GO:0016458	GO:0048285	GO:0006325	GO:0006342	GO:0009213
GO:0072401	GO:0006303	GO:0006261	GO:0051321	GO:0065007	GO:0010556
GO:0072422	GO:0006301	GO:0034641	GO:0007088	GO:0097421	GO:0044085
GO:0072395	GO:0051129	GO:0000082	GO:0040029	GO:0000003	GO:0010468
GO:0000724	GO:0000726	GO:0006996	GO:0090329	GO:0045005	GO:0048518
GO:0000725	GO:2000104	GO:0044843	GO:0044271	GO:0016570	GO:0051785
GO:0072331	GO:0007095	GO:0045787	GO:0019219	GO:0006270	GO:1901992
GO:0045934	GO:0000729	GO:0033554	GO:0033044	GO:0061640	GO:0010638
GO:0006302	GO:0071174	GO:1901990	GO:0033043	GO:0001889	GO:0031123
GO:2000113	GO:0071173	GO:0043170	GO:0045931	GO:0032465	GO:0070192
GO:0010558	GO:0007094	GO:0140014	GO:2000045	GO:1901989	GO:0006369
GO:0071158	GO:0006297	GO:0034645	GO:0006310	GO:0061008	GO:0033047

GO:0048523	GO:0045841	GO:0009059	GO:0060249	GO:0031145	GO:1902751
GO:0031327	GO:0031577	GO:0090068	GO:1902806	GO:0065004	GO:0090305
GO:0010639	GO:1902100	GO:0007062	GO:1903046	GO:0007098	GO:0000083
GO:0048519	GO:2000816	GO:0006807	GO:0070507	GO:0009162	GO:0016572
GO:0009890	GO:1905819	GO:0071897	GO:0060255	GO:0000722	GO:0009628
GO:2000134	GO:0033048	GO:0044237	GO:0070316	GO:0000910	GO:0031497
GO:0031572	GO:0033046	GO:0000070	GO:0031100	GO:0000079	GO:0000018
GO:0006977	GO:0051985	GO:0071840	GO:0009314	GO:0010212	GO:0006405
GO:1902400	GO:0008156	GO:0071704	GO:0010389	GO:0007010	GO:0006464
GO:1902402	GO:0006296	GO:1901576	GO:2000241	GO:0051716	GO:0036211
GO:1902403	GO:0042769	GO:0009058	GO:0090307	GO:0031023	GO:0051173
GO:0072413	GO:0033683	GO:0044238	GO:0051054	GO:0031099	GO:0009411
GO:1902807	GO:0000122	GO:0016043	GO:0050794	GO:1904029	GO:0018205
GO:0072431	GO:0031936	GO:0044249	GO:0032886	GO:0007064	GO:0031124
GO:0071156	GO:0007049	GO:0000723	GO:0016569	GO:0001832	GO:0031326
GO:0070317	GO:0022402	GO:0032200	GO:0046605	GO:0071824	GO:0043412
GO:2001251	GO:1903047	GO:0009987	GO:0051302	GO:0071478	GO:0033045
GO:0031571	GO:0000278	GO:0051783	GO:0140013	GO:0006282	GO:0001701
GO:0051130	GO:0009889	GO:0010604	GO:0044267	GO:0009262	GO:0042592
GO:0006259	GO:0051168	GO:0030071	GO:0033260	GO:0045740	GO:0019985
GO:0051726	GO:0009204	GO:1902099	GO:0044786	GO:0006351	GO:0045144
GO:0044770	GO:1905456	GO:2000112	GO:0051052	GO:0097659	GO:0006974
GO:0006260	GO:0051177	GO:1902275	GO:1901362	GO:0009211	GO:0044819
GO:0051276	GO:1990823	GO:0045840	GO:0051171	GO:0006333	GO:0044783
GO:0034654	GO:1990830	GO:0072527	GO:0022616	GO:2000973	GO:0045892
GO:0051225	GO:0008150	GO:0045132	GO:0032201	GO:0032774	GO:0051253
GO:1903507					

10.3 Abbreviations

ACK	Ammonium-Chloride-Potassium
AKT	protein kinase B
ALDOC	aldolase C, fructose-bisphosphate
ANGPTL4	angiopoietin-like 4
APC	antigen presenting cells
ATP	Adenosine triphosphate
BA	breast ascites
BCL2	B-cell lymphoma 2
BER	base excision repair
BFP	blue fluorescent protein
BH	Benjamini - Hochberg
BMP7	bone morphogenic protein 7
BPE	breast pleural effusion
CASP4	caspase 4
cDNA	complementary DNA
COX2	cyclooxygenase 2
CSC	cancer stem cell
CSF1	colony stimulating factor 1
csf1	colony stimulation factor 1
CT value	threshold cycle
CTC	circulating tumor cell
CXCL	C-X-C motif ligand
DAPI	4′,6-Diamidin-2-phenylindol
DKFZ	German cancer research center
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	dimethyl Sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxy Nukleosidtriphosphate
Doxo+Cyclo	Doxorubicin and Cyclophosphamid
dRN	delta normalized reporter signal
DTC	disseminated tumor cell
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
EGFR	epidermal growth factor receptor
EGLN3	Egl9-family hypoxia inducible factor 3
EMT	epithelial to mesenchymal transition
ENO	enolase
ER	estrogen receptor
ERK	extracellular-signal-regulated kinase
FACS	fluorescence activated cell sorting
FBP1	fructose- 2,6 biphosphatase
FCS	fatal calf serum
FDR	false discovery rates
FHIT	fragile histidine triad
FN1	fibronectin1
FSC	forward scatter

GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GFP	green fluorescent protein
GLUT1	glucose transporter type 1
GO	gene ontology
GPCR	G-protein coupled receptor
GSEA	gene set enrichment analysis
H&E	Haematoxilin and Eosin
H2kD	histocompatibility 2, K1
HDR	homology derived repair
HER2	human epidermal growth factor receptor 2
HIF	hypoxia induced factor
HK	hexokinase
HLA	human leukocyte antigen
hTBP	human TATA box binding protein
IF	immuno fluorescence
IGFBP3	insulin like growth factor binding protein 3
IGF-II	insulin growth factor II
IGFR	insulin growth factor receptor
IL-6	interleukin 6
IP	intraperitoneal
ITGA5	integrin alpha 5
ITPR1	inositol 1,4,5-Trisphosphate Receptor Type 1
JNK	c-Jun N-terminal kinase
KEGG	Kyoto Encyclopedia of Genes and Genomes
kras	kirsten rat sarcoma 2 viral oncogene homolog
LCC	latency competent cell
LGR5	leucine-rich repeat-containing G protein-coupled receptor 5
LOX	lysyl oxidase
mB2M	mouse Beta-2-Microglobulin
MDSC	myloid derived suppressor cells
MET	mesenchymal to epithelial transition
MHC	major histocompatibility complex
MMP	matrix-metallo protease
mRNA	messanger RNA
MMR	miss match repair
MT	mammary tumor
mTOR	mechanistic Target of Rapamycin
NaCl	sodium chlorid
NADPH	nicotinamidadenindinukleotidphosphat
NANOG	Nanog Homeobox
NCT	National Center for Tumor Diseases
NEB	New England Biolabs
NER	nucleotide excition repair
NES	normalized enrichment score
NET	neutrophil extracelullar traps
NFkB	nuclear factor kappa-light-chain-enhancer of activated B cells
NHEJ	non-homologous end joining
NK cell	natural killer cell

NSG	non-obese diabetic-severe (NOD) scid gamma null
OCT4	POU Class 5 Hemeobox 1
PBS	phosphate-buffered saline
PCA	principal component analysis
PDK1	pyruvate dehydrogenase kinase 1
PD-L1	programmed death receptor ligand 1
PDX	patient derived xenograft
PFKFB	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase
PGK	phosphoglycerate kinase
PI3K	phosphoinositide-3-kinase
POSTN	periostin
PPP1R15A	protein phosphatase 1 regulatory subunit 15A
PR	progesteron receptro
PRKCD	protein kinase C delta
qRT-PCR	quantitative realt time - poly chain reaction
RMA	robust multichip average
RNA	ribonucleic acid
ROS	reactive oxygen species
RT	room temperature
RTK	receptor tyrosin kinase
SDC	syndecan
shRNA	shott hairpin
SNAI2	snail family transcriptional repressor 2
SOX2	SRY-Box 2
SPARC	secreted protein acidic and rich in cysteine
SSC	sideward scatter
TCA	tricarboxylic acid cycle
TF	transcription factor
TGFb	transforming growth factor beta
THBS1	thrombospondin 1
TIMP3	TIMP Metallopeptidase Inhibitor 3
TNC	tenascin C
TNFa	tumor necrosis factor alpha
TNFRSF10D	TNF Receptor Superfamily F Decoy
TRAIL	TNF-related apoptosis-inducing ligand
TSPAN	testraspanin
UK	United Kingdom
uPAR	urokinase plasminogen activator receptor
VEGF	vascular epithelial growth factor
VEGFR1	vascular epithelial growth factor receptor 1
VIM	vimentin
WHO	world health organisation
YAP1	yes-associated protein 1
ZEB	zinc-finger E-box binding homeobox

