TESE DE DOUTORAMENTO

FUNCTIONAL CHARACTERIZATION OF CIRCULATING TUMOUR CELL (CTC) CLUSTERS IN BREAST CANCER

Inés Martínez Pena

ESCOLA DE DOUTORAMENTO INTERNACIONAL DA UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

PROGRAMA DE DOUTORAMENTO EN MEDICINA MOLECULAR

SANTIAGO DE COMPOSTELA

2021

DECLARACIÓN DO AUTOR/A DA TESE

D./Dna. **Inés Martínez Pena**

Título da tese: **Functional Characterization of Circulating Tumour Cell (CTC) Clusters in Breast Cancer**

Presento a miña tese, seguindo o procedemento axeitado ao regulamento, e declaro que:

1) A tese abarca os resultados da elaboración do meu traballo.
2) De ser o caso, na tese faise referencia ás colaboracións que tivo este traballo.
3) Confirmo que a tese non incorre en ningún tipo de plaxio doutros autores nin de traballos presentados por min para a obtención doutros títulos.
4) A tese é a versión definitiva presentada para a súa defensa e coincide a versión impresa coa presentada en formato electrónico.

E comprométome a presentar o Compromiso Documental de Supervisión no caso de que o orixinal non estea na Escola.

En **Santiago de Compostela, 05 de Agosto de 2021.**

Sinatura electrónica

AUTORIZACIÓN DO DIRECTOR / TITOR DA TESE

Functional Characterization of Circulating Tumour Cell (CTC) Clusters in Breast Cancer

D./Dna. Roberto Piñeiro Cid

D./Dna. Rafael López López

INFORMA/N:

Que a presente tese, correspóndese co traballo realizado por D/Dna. Inés Martínez Pena, baixo a miña dirección/titorización, e autorizo a súa presentación, considerando que reúne os requisitos esixidos no Regulamento de Estudos de Doutoramento da USC, e que como director desta non incorre nas causas de abstención establecidas na Lei 40/2015.

De acordo co indicado no Regulamento de Estudos de Doutoramento, declara tamén que a presente tese de doutoramento é idónea para ser defendida en base á modalidade de Monográfica con reprodución de publicacions, nos que a participación da doutoranda foi decisiva para a súa elaboración e as publicacións se axustan ao Plan de Investigación.

En Santiago de Compostela, 5 de Agosto de 2021

I, Inés Martínez Pena, author of this thesis, have no conflict of interest to declare.
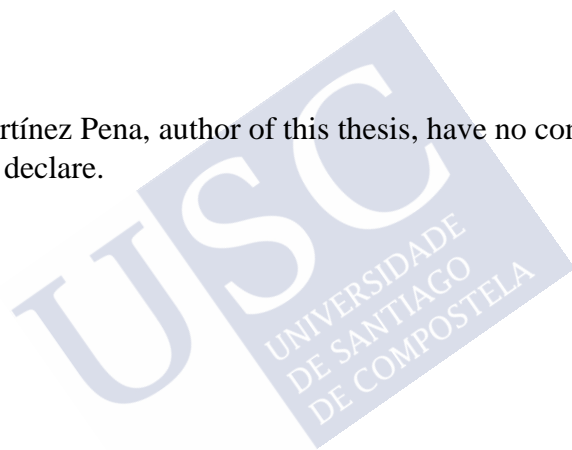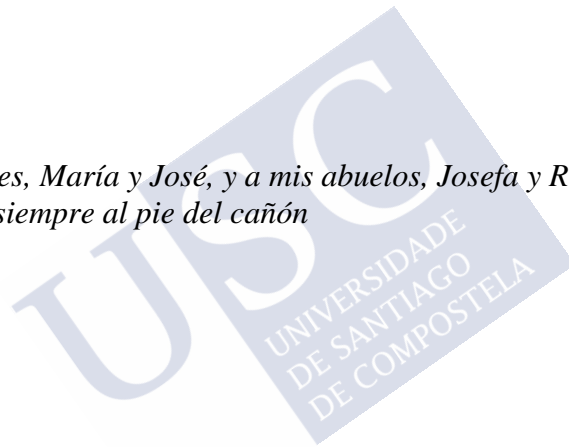
A mi padres, María y José, y a mis abuelos, Josefa y Ricardo,
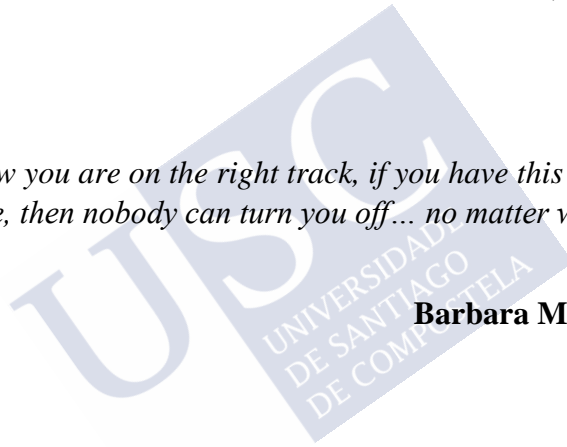*por estar siempre al pie del cañón*

Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we fear less.

Marie curie

If you know you are on the right track, if you have this inner knowledge, then nobody can turn you off… no matter what they say.

Barbara McClintock

INDEX

INDEX

ABBREVIATIONS

ABBREVIATIONS

AP Anatomopathological

BC Breast Cancer

bFGF basic Fibroblast Growth Factor

cDNA complementary DNA

CI Confidence Interval

CK cytokeratin

CTC Circulating Tumour Cell

DEP Dielectrophoresis

DLA Diagnostic Leukapheresis

dpi days post-injection

DoC Duct of Cuvier

ECM Extracellular Matrix

ECOG Eastern Cooperative Oncology Group

EGF Epidermal Growth Factor

eGFP enhanced Green Fluorescent Protein

EGFR Epidermal Growth Factor Receptor

EMT Epithelial to Mesenchymal transition

EpCAM Epithelial Cell Adhesion Molecule

ER Estrogen Receptor

FBS Fetal Bovine Serum

FDA Food and Drug Administration

FOV Field Of View

FPKM Fragments per Kilobase of transcript per Million mapped reads

FSS Fluid Shear Stress

GAPDH Glycerlaldehyde-3-Phosphate Dehydrogenase

GEP Gene Expression Profiling

GFR Growth Factor Reduced

He expected Heterozygosity

HER2 Human Epidermal growth factor Receptor 2

hpf hours post-fertilization

hpi hours post-injection

HR Hazard Ratio

HR Hormone Receptor

IHC Immunohistochemistry

Luc Luciferase

mBC metastatic Breast Cancer

mCTC murine model CTC derived

MET Mesenchymal to Epithelial Transition

M-MLV RT Moloney Murine Leukaemia Virus Reverse Transcriptase

MMP Matrix Metalloproteinase

MNCs Mononuclear Cells

MTA Material Transfer Agreement

NGS Next Generation Sequencing

NHG Nottingham Histological Grade

OS Overall Survival

P Passage

P/S Penicillin/Streptomycin

PBMC Peripheral Blood Mononuclear Cells

PBS Phosphate Buffered Saline

PDX Patient Derived Xenograft

PET Polyethylene terephthalate

PFA Paraformaldehyde

PFS Progression-free survival

PP polypropylene

PR Progesterone Receptor

PT Primary Tumour

PTU N-phenylthiourea

PVP Polyvinylpyrrolidone

qPCR/RT-qPCR quantitative PCR

RPKM Reads per Kilobase of transcript per Million mapped reads

RT Room Temperature

SNV Single Nucleotide Variant

SSIII SuperScript III Reverse Transcriptase

TNBC Triple Negative Breast Cancer

TNM Tumour-Node-Metastasis

TPM Transcripts per Kilobase Million

VAF Variant Allele
Fraction

WBC White Blood Cell

WES Whole Exome
Sequencing

WGA Whole Genome
Sequencing

ZF zebrafish

B2M Beta-2-Microglobulin

RESUMO

RESUMO

O cancro é un conxunto de enfermidades con base xenética complexa. O cancro de mama é o tipo de tumor maligno máis frecuente en mulleres e constitúe a segunda causa das mortes relacionadas con cancro, tras o cancro de pulmón e bronquios. Aínda que existen diferentes factores xenéticos e ambientais coñecidos que contribúen a incrementar o risco de desenvolver cancro de mama, na maioría dos casos descoñécese a súa etioloxía. O cancro de mama correspóndese cun conxunto heteroxéneo de enfermidades que inclúe diferentes subtipos. Por tanto, existen diferencias en termos de incidencia, prognose, progresión ou resposta a terapia entre os distintos subtipos de cancro de mama. Na actualidade, diferéncianse catro subtipos de cancro de mama, en base o diagnóstico molecular, segundo a súa expresión do receptor de proxesterona (PR), do receptor de estróxenos (ER) e do receptor 2 do factor de crecemento epidérmico (HER2): luminal A ($PR^+/ER^+/HER2^-$), luminal B ($PR^{+/-}/ER^+/HER2^{+/-}$), sobre-expresión de HER2 ($PR^-/ER^-/HER2^+$) e 'basal-like' ou triple-negativo ($PR^-/ER^-/HER2^-$).

A maior parte das mortes relacionadas con cancro débense a diseminación tumoral ou metástase. A metástase é o proceso polo cal se produce a propagación das células tumorais cara localizacións secundarias afastadas do foco tumoral primario (tumor de mama). Deste xeito, o cancro pasa de ser unha enfermidade localizada a ser unha doenza sistémica. A metástase é un proceso complexo, constituído por múltiples fases, como migración e invasión dende o tumor primario, intravasación, supervivencia das células tumorais circulantes (CTC) no torrente sanguíneo, extravasación nunha localización secundaria, supervivencia e proliferación, dando lugar a un novo foco tumoral ou lesión metastática. A pesar dos grandes avances realizados na prevención, diagnose e tratamento do cancro de mama, a metástase continúa a ser unha enfermidade incurable. De feito, estímase que entre o 5% e o 10% dos pacientes presentarán metástase no momento do diagnóstico e que aproximadamente o

30% dos pacientes con cancro de mama acabará progresando cara un estado diseminado da enfermidade ó longo do tratamento. Por tanto, incrementar o coñecemento e comprensión sobre o proceso de metástase é esencial para mellorar o tratamento dos pacientes. Neste contexto, as células tumorais circulantes (CTC) desenvolven un papel esencial, xa que son as responsables da formación das metástases. As CTC correspóndense con aquelas células tumorais que son liberadas á circulación sanguínea dende o tumor primario, ou ben dende as lesións metastáticas preexistentes. As CTC poden circular no torrente sanguíneo como células individuais, ou ben como pequenas agrupacións oligoclonais de células, chamadas clústeres de CTC. As CTC, tanto individuais como en forma de clústeres de CTC, poden ser detectadas a partir de mostras de sangue periférica ou 'biopsias líquidas'. A 'biopsia líquida', en contraposición á biopsia sólida tisular convencional, constitúe unha técnica minimamente invasiva que aporta información actualizada e a tempo real do estado de progresión do tumor. Na actualidade, é posible detectar e estudar as CTC presentes nas mostras de biopsia líquida, grazas o desenvolvemento e mellora das tecnoloxías de enriquecemento e illamento de CTC. As diversas metodoloxías de enriquecemento de CTC clasifícanse en dous grandes grupos: i) enriquecemento baseado en propiedades biolóxicas; ii) enriquecemento baseado en propiedades físicas.

Os sistemas de enriquecemento fundamentados en propiedades biolóxicas baséanse no uso de anticorpos para a detección de diferentes marcadores. Estes marcadores poden ser expresados polas CTC (enriquecemento positivo), ou ben ser expresados polo resto de células da mostra pero non polas CTC (enriquecemento negativo). Entre os sistemas de enriquecemento baseados en características biolóxicas destaca o sistema CellSearch®, o cal se basea na detección do antíxeno EpCAM (Epithelial Cell Adhesion Molecule) para a identificación de CTC. CellSearch® é o único sistema de enriquecemento de CTC aprobado pola FDA (Food and Drug Administration) para o seu uso clínico en cancro de mama, próstata e en cancro colorrectal. Este tipo de sistemas de identificación de CTC posúe unha gran utilidade. Non obstante, a detección de CTC baseada en EpCAM

unicamente permite detectar as CTC de fenotipo epitelial, perdendo aquelas CTC de fenotipo máis mesenquimal. Ademais, CellSearch® non permite a recuperación das CTC identificadas en estado viable, restrinxindo a variedade de análises posteriores. Estas limitacións propiciaron o desenvolvemento dos sistemas de illamento baseados en propiedades físicas e por tanto, independentes da expresión de marcadores por parte das CTC. Estes sistemas inmuno-independentes teñen a capacidade potencial de identificar unha poboación de CTC fenotípicamente máis heteroxénea. Estas metodoloxías están baseadas en diferentes propiedades físicas, como densidade (gradiente de densidade), tamaño (microfiltración), tamaño e deformabilidade (microfluídica), propiedades eléctricas (dielectroforese). Así, por exemplo, a leucoaférese diagnóstica (DLA) é un proceso que emprega a densidade como parámetro para o illamento de CTC. A DLA é un procedemento estándar empregado no ámbito clínico para o illamento das células mononucleares mediante a centrifugación continua do sangue. Recentemente, demostrouse que as CTC poden ser co-illadas conxuntamente coas células mononucleares durante a centrifugación continua, xa que posúen densidades similares. Ademais, en comparación coas mostras de sangue periférica, os produtos de DLA teñen a vantaxe de que derivan de grandes volumes de sangue, maximizando a probabilidade de detección de CTC.