10.4 List of figures

10.5 List of table

11 Bibliography

1. World Health Organization, D.o.I., Evidence and Research, *Cancer Mortality database*. 2016.
2. Chaffer, C.L. and R.A. Weinberg, *A perspective on cancer cell metastasis.* Science, 2011. **331**(6024): p. 1559-64.
3. Lambert, A.W., D.R. Pattabiraman, and R.A. Weinberg, *Emerging Biological Principles of Metastasis.* Cell, 2017. **168**(4): p. 670-691.
4. Loberg, R.D., et al., *The lethal phenotype of cancer: the molecular basis of death due to malignancy.* CA Cancer J Clin, 2007. **57**(4): p. 225-41.
5. Ripamonti, C., *Management of dyspnea in advanced cancer patients.* Support Care Cancer, 1999. **7**(4): p. 233-43.
6. Muscaritoli, M., et al., *Prevention and treatment of cancer cachexia: new insights into an old problem.* Eur J Cancer, 2006. **42**(1): p. 31-41.
7. Sallah, S., J.Y. Wan, and N.P. Nguyen, *Venous thrombosis in patients with solid tumors: determination of frequency and characteristics.* Thromb Haemost, 2002. **87**(4): p. 575-9.
8. AJCC, *AJCC Cancer staging manual - sixth edition*. Vol. sixth Edition. 2002, Springer.
9. UK, C.R., *Statistics by cancer type*. 2018.
10. UK, C.R., *breast cancer survival statistics*. 2002-2006.
11. UK, C.R., *Lung cancer survival statistics*. 2003-2006.
12. Valastyan, S. and R.A. Weinberg, *Tumor metastasis: molecular insights and evolving paradigms.* Cell, 2011. **147**(2): p. 275-92.
13. Aguirre-Ghiso, J.A., *How dormant cancer persists and reawakens.* Science, 2018. **361**(6409): p. 1314-1315.
14. Husemann, Y., et al., *Systemic spread is an early step in breast cancer.* Cancer Cell, 2008. **13**(1): p. 58-68.
15. Kang, Y., et al., *A multigenic program mediating breast cancer metastasis to bone.* Cancer Cell, 2003. **3**(6): p. 537-49.
16. Minn, A.J., et al., *Genes that mediate breast cancer metastasis to lung.* Nature, 2005. **436**(7050): p. 518-24.
17. Riedl, C.C., et al., *Retrospective analysis of 18F-FDG PET/CT for staging asymptomatic breast cancer patients younger than 40 years.* J Nucl Med, 2014. **55**(10): p. 1578-83.
18. Coumans, F.A., S. Siesling, and L.W. Terstappen, *Detection of cancer before distant metastasis.* BMC Cancer, 2013. **13**: p. 283.
19. Schmidt-Kittler, O., et al., *From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression.* Proc Natl Acad Sci U S A, 2003. **100**(13): p. 7737-42.
20. Podsypanina, K., et al., *Seeding and propagation of untransformed mouse mammary cells in the lung.* Science, 2008. **321**(5897): p. 1841-4.
21. Klein, C.A., *Parallel progression of primary tumours and metastases.* Nat Rev Cancer, 2009. **9**(4): p. 302-12.
22. Bross, I.D., E. Viadana, and J. Pickren, *Do generalized metastases occur directly from the primary?* J Chronic Dis, 1975. **28**(3): p. 149-59.
23. Tozzi, R., et al., *Laparoscopic treatment of early ovarian cancer: surgical and survival outcomes.* Gynecol Oncol, 2004. **93**(1): p. 199-203.
24. D'Amico, T.A., *Outcomes after surgery for esophageal cancer.* Gastrointest Cancer Res, 2007. **1**(5): p. 188-96.
25. Houg, D.S. and M.F. Bijlsma, *The hepatic pre-metastatic niche in pancreatic ductal adenocarcinoma.* Mol Cancer, 2018. **17**(1): p. 95.
26. Peinado, H., et al., *Pre-metastatic niches: organ-specific homes for metastases.* Nat Rev Cancer, 2017. **17**(5): p. 302-317.
27. Navin, N., et al., *Tumour evolution inferred by single-cell sequencing.* Nature, 2011. **472**(7341): p. 90-4.
28. Ding, L., et al., *Genome remodelling in a basal-like breast cancer metastasis and xenograft.* Nature, 2010. **464**(7291): p. 999-1005.
29. Kim, T.M., et al., *Subclonal Genomic Architectures of Primary and Metastatic Colorectal Cancer Based on Intratumoral Genetic Heterogeneity.* Clin Cancer Res, 2015. **21**(19): p. 4461-72.
30. Yachida, S., et al., *Distant metastasis occurs late during the genetic evolution of pancreatic cancer.* Nature, 2010. **467**(7319): p. 1114-7.

31. Campbell, P.J., et al., *The patterns and dynamics of genomic instability in metastatic pancreatic cancer.* Nature, 2010. **467**(7319): p. 1109-13.
32. Robinson, D., et al., *Integrative Clinical Genomics of Advanced Prostate Cancer.* Cell, 2015. **162**(2): p. 454.
33. Jesinghaus, M., et al., *Distinctive Spatiotemporal Stability of Somatic Mutations in Metastasized Microsatellite-stable Colorectal Cancer.* Am J Surg Pathol, 2015. **39**(8): p. 1140-7.
34. Turajlic, S. and C. Swanton, *Metastasis as an evolutionary process.* Science, 2016. **352**(6282): p. 169-75.
35. Butler, T.P. and P.M. Gullino, *Quantitation of cell shedding into efferent blood of mammary adenocarcinoma.* Cancer Res, 1975. **35**(3): p. 512-6.
36. Thiery, J.P., *Epithelial-mesenchymal transitions in tumour progression.* Nat Rev Cancer, 2002. **2**(6): p. 442-54.
37. Braun, S., et al., *A pooled analysis of bone marrow micrometastasis in breast cancer.* N Engl J Med, 2005. **353**(8): p. 793-802.
38. Zeidman, I., C.M. Mc, and D.R. Coman, *Factors affecting the number of tumor metastases; experiments with a transplantable mouse tumor.* Cancer Res, 1950. **10**(6): p. 357-9.
39. Fidler, I.J., *Metastasis: quantitative analysis of distribution and fate of tumor emboli labeled with 125 I-5-iodo-2'-deoxyuridine.* J Natl Cancer Inst, 1970. **45**(4): p. 773-82.
40. Fidler, I.J. and M.L. Kripke, *Metastasis results from preexisting variant cells within a malignant tumor.* Science, 1977. **197**(4306): p. 893-5.
41. Gonzales, K.A.U. and E. Fuchs, *Skin and Its Regenerative Powers: An Alliance between Stem Cells and Their Niche.* Dev Cell, 2017. **43**(4): p. 387-401.
42. Clevers, H., *The intestinal crypt, a prototype stem cell compartment.* Cell, 2013. **154**(2): p. 274-84.
43. Doulatov, S., et al., *Hematopoiesis: a human perspective.* Cell Stem Cell, 2012. **10**(2): p. 120-36.
44. Batlle, E. and H. Clevers, *Cancer stem cells revisited.* Nat Med, 2017. **23**(10): p. 1124-1134.
45. Al-Hajj, M., et al., *Prospective identification of tumorigenic breast cancer cells.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3983-8.
46. Dalerba, P., et al., *Phenotypic characterization of human colorectal cancer stem cells.* Proc Natl Acad Sci U S A, 2007. **104**(24): p. 10158-63.
47. Ricci-Vitiani, L., et al., *Identification and expansion of human colon-cancer-initiating cells.* Nature, 2007. **445**(7123): p. 111-5.
48. Shimokawa, M., et al., *Visualization and targeting of LGR5(+) human colon cancer stem cells.* Nature, 2017. **545**(7653): p. 187-192.
49. Kozar, S., et al., *Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas.* Cell Stem Cell, 2013. **13**(5): p. 626-33.
50. Barker, N., et al., *Identification of stem cells in small intestine and colon by marker gene Lgr5.* Nature, 2007. **449**(7165): p. 1003-7.
51. Fu, N.Y., et al., *Identification of quiescent and spatially restricted mammary stem cells that are hormone responsive.* Nat Cell Biol, 2017. **19**(3): p. 164-176.
52. Oskarsson, T., et al., *Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs.* Nat Med, 2011. **17**(7): p. 867-74.
53. de Sousa e Melo, F., et al., *A distinct role for Lgr5(+) stem cells in primary and metastatic colon cancer.* Nature, 2017. **543**(7647): p. 676-680.
54. Reya, T. and H. Clevers, *Wnt signalling in stem cells and cancer.* Nature, 2005. **434**(7035): p. 843-50.
55. Reya, T., et al., *Stem cells, cancer, and cancer stem cells.* Nature, 2001. **414**(6859): p. 105-11.
56. van 't Veer, L.J., et al., *Gene expression profiling predicts clinical outcome of breast cancer.* Nature, 2002. **415**(6871): p. 530-6.
57. Ramaswamy, S., et al., *A molecular signature of metastasis in primary solid tumors.* Nat Genet, 2003. **33**(1): p. 49-54.
58. Ma, X.J., et al., *Gene expression profiles of human breast cancer progression.* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 5974-9.
59. Weigelt, B., J.L. Peterse, and L.J. van 't Veer, *Breast cancer metastasis: markers and models.* Nat Rev Cancer, 2005. **5**(8): p. 591-602.
60. Hill, R.P., et al., *Dynamic heterogeneity: rapid generation of metastatic variants in mouse B16 melanoma cells.* Science, 1984. **224**(4652): p. 998-1001.
61. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
62. Nguyen, D.X., P.D. Bos, and J. Massague, *Metastasis: from dissemination to organ-specific colonization.* Nat Rev Cancer, 2009. **9**(4): p. 274-84.

63. Guo, W. and F.G. Giancotti, *Integrin signalling during tumour progression.* Nat Rev Mol Cell Biol, 2004. **5**(10): p. 816-26.
64. Nieto, M.A., et al., *Emt: 2016.* Cell, 2016. **166**(1): p. 21-45.
65. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.
66. Arnoux, V., et al., *Erk5 controls Slug expression and keratinocyte activation during wound healing.* Mol Biol Cell, 2008. **19**(11): p. 4738-49.
67. Ahmed, N., et al., *Molecular pathways regulating EGF-induced epithelio-mesenchymal transition in human ovarian surface epithelium.* Am J Physiol Cell Physiol, 2006. **290**(6): p. C1532-42.
68. Greenburg, G. and E.D. Hay, *Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells.* J Cell Biol, 1982. **95**(1): p. 333-9.
69. Puisieux, A., T. Brabletz, and J. Caramel, *Oncogenic roles of EMT-inducing transcription factors.* Nat Cell Biol, 2014. **16**(6): p. 488-94.
70. Burdsal, C.A., C.H. Damsky, and R.A. Pedersen, *The role of E-cadherin and integrins in mesoderm differentiation and migration at the mammalian primitive streak.* Development, 1993. **118**(3): p. 829-44.
71. Batlle, E., et al., *The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells.* Nat Cell Biol, 2000. **2**(2): p. 84-9.
72. Cano, A., et al., *The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression.* Nat Cell Biol, 2000. **2**(2): p. 76-83.
73. Comijn, J., et al., *The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion.* Mol Cell, 2001. **7**(6): p. 1267-78.
74. Eger, A., et al., *DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells.* Oncogene, 2005. **24**(14): p. 2375-85.
75. Perez-Moreno, M.A., et al., *A new role for E12/E47 in the repression of E-cadherin expression and epithelial-mesenchymal transitions.* J Biol Chem, 2001. **276**(29): p. 27424-31.
76. Yang, J., et al., *Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis.* Cell, 2004. **117**(7): p. 927-39.
77. Iwano, M., et al., *Evidence that fibroblasts derive from epithelium during tissue fibrosis.* J Clin Invest, 2002. **110**(3): p. 341-50.
78. Nitta, T., et al., *Murine cirrhosis induces hepatocyte epithelial mesenchymal transition and alterations in survival signaling pathways.* Hepatology, 2008. **48**(3): p. 909-19.
79. Yanez-Mo, M., et al., *Peritoneal dialysis and epithelial-to-mesenchymal transition of mesothelial cells.* N Engl J Med, 2003. **348**(5): p. 403-13.
80. Brabletz, T., et al., *Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10356-61.
81. Prall, F., *Tumour budding in colorectal carcinoma.* Histopathology, 2007. **50**(1): p. 151-62.
82. Moody, S.E., et al., *The transcriptional repressor Snail promotes mammary tumor recurrence.* Cancer Cell, 2005. **8**(3): p. 197-209.
83. Gabbert, H., et al., *Tumor dedifferentiation: an important step in tumor invasion.* Clin Exp Metastasis, 1985. **3**(4): p. 257-79.
84. Ocana, O.H., et al., *Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1.* Cancer Cell, 2012. **22**(6): p. 709-24.
85. Chaffer, C.L., et al., *Mesenchymal-to-epithelial transition facilitates bladder cancer metastasis: role of fibroblast growth factor receptor-2.* Cancer Res, 2006. **66**(23): p. 11271-8.
86. Jorda, M., et al., *Upregulation of MMP-9 in MDCK epithelial cell line in response to expression of the Snail transcription factor.* J Cell Sci, 2005. **118**(Pt 15): p. 3371-85.
87. Miyoshi, A., et al., *Snail and SIP1 increase cancer invasion by upregulating MMP family in hepatocellular carcinoma cells.* Br J Cancer, 2004. **90**(6): p. 1265-73.
88. Vega, S., et al., *Snail blocks the cell cycle and confers resistance to cell death.* Genes Dev, 2004. **18**(10): p. 1131-43.
89. Mani, S.A., et al., *The epithelial-mesenchymal transition generates cells with properties of stem cells.* Cell, 2008. **133**(4): p. 704-15.
90. Ye, X., et al., *Distinct EMT programs control normal mammary stem cells and tumour-initiating cells.* Nature, 2015. **525**(7568): p. 256-60.
91. Mejlvang, J., et al., *Direct repression of cyclin D1 by SIP1 attenuates cell cycle progression in cells undergoing an epithelial mesenchymal transition.* Mol Biol Cell, 2007. **18**(11): p. 4615-24.
92. Chaffer, C.L., et al., *Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity.* Cell, 2013. **154**(1): p. 61-74.