Os sistemas microfluídicos permiten identificar as CTC presentes nunha mostra, fundamentándose no seu maior tamaño e menor deformabilidade en comparación co resto de células sanguíneas. Os dispositivos microfluídicos son uns dos máis empregados na actualidade, xa que realizan múltiples pasos dentro dun chip ('lab on a chip'), permitindo un procesamento continuo que minimiza a perda de mostra durante o proceso. Parsortix™ é un sistema microfluídico semiautomatizado capaz de capturar CTC en diversos tipos de mostras (sangue periférico, ascite, medula ósea). Trátase dun sistema coste-efectivo, versátil e sinxelo que permite recuperar as CTC identificadas nun estado viable, o que posibilita o seu posterior cultivo *in vitro*, ou caracterización transcriptómica, entre outros.

A información máis elemental que se pode obter a partir dunha mostra de biopsia líquida (aproximadamente 7,5 mL de sangue) é o reconto de CTC. A pesar de ser un parámetro aparentemente sinxelo de determinar, o reconto de CTC posúe un gran valor prognóstico. De feito, estableceuse que a detección por CellSearch® de ≥5 CTC/7,5 mL de sangue periférica constitúe un factor de prognóstico negativo en pacientes con cancro de mama, mentres que aqueles pacientes que posúen <5 CTC/7,5 mL presentan un prognóstico favorable. Por tanto, as CTC constitúen unha ferramenta esencial e de gran potencial para impulsar o desenvolvemento dunha medicina personalizada, orientada á monitorización da progresión tumoral durante a terapia.

Pola súa parte, os clústeres de CTC correspóndense cunha subpoboación minoritaria, xa que unicamente representan entre o 1% e o 30% de todas as CTC. A pesar da súa reducida frecuencia en sangue, pénsase que os clústeres de CTC posúen un potencial metastático entre 23 e 50 veces maior que as CTC individuais. De feito, estímase que os clústeres de CTC son os responsables da formación da maioría das metástases e a súa presenza en mostras de biopsia líquida correlaciónase con unha prognose adversa.

A maior supervivencia e maior resistencia a anoikis (morte celular por ausencia de anclaxe) dos clústeres de CTC na circulación sanguínea, a súa menor incidencia de apoptose na localización secundaria e a expresión de marcadores de tipo 'stem' son algúns dos factores que explican parcialmente o maior potencial metastático dos clústeres de CTC. Ademais, a enumeración dos clústeres de CTC constitúe un factor prognóstico independente, que aporta información adicional o reconto de CTC por si só. A pesar da gran relevancia dos clústeres de CTC no proceso de progresión tumoral e formación de metástases, existe un coñecemento limitado sobre a súa orixe, a súa bioloxía, ou sobre as características fenotípicas que os fan máis metastáticos que as CTC individuais. Isto débese a súa reducida frecuencia no sangue, así como a limitada capacidade das tecnoloxías de illamento actuais para identificar e illar os clústeres de CTC sen alterar a súa integridade física, o que restrinxe de maneira significativa a cantidade de material

biolóxico para o seu estudo. Polo tanto, faise necesario o desenvolvemento de ferramentas alternativas que contribúan a superar estas limitacións. Neste senso, a xeración de modelos experimentais de clústeres de CTC permite incrementar a cantidade de material biolóxico dispoñible para o seu estudo.

A presente tese de doutoramento ten por obxectivo principal a realización dun estudo comparativo entre as CTC individuais e os clústeres de CTC, coa fin de profundar na bioloxía dos clústeres de CTC, determinar as características diferenciais que lle aportan maior potencial metastático e, en última instancia contribuír a determinar o seu papel durante a progresión tumoral e a formación das metástases. Para acadar este obxectivo, realizamos o illamento de CTC individuais e clusters de CTC a partir de mostras biopsia líquida procedentes de pacientes con cancro de mama metastático, avaliamos a capacidade prognóstica das CTC e dos clústeres de CTC e optimizamos as condicións de illamento inmuno-independente de clústeres de CTC a partir de diferentes mostras biolóxicas de partida (biopsia líquida, produtos de DLA). Ademais, desenvolvemos modelos de CTC individuais e de clústeres de CTC empregando liñas celulares humanas de cancro de mama e, finalmente, caracterizamos, tanto funcional, como molecularmente os modelos para comprobar a súa idoneidade e o seu potencial como ferramenta de estudo da bioloxía dos clústeres de CTC.

Para avaliar a capacidade prognóstica dos clústeres de CTC, recolléronse mostras de sangue periférico dunha cohorte de pacientes (N = 54) con distintos subtipos de cancro de mama metastático. As mostras foron tomadas en diferentes momentos da enfermidade, no momento basal, previo ó inicio do tratamento quimioterápico, e no momento de progresión a este. As mostras de biopsia líquida foron empregadas para o reconto de CTC individuais e de clústeres de CTC mediante o sistema CellSearch®. Esta enumeración permitiunos confirmar que no momento basal o reconto dos clústeres de CTC constitúe un factor prognóstico independente que aporta valor adicional o reconto de CTC individuais, particularmente en pacientes con elevados recontos de CTC individuais. Ademais, a presenza continuada o

longo do tratamento de clústeres de CTC no sangue periférico dos pacientes asóciase cun prognóstico adverso, xa que se correlaciona cunha redución da supervivencia.

A detección unicamente de aquelas CTC de fenotipo epitelial (que expresan EpCAM), a imposibilidade de recuperar as CTC detectadas en estado viable e o feito de tratarse dun sistema que non permite a optimización de protocolos para o illamento de clústeres de CTC son algunhas das grandes limitacións do CellSearch®, así como doutros sistemas de enriquecemento inmuno-dependentes. Para superar estas limitacións, levamos a cabo a detección inmuno-independente de CTC e especialmente de clústeres de CTC a partir de mostras de sangue periférico obtidas de pacientes con cancro de mama metastático. O procesamento levouse a cabo no sistema microfluídico Parsortix™, aplicando un protocolo de separación con menor velocidade de fluxo e menor presión, especialmente deseñado para preservar a integridade dos clústeres de CTC. Ademais, a optimización dunha inmuno-tinguidura sobre células vivas permitiu identificar tanto CTC individuais como clústeres de CTC nestas mostras. A optimización deste fluxo de traballo permite a recuperación de CTC en estado viable que poden ser posteriormente empregadas para diferentes análises que requiren de células vivas, como o cultivo *in vitro* ou estudos transcricionais a nivel de célula individual (single-cell analysis). Esta metodoloxía posibilita un estudo máis detallado da bioloxía das CTC e dos clústeres de CTC, permitindo afondar na comprensión do proceso de metástase.

Adicionalmente e co propósito de maximizar a frecuencia de detección de CTC e especialmente de clústeres de CTC, tamén se verificou se é factíbel o uso de produtos de DLA para o illamento de CTC e clústeres de CTC no Parsortix™. A optimización dos parámetros máis relevantes do fluxo de traballo, como a filtración previa da mostra ou a selección do protocolo de separación máis axeitado, permitiunos identificar, illar e recuperar tanto CTC individuais, como clústeres de CTC. A pesar da súa elevada celularidade, demostramos que é factíbel usar mostras de DLA para o illamento de clústeres de CTC. Aínda que serán necesarias

probas adicionais, os datos indican que os produtos DLA constitúen un tipo de mostra con gran potencial para incrementar a capacidade de detección de clústeres de CTC, contribuíndo a superar a limitación da baixa frecuencia de clústeres de CTC detectados nas mostras comúns de sangue periférico (7,5 mL).

Por outra banda, desenvolvemos modelos *in vitro* de CTC individuais e de clústeres de CTC, mediante o uso de liñas celulares humanas de cancro de mama. Utilizáronse dúas liñas celulares diferentes, cada unha pertencente a un subtipo de cancro de mama distinto e con diferentes características fenotípicas. Os modelos de CTC son recursos alternativos axeitados para vencer a limitación da baixa frecuencia de CTC e, sobre todo de clústeres de CTC no sangue dos pacientes. Tras a súa xeración, estes modelos foron caracterizados funcional e molecularmente, coa fin de comprobar a súa idoneidade. Os ensaios funcionais *in vitro* foron deseñados para simular diferentes etapas do proceso de metástase. Os datos obtidos cos ensaios *in vitro* foron confirmados polos ensaios funcionais *in vivo*, realizados sobre embrións de peixe cebra (*Danio rerio*) e sobre modelos murinos (*Mus musculus*). A caracterización funcional *in vitro* e no rato mostrou que os modelos de clústeres de CTC posúen maior capacidade de migración, invasión, supervivencia en circulación, formación de colonias e maior capacidade de colonización para xerar lesións metastáticas, en comparación co modelo de CTC individuais. Ademais, nos peixes inxectados co modelo de clústeres de CTC observouse menor diseminación celular pero maior capacidade de supervivencia e proliferación das células diseminadas en comparación cos inxectados con CTC individuais, suxerindo unha maior capacidade de supervivencia dos clústeres de CTC na circulación do peixe. Estes resultados están apoiados pola caracterización molecular, a cal mostrou que as células diseminadas nos peixes inxectados co modelo de clústeres de CTC posúen unha maior expresión de xenes tipo 'stem' (CD44, ITGA6), xenes relacionados coa regulación do ciclo celular e a proliferación (CDK4, E2F4) e unha maior expresión do xene de supervivencia PLAU, en comparación coas células diseminadas nos peixes inxectados con CTC individuais. Concordemente, tamén se observou unha menor expresión do

xene pro-apopótico BAX nas células diseminadas nos xenotrasplante de clústeres de CTC. Non só existe unha concordancia entre as observacións funcionais e moleculares, senón que estes resultados tamén están apoiados polas observacións publicadas na literatura, onde se suxire unha posible maior capacidade de supervivencia dos clústeres de CTC para explicar parte do seu maior potencial metastático. Consecuentemente, o modelo experimental de clústeres de CTC xerado a partir de liñas celulares de cancro de mama representa de xeito fidedigno os trazos fenotípicos dos clústeres de CTC illados do sangue de pacientes, sendo unha ferramenta útil para o estudo da súa bioloxía.

O uso dun modelo ortotópico murino de cancro de mama metastático humano, de subtipo triple negativo, permitiunos monitorizar o proceso de progresión tumoral completo, dende a formación do tumor primario ata a aparición das metástases, nun sistema biolóxico complexo. Este modelo tamén posibilitou a captura e illamento de CTC e de clústeres de CTC a partir de mostras de sangue obtidas mediante punción cardíaca. O procesamento destas mostras no sistema microfluídico Parsortix™ e o posterior reconto do número de células illadas permitiunos corroborar que os clústeres de CTC son unha poboación minoritaria. Concretamente, neste modelo os clústeres de CTC representaron aproximadamente o 12,7% de todas as CTC. O procesamento de mostras de sangue derivadas deste modelo ortotópico empregando un kit de enriquecemento negativo posibilitou o illamento de CTC vivas que foron posteriormente postas en cultivo *in vitro*. Esta liña derivada das CTC do modelo murino, chamada mCTC, medra *in vitro* en condicións de baixa adherencia, en forma de pequenas agrupacións celulares, polo que podería constituír potencialmente un modelo de clústeres de CTC máis realista, dende o punto de vista fisiolóxico. A caracterización funcional preliminar da liña mCTC indica que posúe unha maior resistencia a estrés fluídico *in vitro*, así como unha maior capacidade de diseminación e supervivencia no peixe cebra, en comparación co modelo de clústeres de CTC. Os datos transcriptómicos (RNAseq) preliminares mostran a existencia de xenes diferencialmente

expresados entre a liña mCTC e as células control. Aínda que será necesaria unha validación posterior, algúns destes xenes poderían estar implicados nas características metastáticas da liña mCTC. Por tanto, a liña mCTC posúe un elevado potencial como modelo de clústeres de CTC fisioloxicamente máis realista, aínda que se necesitará unha caracterización máis exhaustiva para confirmar esta hipótese. Adicionalmente, as células tumorais illadas no Parsortix^{TM}, tanto CTC individuais como clústeres de CTC foron recuperadas e empregadas para a realización de estudos xenómicos de célula individual coa fin de estudar o perfil mutacional das CTC e dos clústeres de CTC. Os datos xenómicos preliminares mostran que é posible inferir mutacións a partir de CTC individuais e de clústeres e que os clústeres de CTC parecen posuír unha maior variabilidade xenética que as CTC individuais, amosando SNV (Single Nucleotide Variants) específicos que non existen nas CTC individuais. Estes resultados poderían contribuír a explicar a policlonalidade dos clústeres de CTC e a entender o seu maior potencial metastático. Non obstante, requiriranse estudos adicionais para identificar os trazos xenómicos das CTC e dos clústeres de CTC responsables da xeración das metástases.