93. Schmidt, J.M., et al., *Stem-cell-like properties and epithelial plasticity arise as stable traits after transient Twist1 activation.* Cell Rep, 2015. **10**(2): p. 131-9.
94. Pastushenko, I., et al., *Identification of the tumour transition states occurring during EMT.* Nature, 2018. **556**(7702): p. 463-468.
95. Hong, T., et al., *An Ovol2-Zeb1 Mutual Inhibitory Circuit Governs Bidirectional and Multi-step Transition between Epithelial and Mesenchymal States.* PLoS Comput Biol, 2015. **11**(11): p. e1004569.
96. Zhang, J., et al., *TGF-beta-induced epithelial-to-mesenchymal transition proceeds through stepwise activation of multiple feedback loops.* Sci Signal, 2014. **7**(345): p. ra91.
97. Wicki, A., et al., *Tumor invasion in the absence of epithelial-mesenchymal transition: podoplanin-mediated remodeling of the actin cytoskeleton.* Cancer Cell, 2006. **9**(4): p. 261-72.
98. Wyckoff, J.B., et al., *Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors.* Cancer Res, 2007. **67**(6): p. 2649-56.
99. Labernadie, A., et al., *A mechanically active heterotypic E-cadherin/N-cadherin adhesion enables fibroblasts to drive cancer cell invasion.* Nat Cell Biol, 2017. **19**(3): p. 224-237.
100. Mu, D., et al., *The integrin alpha(v)beta8 mediates epithelial homeostasis through MT1-MMP-dependent activation of TGF-beta1.* J Cell Biol, 2002. **157**(3): p. 493-507.
101. Munger, J.S., et al., *The integrin alpha v beta 6 binds and activates latent TGF beta 1: a mechanism for regulating pulmonary inflammation and fibrosis.* Cell, 1999. **96**(3): p. 319-28.
102. Haeger, A., et al., *Collective cell migration: guidance principles and hierarchies.* Trends Cell Biol, 2015. **25**(9): p. 556-66.
103. Aceto, N., et al., *Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis.* Cell, 2014. **158**(5): p. 1110-1122.
104. Cheung, K.J., et al., *Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters.* Proc Natl Acad Sci U S A, 2016. **113**(7): p. E854-63.
105. Zajac, O., et al., *Tumour spheres with inverted polarity drive the formation of peritoneal metastases in patients with hypermethylated colorectal carcinomas.* Nat Cell Biol, 2018. **20**(3): p. 296-306.
106. Gkountela, S., et al., *Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding.* Cell, 2019. **176**(1-2): p. 98-112 e14.
107. Goswami, S., et al., *Macrophages promote the invasion of breast carcinoma cells via a colony-stimulating factor-1/epidermal growth factor paracrine loop.* Cancer Res, 2005. **65**(12): p. 5278-83.
108. Wyckoff, J., et al., *A paracrine loop between tumor cells and macrophages is required for tumor cell migration in mammary tumors.* Cancer Res, 2004. **64**(19): p. 7022-9.
109. Dumont, N., et al., *Breast fibroblasts modulate early dissemination, tumorigenesis, and metastasis through alteration of extracellular matrix characteristics.* Neoplasia, 2013. **15**(3): p. 249-62.
110. Lochter, A., et al., *Matrix metalloproteinase stromelysin-1 triggers a cascade of molecular alterations that leads to stable epithelial-to-mesenchymal conversion and a premalignant phenotype in mammary epithelial cells.* J Cell Biol, 1997. **139**(7): p. 1861-72.
111. Sternlicht, M.D., et al., *The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis.* Cell, 1999. **98**(2): p. 137-46.
112. Gaggioli, C., et al., *Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells.* Nat Cell Biol, 2007. **9**(12): p. 1392-400.
113. Joyce, J.A. and J.W. Pollard, *Microenvironmental regulation of metastasis.* Nat Rev Cancer, 2009. **9**(4): p. 239-52.
114. Sleeman, J.P., *The lymph node as a bridgehead in the metastatic dissemination of tumors.* Recent Results Cancer Res, 2000. **157**: p. 55-81.
115. Brown, M., et al., *Lymph node blood vessels provide exit routes for metastatic tumor cell dissemination in mice.* Science, 2018. **359**(6382): p. 1408-1411.
116. Pereira, E.R., et al., *Lymph node metastases can invade local blood vessels, exit the node, and colonize distant organs in mice.* Science, 2018. **359**(6382): p. 1403-1407.
117. Chiang, S.P., R.M. Cabrera, and J.E. Segall, *Tumor cell intravasation.* Am J Physiol Cell Physiol, 2016. **311**(1): p. C1-C14.
118. Huang, R.L., et al., *ANGPTL4 modulates vascular junction integrity by integrin signaling and disruption of intercellular VE-cadherin and claudin-5 clusters.* Blood, 2011. **118**(14): p. 3990-4002.
119. Roh-Johnson, M., et al., *Macrophage contact induces RhoA GTPase signaling to trigger tumor cell intravasation.* Oncogene, 2014. **33**(33): p. 4203-12.
120. Harney, A.S., et al., *Real-Time Imaging Reveals Local, Transient Vascular Permeability, and Tumor Cell Intravasation Stimulated by TIE2hi Macrophage-Derived VEGFA.* Cancer Discov, 2015. **5**(9): p. 932-43.

121. Deryugina, E.I. and W.B. Kiosses, *Intratumoral Cancer Cell Intravasation Can Occur Independent of Invasion into the Adjacent Stroma.* Cell Rep, 2017. **19**(3): p. 601-616.
122. Tien, Y.W., et al., *Tumor angiogenesis and its possible role in intravasation of colorectal epithelial cells.* Clin Cancer Res, 2001. **7**(6): p. 1627-32.
123. Yamamura, T., et al., *Morphologic analysis of microvessels in colorectal tumors with respect to the formation of liver metastases.* J Surg Oncol, 2001. **78**(4): p. 259-64.
124. McDonald, D.M. and P. Baluk, *Significance of blood vessel leakiness in cancer.* Cancer Res, 2002. **62**(18): p. 5381-5.
125. Im, J.H., et al., *Coagulation facilitates tumor cell spreading in the pulmonary vasculature during early metastatic colony formation.* Cancer Res, 2004. **64**(23): p. 8613-9.
126. Nieswandt, B., et al., *Lysis of tumor cells by natural killer cells in mice is impeded by platelets.* Cancer Res, 1999. **59**(6): p. 1295-300.
127. Menczer, J., et al., *Ovarian carcinoma associated thrombocytosis. Correlation with prognostic factors and with survival.* Eur J Gynaecol Oncol, 1998. **19**(1): p. 82-4.
128. O'Keefe, S.C., et al., *Thrombocytosis is associated with a significant increase in the cancer specific death rate after radical nephrectomy.* J Urol, 2002. **168**(4 Pt 1): p. 1378-80.
129. Spigel, S.C. and L.R. Mooney, *Extreme thrombocytosis associated with malignancy.* Cancer, 1977. **39**(1): p. 339-41.
130. Nash, G.F., et al., *Platelets and cancer.* Lancet Oncol, 2002. **3**(7): p. 425-30.
131. Frisch, S.M. and H. Francis, *Disruption of epithelial cell-matrix interactions induces apoptosis.* J Cell Biol, 1994. **124**(4): p. 619-26.
132. Meng, S., et al., *Circulating tumor cells in patients with breast cancer dormancy.* Clin Cancer Res, 2004. **10**(24): p. 8152-62.
133. Cohen, S.J., 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**(19): p. 3213-21.
134. Cristofanilli, M., et al., *Circulating tumor cells, disease progression, and survival in metastatic breast cancer.* N Engl J Med, 2004. **351**(8): p. 781-91.
135. de Bono, J.S., et al., *Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer.* Clin Cancer Res, 2008. **14**(19): p. 6302-9.
136. Moreno, J.G., et al., *Changes in circulating carcinoma cells in patients with metastatic prostate cancer correlate with disease status.* Urology, 2001. **58**(3): p. 386-92.
137. Qiao, Y., et al., *Prognostic value of circulating tumor cells in the peripheral blood of patients with esophageal squamous cell carcinoma.* Onco Targets Ther, 2017. **10**: p. 1363-1373.
138. Yu, M., et al., *Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition.* Science, 2013. **339**(6119): p. 580-4.
139. Mohamed, H., et al., *Isolation of tumor cells using size and deformation.* J Chromatogr A, 2009. **1216**(47): p. 8289-95.
140. Vona, G., et al., *Isolation by size of epithelial tumor cells : a new method for the immunomorphological and molecular characterization of circulatingtumor cells.* Am J Pathol, 2000. **156**(1): p. 57-63.
141. Nagrath, S., et al., *Isolation of rare circulating tumour cells in cancer patients by microchip technology.* Nature, 2007. **450**(7173): p. 1235-9.
142. Yu, M., et al., *Circulating tumor cells: approaches to isolation and characterization.* J Cell Biol, 2011. **192**(3): p. 373-82.
143. Al-Mehdi, A.B., et al., *Intravascular origin of metastasis from the proliferation of endothelium-attached tumor cells: a new model for metastasis.* Nat Med, 2000. **6**(1): p. 100-2.
144. Gupta, G.P., et al., *Mediators of vascular remodelling co-opted for sequential steps in lung metastasis.* Nature, 2007. **446**(7137): p. 765-70.
145. Padua, D., et al., *TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4.* Cell, 2008. **133**(1): p. 66-77.
146. Qian, B.Z., et al., *CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis.* Nature, 2011. **475**(7355): p. 222-5.
147. Zhang, X.H., et al., *Latent bone metastasis in breast cancer tied to Src-dependent survival signals.* Cancer Cell, 2009. **16**(1): p. 67-78.
148. Massague, J. and A.C. Obenauf, *Metastatic colonization by circulating tumour cells.* Nature, 2016. **529**(7586): p. 298-306.
149. Eyles, J., et al., *Tumor cells disseminate early, but immunosurveillance limits metastatic outgrowth, in a mouse model of melanoma.* J Clin Invest, 2010. **120**(6): p. 2030-9.

150. Bidwell, B.N., et al., *Silencing of Irf7 pathways in breast cancer cells promotes bone metastasis through immune escape.* Nat Med, 2012. **18**(8): p. 1224-31.
151. Wang, M., et al., *Mechanism of immune evasion in breast cancer.* Onco Targets Ther, 2017. **10**: p. 1561-1573.
152. Seliger, B., M.J. Maeurer, and S. Ferrone, *Antigen-processing machinery breakdown and tumor growth.* Immunol Today, 2000. **21**(9): p. 455-64.
153. Bertucci, F., et al., *PDL1 expression in inflammatory breast cancer is frequent and predicts for the pathological response to chemotherapy.* Oncotarget, 2015. **6**(15): p. 13506-19.
154. Wang, X., et al., *CTC immune escape mediated by PD-L1.* Med Hypotheses, 2016. **93**: p. 138-9.
155. Liu, Y. and X. Cao, *Characteristics and Significance of the Pre-metastatic Niche.* Cancer Cell, 2016. **30**(5): p. 668-681.
156. Kaplan, R.N., et al., *VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche.* Nature, 2005. **438**(7069): p. 820-7.
157. Hiratsuka, S., et al., *The S100A8-serum amyloid A3-TLR4 paracrine cascade establishes a pre-metastatic phase.* Nat Cell Biol, 2008. **10**(11): p. 1349-55.
158. Deng, J., et al., *S1PR1-STAT3 signaling is crucial for myeloid cell colonization at future metastatic sites.* Cancer Cell, 2012. **21**(5): p. 642-54.
159. Chafe, S.C., et al., *Carbonic anhydrase IX promotes myeloid-derived suppressor cell mobilization and establishment of a metastatic niche by stimulating G-CSF production.* Cancer Res, 2015. **75**(6): p. 996-1008.
160. Costa-Silva, B., et al., *Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver.* Nat Cell Biol, 2015. **17**(6): p. 816-26.
161. Zhang, L., et al., *Microenvironment-induced PTEN loss by exosomal microRNA primes brain metastasis outgrowth.* Nature, 2015. **527**(7576): p. 100-104.
162. Fremder, E., et al., *Tumor-derived microparticles induce bone marrow-derived cell mobilization and tumor homing: a process regulated by osteopontin.* Int J Cancer, 2014. **135**(2): p. 270-81.
163. Clever, D., et al., *Oxygen Sensing by T Cells Establishes an Immunologically Tolerant Metastatic Niche.* Cell, 2016. **166**(5): p. 1117-1131 e14.
164. Huang, Y., et al., *Pulmonary vascular destabilization in the premetastatic phase facilitates lung metastasis.* Cancer Res, 2009. **69**(19): p. 7529-37.
165. Lee, E., et al., *Breast cancer cells condition lymphatic endothelial cells within pre-metastatic niches to promote metastasis.* Nat Commun, 2014. **5**: p. 4715.
166. Chu, G.C., et al., *RANK- and c-Met-mediated signal network promotes prostate cancer metastatic colonization.* Endocr Relat Cancer, 2014. **21**(2): p. 311-26.
167. Wu, Z., et al., *TPO-Induced Metabolic Reprogramming Drives Liver Metastasis of Colorectal Cancer CD110+ Tumor-Initiating Cells.* Cell Stem Cell, 2015. **17**(1): p. 47-59.
168. Scherz-Shouval, R., et al., *The reprogramming of tumor stroma by HSF1 is a potent enabler of malignancy.* Cell, 2014. **158**(3): p. 564-78.
169. Willis, R.A., *The Spread of Tumours in the Human Body.* J. & A. Churchill, London, 1934.
170. Klein, C.A., *Framework models of tumor dormancy from patient-derived observations.* Curr Opin Genet Dev, 2011. **21**(1): p. 42-9.
171. Demicheli, R., et al., *Breast cancer recurrence dynamics following adjuvant CMF is consistent with tumor dormancy and mastectomy-driven acceleration of the metastatic process.* Ann Oncol, 2005. **16**(9): p. 1449-57.
172. Hoffman, P.C., A.M. Mauer, and E.E. Vokes, *Lung cancer.* Lancet, 2000. **355**(9202): p. 479-85.
173. Dasgupta, A., A.R. Lim, and C.M. Ghajar, *Circulating and disseminated tumor cells: harbingers or initiators of metastasis?* Mol Oncol, 2017. **11**(1): p. 40-61.
174. Sosa, M.S., P. Bragado, and J.A. Aguirre-Ghiso, *Mechanisms of disseminated cancer cell dormancy: an awakening field.* Nat Rev Cancer, 2014. **14**(9): p. 611-22.
175. Holmgren, L., M.S. O'Reilly, and J. Folkman, *Dormancy of micrometastases: balanced proliferation and apoptosis in the presence of angiogenesis suppression.* Nat Med, 1995. **1**(2): p. 149-53.
176. Baeriswyl, V. and G. Christofori, *The angiogenic switch in carcinogenesis.* Semin Cancer Biol, 2009. **19**(5): p. 329-37.
177. Bielenberg, D.R. and B.R. Zetter, *The Contribution of Angiogenesis to the Process of Metastasis.* Cancer J, 2015. **21**(4): p. 267-73.
178. Folkman, J., *Tumor angiogenesis: therapeutic implications.* N Engl J Med, 1971. **285**(21): p. 1182-6.
179. Kienast, Y., et al., *Real-time imaging reveals the single steps of brain metastasis formation.* Nat Med, 2010. **16**(1): p. 116-22.

180. Muller-Hermelink, N., et al., *TNFR1 signaling and IFN-gamma signaling determine whether T cells induce tumor dormancy or promote multistage carcinogenesis.* Cancer Cell, 2008. **13**(6): p. 507-18.
181. Aguirre-Ghiso, J.A., et al., *ERK(MAPK) activity as a determinant of tumor growth and dormancy; regulation by p38(SAPK).* Cancer Res, 2003. **63**(7): p. 1684-95.
182. Nash, K.T., et al., *Requirement of KISS1 secretion for multiple organ metastasis suppression and maintenance of tumor dormancy.* J Natl Cancer Inst, 2007. **99**(4): p. 309-21.
183. Xu, L., et al., *GPR56, an atypical G protein-coupled receptor, binds tissue transglutaminase, TG2, and inhibits melanoma tumor growth and metastasis.* Proc Natl Acad Sci U S A, 2006. **103**(24): p. 9023-8.
184. Kim, R.S., et al., *Dormancy signatures and metastasis in estrogen receptor positive and negative breast cancer.* PLoS One, 2012. **7**(4): p. e35569.
185. Coller, H.A., L. Sang, and J.M. Roberts, *A new description of cellular quiescence.* PLoS Biol, 2006. **4**(3): p. e83.
186. Kobayashi, A., et al., *Bone morphogenetic protein 7 in dormancy and metastasis of prostate cancer stem-like cells in bone.* J Exp Med, 2011. **208**(13): p. 2641-55.
187. White, A.C., et al., *Stem cell quiescence acts as a tumour suppressor in squamous tumours.* Nat Cell Biol, 2014. **16**(1): p. 99-107.
188. Balz, L.M., et al., *The interplay of HER2/HER3/PI3K and EGFR/HER2/PLC-gamma1 signalling in breast cancer cell migration and dissemination.* J Pathol, 2012. **227**(2): p. 234-44.
189. Humtsoe, J.O. and R.H. Kramer, *Differential epidermal growth factor receptor signaling regulates anchorage-independent growth by modulation of the PI3K/AKT pathway.* Oncogene, 2010. **29**(8): p. 1214-26.
190. Jo, H., et al., *Cancer cell-derived clusterin modulates the phosphatidylinositol 3'-kinase-Akt pathway through attenuation of insulin-like growth factor 1 during serum deprivation.* Mol Cell Biol, 2008. **28**(13): p. 4285-99.
191. Schewe, D.M. and J.A. Aguirre-Ghiso, *ATF6alpha-Rheb-mTOR signaling promotes survival of dormant tumor cells in vivo.* Proc Natl Acad Sci U S A, 2008. **105**(30): p. 10519-24.
192. Aguirre-Ghiso, J.A., et al., *Urokinase receptor and fibronectin regulate the ERK(MAPK) to p38(MAPK) activity ratios that determine carcinoma cell proliferation or dormancy in vivo.* Mol Biol Cell, 2001. **12**(4): p. 863-79.
193. Ghajar, C.M., et al., *The perivascular niche regulates breast tumour dormancy.* Nat Cell Biol, 2013. **15**(7): p. 807-17.
194. Indraccolo, S., et al., *Cross-talk between tumor and endothelial cells involving the Notch3-Dll4 interaction marks escape from tumor dormancy.* Cancer Res, 2009. **69**(4): p. 1314-23.
195. Becker, S., et al., *Detection of cytokeratin-positive cells in the bone marrow of breast cancer patients undergoing adjuvant therapy.* Breast Cancer Res Treat, 2006. **97**(1): p. 91-6.
196. Hart, I.R. and I.J. Fidler, *Role of organ selectivity in the determination of metastatic patterns of B16 melanoma.* Cancer Res, 1980. **40**(7): p. 2281-7.
197. Paget, S., *The distribution of secondary growths in cancer of the breast. 1889.* Cancer Metastasis Rev, 1989. **8**(2): p. 98-101.
198. Fidler, I.J., *The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited.* Nat Rev Cancer, 2003. **3**(6): p. 453-8.
199. Albrengues, J., et al., *Neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in mice.* Science, 2018. **361**(6409).
200. Barkan, D., et al., *Inhibition of metastatic outgrowth from single dormant tumor cells by targeting the cytoskeleton.* Cancer Res, 2008. **68**(15): p. 6241-50.
201. Grivennikov, S.I., F.R. Greten, and M. Karin, *Immunity, inflammation, and cancer.* Cell, 2010. **140**(6): p. 883-99.
202. Quail, D.F. and J.A. Joyce, *Microenvironmental regulation of tumor progression and metastasis.* Nat Med, 2013. **19**(11): p. 1423-37.
203. Balkwill, F. and A. Mantovani, *Inflammation and cancer: back to Virchow?* Lancet, 2001. **357**(9255): p. 539-45.
204. Soucek, L., et al., *Mast cells are required for angiogenesis and macroscopic expansion of Myc-induced pancreatic islet tumors.* Nat Med, 2007. **13**(10): p. 1211-8.
205. Malanchi, I., et al., *Interactions between cancer stem cells and their niche govern metastatic colonization.* Nature, 2011. **481**(7379): p. 85-9.
206. Chen, W.J., et al., *Cancer-associated fibroblasts regulate the plasticity of lung cancer stemness via paracrine signalling.* Nat Commun, 2014. **5**: p. 3472.