Os clústeres de CTC desenrolan un papel fundamental na diseminación tumoral, xa que son os responsables da formación da maioría das metástases. En conxunto, o presente proxecto de tese confirma a relevancia clínica e o valor prognóstico dos clústeres de CTC e, por tanto, da biopsia líquida, nunha cohorte non seleccionada de pacientes con cancro de mama metastático. A pesar da súa reducida frecuencia en sangue, o reconto de clústeres de CTC aporta valor informativo adicional á enumeración das CTC individuais por si soas. Ademais, a optimización das condicións de enriquecemento de CTC en sistemas de illamento inmuno-independente, así como o uso de mostras de biopsia líquida alternativas a 7,5 mL de sangue periférico, como os produtos de DLA, non só poden permitir maximizar a capacidade de detección de CTC e clústeres de CTC, senón que posibilita potencialmente a captación dunha maior heteroxeneidade nas CTC, representando de xeito máis realista a poboación de CTC presente no sangue dos pacientes. Por outra banda, os modelos de clústeres de CTC desenvoltos no presente

proxecto recapitulan de maneira realista os trazos fenotípicos dos clústeres de CTC illados de paciente con cancro de mama. Por tanto, constitúen ferramentas con un gran potencial para o estudo da súa bioloxía. A combinación das distintas ferramentas e recursos mostrados neste proxecto de tese poden contribuír a superar as limitación impostas pola baixa frecuencia de clústeres de CTC no sangue dos pacientes e posibilita o aumento da dispoñibilidade de mostra inicial, o que permitirá afondar no coñecemento sobre o papel que desenvolven os clústeres de CTC durante a progresión tumoral e a formación de metástases. Este coñecemento resulta esencial para impulsar o avance da medicina personalizada e a oncoloxía de precisión.

RESUMEN

RESUMEN

El cáncer es un conjunto de enfermedades con base genética compleja. El cáncer de mama es el tipo de tumor maligno más común en la población femenina y es la segunda causa principal de muerte relacionada con el cáncer, después del cáncer de pulmón y bronquios. Aunque existen diferentes factores genéticos y ambientales conocidos que contribuyen a un mayor riesgo de desarrollar cáncer de mama, en la mayoría de los casos se desconoce su etiología. El cáncer de mama se corresponde con un conjunto heterogéneo de enfermedades y que por tanto, incluye diferentes subtipos. Existen diferencias en términos de incidencia, pronóstico, progresión o respuesta a la terapia entre los diferentes subtipos de cáncer de mama. Actualmente, se diferencian cuatro subtipos de cáncer de mama, en base al diagnóstico molecular, según la expresión del receptor de progesterona (PR), del receptor de estrógenos (ER) y del receptor 2 del factor de crecimiento epidérmico (HER2): luminal A ($PR^+/ER^+/HER2^-$), luminal B ($PR^{+/-}/ER^+/HER2^{+/-}$), sobreexpresión de HER2 ($PR^-/ER^-/HER2^+$) y 'basal-like' o triple negativo ($PR^-/ER^-/HER2^-$).

La mayoría de las muertes relacionadas con cáncer se deben a la diseminación tumoral o metástasis. La metástasis es el proceso por el cual las células tumorales diseminan a localizaciones secundarias, alejadas del foco tumoral primario (tumor de mama). De esta forma, el cáncer pasa de ser una enfermedad localizada a ser una enfermedad sistémica. La metástasis es un proceso complejo, constituido por múltiples fases: migración e invasión desde el tumor primario, intravasación, supervivencia de las células tumorales circulantes (CTC) en el torrente sanguíneo, extravasación en una localización secundaria, supervivencia y proliferación, lo que genera un nuevo foco tumoral, o lesión metastásica. A pesar de los grandes avances realizados en la prevención, el diagnóstico y el tratamiento del cáncer de mama, la metástasis sigue siendo una enfermedad incurable. De hecho, se estima que entre el 5% y el 10% de los

pacientes presentarán metástasis en el momento del diagnóstico y que aproximadamente el 30% de los pacientes con cáncer de mama progresarán a un estado diseminado de la enfermedad a lo largo del tratamiento. Por tanto, incrementar el grado de conocimiento y comprensión sobre el proceso de metástasis resulta fundamental para mejorar el tratamiento de los pacientes. En este contexto, las células tumorales circulantes (CTC) desempeñan un papel fundamental, ya que son las responsables de la formación de las metástasis. Las CTC se corresponden con las células tumorales liberadas a la circulación sanguínea desde el tumor primario, o bien desde las lesiones metastásicas preexistentes. Las CTC pueden circular en el torrente sanguíneo como células individuales, o como pequeños grupos oligoclonales de células, llamados clústeres de CTC. Las CTC, tanto individuales como en forma de clústeres de CTC, pueden detectarse a partir de muestras de sangre periférica o "biopsias líquidas". La "biopsia líquida", en contraposición a la biopsia de tejido sólido convencional, es una técnica mínimamente invasiva que proporciona información actualizada y a tiempo real sobre la progresión del tumor. En la actualidad, es posible detectar y estudiar las CTC presentes en muestras de biopsia líquida, gracias al desarrollo y mejora de las tecnologías de enriquecimiento y aislamiento de CTC. Las diversas metodologías de enriquecimiento de CTC se clasifican en dos grandes grupos: i) enriquecimiento basado en propiedades biológicas; ii) enriquecimiento basado en propiedades físicas.

Los sistemas de enriquecimiento basados en propiedades biológicas se fundamentan en el uso de anticuerpos para la detección de diferentes marcadores. Estos marcadores pueden ser expresados por las CTC (enriquecimiento positivo), o ser expresados por el resto de las células de la muestra pero no por las CTC (enriquecimiento negativo). Entre los sistemas de enriquecimiento basados en características biológicas, destaca el sistema CellSearch®, que se basa en la detección del antígeno EpCAM (Epithelial Cell Adhesion Molecule) para la identificación de CTC. CellSearch® es el único sistema de enriquecimiento de CTC aprobado por la FDA (Food and Drug Administration) para su uso clínico en cáncer de mama, próstata

y colorrectal. Este tipo de sistemas de identificación de CTC son muy útiles. Sin embargo, la detección de CTC basada en EpCAM solo permite la identificación de CTC de fenotipo epitelial, perdiendo las CTC más mesenquimales. Además, CellSearch® no permite la recuperación de las CTC identificadas en estado viable, restringiendo la variedad de análisis posteriores. Estas limitaciones han llevado al desarrollo de sistemas de aislamiento basados en propiedades físicas y, por tanto, independientes de la expresión de marcadores por parte de las CTC. Estos sistemas inmuno-independientes tienen la capacidad potencial de identificar una población de CTC fenotípicamente más heterogénea. Estas metodologías se basan en diferentes propiedades físicas, como densidad (gradiente de densidad), tamaño (microfiltración), tamaño y deformabilidad (microfluídica), propiedades eléctricas (dielectroforesis). Así, por ejemplo, la leucaféresis diagnóstica (DLA) es un proceso que utiliza la densidad como parámetro para el aislamiento de CTC. La DLA es un procedimiento estándar utilizado en el ámbito clínico para el aislamiento de células mononucleares mediante la centrifugación continua de la sangre. Recientemente, se ha demostrado que es posible co-aislar CTC conjuntamente con las células mononucleares durante la centrifugación continua, ya que poseen una densidad similar. Además, en comparación con las muestras de sangre periférica, los productos de DLA tienen la ventaja de que derivan de grandes volúmenes de sangre, lo que maximiza la probabilidad de detección de CTC.

Los sistemas de microfluídica permiten identificar las CTC presentes en una muestra, en función de su mayor tamaño y menor deformabilidad en comparación con el resto de las células sanguíneas. Los dispositivos de microfluídica son unos de los más utilizados en la actualidad, ya que realizan varios pasos dentro de un chip ("lab on a chip"), lo que permite un procesamiento continuo que minimiza la pérdida de muestra durante el proceso. Parsortix^{TM} es un sistema de microfluídica semiautomatizado, capaz de capturar CTC en diversos tipos de muestra (sangre periférica, ascitis, médula ósea). Se trata de un sistema coste-efectivo, versátil y sencillo que permite recuperar las CTC

identificadas en un estado viable, lo que posibilita su posterior cultivo *in vitro*, o su caracterización transcriptómica, entre otros.

La información más básica que se puede obtener de una muestra de biopsia líquida (aproximadamente 7,5 mL de sangre) es el recuento de CTC. A pesar de ser un parámetro aparentemente sencillo de determinar, el recuento de CTC posee un gran valor pronóstico. De hecho, se ha establecido que la detección mediante CellSearch® de ≥5 CTC/7,5 mL de sangre periférica es un factor pronóstico negativo en pacientes con cáncer de mama, mientras que aquellos pacientes con <5 CTC/7,5 mL tienen un pronóstico favorable. Por lo tanto, las CTC constituyen una herramienta esencial y de elevado potencial para impulsar el desarrollo de una medicina personalizada, dirigida a monitorizar la progresión tumoral durante la terapia.

Por su parte, los clústeres de CTC se corresponden con una subpoblación minoritaria, ya que solo representan entre el 1% y el 30% del total de CTC. A pesar de su baja frecuencia en sangre, se cree que los clústeres de CTC tienen un potencial metastásico entre 23 y 50 veces mayor que las CTC individuales. De hecho, se estima que los clústeres de CTC son los responsables de la formación de la mayoría de las metástasis y su presencia en muestras de biopsia líquida se correlaciona con un pronóstico adverso.

La mayor supervivencia y resistencia a anoikis (muerte celular por ausencia de anclaje) de los clústeres de CTC en el torrente sanguíneo, su menor incidencia de apoptosis en la localización secundaria y la expresión de marcadores de tipo "stem" son algunos de los factores que explican parcialmente el mayor potencial metastásico de los clústeres de CTC. Además, la enumeración de clústeres de CTC constituye un factor pronóstico independiente, que proporciona información adicional al recuento de CTC individuales por sí solo. A pesar de la gran relevancia de los clústeres de CTC en el proceso de progresión tumoral y formación de metástasis, existe un conocimiento limitado sobre su origen, su biología o sobre las características fenotípicas que los hacen más metastásicos que las CTC individuales. Esto se

debe a su baja frecuencia en sangre, así como a la capacidad limitada de las tecnologías de aislamiento actuales para identificar y aislar los clústeres de CTC sin alterar su integridad física, lo que restringe significativamente la cantidad de material biológico inicial. Por tanto, se genera la necesidad de desarrollar herramientas alternativas que contribuyan a superar estas limitaciones. En este sentido, la generación *in vitro* de modelos de clústeres de CTC permite incrementar la cantidad de material biológico disponible para su estudio.

El objetivo principal de la presente tesis doctoral es realizar un estudio comparativo entre las CTC individuales y los clústeres de CTC, con el fin de profundizar en la biología de los clústeres de CTC, determinar las características diferenciales que les aportan un mayor potencial metastásico y, en última instancia, contribuir a determinar su papel durante la progresión tumoral y la formación de metástasis. Para alcanzar este objetivo, se realizó el aislamiento de CTC individuales y clústeres de CTC a partir de muestras de biopsia líquida de pacientes con cáncer de mama metastásico, se evaluó la capacidad pronóstica de las CTC individuales y de los clústeres de CTC y se optimizaron las condiciones de aislamiento inmuno-independiente de los clústeres de CTC a partir de diferentes tipos de muestras biológicas (biopsia líquida, productos DLA). Además, se desarrollaron modelos *in vitro* de CTC individuales y clústeres de CTC, utilizando líneas celulares humanas de cáncer de mama y, finalmente, se llevó a cabo la caracterización, tanto funcional como molecular, de los modelos de CTC individuales y clústeres de CTC para comprobar su idoneidad y su potencial como herramienta para el estudio de la biología de los clústeres de CTC.