207. Park, B.K., et al., *NF-kappaB in breast cancer cells promotes osteolytic bone metastasis by inducing osteoclastogenesis via GM-CSF.* Nat Med, 2007. **13**(1): p. 62-9.
208. Tabaries, S., et al., *Claudin-2 is selectively enriched in and promotes the formation of breast cancer liver metastases through engagement of integrin complexes.* Oncogene, 2011. **30**(11): p. 1318-28.
209. Bos, P.D., et al., *Genes that mediate breast cancer metastasis to the brain.* Nature, 2009. **459**(7249): p. 1005-9.
210. Tabaries, S., et al., *Granulocytic immune infiltrates are essential for the efficient formation of breast cancer liver metastases.* Breast Cancer Res, 2015. **17**: p. 45.
211. Edlund, M., S.Y. Sung, and L.W. Chung, *Modulation of prostate cancer growth in bone microenvironments.* J Cell Biochem, 2004. **91**(4): p. 686-705.
212. Hess, K.R., et al., *Metastatic patterns in adenocarcinoma.* Cancer, 2006. **106**(7): p. 1624-33.
213. Patanaphan, V., O.M. Salazar, and R. Risco, *Breast cancer: metastatic patterns and their prognosis.* South Med J, 1988. **81**(9): p. 1109-12.
214. Boire, A., et al., *Complement Component 3 Adapts the Cerebrospinal Fluid for Leptomeningeal Metastasis.* Cell, 2017. **168**(6): p. 1101-1113 e13.
215. Minn, A.J., et al., *Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors.* J Clin Invest, 2005. **115**(1): p. 44-55.
216. D'Arrigo, A., et al., *Metastatic transcriptional pattern revealed by gene expression profiling in primary colorectal carcinoma.* Int J Cancer, 2005. **115**(2): p. 256-62.
217. Budhu, A.S., et al., *The molecular signature of metastases of human hepatocellular carcinoma.* Oncology, 2005. **69 Suppl 1**: p. 23-7.
218. Erler, J.T., et al., *Hypoxia-induced lysyl oxidase is a critical mediator of bone marrow cell recruitment to form the premetastatic niche.* Cancer Cell, 2009. **15**(1): p. 35-44.
219. Wang, H., et al., *The osteogenic niche promotes early-stage bone colonization of disseminated breast cancer cells.* Cancer Cell, 2015. **27**(2): p. 193-210.
220. Sethi, N., et al., *Tumor-derived JAGGED1 promotes osteolytic bone metastasis of breast cancer by engaging notch signaling in bone cells.* Cancer Cell, 2011. **19**(2): p. 192-205.
221. Muller, A., et al., *Involvement of chemokine receptors in breast cancer metastasis.* Nature, 2001. **410**(6824): p. 50-6.
222. Paku, S., et al., *Organ-specificity of the extravasation process: an ultrastructural study.* Clin Exp Metastasis, 2000. **18**(6): p. 481-92.
223. Vanharanta, S. and J. Massague, *Origins of metastatic traits.* Cancer Cell, 2013. **24**(4): p. 410-21.
224. Weis, S., et al., *Endothelial barrier disruption by VEGF-mediated Src activity potentiates tumor cell extravasation and metastasis.* J Cell Biol, 2004. **167**(2): p. 223-9.
225. Kopp, H.G., et al., *The bone marrow vascular niche: home of HSC differentiation and mobilization.* Physiology (Bethesda), 2005. **20**: p. 349-56.
226. Weil, R.J., et al., *Breast cancer metastasis to the central nervous system.* Am J Pathol, 2005. **167**(4): p. 913-20.
227. Wang, H., et al., *Tumor cell alpha3beta1 integrin and vascular laminin-5 mediate pulmonary arrest and metastasis.* J Cell Biol, 2004. **164**(6): p. 935-41.
228. Gilbert, L.A. and M.T. Hemann, *DNA damage-mediated induction of a chemoresistant niche.* Cell, 2010. **143**(3): p. 355-66.
229. Wiedswang, G., et al., *Isolated tumor cells in bone marrow three years after diagnosis in disease-free breast cancer patients predict unfavorable clinical outcome.* Clin Cancer Res, 2004. **10**(16): p. 5342-8.
230. Kang, H.C., et al., *Identification of genes with differential expression in acquired drug-resistant gastric cancer cells using high-density oligonucleotide microarrays.* Clin Cancer Res, 2004. **10**(1 Pt 1): p. 272-84.
231. Engelman, J.A., et al., *MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling.* Science, 2007. **316**(5827): p. 1039-43.
232. Liu, H., L. Lv, and K. Yang, *Chemotherapy targeting cancer stem cells.* Am J Cancer Res, 2015. **5**(3): p. 880-93.
233. Volk-Draper, L., et al., *Paclitaxel therapy promotes breast cancer metastasis in a TLR4-dependent manner.* Cancer Res, 2014. **74**(19): p. 5421-34.
234. Carlson, P., et al., *Targeting the perivascular niche sensitizes disseminated tumour cells to chemotherapy.* Nat Cell Biol, 2019. **21**(2): p. 238-250.
235. Steeg, P.S., *Targeting metastasis.* Nat Rev Cancer, 2016. **16**(4): p. 201-18.

236. Palmieri, D., et al., *Medroxyprogesterone acetate elevation of Nm23-H1 metastasis suppressor expression in hormone receptor-negative breast cancer.* J Natl Cancer Inst, 2005. **97**(9): p. 632-42.
237. Koo, B.K., et al., *Porcupine inhibitor suppresses paracrine Wnt-driven growth of Rnf43;Znrf3-mutant neoplasia.* Proc Natl Acad Sci U S A, 2015. **112**(24): p. 7548-50.
238. Steinhart, Z., et al., *Genome-wide CRISPR screens reveal a Wnt-FZD5 signaling circuit as a druggable vulnerability of RNF43-mutant pancreatic tumors.* Nat Med, 2017. **23**(1): p. 60-68.
239. Junttila, M.R., et al., *Targeting LGR5+ cells with an antibody-drug conjugate for the treatment of colon cancer.* Sci Transl Med, 2015. **7**(314): p. 314ra186.
240. Feng, Y., et al., *Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis.* Genes Dis, 2018. **5**(2): p. 77-106.
241. Polyak, K., *Breast cancer: origins and evolution.* J Clin Invest, 2007. **117**(11): p. 3155-63.
242. Brennan, K., et al., *Tight junctions: a barrier to the initiation and progression of breast cancer?* J Biomed Biotechnol, 2010. **2010**: p. 460607.
243. Horton, J.K., et al., *Breast Cancer Biology: Clinical Implications for Breast Radiation Therapy.* Int J Radiat Oncol Biol Phys, 2018. **100**(1): p. 23-37.
244. Baldassarre, G. and B. Belletti, *Molecular biology of breast tumors and prognosis.* F1000Res, 2016. **5**.
245. Fallahpour, S., et al., *Breast cancer survival by molecular subtype: a population-based analysis of cancer registry data.* CMAJ Open, 2017. **5**(3): p. E734-E739.
246. Veronesi, U., et al., *Breast cancer.* Lancet, 2005. **365**(9472): p. 1727-41.
247. Kennecke, H., et al., *Metastatic behavior of breast cancer subtypes.* J Clin Oncol, 2010. **28**(20): p. 3271-7.
248. Wu, Q., et al., *Breast cancer subtypes predict the preferential site of distant metastases: a SEER based study.* Oncotarget, 2017. **8**(17): p. 27990-27996.
249. Fisher, B., et al., *Effect of preoperative chemotherapy on the outcome of women with operable breast cancer.* J Clin Oncol, 1998. **16**(8): p. 2672-85.
250. Cicatiello, L., et al., *Estrogens and progesterone promote persistent CCND1 gene activation during G1 by inducing transcriptional derepression via c-Jun/c-Fos/estrogen receptor (progesterone receptor) complex assembly to a distal regulatory element and recruitment of cyclin D1 to its own gene promoter.* Mol Cell Biol, 2004. **24**(16): p. 7260-74.
251. Liu, S., et al., *HER2 Overexpression Triggers an IL1alpha Proinflammatory Circuit to Drive Tumorigenesis and Promote Chemotherapy Resistance.* Cancer Res, 2018. **78**(8): p. 2040-2051.
252. Vogel, C.L., et al., *Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer.* J Clin Oncol, 2002. **20**(3): p. 719-26.
253. Goldhirsch, A., et al., *Meeting highlights: updated international expert consensus on the primary therapy of early breast cancer.* J Clin Oncol, 2003. **21**(17): p. 3357-65.
254. Ponomarev, V., et al., *A novel triple-modality reporter gene for whole-body fluorescent, bioluminescent, and nuclear noninvasive imaging.* Eur J Nucl Med Mol Imaging, 2004. **31**(5): p. 740-51.
255. DeRose, Y.S., et al., *Patient-derived models of human breast cancer: protocols for in vitro and in vivo applications in tumor biology and translational medicine.* Curr Protoc Pharmacol, 2013. **Chapter 14**: p. Unit14 23.
256. Fellmann, C., et al., *An optimized microRNA backbone for effective single-copy RNAi.* Cell Rep, 2013. **5**(6): p. 1704-13.
257. Knott, S.R.V., et al., *A computational algorithm to predict shRNA potency.* Mol Cell, 2014. **56**(6): p. 796-807.
258. Kallio, M.A., et al., *Chipster: user-friendly analysis software for microarray and other high-throughput data.* BMC Genomics, 2011. **12**: p. 507.
259. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.
260. Roesli, C., D. Neri, and J.N. Rybak, *In vivo protein biotinylation and sample preparation for the proteomic identification of organ- and disease-specific antigens accessible from the vasculature.* Nat Protoc, 2006. **1**(1): p. 192-9.
261. Lawson, D.A., et al., *Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells.* Nature, 2015. **526**(7571): p. 131-5.
262. Comsa, S., A.M. Cimpean, and M. Raica, *The Story of MCF-7 Breast Cancer Cell Line: 40 years of Experience in Research.* Anticancer Res, 2015. **35**(6): p. 3147-54.

263. Kim, J.Y., et al., *The value of phosphohistone H3 as a proliferation marker for evaluating invasive breast cancers: A comparative study with Ki67.* Oncotarget, 2017. **8**(39): p. 65064-65076.
264. Ma, J., et al., *SPARC inhibits breast cancer bone metastasis and may be a clinical therapeutic target.* Oncol Lett, 2017. **14**(5): p. 5876-5882.
265. Crowley, L.C. and N.J. Waterhouse, *Detecting Cleaved Caspase-3 in Apoptotic Cells by Flow Cytometry.* Cold Spring Harb Protoc, 2016. **2016**(11).
266. Rokavec, M., D. Horst, and H. Hermeking, *Cellular Model of Colon Cancer Progression Reveals Signatures of mRNAs, miRNA, lncRNAs, and Epigenetic Modifications Associated with Metastasis.* Cancer Res, 2017. **77**(8): p. 1854-1867.
267. Jechlinger, M., et al., *Expression profiling of epithelial plasticity in tumor progression.* Oncogene, 2003. **22**(46): p. 7155-69.
268. Aigner, K., et al., *The transcription factor ZEB1 (deltaEF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity.* Oncogene, 2007. **26**(49): p. 6979-88.
269. Warburg, O., *On the origin of cancer cells.* Science, 1956. **123**(3191): p. 309-14.
270. Zhang, S., et al., *The pivotal role of pyruvate dehydrogenase kinases in metabolic flexibility.* Nutr Metab (Lond), 2014. **11**(1): p. 10.
271. Mootha, V.K., et al., *PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.* Nat Genet, 2003. **34**(3): p. 267-73.
272. Schagger, H. and K. Pfeiffer, *Supercomplexes in the respiratory chains of yeast and mammalian mitochondria.* EMBO J, 2000. **19**(8): p. 1777-83.
273. Wheaton, W.W. and N.S. Chandel, *Hypoxia. 2. Hypoxia regulates cellular metabolism.* Am J Physiol Cell Physiol, 2011. **300**(3): p. C385-93.
274. Elvidge, G.P., et al., *Concordant regulation of gene expression by hypoxia and 2-oxoglutarate-dependent dioxygenase inhibition: the role of HIF-1alpha, HIF-2alpha, and other pathways.* J Biol Chem, 2006. **281**(22): p. 15215-26.
275. Fardin, P., et al., *A biology-driven approach identifies the hypoxia gene signature as a predictor of the outcome of neuroblastoma patients.* Mol Cancer, 2010. **9**: p. 185.
276. Semenza, G.L., *Hypoxia-inducible factor 1: oxygen homeostasis and disease pathophysiology.* Trends Mol Med, 2001. **7**(8): p. 345-50.
277. Schofield, C.J. and P.J. Ratcliffe, *Oxygen sensing by HIF hydroxylases.* Nat Rev Mol Cell Biol, 2004. **5**(5): p. 343-54.
278. Mantegazza, A.R., et al., *Presentation of phagocytosed antigens by MHC class I and II.* Traffic, 2013. **14**(2): p. 135-52.
279. Laboratory, T.J., https://www.jax.org/jax-mice-and-services/find-and-order-jax-mice/nsg-portfolio.
280. Shen, Z. and B.Z. Stanger, *YAP regulates S-phase entry in endothelial cells.* PLoS One, 2015. **10**(1): p. e0117522.
281. Janse van Rensburg, H.J., et al., *The Hippo Pathway Component TAZ Promotes Immune Evasion in Human Cancer through PD-L1.* Cancer Res, 2018. **78**(6): p. 1457-1470.
282. Moroishi, T., et al., *The Hippo Pathway Kinases LATS1/2 Suppress Cancer Immunity.* Cell, 2016. **167**(6): p. 1525-1539 e17.
283. Zhang, Y., H. Zhang, and B. Zhao, *Hippo Signaling in the Immune System.* Trends Biochem Sci, 2018. **43**(2): p. 77-80.
284. Meng, Z., T. Moroishi, and K.L. Guan, *Mechanisms of Hippo pathway regulation.* Genes Dev, 2016. **30**(1): p. 1-17.
285. Zhao, B., et al., *Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control.* Genes Dev, 2007. **21**(21): p. 2747-61.
286. Balko, J.M., et al., *Profiling of residual breast cancers after neoadjuvant chemotherapy identifies DUSP4 deficiency as a mechanism of drug resistance.* Nat Med, 2012. **18**(7): p. 1052-9.
287. Stipp, C.S., T.V. Kolesnikova, and M.E. Hemler, *Functional domains in tetraspanin proteins.* Trends Biochem Sci, 2003. **28**(2): p. 106-12.
288. Fentiman, I.S., R.D. Rubens, and J.L. Hayward, *Ascites in breast cancer.* Br Med J (Clin Res Ed), 1983. **287**(6398): p. 1023.
289. Zeng, A., et al., *Prospectively Isolated Tetraspanin(+) Neoblasts Are Adult Pluripotent Stem Cells Underlying Planaria Regeneration.* Cell, 2018. **173**(7): p. 1593-1608 e20.
290. Reynolds, B.A. and S. Weiss, *Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system.* Science, 1992. **255**(5052): p. 1707-10.
291. Stingl, J., et al., *Purification and unique properties of mammary epithelial stem cells.* Nature, 2006. **439**(7079): p. 993-7.
292. Smith, B.A., et al., *A Human Adult Stem Cell Signature Marks Aggressive Variants across Epithelial Cancers.* Cell Rep, 2018. **24**(12): p. 3353-3366 e5.

293. Pece, S., et al., *Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content.* Cell, 2010. **140**(1): p. 62-73.
294. Suzuki, M., et al., *Dormant cancer cells retrieved from metastasis-free organs regain tumorigenic and metastatic potency.* Am J Pathol, 2006. **169**(2): p. 673-81.
295. Putz, E., et al., *Phenotypic characteristics of cell lines derived from disseminated cancer cells in bone marrow of patients with solid epithelial tumors: establishment of working models for human micrometastases.* Cancer Res, 1999. **59**(1): p. 241-8.
296. Malladi, S., et al., *Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT.* Cell, 2016. **165**(1): p. 45-60.
297. Kumar, S., et al., *Intra-Tumoral Metabolic Zonation and Resultant Phenotypic Diversification Are Dictated by Blood Vessel Proximity.* Cell Metab, 2019.
298. Gerlee, P., *The model muddle: in search of tumor growth laws.* Cancer Res, 2013. **73**(8): p. 2407-11.
299. Baudoux, T.E.R., et al., *Donor Cancer Transmission in Kidney Transplantation.* Kidney Int Rep, 2017. **2**(2): p. 134-137.
300. Chapman, J.R., A.C. Webster, and G. Wong, *Cancer in the transplant recipient.* Cold Spring Harb Perspect Med, 2013. **3**(7).
301. Desai, R., et al., *Cancer transmission from organ donors-unavoidable but low risk.* Transplantation, 2012. **94**(12): p. 1200-7.
302. Matser, Y.A.H., et al., *Transmission of breast cancer by a single multiorgan donor to 4 transplant recipients.* Am J Transplant, 2018. **18**(7): p. 1810-1814.
303. Penn, I., *Transmission of cancer from organ donors.* Ann Transplant, 1997. **2**(4): p. 7-12.
304. MacKie, R.M., R. Reid, and B. Junor, *Fatal melanoma transferred in a donated kidney 16 years after melanoma surgery.* N Engl J Med, 2003. **348**(6): p. 567-8.
305. Baker, S.M., R.W. Buckheit, 3rd, and M.M. Falk, *Green-to-red photoconvertible fluorescent proteins: tracking cell and protein dynamics on standard wide-field mercury arc-based microscopes.* BMC Cell Biol, 2010. **11**: p. 15.
306. Stathopoulou, A., et al., *Molecular detection of cytokeratin-19-positive cells in the peripheral blood of patients with operable breast cancer: evaluation of their prognostic significance.* J Clin Oncol, 2002. **20**(16): p. 3404-12.
307. Braun, S., et al., *Lack of effect of adjuvant chemotherapy on the elimination of single dormant tumor cells in bone marrow of high-risk breast cancer patients.* J Clin Oncol, 2000. **18**(1): p. 80-6.
308. Wiedswang, G., et al., *Detection of isolated tumor cells in bone marrow is an independent prognostic factor in breast cancer.* J Clin Oncol, 2003. **21**(18): p. 3469-78.
309. Plaks, V., C.D. Koopman, and Z. Werb, *Cancer. Circulating tumor cells.* Science, 2013. **341**(6151): p. 1186-8.
310. Chen, W., et al., *Organotropism: new insights into molecular mechanisms of breast cancer metastasis.* NPJ Precis Oncol, 2018. **2**(1): p. 4.
311. Echeverria, G.V., et al., *High-resolution clonal mapping of multi-organ metastasis in triple negative breast cancer.* Nat Commun, 2018. **9**(1): p. 5079.
312. Pascual, G., D. Dominguez, and S.A. Benitah, *The contributions of cancer cell metabolism to metastasis.* Dis Model Mech, 2018. **11**(8).
313. Sancho, P., et al., *MYC/PGC-1alpha Balance Determines the Metabolic Phenotype and Plasticity of Pancreatic Cancer Stem Cells.* Cell Metab, 2015. **22**(4): p. 590-605.
314. Viale, A., et al., *Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function.* Nature, 2014. **514**(7524): p. 628-32.
315. Viale, A., D. Corti, and G.F. Draetta, *Tumors and mitochondrial respiration: a neglected connection.* Cancer Res, 2015. **75**(18): p. 3685-6.
316. Basnet, H., et al., *Flura-seq identifies organ-specific metabolic adaptations during early metastatic colonization.* Elife, 2019. **8**.
317. Semenza, G.L., *Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics.* Oncogene, 2010. **29**(5): p. 625-34.
318. Dong, C., et al., *Loss of FBP1 by Snail-mediated repression provides metabolic advantages in basal-like breast cancer.* Cancer Cell, 2013. **23**(3): p. 316-31.
319. Lacroix, M. and G. Leclercq, *Relevance of breast cancer cell lines as models for breast tumours: an update.* Breast Cancer Res Treat, 2004. **83**(3): p. 249-89.
320. Jo, M., et al., *Reversibility of epithelial-mesenchymal transition (EMT) induced in breast cancer cells by activation of urokinase receptor-dependent cell signaling.* J Biol Chem, 2009. **284**(34): p. 22825-33.