Con el fin de evaluar la capacidad pronóstica de los clústeres de CTC, se recolectaron muestras de sangre periférica de una cohorte de pacientes (N = 54) con diferentes subtipos de cáncer de mama metastásico. Se tomaron muestras en diferentes momentos de la enfermedad, en el momento basal, previo al inicio del tratamiento quimioterápico, y en el momento de progresión a este. Las muestras de biopsia líquida se utilizaron para el recuento de CTC individuales y de clústeres de CTC, usando el sistema

CellSearch®. Este recuento nos permitió confirmar que en el momento basal la enumeración de los clústeres de CTC constituye un factor pronóstico independiente que aporta valor informativo adicional al recuento de CTC individuales, particularmente en pacientes con elevados recuentos de CTC individuales. Además, la presencia continuada a lo largo del tratamiento de clústeres de CTC en la sangre periférica de los pacientes se asocia con un pronóstico adverso, ya que se correlaciona con una reducción de la supervivencia.

La detección de aquellas CTC de fenotipo epitelial que expresan EpCAM, la incapacidad de recuperar las CTCs detectadas en un estado viable, así como el hecho de tratarse de un sistema que no permite la optimización de protocolos para el aislamiento de clústeres de CTC, son algunas de las principales limitaciones del CellSearch®, así como de otros sistemas de enriquecimiento inmuno-dependientes. Para superar esta limitación, llevamos a cabo la detección inmuno-independiente de CTC y especialmente de clústeres de CTC, a partir de muestras de sangre periférica derivadas de pacientes con cáncer de mama metastásico. El procesamiento se realizó en el sistema microfluídico Parsortix™, aplicando un protocolo de separación con menor flujo y menor presión, especialmente diseñado para preservar la integridad de los clústeres de CTC. Además, la optimización de una inmunotinción sobre célula viva hizo posible identificar tanto CTC individuales, como de clústeres de CTC en estas muestras. La optimización de este flujo de trabajo posibilita la recuperación de CTC viables que pueden ser utilizadas posteriormente para diferentes análisis que requieran de células vivas, como el cultivo *in vitro* o análisis transcipcionales a nivel de célula única. Esta metodología posibilita la realización de un estudio más detallado de la biología de las CTC y de los clústeres de CTC, lo que permitirá una comprensión más profunda del proceso de metástasis.

Adicionalmente y con el propósito de maximizar la frecuencia de detección de CTC y de clústeres de CTC, también se testó el posible uso de muestras de DLA para el aislamiento de CTC y de clústeres de CTC en el Parsortix™. La optimización de

los parámetros más relevantes del flujo de trabajo, como la filtración previa de la muestra o la selección del protocolo de separación más apropiado, nos permitió identificar, aislar y recuperar tanto CTC individuales, como clústeres de CTC. A pesar de su elevada celularidad, ha sido posible demostrar que es factible utilizar muestras de DLA para el aislamiento de clústeres de CTC. Aunque será necesaria la realización de pruebas adicionales, los datos indican que los productos de DLA constituyen un tipo de muestra con gran potencial para incrementar la capacidad de detectar clústeres de CTC, contribuyendo a superar la limitación de la baja frecuencia de clústeres de CTC en muestras de sangre periférica (7,5 mL).

Por otra parte, se han desarrollado modelos *in vitro* de CTC individuales y clústeres de CTC, mediante el uso de líneas celulares humanas de cáncer de mama. Se utilizaron dos líneas celulares diferentes, cada una perteneciente a un subtipo de cáncer de mama distinto y con diferentes características fenotípicas. Los modelos de CTC son recursos alternativos adecuados para superar la limitación de la baja frecuencia de CTC y especialmente de clústeres de CTC en la sangre de los pacientes. Tras su generación, estos modelos fueron caracterizados funcional y molecularmente, con el fin de comprobar su idoneidad. Se diseñaron diferentes ensayos funcionales *in vitro* para simular las distintas etapas del proceso de metástasis. Los datos obtenidos con los ensayos *in vitro* fueron respaldados mediante ensayos funcionales *in vivo*, realizados en embriones de pez cebra (*Danio rerio*) y en modelos murinos (*Mus musculus*). La caracterización funcional *in vitro* y en ratón mostró que los modelos de clústeres de CTC tienen mayor capacidad de migración, invasión, supervivencia en circulación, formación de colonias y mayor capacidad de colonización para generar lesiones metastásicas, en comparación con el modelo de CTC individuales. Además, en los peces inyectados con el modelo de clústeres de CTC se observó una menor diseminación celular hacia la cola pero una mayor capacidad de supervivencia y proliferación de las células diseminadas en comparación con los inyectados con CTC individuales, lo que sugiere una mayor capacidad de supervivencia de los clústeres de CTC en la

circulación de los peces. Estos resultados están respaldados por la caracterización molecular, la cual mostró que las células diseminadas en peces inyectados con el modelo de clústeres de CTC poseen mayor expresión de genes de tipo 'stem' (CD44, ITGA6), genes relacionados con la regulación del ciclo celular y con proliferación (CDK4, E2F4) y mayor expresión del gen de supervivencia PLAU, en comparación con las células diseminadas en peces inyectados con CTC individuales. Asimismo, también se observó una menor expresión del gen pro-apoptótico BAX en células diseminadas en los xenotrasplantes de clústeres de CTC. No solo existe una concordancia entre las observaciones funcionales y moleculares, sino que estos resultados también están respaldados por observaciones publicadas en la literatura, donde se sugiere una mayor capacidad de supervivencia de los clústeres de CTC para explicar parcialmente su mayor potencial metastásico. Consecuentemente, el modelo de clústeres de CTC generado a partir de líneas celulares de cáncer de mama es representativo de las características fenotípicas de los clústeres de CTC aislados de la sangre de paciente, siendo una herramienta útil para el estudio de su biología.

El uso de un modelo ortotópico murino de cáncer de mama metastásico humano, de subtipo triple negativo, permitió monitorizar el proceso de progresión tumoral completo, desde la formación del tumor primario hasta la aparición de metástasis en un sistema biológico complejo. Este modelo también permitió capturar y aislar CTC y clústeres de CTC, a partir de muestras de sangre obtenidas por punción cardíaca. El procesamiento de estas muestras en el sistema de microfluídica Parsortix^{TM} y el posterior recuento del número de células aisladas nos permitió corroborar que los clústeres de CTC son una población minoritaria. Específicamente, los clústeres de CTC representaron aproximadamente el 12,7% de todas las CTC en este modelo. El procesamiento de muestras de sangre derivadas de este modelo ortotópico mediante un kit de enriquecimiento negativo permitió el aislamiento de CTC vivas que posteriormente fueron puestas en cultivo *in vitro*. Esta línea derivada de las CTC del modelo murino, denominada mCTC, crece *in vitro* en condiciones de baja

adherencia, en forma de pequeños agregados celulares, por lo que potencialmente podría constituir un modelo de clústeres de CTC más realista, desde un punto de vista fisiológico. La caracterización funcional de la línea mCTC indica que esta línea posee una mayor resistencia al estrés fluídico *in vitro*, así como una mayor capacidad de diseminación y supervivencia en el pez cebra, en comparación con el modelo *in vitro* de clústeres de CTC. Los datos transcriptómicos preliminares (RNAseq) muestran la existencia de genes diferencialmente expresados entre la línea mCTC y las células control. Aunque será necesaria una validación posterior, algunos de estos genes podrían estar implicados en las características metastásicas de la línea mCTC. Por lo tanto, la línea mCTC tiene un alto potencial como modelo de clústeres de CTC fisiológicamente más realista, aunque se necesitará una caracterización más completa para confirmar esta hipótesis. Además, las células detectadas, tanto CTC individuales como clústeres de CTC, fueron recuperadas y posteriormente empleadas para la realización de estudios genómicos de célula individual, con el fin de estudiar el perfil mutacional de las CTC y de los clústeres de CTC. Los datos genómicos preliminares mostraron que es posible inferir mutaciones a partir de CTC individuales y de clústeres de CTC y que los clústeres de CTC parecen poseer una mayor variabilidad genética que las CTC individuales, mostrando SNV (Single Nucleotide Variants) específicas que no existen en las CTC individuales. Estos resultados podrían contribuir a explicar la policlonalidad de los clústeres de CTC y a entender su mayor potencial metastásico. Sin embargo, serán necesarios estudios adicionales para identificar los rasgos genómicos de las CTC y los clústeres de CTC responsables de la generación de metástasis.

Los clústeres de CTC desempeñan un papel clave en la diseminación tumoral, ya que son los responsables de la formación de la mayoría de las metástasis. En conjunto, el presente proyecto confirma la relevancia clínica y el valor pronóstico de los clústeres de CTC y, por lo tanto, de la biopsia líquida, en una cohorte no seleccionada de pacientes con cáncer de mama metastásico. A pesar de su baja frecuencia en sangre, el recuento de clústeres de CTC proporciona información adicional

a la enumeración de CTC individuales por sí sola. Además, la optimización de las condiciones de enriquecimiento de CTC en sistemas de aislamiento inmuno-independientes, así como el uso de muestras de biopsia líquida alternativas a 7,5 mL de sangre periférica, como los productos DLA, no solo pueden maximizar la capacidad de detección de CTC y clústeres de CTC, si no que permiten potencialmente la identificación de una mayor heterogeneidad en las CTC, representando de forma más realista la población de CTC presentes en la sangre de los pacientes. Por otro lado, los modelos de clústeres de CTC desarrollados en el presente proyecto recapitulan de manera fidedigna las características fenotípicas de los clústeres de CTC aislados de pacientes con cáncer de mama. Por tanto, constituyen herramientas con gran potencial para el estudio de su biología. La combinación de las diferentes herramientas y recursos desarrollados en este proyecto de tesis pueden contribuir a superar las limitaciones impuestas por la baja frecuencia de clústeres de CTC en la sangre de los pacientes e incrementan la disponibilidad de muestra inicial, lo que posibilita un mayor estudio y comprensión sobre el papel de los clústeres de CTC durante la progresión tumoral y la formación de metástasis. Este conocimiento es esencial para impulsar el avance de la medicina personalizada y la oncología de precisión.

SUMMARY

SUMMARY

The term 'cancer' includes a group of genetically complex diseases. Breast cancer (BC) is the most common type of malignancy in women and is the second leading cause of cancer-related deaths, after lung and bronchus cancer. Although there are different known genetic and environmental factors that increase the risk of developing BC, the aetiology of most BC cases is still unknown. BC is a heterogeneous group of diseases, which show differences in terms of incidence, prognosis, progression, or response to therapy. There are four different subtypes of BC, based on the molecular diagnosis, according to their expression of the progesterone receptor (PR), estrogen receptor (ER) and epidermal growth factor receptor 2 (HER2): luminal A (PR^+/ER^+/$HER2^-$), luminal B ($PR^{+/-}$/ER^+/$HER2^{+/-}$), HER2 overexpression (PR^-/ER^-/$HER2^+$), and 'basal-like' or triple-negative (PR^-/ER^-/$HER2^-$).

The majority of cancer-related deaths are due to the development of a disseminated state or metastasis. Metastasis is a process by which tumour cells spread to secondary locations far away from the primary tumour, hence becoming a systemic disease. Metastasis is a complex, multi-step process that includes different steps: migration and invasion from the primary tumour, intravasation, survival into the bloodstream, extravasation in a secondary site, survival and proliferation generating a new tumour lesion or metastasis. Despite the great advances made in the prevention, diagnosis and treatment of BC, metastasis remains an incurable disease. In fact, between 5% and 10% of patients will show metastasis at the time of diagnosis, and approximately 30% of BC patients will eventually progress to a disseminated state of the disease throughout treatment. Therefore, increasing knowledge and understanding about the metastatic process is essential to improve patient care. In this context, circulating tumour cells (CTCs) play an essential role, as they are the responsible for the formation of metastases. CTCs are those tumour cells that are released into the bloodstream from the

primary tumour (breast cancer) or the pre-existing metastatic lesions. CTCs can travel into the bloodstream as individual cells, or as small oligoclonal group of cells, called CTC clusters. CTCs, both individual and CTC clusters, can be detected from peripheral blood samples or 'liquid biopsies'. In contrast to conventional tissue biopsies (or 'solid biopsies'), 'liquid biopsy' is a minimally invasive technique that provides up-to-date, real-time information about the progression status of the tumour. It is possible to detect and study CTCs present in liquid biopsy samples due to the development and improvement of CTC enrichment and isolation technologies. The different CTCs enrichment methodologies are classified into two different groups: i) enrichment based on biological properties; ii) enrichment based on physical properties.

CTC enrichment systems based on biological properties are based on the use of antibodies for the detection of different markers. These markers can be expressed by CTCs (positive enrichment) or by the rest of the cells in the sample but not by CTCs (negative enrichment). Among the enrichment systems based on biological characteristics, the CellSearch® is one of the most used. CellSearch® is based on the expression of the EpCAM (Epithelial Cell Adhesion Molecule) antigen for the identification of CTCs. CellSearch® is the only system approved by the FDA (Food and Drug Administration) for clinical use in breast, prostate, and colorectal cancer. Despite its utility, EpCAM-based detection of CTCs only allows the detection of epithelial CTCs, losing those CTCs with mesenchymal (or stem-like) traits. Moreover, CellSearch® does not allow the recovery of CTCs in a viable state, hence restricting the variety of downstream analyses. These limitations have led to the development of isolation systems based on physical properties and therefore independent of the expression of markers by CTCs. These immune-independent systems have the potential ability to identify a phenotypically more heterogeneous population of CTCs. These methodologies are based on different physical properties, such as density (density gradient), size (microfiltration), size and deformability (microfluidic), electrical properties (dielectrophoresis). For instance, diagnostic leukapheresis (DLA) is a procedure that uses density as a parameter for the isolation of

CTCs. DLA is a standard procedure used in the clinical setting for the isolation of mononuclear cells by continuous centrifugation of blood. It was recently proposed that CTCs can be co-isolated together with mononuclear cells during continuous centrifugation, as they have similar densities. Besides, DLA products derive from large volumes of blood, maximizing the likelihood of CTC detection, in comparison with peripheral blood samples.

Microfluidic systems allow the detection of CTCs based on their larger size and lower deformability compared to the rest of the blood cells. Microfluidic devices are nowadays one of the most used, as they perform multiple steps within a chip ('lab on a chip'), allowing continuous processing that minimizes sample loss during the process. Parsortix^{TM} is a semi-automated microfluidic system capable of capturing CTCs in different types of samples (peripheral blood, ascites, bone marrow). It is a cost-effective, versatile, and simple system that allows the recovery of the isolated CTCs in a viable for later *in vitro* culture, or transcriptomic characterization, among others.

The essential information that can be obtained from a liquid biopsy sample (approximately 7.5 mL of peripheral blood) is the CTC count. Despite being a seemingly simple parameter to determine, the CTC enumeration has great prognostic value. In fact, it was established that the detection of $\geq$5CTCs/7.5 mL of peripheral blood by CellSearch$^{®}$ is a negative prognostic factor in BC patients, whereas those patients with <5CTCs/7.5 mL show a favourable prognosis. Therefore, CTCs constitute an essential and high-potential tool to drive the development of a personalized medicine and to track tumour progression during therapy.

On the other hand, CTC clusters are a minority subpopulation, as they only represent between 1% and 30% of all CTCs. Despite their low frequency in blood, CTC clusters have a metastatic potential between 23 and 50 times greater than individual CTCs. In fact, it is believed that CTC clusters are responsible for the majority of metastases, and their presence in liquid biopsy samples correlate with an adverse prognosis.

Their greater survival and resistance to anoikis (cell death due to the lack of anchorage signals) into the bloodstream, their lower incidence of apoptosis at the secondary location, and the expression of 'stem-like' markers are some of the factors that partially explain the higher metastatic potential of CTC clusters. Additionally, the enumeration of CTC clusters is an independent prognostic factor that provides additional information to the CTCs count alone. Despite the great relevance of CTC clusters in the process of tumour progression and metastasis formation, there is limited knowledge about their origin, their biology, or about the phenotypic traits that make them more metastatic than individual CTCs. This is due to their low frequency in the blood, as well as the limited ability of current isolation technologies to identify and isolate CTC clusters without altering their physical integrity, which restricts the amount of starting biological material for their study. Hence, this generates the necessity of developing alternative tools to overcome these limitations. In this sense, the generation of *in vitro* CTC cluster models may permit to increase the amount of biological material available for their study.

The main objective of this doctoral thesis is to conduct a comparative study between individual CTCs and CTC clusters, in order to further study the biology of CTC clusters, to determine the differential characteristics that provide them with greater metastatic potential, and ultimately to seek the role of CTC clusters during tumour progression and metastasis formation. To fulfil this purpose, the isolation of individual CTCs and CTC clusters from liquid biopsy samples derived from metastatic BC patients was performed, the prognostic capacity of individual CTCs and CTC clusters was assessed, and the immune-independent isolation conditions of CTC clusters using different biological samples (liquid biopsy, DLA products) was optimized. In addition, experimental models of individual CTCs and CTC clusters were developed using human BC cell lines, and finally, a functional and molecular characterization of the individual CTCs and CTC cluster models were performed to test their suitability and their potential as a tool for studying the biology of CTC clusters.

To assess the prognostic capacity of CTC clusters, peripheral blood samples were collected from a cohort of patients (N = 54) with different metastatic BC subtypes. Blood samples were taken at different times of the disease, at baseline, previously to the start of the chemotherapy, and at the moment of progression. Liquid biopsy samples were used for individual CTC and CTC cluster enumeration using the CellSearch® system. This enumeration allowed us to confirm that the enumeration of CTC clusters is an independent prognostic factor at baseline that adds value to the count of individual CTCs, particularly in patients with high counts of individual CTCs. Moreover, the continued presence of CTC clusters in the peripheral blood of patients, was associated with an adverse prognosis, as it correlated with a shorter survival.

Some of the major limitations of the CellSearch®, as well as the other immune-dependent enrichment systems, are the detection of those epithelial CTCs that express EpCAM, the inability to recover CTCs in a viable state, as well as the inability to optimize protocols for CTC cluster isolation in the CellSearch®. To overcome these limitations, we performed the immune-independent detection of CTCs and especially CTC clusters from peripheral blood samples derived from patients with metastatic BC. The detection of CTC was performed using the Parsortix™ system. Furthermore, a separation protocol with a lower flow rate and lower pressure was used. This protocol was specially designed to preserve the integrity of CTC clusters. Besides, the optimization of a live cell immunostaining made it possible to identify both individual CTCs and CTC clusters in these samples. This workflow allowed the recovery of CTCs in a viable state that could be later used for different analyses that require alive cells, such as *in vitro* culture, or single-cell transcriptomic analysis. Hence, this workflow enables a deeper study of the biology of CTCs and CTC clusters, allowing a better understanding of the metastasis process.

With the purpose of maximizing the frequency of detection of CTCs and especially CTC clusters, we evaluated the feasibility of using DLA samples for the isolation of CTCs and CTC clusters in the Parsortix™. The optimization of the most relevant steps of

the workflow, such as pre-filtering the sample or selecting the most appropriate separation protocol, allowed us to identify, isolate, and recover both individual CTCs and CTC clusters. Despite the high cellularity of DLA samples, we demonstrated that it is feasible to use DLA products for the isolation of CTC clusters. Although further assays will be required, our data indicated that DLA products constitute a type of sample with great potential to increase the likelihood of CTC cluster detection. This will help overcome the limitation of the low frequency of CTC clusters in conventional peripheral blood samples (7.5 mL).

On the other hand, we developed *in vitro* models of individual CTCs and CTC clusters by using human BC cell lines. Two different cell lines were used, each one belonging to a different BC subtype and with different phenotypic traits. CTC models are a suitable alternative to overcome the limitation of the low frequency of CTCs and especially of CTC clusters in the blood of patients. After their generation, these models were functionally and molecularly characterized to test their suitability. *In vitro* functional assays were designed to mimic different steps of the metastatic process. The results obtained with the *in vitro* assays were confirmed by *in vivo* functional assays, performed in zebrafish embryos (*Danio rerio*) and murine models (*Mus musculus*). Functional characterization *in vitro* and in mice showed that CTC cluster models had higher migration, invasion, survival into the bloodstream, ability to form colonies and higher colonization capacity to generate metastatic lesions, compared to the model of individual CTCs. In addition, in zebrafish experiments, it was observed that those fish injected with the CTC clusters showed lower cell dissemination, but higher survival and proliferation of the disseminated cells compared to the xenografts injected with individual CTCs, suggesting a higher survival capacity of CTC clusters in the circulation of the fish. These results were supported by molecular characterization, which showed that disseminated cells in those fish injected with CTC clusters had a higher expression of stem-like genes (CD44, ITGA6), genes related to cell cycle regulation and proliferation (CDK4, E2F4) and higher expression of the PLAU survival gene, compared with cells disseminated in the fish injected with

individual CTCs. Consistently, lower expression of the pro-apoptotic BAX gene was also observed in cells disseminated in the CTC cluster xenografts. Besides the concordance observed between functional and molecular results, these data are also supported by the findings reported in the literature, where a potential greater survival of CTC clusters was suggested to partially explain their higher metastatic potential. Consequently, this CTC cluster model generated from BC cell lines reliably represents the phenotypic traits of CTC clusters isolated from the blood of patients, being a useful tool for the study of their biology.

The use of a murine orthotopic model of human TNBC allowed us to track the entire process of tumour progression in a complex biological system, from the formation of the primary tumour to the onset of metastases. This model also allowed us to obtain CTCs and CTC clusters from blood samples collected by cardiac puncture. Blood samples were processed in the Parsortix^{TM} system. The subsequent counting of the number of isolated cells corroborated that CTC clusters were a minority population. Specifically, in this model CTC clusters accounted for approximately 12.7% of all CTCs. The enrichment of CTCs using a negative enrichment kit enabled the isolation of alive CTCs that were later put in culture. This cell line derived from the CTCs of the murine model, named 'mCTC', grows *in vitro* under low attachment conditions as small cellular aggregates. Thus, it could potentially be a more realistic CTC cluster model, from a physiological point of view. Functional characterization of the mCTC cell line evinced that it had higher resistance to fluid stress *in vitro*, as well as a greater capacity for dissemination and survival in the zebrafish, compared to the cluster model of CTCs. Preliminary transcriptomic data (RNAseq) showed the existence of differentially expressed genes between the mCTC line and the control cells. Although further validation will be required, some of these genes could potentially be implicated in the metastatic characteristics of the mCTC cell line. Therefore, the mCTC line has high potential as a more physiologically realistic model of CTC clusters, although a deeper characterization will be needed to confirm this hypothesis. Moreover, harvested cells from the Parsortix^{TM} were used to perform genomic studies at a single-cell

level to evaluate the mutational profile of single CTCs and CTC clusters. Preliminary genomic data showed that it was possible to infer mutations from single CTCs and CTC clusters and that CTC clusters seemed to have a higher genetic variability than individual CTCs, showing specific SNVs (Single Nucleotide Variants) that were not found in individual CTCs. These results could help to explain the possible polyclonal origin of CTC clusters and to better understand their higher metastatic potential. However, further studies will be required to identify the genomic traits of those CTCs and CTC clusters responsible for the seeding of metastasis.

CTCs clusters play a key role in tumour dissemination, as they are responsible for the formation of the majority of metastases. Taken together, this project confirmed the clinical relevance and the prognostic value of CTC clusters and of liquid biopsy in an unselected cohort of patients with metastatic BC. Despite their low frequency in blood, the enumeration of CTC clusters provided additional and valuable information to the enumeration of individual CTCs alone. Furthermore, the optimization of CTC and especially CTC cluster enrichment conditions in the Parsortix^{TM}, as well as the use of liquid biopsy samples, such as DLA products, as alternatives to 7.5 mL of peripheral blood samples not only could maximize the detection capacity of CTCs and CTC clusters but also could increase the ability to detect a higher heterogeneity within isolated CTCs. This greater variability would provide a more realistic landscape of the CTC population existing in the blood of BC patients. On the other hand, the CTC cluster models developed in this project recapitulated the phenotypic features of CTC clusters isolated from BC patients. Therefore, these models are tools with enormous potential for the study of their biology. The combination of the different tools and resources presented in this thesis project can contribute to overcome the restrictions derived from the low frequency of CTC clusters in the blood of patients and would increase the availability of tumour material, which would allow a deeper understanding of the role of CTC clusters during tumour progression and metastasis formation. This

knowledge is essential to prompt the advance of personalized medicine and precision oncology.

INTRODUCTION

INTRODUCTION

1. BREAST CANCER

1.1. Epidemiology and aetiology

Breast cancer (BC) corresponds to the tumour with the highest incidence rate, and it is the second leading cause of cancer-related deaths in women, after lung and bronchus cancer [1]. With more than 2 million new cases of BC diagnosed per year worldwide [2], it currently represents about 30% of all female cancers [2]. The relevant advances in prevention, early diagnosis and treatment allowed the sustention of a declining trend in the BC death rate since 1989 [1]. Nevertheless, it is estimated that approximately 627,000 women died from BC during the year 2018, which is equivalent to 15% of all female cancer deaths (WHO) [3].