321. Kim, J., et al., *EGF induces epithelial-mesenchymal transition through phospho-Smad2/3-Snail signaling pathway in breast cancer cells.* Oncotarget, 2016. **7**(51): p. 85021-85032.
322. Biddle, A., et al., *Cancer stem cells in squamous cell carcinoma switch between two distinct phenotypes that are preferentially migratory or proliferative.* Cancer Res, 2011. **71**(15): p. 5317-26.
323. Yu, J.S. and W. Cui, *Proliferation, survival and metabolism: the role of PI3K/AKT/mTOR signalling in pluripotency and cell fate determination.* Development, 2016. **143**(17): p. 3050-60.
324. Watts, C., *Capture and processing of exogenous antigens for presentation on MHC molecules.* Annu Rev Immunol, 1997. **15**: p. 821-50.
325. Schmidt, M. and J.R. Lill, *MHC class I presented antigens from malignancies: A perspective on analytical characterization & immunogenicity.* J Proteomics, 2019. **191**: p. 48-57.
326. Marincola, F.M., et al., *Escape of human solid tumors from T-cell recognition: molecular mechanisms and functional significance.* Adv Immunol, 2000. **74**: p. 181-273.
327. Garrido, F. and I. Algarra, *MHC antigens and tumor escape from immune surveillance.* Adv Cancer Res, 2001. **83**: p. 117-58.
328. Pantel, K., et al., *Frequent down-regulation of major histocompatibility class I antigen expression on individual micrometastatic carcinoma cells.* Cancer Res, 1991. **51**(17): p. 4712-5.
329. Zia, A., F.W. Schildberg, and I. Funke, *MHC class I negative phenotype of disseminated tumor cells in bone marrow is associated with poor survival in R0M0 breast cancer patients.* Int J Cancer, 2001. **93**(4): p. 566-70.
330. Thibodeau, J., M.C. Bourgeois-Daigneault, and R. Lapointe, *Targeting the MHC Class II antigen presentation pathway in cancer immunotherapy.* Oncoimmunology, 2012. **1**(6): p. 908-916.
331. Oldford, S.A., et al., *HLA-DRB alleles are differentially expressed by tumor cells in breast carcinoma.* Int J Cancer, 2004. **112**(3): p. 399-406.
332. Nizar, S., et al., *T-regulatory cell modulation: the future of cancer immunotherapy?* Br J Cancer, 2009. **100**(11): p. 1697-703.
333. Marty Pyke, R., et al., *Evolutionary Pressure against MHC Class II Binding Cancer Mutations.* Cell, 2018. **175**(7): p. 1991.
334. de Charette, M., A. Marabelle, and R. Houot, *Turning tumour cells into antigen presenting cells: The next step to improve cancer immunotherapy?* Eur J Cancer, 2016. **68**: p. 134-147.
335. Kalbasi, A. and A. Ribas, *Antigen Presentation Keeps Trending in Immunotherapy Resistance.* Clin Cancer Res, 2018. **24**(14): p. 3239-3241.
336. Liao, W., et al., *KRAS-IRF2 Axis Drives Immune Suppression and Immune Therapy Resistance in Colorectal Cancer.* Cancer Cell, 2019. **35**(4): p. 559-572 e7.
337. Miao, Y., et al., *Adaptive Immune Resistance Emerges from Tumor-Initiating Stem Cells.* Cell, 2019. **177**(5): p. 1172-1186 e14.
338. Guo, X., et al., *Single tumor-initiating cells evade immune clearance by recruiting type II macrophages.* Genes Dev, 2017. **31**(3): p. 247-259.
339. Kim, W., et al., *Hepatic Hippo signaling inhibits protumoural microenvironment to suppress hepatocellular carcinoma.* Gut, 2018. **67**(9): p. 1692-1703.
340. Taha, Z., H.J. Janse van Rensburg, and X. Yang, *The Hippo Pathway: Immunity and Cancer.* Cancers (Basel), 2018. **10**(4).
341. Hemler, M.E., *Tetraspanin proteins promote multiple cancer stages.* Nat Rev Cancer, 2014. **14**(1): p. 49-60.
342. Zoller, M., *Tetraspanins: push and pull in suppressing and promoting metastasis.* Nat Rev Cancer, 2009. **9**(1): p. 40-55.
343. van Deventer, S.J., V.E. Dunlock, and A.B. van Spriel, *Molecular interactions shaping the tetraspanin web.* Biochem Soc Trans, 2017. **45**(3): p. 741-750.
344. Seigneuret, M., et al., *Structure of the tetraspanin main extracellular domain. A partially conserved fold with a structurally variable domain insertion.* J Biol Chem, 2001. **276**(43): p. 40055-64.
345. Tejera, E., et al., *CD81 regulates cell migration through its association with Rac GTPase.* Mol Biol Cell, 2013. **24**(3): p. 261-73.
346. Termini, C.M., K.A. Lidke, and J.M. Gillette, *Tetraspanin CD82 Regulates the Spatiotemporal Dynamics of PKCalpha in Acute Myeloid Leukemia.* Sci Rep, 2016. **6**: p. 29859.
347. Cailleteau, L., et al., *alpha2beta1 integrin controls association of Rac with the membrane and triggers quiescence of endothelial cells.* J Cell Sci, 2010. **123**(Pt 14): p. 2491-501.
348. Lapalombella, R., et al., *Tetraspanin CD37 directly mediates transduction of survival and apoptotic signals.* Cancer Cell, 2012. **21**(5): p. 694-708.
349. Bonnet, M., et al., *Targeting the Tetraspanins with Monoclonal Antibodies in Oncology: Focus on Tspan8/Co-029.* Cancers (Basel), 2019. **11**(2).

350. Gesierich, S., et al., *Systemic induction of the angiogenesis switch by the tetraspanin D6.1A/CO-029.* Cancer Res, 2006. **66**(14): p. 7083-94.
351. Herlevsen, M., et al., *The association of the tetraspanin D6.1A with the alpha6beta4 integrin supports cell motility and liver metastasis formation.* J Cell Sci, 2003. **116**(Pt 21): p. 4373-90.
352. Kuhn, S., et al., *A complex of EpCAM, claudin-7, CD44 variant isoforms, and tetraspanins promotes colorectal cancer progression.* Mol Cancer Res, 2007. **5**(6): p. 553-67.
353. Huerta, S., et al., *Gene expression profile of metastatic colon cancer cells resistant to cisplatin-induced apoptosis.* Int J Oncol, 2003. **22**(3): p. 663-70.
354. Kanetaka, K., et al., *Possible involvement of tetraspanin CO-029 in hematogenous intrahepatic metastasis of liver cancer cells.* J Gastroenterol Hepatol, 2003. **18**(11): p. 1309-14.
355. Yue, S., et al., *The tetraspanins CD151 and Tspan8 are essential exosome components for the crosstalk between cancer initiating cells and their surrounding.* Oncotarget, 2015. **6**(4): p. 2366-84.
356. Zhao, K., et al., *Tspan8 and Tspan8/CD151 knockout mice unravel the contribution of tumor and host exosomes to tumor progression.* J Exp Clin Cancer Res, 2018. **37**(1): p. 312.
357. Lu, J., et al., *Exosomal tetraspanins mediate cancer metastasis by altering host microenvironment.* Oncotarget, 2017. **8**(37): p. 62803-62815.
358. Li, L., et al., *Quantitative proteomics analysis of the role of tetraspanin-8 in the drug resistance of gastric cancer.* Int J Oncol, 2018. **52**(2): p. 473-484.
359. Park, C.S., et al., *Therapeutic targeting of tetraspanin8 in epithelial ovarian cancer invasion and metastasis.* Oncogene, 2016. **35**(34): p. 4540-8.
360. Ailane, N., et al., *Effect of an anti-human Co-029/tspan8 mouse monoclonal antibody on tumor growth in a nude mouse model.* Front Physiol, 2014. **5**: p. 364.
361. Wollscheid, V., et al., *Identification of a new proliferation-associated protein NET-1/C4.8 characteristic for a subset of high-grade cervical intraepithelial neoplasia and cervical carcinomas.* Int J Cancer, 2002. **99**(6): p. 771-5.
362. Scholz, C.J., et al., *Tspan-1 is a tetraspanin preferentially expressed by mucinous and endometrioid subtypes of human ovarian carcinomas.* Cancer Lett, 2009. **275**(2): p. 198-203.
363. Chen, L., et al., *Association of NET-1 gene expression with human hepatocellular carcinoma.* Int J Surg Pathol, 2007. **15**(4): p. 346-53.
364. Munkley, J., et al., *The cancer-associated cell migration protein TSPAN1 is under control of androgens and its upregulation increases prostate cancer cell migration.* Sci Rep, 2017. **7**(1): p. 5249.
365. Wang, Y., et al., *Tetraspanin 1 promotes epithelial-to-mesenchymal transition and metastasis of cholangiocarcinoma via PI3K/AKT signaling.* J Exp Clin Cancer Res, 2018. **37**(1): p. 300.
366. Zhang, X., et al., *TSPAN1 upregulates MMP2 to promote pancreatic cancer cell migration and invasion via PLCgamma.* Oncol Rep, 2019.
367. Hou, F.Q., et al., *Tetraspanin 1 is involved in survival, proliferation and carcinogenesis of pancreatic cancer.* Oncol Rep, 2015. **34**(6): p. 3068-76.
368. Desouki, M.M., et al., *Identification of metastasis-associated breast cancer genes using a high-resolution whole genome profiling approach.* J Cancer Res Clin Oncol, 2011. **137**(5): p. 795-809.
369. Voglstaetter, M., et al., *Tspan8 is expressed in breast cancer and regulates E-cadherin / catenin signalling and metastasis accompanied by increased circulating extracellular vesicles.* J Pathol, 2019.
370. Nazarenko, I., et al., *Cell surface tetraspanin Tspan8 contributes to molecular pathways of exosome-induced endothelial cell activation.* Cancer Res, 2010. **70**(4): p. 1668-78.
371. Radisky, E.S., M. Raeeszadeh-Sarmazdeh, and D.C. Radisky, *Therapeutic Potential of Matrix Metalloproteinase Inhibition in Breast Cancer.* J Cell Biochem, 2017. **118**(11): p. 3531-3548.
372. Lowy, C.M. and T. Oskarsson, *Tenascin C in metastasis: A view from the invasive front.* Cell Adh Migr, 2015. **9**(1-2): p. 112-24.
373. Insua-Rodriguez, J., et al., *Stress signaling in breast cancer cells induces matrix components that promote chemoresistant metastasis.* EMBO Mol Med, 2018. **10**(10).
374. Aguirre-Ghiso, J.A., P. Bragado, and M.S. Sosa, *Metastasis awakening: targeting dormant cancer.* Nat Med, 2013. **19**(3): p. 276-7.
375. Stoecklein, N.H., et al., *Direct genetic analysis of single disseminated cancer cells for prediction of outcome and therapy selection in esophageal cancer.* Cancer Cell, 2008. **13**(5): p. 441-53.
376. Ghajar, C.M., *Metastasis prevention by targeting the dormant niche.* Nat Rev Cancer, 2015. **15**(4): p. 238-47.
377. Lim, E., et al., *Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways.* Breast Cancer Res, 2010. **12**(2): p. R21.