Taking into consideration that breast carcinoma is a genetically complex disease, several environmental factors have been reported to be associated with an increased risk of developing BC. These risk factors comprise female gender, increased patient age, early menarche age, late menopause, late age at first childbirth, familial occurrence of BC at a young age, increased mammographic breast density, presence of benign breast disease, exposure to chest radiation, hormonal and alcohol intake as well as obesity [3–5]. Furthermore, it was widely demonstrated that genetic mutations, such as mutations in *BRCA1* and *BRCA2* genes are associated with an increased risk of BC. Nevertheless, they only represent a minority of BC cases [6]. This biological complexity contributes to explain the fact that the aetiology in the majority of BC cases is unknown.

BC is a highly heterogeneous group of diseases, which involves various BC subtypes with phenotypic differences. Consequently, there are variations in terms of incidence, prognosis, progression and response to therapy between them [7–9].

Conventionally, BC subtypes were determined by using classical immunohistochemistry (IHC) based on different biomarkers, such as estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2), combined with classic clinical and histopathological parameters, which include cell proliferation index (Ki67), tumour size, tumour grade and nodal involvement. These criteria are usually used in the clinic for patient prognosis and management [10].

1.2. Subclassification of BC tumours

Clinically, the tumour IHC status stratifies BC tumours into three main groups: luminal, HER2 overexpression subtype and triple-negative breast cancer (TNBC) [8]. Nevertheless, with the development of the molecular techniques and the tissue microarray technology at the beginning of the XXI century, it was possible to apply gene expression profiling (GEP) to assess BC heterogeneity in a more accurate manner [9,10]. One example of the combined approach of IHC status and GEP was the study performed by Sørlie and colleagues in 2001, in which they described a 'molecular portrait' of BC by using 456 cDNA clones. According to their research, BC tumours were classified into five subtypes with diverse clinical outcomes: luminal A, luminal B, HER2 overexpression, basal (or TNBC) and normal-like tumours [11]. Each subtype proposed by Sørlie was associated with a specific IHC status, except for normal-like, which shares IHC status with luminal A differing only in their expression pattern, with the normal-like resembling the normal breast profiling [10]. For this reason, the results of Sørlie set the standard for BC tumour stratification. Nevertheless, there are alternative classifications, such as the one proposed by Fan and colleagues, who classified BC in 4 different categories, where the normal-like subtype was not identified according to Sørlie's results [10].

The luminal group is divided into two different subtypes, luminal A and luminal B. Luminal A is positive for both ER and PR but negative for HER2 marker. It also has a low proliferation index (Ki67 < 14%). On the other hand, luminal B is characterised by positive expression of ER, variable expression of PR and high

expression of Ki67 ($\geq$ 14%). The luminal group is the most frequent subtype of BC and it is associated with a good prognosis [12,13]. The HER2 overexpression subtype shows the negative expression of hormonal receptors, high expression of HER2 as well as high proliferation level (Ki67 $\geq$ 14%)[7]. Basal like or TNBC is characterised by the absence of hormonal receptors and HER2 expression and it shows an elevated expression of Ki67 ($\geq$ 14%). TNBC is a highly aggressive BC subtype, and it lacks a specific treatment. Therefore, TNBC is linked to an adverse disease outcome [14] (table 1).

Table 1. BC subtype classification according to IHC status.

IHC status	Luminal A	Luminal B	HER2 overexpression	Basal-like (or TNBC)
ER	+++	+	-	-
PR	++ (>20%)	+/-	-	-
HER2	-	-/+	+++	-
Ki67 (%)	Low (<14%)	High (>14%)	High	High
Outcome	Good	Intermediate	Poor	Poor

Abbreviations: IHC immunohistochemistry; ER Estrogen Receptor; PR Progesterone Receptor; HER2 Human Epidermal growth factor Receptor 2.

1.3. BC stages

The Tumour-Node-Metastasis (TNM) staging system is the most effective method to divide tumour progression into different categories of clinical usefulness for patient treatment recommendations and prognosis, although there are several other cancer-staging approaches currently available [15]. The TNM system was originally proposed by Pierre Denoix in 1943-1952 and, in 1958 the first recommendations of breast and larynx cancer staging were published [16]. Afterwards, since the origin of the American Joined Committee on Cancer (AJCC) in 1977, it has been developed a staging system grounded in the TNM method. The staging approach suggested by AJCC has evolved throughout time, updating this system as new knowledge regarding cancer-staging and prognostic factors are published [17].

TNM cancer-staging system is based on anatomic traits, such as tumour size (T), nodal status (N) and metastases (M) [17]. The TNM staging method classifies the primary tumour (T) in different categories (TX, Tis, T0-T4), according to the existence, size and extent of the primary tumour. The lymph node status (N) is divided into five groups (NX, N0-N3), determined by the presence and the regional extent of nodal invasion. The metastatic state (M) is subdivided into two categories, M0 and M1, which represent the presence or absence of metastatic lesions, respectively [18,19] (table 2).

Table 2. TNM status (based on AJCC Cancer Staging Manual).

	Primary Tumour (P)
TX	Primary tumour cannot be assessed
T0	No evidence of primary tumour
Tis	Carcinoma in situ (early cancer that has not spread to neighbouring tissue)
T1	Tumour ≤ 20mm in greatest dimension
T2	Tumour > 20mm but ≤ 50mm in greatest dimension
T3	Tumour > 50mm in greatest dimension
T4	Tumour of any size with direct extension to the chest wall and/or to the skin (ulceration or macroscopic nodules); invasion of the dermis alone does not qualify as T4
	Lymph Nodes (N)
NX	Regional lymph nodes cannot be evaluated
N0	No regional lymph node involvement
N1	Metastases to movable ipsilateral level I (low-axilla) and II (mid-axilla) axillary lymph node(s)
N2	• Metastases in ipsilateral level I (low-axilla) and II (mid-axilla) axillary lymph nodes that are clinically fixed or matted … or … • Metastases in ipsilateral internal mammary lymph nodes in the absence of axillary lymph node metastases
N3	• Metastases in ipsilateral infraclavicular (level III axillary) lymph node(s) with or without level I (low-axilla) and II (mid-axilla) axillary lymph node involvement … or … • Metastases in ipsilateral internal mammary lymph node(s) with level I (low-axilla) and II (mid-axilla) axillary lymph node metastases … or …

- Metastases in ipsilateral supraclavicular lymph node(s) with or without axillary or internal mammary lymph node involvement

Metastases (M)	
M0	No distant metastasis
M1	Distant metastasis detected

In 2017, it was published the 8th Edition of the AJCC TNM classification of malignant tumours, which maintains its basis on the traditional TNM system but it incorporates biomarkers into the anatomical cancer-staging method, such as ER, PR, HER2, histologic grade and commercial multigene assays [17,19]. In the 8th edition of AJCC Staging Manual, patients are clinically classified by the classical TNM anatomic system as well as by the information provided by the biomarkers. This information generates a Clinical Prognostic Stage Group, which should be determined by initial assessment before any systemic therapy.

2. METASTASIS

Metastasis is a complex and multi-step process by which new tumour lesions are generated in locations far from the primary tumour, hence becoming a systemic disease [20]. Metastasis is responsible for 90% of all cancer-related deaths [21]. Despite the relevant advances that have been made in prevention, diagnosis and treatment, BC metastasis continues to be an incurable disease. In fact, between 5% and 10% of BC patients show metastasis at the time of diagnosis and about 30% of all BC patients will develop metastasis throughout the time of treatment [22].

The metastatic process can be divided into different stages of development: individual or collective migration and invasion; intravasation and survival into the bloodstream; extravasation; colonization, survival and growth in a distant tissue/organ (secondary site) [23] (figure 1).
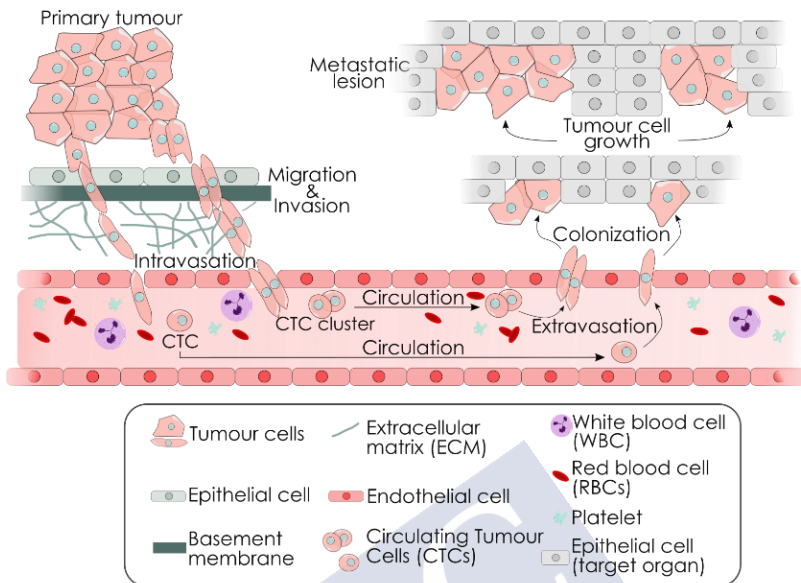
Figure 1. Overview of the multiple steps of the metastatic cascade.

2.1. Migration and Invasion

The metastatic cascade starts with the migration of tumour cells and the invasion of surrounding tissues. The migration step could be performed as single-cell migration or as a collective migration process, if a group of connected cells move together fronted by a small number of leader cells [24]. This step involves the modification of cell-cell interactions, as well as the interaction between tumour cells and the extracellular matrix (ECM) [25,26].

Cadherins are a group of proteins that play an important role in mediating cell-cell adhesion [27,28]. It has been largely demonstrated that alterations in the expression levels and expression patterns of cadherins occur during tumour spread [27]. For instance, E-cadherin mediates cell to cell adhesion and is responsible for keeping the polarity and the integrity of epithelial cells [29,30]. It has been shown that the downregulation of E-cadherin not only reduces the adhesion between epithelial cells, increasing cell motility, but also can activate diverse signalling pathways. These signalling pathways confer a higher migratory

and invasive phenotype, such as E-cadherin/β-catenin/T cell factor (TCF)/lymphocyte enhanced factor (LEF) pathway, which activates downstream transcription factors implicated in cancer progression. This can enhance the ability of tumour cells to cross the basement membrane [30]. Nevertheless, the role of E-cadherin in BC metastasis is highly complex. In fact, it has been observed that E-cadherin promotes metastasis in different preclinical models of carcinoma. A study performed by Padmanaban and colleagues showed that the loss of E-cadherin not only increased invasion but also reduced tumour cell proliferation, cell survival, circulating tumour cell number, seeding of cancer cells in distant organs and metastasis development [31].

On the other hand, N-cadherin is not expressed in normal epithelial cells and its expression is associated with a mesenchymal phenotype [32]. N-cadherin promotes the activation of different downstream signalling pathways, like PLCγ, PI3K and MAPK. These signalling pathways boost the detachment of tumour cells as well as their ability to migrate and invade. The upregulation of N-cadherin together with the downregulation of E-cadherin is a relevant event related to the Epithelial to Mesenchymal Transition (EMT). EMT is the process by which epithelial cells lose their epithelial traits, acquiring a more undifferentiated mesenchymal-like phenotype [33,34]. EMT promotes the loss of apical-basal cell polarity, the cell-cell junctions, and the adherence of cells to their ECM. Hence, EMT has a relevant role in tumour progression as it increases the ability of cells to invade, intravasate and produce proteases involved in ECM degradation [35]. The ECM degradation is mainly performed by matrix metalloproteinases (MMPs), whose proteolytic activity increases the accessibility to blood vessels and the lymphatic system [36,37].

The adherence of tumour cells to the ECM is mediated by integrins [28]. It has been observed that aberrant integrin activity is associated with different diseases, including cancer [38–40]. Integrins contribute to upregulate the proteolytic activity of MMPs, thus altering the cell adhesion dynamics of tumour cells

[41]. Therefore, integrins are being used as potential therapeutic targets due to their implication in tumour progression [40,42–44].

2.2. Intravasation, extravasation, and survival into the bloodstream

Developing the ability to invade and migrate allow tumour cells to reach access to blood or lymphatic vessels. Intravasation is an early step of metastasis by which tumour cells cross the endothelial barrier and enter the lumen of blood or lymphatic vessels (hematogenous and lymphatic dissemination, respectively). In this way, intravasated tumour cells become Circulating Tumour Cells (CTCs). CTCs can travel as either, individual cells, or as a small oligoclonal group of cells, called CTC cluster [45]. BC can generate metastasis through the axillary lymph nodes. However, hematogenous dissemination is the main route for BC tumour spread [46]. The mechanisms of intravasation are the same as those that drive extravasation [47,48]. Hematogenous extravasation can be an active or a passive process. Thus, CTCs could be physically trapped in capillaries of reduced diameter, or they can adhere to the endothelium through the expression of complementary ligands and receptors in both, CTCs and endothelial cells [49]. This depends on the tumour type, the blood vessel structure and the conditions present in the tumour microenvironment. This step involves complex interactions between tumour cells and endothelial cells in order to pass through the endothelial cell junctions and exit the bloodstream. These interactions include different molecule types, such as proteins, lipids, and carbohydrates [50]. It has been reported two different mechanisms of CTC extravasation: diapedesis and angiopellosis. Diapedesis is a multistep process similar to the mechanism used by white blood cells (WBCs) to exit the bloodstream. It consists of different steps (cell rolling, adhesion to endothelium and transmigration), and requires the presence of complementary receptors and ligands in CTCs and endothelial cells, as well as the deformation of the CTCs to pass through the wall of the blood vessel [51]. Angiopellosis also needs the attachment between CTCs and endothelial cells. However, this extravasation process is mainly carried out by the active

deformation of the endothelial cells, that surround CTCs for their later expulsion [52,53].

During circulation, CTCs should survive the environmental stress of the bloodstream to successfully develop metastasis. In this step, CTCs are exposed to different stress sources, such as the absence of cell anchorage that promotes anoikis. Anoikis is a form of cell death due to the loss of cell-cell, and cell-ECM adhesions. In this regard, CTCs that have undergone EMT could have a survival advantage compared to the epithelial CTCs, since they are less dependent on cell-cell and cell-ECM adhesions, and this also explains the presence of CTCs with mesenchymal traits in BC patients [49,54]. Besides, CTC clusters could have a greater resistance to anoikis due to the existence of strong cell-cell junctions within the cells in the cluster [55,56]. CTCs are also exposed to the action of the immune system, oxidative stress, and the impact with the blood cells [20,57]. The shear stress generated by hemodynamic forces of the bloodstream over CTCs is also one of the major stress sources that compromises CTC survival during circulation. In this regard, biophysical characteristics of CTCs play an essential role in cell survival during dissemination [58–60]. Altogether, these different stress factors present in the circulation explain the fact that the majority of CTCs die within the bloodstream [61,62]. Therefore, this highlights that metastasis is an inefficient process and that survival into the bloodstream is an essential and limiting step of the metastatic cascade.

2.3. Colonization, and growth in a distant organ

After circulation and extravasation in a secondary location, tumour cells should survive in the new niche (called metastatic niche) to properly develop a metastatic lesion. This includes avoiding the action of the immune system, as well as acquiring the ability to proliferate. Thus, those CTCs that have undergone EMT, should revert this process through the Mesenchymal to Epithelial Transition (MET) to re-acquire epithelial traits and the capacity to grow [63,64]. The adaptation of tumour cells to the metastatic niche relies on complex interactions of tumour cells with the cells of the target tissue in the secondary site, the ECM,

and the immune cells infiltrated in the secondary site [65]. In 1889, Stephen Paget developed the 'seed and soil' theory to explain tumour dissemination and the colonization of secondary sites by tumour cells. The 'seed and soil' theory claimed that tumour cells, or 'seeds', can only proliferate and generate metastasis in a suitable microenvironment, or 'soil' [66]. This theory highlighted the relevance of interactions between tumour cells and the tumour microenvironment in the secondary site. Although these interactions remain poorly understood, it is well known that tumour cells can pre-adapt the metastatic niche before colonization to later favour their survival, a concept known as the 'pre-metastatic niche' [67]. The secretion of different systemic signals (cytokines, ECM remodelling enzymes, exosomes) by tumour cells allows the pre-formation of a more permissive microenvironment that will later favour tumour cell growth and the development of metastases [20,25,67,68].

3. CIRCULATING TUMOUR CELLS (CTCs)

3.1. Relevance of CTCs in tumour progression

CTCs were firstly described in 1869 by the doctor Thomas Ashworth when he analysed samples from the autopsy of a patient with metastatic breast cancer (mBC). Thomas Ashworth described the presence of cells with the same traits as tumour cells in a blood sample of the patient. Due to this observation, he claimed that these cells were derived from an existing tumour and that they must have circulated through a large part of the bloodstream [69,70].

CTCs are defined as tumour cells released into the bloodstream from the primary tumour or/and the metastases, and they are responsible for tumour spread [71,72]. The presence of CTCs in the peripheral blood is associated with a poor disease outcome [73]. However, the mechanisms of CTC shedding are still not fully understood, as different mechanisms have been proposed [65]. During the last decades, CTCs isolated from peripheral blood samples, or liquid biopsies, were gaining

relevance in the understanding of the process of seeding metastasis. Unlike classic tissue biopsies, or 'solid biopsies', liquid biopsy is a non-invasive technique that allows CTC isolation to provide real-time information of the tumour status [74]. Nevertheless, technical challenges were limiting the advance in this field, mainly due to the low frequency of CTCs in the blood of cancer patients, and the limitations of the CTC isolation technologies [75]. Recently, the development of a wide range of technologies for CTC enrichment is allowing to gain insight into the essential role of CTCs within the metastatic cascade.

3.2. CTC enrichment technologies

CTC enrichment technologies have the objective of specifically detect and isolate CTCs, distinguishing them from other contaminating cells within a complex sample, such as whole peripheral blood samples. Thus, current CTC enrichment methods are classified into two main categories, according to the principle used to fulfil this purpose: based on biological properties or based on biophysical properties of cells.

3.2.1. CTC enrichment based on biological properties

These CTC enrichment methods are based on the use of antibodies against cell surface markers. The antigens used for CTC immunoisolation could be expressed by the CTCs (positive enrichment) or could be present in the rest of blood cells but not in CTCs (negative selection).

3.2.1.1. Positive enrichment

The CellSearch® system (Menarini Silicon Biosystems) is one of the most used devices for CTC isolation. The CellSearch® is the only system approved by the Food and Drug Administration (FDA) for clinical use in metastatic colorectal, prostate and breast cancer [76,77]. CellSearch® is based on the use of metallic beads functionalized with antibodies against the Epithelial Cell Adhesion Molecule (EpCAM) to detect and isolate CTCs from peripheral blood samples. Moreover, CellSearch® also performs immunostaining of those EpCAM+

cells against cytokeratins (CKs) 8, 18 and 19, the leukocyte marker CD45, and the nuclear staining DAPI (4',6-diamidino-2-phenylindole positivity). Hence, a CTC is defined as a cell EpCAM$^+$/CKs$^+$/DAPI$^+$/CD45$^-$, with a minimum size of 4x4 µm^2 [76,78]. Sample processing in the CellSearch® is mainly used for CTC enumeration, and it allows downstream genomic analysis of the captured cells. However, CellSearch® CTC detection uses whole blood collected in tubes with preservative reagents, which do not allow downstream transcriptomic analysis, neither later recovery of cells in a viable state to perform functional characterization, among other downstream procedures [78]. The CellSearch® also has the limitation of detecting only CTCs with an epithelial phenotype. Hence, it is not capable of detecting CTCs with a more undifferentiated (or stem) mesenchymal traits where low or no EpCAM expression is found. Despite its limitations, CellSearch® has great clinical utility, as a correlation between CTC enumeration and patient outcome was found. For instance, an enumeration of ≥5 CTC/7.5mL of blood correlates with a worse prognosis in metastatic breast and prostate cancer patients, while a cut-off of ≥3 CTC/7.5mL correlates with a negative prognosis in metastatic colorectal cancer patients [79–81].

Unlike CellSearch®, other immune-dependent enrichment methods allow the recovery of CTCs in a viable state, permitting gene expression studies. For instance, CellCollector® (Gilupi) is a positive CTC enrichment methodology based on the detection of EpCAM. CellCollector® is a CE certified medical device that isolates CTCs *in vivo*, directly from the blood of the patient. It consists of a medical stainless steel wire functionalized with antibodies against EpCAM that is placed as a catheter. CellCollector® is able to detect CTCs even in patients with early BC and with no diagnosed metastasis. However, it requires a manual screening for the identification of CTCs [78,82].

CELLection^{TM} Epithelial Enrich Dynabeads^{TM} (ThermoFisher Scientific) is also a positive CTC enrichment kit based on the expression of EpCAM. It can be used with whole blood, bone marrow, or with peripheral blood mononuclear cell (PBMC) samples. Dynabeads^{TM} are magnetic beads coated with

antibodies against EpCAM that are directly added to the sample to allow interaction with CTCs for later separation with a magnet. Moreover, Dynabeads[TM] can be customised for detecting different antigens, such as cell surface proteins or ECM components [83].

There are many technologies based on EpCAM expression for CTC isolation. However, the suitability of EpCAM as a marker for CTC enrichment is currently under discussion, as the role of EpCAM in tumour progression is not fully understood. Besides, it has been reported that EpCAM can have a transient expression due to EMT [84]. Consequently, alternative immune-based methods are being developed for detecting CTCs. The AdnaTest is a kit specifically designed for BC or prostate cancer. It is formed by magnetic beads functionalized with a cocktail of antibodies against different markers (EpCAM, HER2, CA 15-3 (MUC1), and optionally against PR and ER). This system uses whole blood for CTC detection and allows gene expression analyses by quantitative multiplexed PCR (qPCR). Thus, AdnaTest can provide useful information regarding the progression of the disease during therapy, as well as the genetic origin of the isolated CTCs [85–87].

3.2.1.2.   *Negative enrichment*

Moreover, negative immune-dependent enrichment techniques have been developed to overcome the limitations of those EpCAM-based technologies [88]. RossetteSep[TM] (EasySep[TM] Direct Human CTC Enrichment Kit, Stemcell Technologies) is a good example and a widely used negative CTC enrichment system. This kit is used over whole blood samples and it contains a cocktail of antibodies that specifically recognise antigens present in blood cells. The density gradient centrifugation allows the separation of the blood components: contaminating cells (settle down to the bottom), plasma (the upper part), and the CTCs, which are located in the interface between the plasma and the lower part of contaminating cells. CTCs recovered with this system are viable and can be used for downstream assays, such as

functional or molecular phenotyping analyses, as it has been previously demonstrated [89,90].

3.2.2. CTC enrichment based on biophysical properties

Biophysical-based CTC enrichment technologies are generating great interest as they do not rely on the identification of cell surface markers. These can potentially detect and isolate a more heterogeneous CTC population. These label-free enrichment technologies can be based on different physical properties. The main physical properties used by these systems are density (density gradient), size (microfiltration), different size and deformability (microfluidics), or electrical charges (dielectrophoresis) [91].

3.2.2.1. Density gradient

Density-based methods were one of the first technologies applied for CTC isolation. These techniques have the advantage of fast sample processing time. However, they also have limitations, as they show low specificity that generates enrichment of CTCs with low purity [76,78].

Diagnostic Leukapheresis (DLA) is a good example of a density-based method for CTC isolation that can overcome the limitation of low CTC frequency in peripheral blood samples. DLA is a standard procedure used in the clinic to collect Mononuclear Cells (MNCs) due to continuous centrifugation of a large volume of blood [92]. It has been recently proposed that CTCs could be co-isolated along with MNCs during DLA centrifutíon, as they have similar densities [93]. Moreover, DLA screens large volumes of blood, thus increasing the likelihood of CTC detection in comparison with peripheral blood samples [92,94]. It is estimated that the processing of DLA in the CellSearch® system increases between 0 and 32 fold the ability to detect CTCs in comparison with peripheral blood samples (7.5 mL) [95]. It has been reported that DLA allows the enrichment of CTCs in a viable state that can be later used for downstream characterization [94]. Besides, CTCs isolated from cryopreserved DLA products

maintain their viability, which makes sample processing more easily manageable [92].

3.2.2.2. Microfiltration

Microfiltration technologies were developed in the 1960s, based on the higher size of CTCs in comparison with blood cells [91]. Nowadays, there are different microfiltration-based devices for CTC isolation. For instance, ISET® (Isolation by Size of Epithelial Tumour cells) (Rarecell, Paris) was the first size-based device for CTC detection. It uses a calibrated track-etched polycarbonate membrane with 8 µm cylindrical pores for CTC enrichment from diluted peripheral blood samples [91,96].

3.2.2.3. Microfluidics

Microfluidic devices are based on the higher size and lower deformability of CTCs compared with blood cells. Currently, microfluidic CTC enrichment techniques are one of the most used immune-independent methods as they perform different steps of sample processing within a chip. Thus, they integrate the 'lab on a chip' concept and allow continuous sample processing to reduce cell loss.