378. Smith, B.A., et al., *A basal stem cell signature identifies aggressive prostate cancer phenotypes.* Proc Natl Acad Sci U S A, 2015. **112**(47): p. E6544-52.

12. Contributions

This work would not have been possible without the help and support of several persons

Jasmin Meier helped with animal affairs and assisted in several *in vivo* and *in vitro* experiments

Dr Maren Pein established and provided the protocols for FACS isolation of breast cancer cells from organs

Dr Arnaud Descot generated the miRE knockdown backbones.

Prof Alana Welm and her lab generously provided the PDX models

The pleural effusion samples were provided by **Prof. Dr. Med. Marc Sütterlin** and **Dr. Med. Saskia Speck** from the University Hospital of Mannheim as well as **Prof. Dr. Med. Andreas Schneeweiss** from the National Center for Tumor disease (NCT) in Heidelberg.

Dr Markus Eich, **Tobias Rubner**, **Klaus Hexel** and **Dr Steffen Schmitt** from the DKFZ flow cytometry core facility supported and performed some of the FACS experiments.

Dr Damir Krunic supported the microscopy experiments and wrote the Macros for picture analysis

Tatjana Schmidt and the team of the gene profiling core facility of the DKFZ performed the gene expression analysis on Affymetrix arrays

Dr Thordur Oskarsson helped designing the experiments

Miriam Butler, Dr Ronny Schmidt and **Dr Thordur Oskarsson** critically proof read this manuscript.

13. Acknowledgement

In the more than four years of my PhD training I grew on a scientific as well as on a personal level and this is not least thanks to the people I was able to spend that time with. I´m deeply grateful for the experiences and for the personal and scientific support I received in the past years.

I would firstly like to thank **Dr Thordur Oskarsson** for choosing me as a Master Student and having enough trust in me to expand my research stay in his lab to my PhD thesis. Thank you that you always supported my work, encouraged me to never give up and always had an open door. I truly believe that the many critical discussions we had, significantly shaped the project.

Special thanks go to **Prof Andreas Trumpp**. With Hi-Stem you really did put together an excellent scientific working environment that you did not only establish but also maintain and expand throughout the years. Thank you for allowing me to be a part of it. Thank you for your valuable scientific input during lab meetings and my TAC meetings and for your interest and believe in my project.

Thank you to **Prof Dr Claus R. Bartram** for being part of my TAC committee, second thesis examiner as well as head of my Defense committee. Your scientific input during the TAC meetings improved and shaped my project significantly. Thank you for your time, even though your retirement would have been more than deserved. Your kind and gentle personality made every meeting very enjoyable.

I also would like to thank **Prof Stefan Wiemann** and **Dr Richard Harbottle** for participating in my PhD defense committee. I know that time is always limited, thus I´m very grateful that you took that time for me.

The most time during the last years I have for sure spent with the *Oskarssons*. I feel very lucky to be able to share my scientific life with such amazing people. **Jasmin**, **Maren**, **Tsunaki**, **Jacob**, **Camille**, **Arnaud**, **Angi**, **Miriam** and **Lena** you are not only great scientists but also amazing personalities and I´m very grateful that I worked, discussed, laughed, and celebrated with all of you during these years. Thank you, **Angi,** for sharing the office with me in the last year. I really appreciated your scientific input and enjoyed our discussions. **Jasmin**, you introduced me to the lab, showed me how things work here and always supported me. Thank you for your support and help especially with mouse regulatory work, organization and cell culture. And especially thank you for keeping the lab up and running. Without you our freezers would be a total mess! Thank you, **Jacob,** for your positive attitude and your cheerful mind. It was always very nice working with you. Your Spanish way of looking at things was more than welcome in the stringent German working day. Thank you for your input in the project and the valuable advice you gave me for many experiments. **Maren**, I very much enjoyed working with you. Not only were your scientific advises always very helpful and you very always very open for collaborations and sharing of protocols but also was your never-giving up attitude a great model for me. Thank you for your support of the last years. **Tsunaki**, I´m very happy that you joined our lab. Thank you for being such an open minded and positive person and for all the fun we had together. I also really appreciated your scientific input and advice and enjoyed discussing science with you. **Arnaud**, you supervised my Master projects and together with Camille we shared a desk for a very long time. I´m very grateful for the time we worked together in the lab and for getting to know you as a

person. You taught me a lot not only about cloning and high throughput experiments but also on a personal level. Thank you for always being supportive even beyond your time in our lab. **Camille**, during the time we spend next to each other, you became a very close friend to me and I wouldn´t want to miss you anymore! I very much enjoy spending time with you. Thank you for all your mental support and help over the years and for always being there!
Although our time in the Oskarsson lab overlapped only for few weeks, you became a really close friend to me, **Miriam**. Often there are weeks or months between our visits but always it feels as if we met yesterday. I enjoy our scientific discussions as much as our private ones. I would also like to especially thank you for putting so much effort in reviewing my thesis on a quite short notice. **Lena**, I enjoyed a lot the time you spent in our lab. Through your positive attitude it was always a lot of fun working with you!

I very much appreciated the collaborative and open atmosphere in whole Hi-Stem and I would like to thank all of you! Every helpful suggestions, nice word and nice, funny and supportive discussion over the years helped me a lot. Just to name a few, I would like to thank **Corinna** for organizing so much around here, taking care of all our orders and supporting me whenever possible. But also thank you for being the straightforward person who always says what she thinks. I really appreciate this attitude! **Franzy**, although you are not part of the *Oskarssons,* you were an essential part of our social group. Thank you for the many discussions and all the fun we had. I enjoyed a lot the time we spent together in and outsite the lab. **Lisa**, with you I shared almost the whole journey. We did our Master thesis in parallel, were in the same selection round and then did our PhD site by site. I enjoyed our time together in the lab and during dinners, ice scatting and painting and really appreciate that I can call you a friend. I would also like to thank **Pia**, **Andrea**, **Sarah Jane**, **Felix** and **Vera** for the nice time we spent together. Special thanks go also to **Erika** and **Dagmar**. You make our lives in the lab so much easier by dealing with all the bureaucratic issues when don't understand. Dagmar, thank you also for supporting me in finding new defense dates over and over again and for never giving up.

I would like to thank the Helmholtz International Graduate School, especially **Lindsay** for supporting me during my entire PhD and for many helpful advises.

Within the DKFZ, the core facilities play an essential role for the great working environment. I would especially like to thank **Steffen**, **Markus**, **Klaus**, **Tobias** and **Florian** from the FACS core facility, **Damir** from the Microscopy core facility and **Tatjana** from the Gene expression profiling core facility for their great help and support during the last years.

Thank you, **Ann-Christin**, **Juliane** and **Kathi**. Together with Lisa we met during the DKFZ selection round and became friends. Our discussions and your support always helped and encouraged me during these years

In addition, I would like to thank the Sciomics team. Already since five years I can call myself part of the "Sciomics family" and I´m very happy about that. **Christoph**, I´m very grateful that you gave me the chance to get to know a different site of science. Thank you for trusting me enough to represent your work during different conferences and meetings and for teaching me so much. I always enjoyed spending time with you and the whole team and had a lot of fun hiking, canoeing, climbing and eating with all of you. Thank you, **Katharina**, **Iuliia**, **Ramona**, **Anne**, **Florian** and **Jacqueline** for always being supportive and great fun. Many of you became close friends.

13. Acknowledgement

Special thanks go to the “awesome potatoes”. Almost since I started my PhD, I have been part of this awesome, crazy theater group. Your company was always the most efficient way to cheer me up after a long and disappointing working day. Thank you, **Veli,** for establishing and maintaining this great community.

Lastly and most importantly I would like to thank my family and closest friends. **Mama**, **Papa**, **André** und **Oma**, ich weiß, dass es nicht immer einfach für euch war zu verstehen was ich hier eigentlich mache und wieso das alles so lange dauert. Aber ich bin euch sehr dankbar, dass ihr immer an mich geglaubt habt und mich bei allem unterstützt, was ich mir in den Kopf gesetzt hab. Es tut sehr gut zu wissen eine so starke und wohlwollende Familie im Rücken zu haben, auf die man sich imme verlassen kann. Danke dass ihr immer für mich da seid. **Patin**, Danke, dass du mich damals bei dir auf genommen hast als ich nach Heidelberg gekommen bin. Du bist meine „Familie in der Ferne“ und ich danke dir für die vielen Gespräche und die schöne gemeinsame Zeit. **Chrissi**, seit 25 Jahren darf ich dich meine Freundin nennen und auch wenn wir uns jetzt nicht mehr täglich sehen, gehörst du zu den wichtigsten Menschen in meinem Leben. Du hast mich über die Jahre bei so vielem unterstütz, hast mir immer zu gehört und warst immer für mich da. Es ist nicht selbst verständlich eine solche Freundin zu haben! Vielen Dank dass ich immer auf dich zählen kann. **Ronny**, ich bin sehr froh dich an meiner Seite zu haben. Du unterstützt mich nicht nur wissenschaftlich als „wandelndes Wikipedia“, sondern bist, und das ist viel wichtiger, auch privat immer für mich. Danke für deine stetige Unterstützung in allen Situation!

.