The Parsortix™ Cell Separation System (Angle plc) is a semiautomated device capable of isolating CTCs from different body fluids, such as blood, bone marrow or ascites [97]. It uses a disposable cassette with serpentine channels that finish in a 6.5 µm gap to capture CTCs, while the rest of the blood cells flow through (figure 2). After blood enrichment, cells can be stained by in-cassette immunofluorescence for CTC identification, and/or recovered by an optional harvesting step. Harvested CTCs are viable, so they can be used to *in vitro* culture, or to perform functional and molecular characterization [98]. Its cost-effectiveness, versatility, simplicity, and label-free technology provide great potential to Parsortix™ and boost its use for a deeper understanding of CTC biology [99,100].
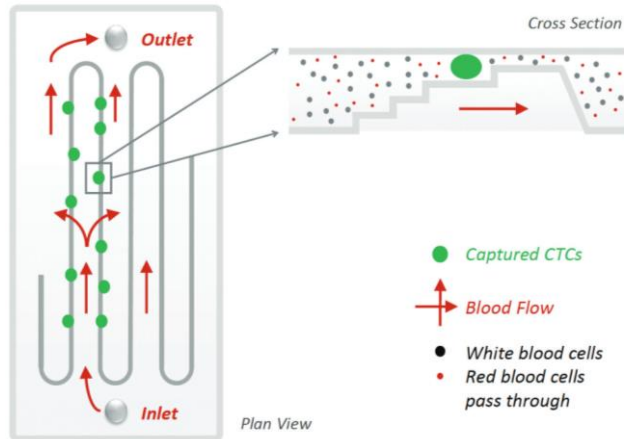
Figure 2. Overview of a ParsortixTM cassette. The sample flows through serpentine channels. These channels have a stair-like silhouette that finishes in a 6.5 μm gap in which CTCs are captured, while the rest of blood cells flow through. Red arrows indicate the direction of the flow inside the cassette.
Image used with permission of Angle plc.

3.2.2.4. *Dielectrophoresis*

Dielectrophoresis (DEP) separates CTCs based on their differential movement in the presence of an electric field, as they are charged particles. CTCs have a differential surface charge in comparison with other cells, thus an electric flow can separate them from the rest of the cells of the sample [101].

DEPArrayTM (Menarini Silicon Biosystems) system combines microfluidics with DEP to detect and isolate individual CTCs based on their electrophoretic properties [76]. CTCs are identified and selected by an image-based analysis. Then, selected cells are isolated inside dielectrophoretic cages and finally, cells are individually recovered for downstream analysis [102]. DEPArrayTM is normally used in combination with other separation techniques, as a secondary isolation method to eliminate the background contaminating cells and increase the purity of the individually isolated CTCs [76,78].

4. CIRCULATING TUMOUR CELL CLUSTERS (CTC CLUSTERS)

CTC clusters (a.k.a circulating tumour microemboli, circulating micrometastasis, circulating tumour aggregates, or tumour cell clumps) are defined as a small group of two or more (>100) tumour cells that travel together through the bloodstream [103]. CTC clusters can be made of tumour cells exclusively (homotypic clusters), or be associated with other non-tumour cells, such as platelets, neutrophils, or PBMCs (heterotypic clusters) [104]. It has been demonstrated that the presence of CTC clusters in the peripheral blood of cancer patients is associated with an adverse outcome. CTC clusters have a higher metastatic potential than individual CTCs [103,104].

CTC clusters were firstly described in 1858 by Rudolph Virchow, a German physician known for being the founder of cellular pathology. Rudolf Virchow postulated the arrest of tumour microemboli in the vasculature as the origin of metastases [105]. However, it was not until 1954 that the first studies regarding the relevance of CTC clusters in tumour progression were published. A work performed by Watanabe demonstrated that CTC clusters were more efficient in metastatic seeding than individual CTCs when intravenously injected in mice at equal numbers [106,107]. Despite knowing the existence of CTC clusters for over a century, the origin of CTC clusters, or the traits that make them more metastatic are still in the process of being elucidated.

4.1. Origin of CTC clusters

The origin of CTC clusters is still under discussion and two main hypotheses have been proposed. One possibility is that CTC clusters could be directly separated from the primary tumour and/or from metastasis. Another possibility is that CTC clusters originate from the aggregation and/or proliferation of individual CTCs within the bloodstream [107] (figure 3).
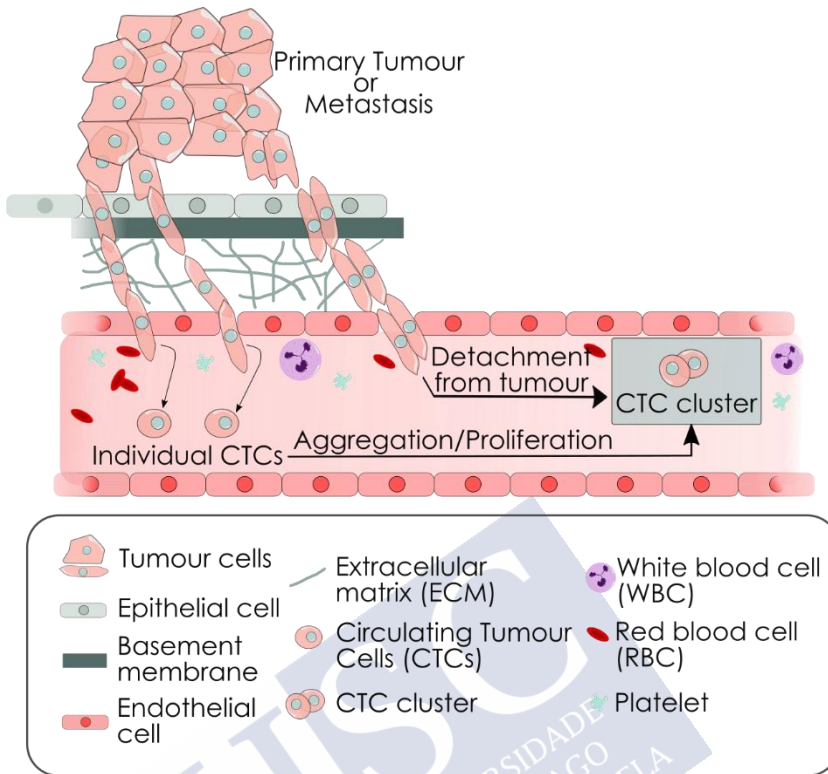
Figure 3. Overview of the possible origin of CTC clusters. CTC clusters can be generated by the direct detachment of a group of cells from the primary tumour or a metastatic lesion. Alternatively, CTC clusters could derive from the aggregation and/or proliferation of individual CTCs within the bloodstream.

Although there is limited evidence about the origin of CTC clusters in BC, some studies point to a direct detachment of a group of cells from the tumour as the origin of CTC clusters and exclude the intravascular aggregation of cells. In fact, the injection of two different fluorescently-labelled tumour cells in the mammary fat pad of opposite flanks of mice demonstrated that BC metastases arise from a collective migration and invasion of a group of cells, and discarded the intravascular aggregation of individual tumour cells [45,108]. Moreover, an *in vitro* assay performed in a platform that mimics the conditions of the bloodstream has shown that this environment does not favour the intravascular aggregation and/or proliferation of single CTCs [107]. In line with this, molecular mechanisms associated with CTC

cluster formation have been reported. For instance, it has been described that plakoglobin (JUP) is essential for CTC cluster formation. In fact, the presence of high expression levels of plakoglobin correlated with an adverse outcome of BC patients and plakoglobin knockdown in murine models reduced CTC cluster formation and metastatic incidence [45]. Plakoglobin is a relevant component of adherens junctions and desmosomes. Hence, plakoglobin can play a fundamental role in keeping the cell-to-cell adhesion of tumour cells within the CTC cluster [45]. The intermediate filament keratin 14 (KRT14) was also found to be associated with CTC clusters. Like plakoglobin, keratin 14 is also a protein related to cell junctions that regulates cell-cell and cell-ECM adhesion. It has been demonstrated that keratin 14 is necessary for collective invasion and dissemination but not for the dissemination of individual cells [108].

On the contrary, a recent work performed with murine Patient-Derived Xenograft (PDX) models of BC showed the feasibility of CTC cluster formation by adhesion of individual cells within the bloodstream. The authors reported that homophilic CD44-CD44 interactions allow cell aggregation to form CTC clusters in circulation, and that CD44 depletion prevented tumour cells from aggregation. Thus CD44 is required for CTC cluster formation by intravascular cell aggregation [109]. In line with this, previous research showed that the primary attachment to endothelia allowed tumour cells to form cellular aggregates during the initial steps of metastasis development [110].

Moreover, it has been proposed the 'cell jamming' as another alternative mechanism for CTC cluster formation during tumour spread. The 'cell jamming' model proposes that the density of the ECM determines the type of invasion (individual or collective) performed by mesenchymal tumour cells. According to this, high density ECM induces cell-cell interactions and collective migration, thus favouring CTC cluster formation [111].

Consequently, the currently available data suggest a possible combination of all of the above-mentioned mechanisms to

explain the formation of CTC clusters. However, further studies are required to shed light on the origin of CTC clusters.

4.2. Metastatic traits of CTC clusters

CTC clusters only represent 1-30% of all CTCs [104]. Despite being a minor subpopulation of CTCs, it has been reported that CTC clusters have a 23-50-fold higher metastatic potential than individual CTCs. Preclinical studies have estimated that CTC clusters are the responsible for 50-97% of metastasis [45,108].

However, the characteristics and mechanisms that make CTC clusters more metastatic than individual CTCs are still not fully understood. In this regard, a study in melanoma demonstrated that tumour cells with low metastatic potential can increase their metastatic ability when they interact and group with tumour cells with a high metastatic capacity [112].

On the contrary, a previous study showed that combining tumour cells with different metastatic potentials within a cluster did not increase the ability to metastasize of the low-metastatic tumour cells. This study showed that the majority of the metastases were developed by those tumour cells with high metastatic potential [113]. Therefore, further work is required to better understand the heterogeneity within CTC clusters and how tumour cells interact with each other in the cluster. Related to this, it should be highlighted that cell heterogeneity can play an important role in tumour spread. CTC clusters can simultaneously exhibit both mesenchymal and epithelial markers [54]. This could be suggesting a possible cooperation between different cell clones within the CTC clusters, although so far it has not been demonstrated. Besides, genomic studies and the use of other OMIC techniques is essential to decipher the potential polyclonality of CTC clusters and the clonal cooperation of cells within CTC clusters, which will enlighten the origin of polyclonal metastases [114].

Moreover, CTC clusters have a higher survival ability in circulation and a higher resistance to apoptosis at the site of metastasis than single CTCs [107]. The survival advantage of CTC

clusters could be partially explained by the existence of strong cell-cell interactions between tumour cells that would make them more resistant to anoikis [56,115].

Furthermore, the higher resistance to apoptosis observed in CTC clusters might also be due to the lower circulation half-life of CTC clusters (6-10 minutes) in comparison with individual CTCs (25-30 minutes) observed in BC murine models [45]. Due to the larger size of CTC clusters, they can be physically entrapped in small capillaries more easily. This would favour CTC cluster survival and later growth to form metastatic lesions [104]. This higher resistance to apoptosis was not only observed in animal models but also in BC patients. A study performed in TNBC by Paoletti and collaborators found that only 0.4% of cells were apoptotic in the CTC cluster population, while in single CTCs this percentage reached 20% [116].

Changes in DNA methylation of cells within CTC clusters isolated from BC patients showed hypomethylation in the binding sites of transcription factors regulating stem-related genes, such as OCT4, NANOG, SOX2, and SIN3A, suggesting the presence of stem-like traits in CTC clusters. This methylation pattern correlated with a higher metastatic capacity of CTC clusters and a worse outcome of a subset of BC patients [117]. The stem-like properties can increase cell plasticity and adaptation capacity to microenvironmental changes, which are characteristics that will favour the development of metastasis [109,117].

CTC clusters have been mainly studied in mBC patients. However, it has been recently reported that it is also possible to detect CTC clusters in early-stage BC, suggesting that dissemination of CTC clusters may be an early event in BC [118,119]. Recently published studies showed that CTC clusters isolated from early-stage BC patients and PDXs had mutations that were not present in the primary tumour and that can potentially provide information about the genes involved in tumour progression [114,120]. Therefore, further studies are required in different BC stages of development to fully understand the role of CTC clusters in tumour spread.

4.3. Clinical value of CTC clusters

It has been largely demonstrated that the enumeration of the CTCs isolated in the CellSearch® system correlates with patient survival in mBC [79,121]. Moreover, CTC enumeration not only has prognostic power in mBC but also in early-stage BC [122]. Generally, a cut-off ≥5 CTCs/7.5 mL of blood is established to differentiate between favourable (<5 CTCs/7.5 mL of blood) and unfavourable (≥5 CTCs/7.5 mL of blood) patient outcomes [121,123].

Unlike for CTCs, CTC cluster enumeration and its potential clinical relevance have not been so widely studied. This could be due to the CTC enrichment technologies, which can potentially affect the CTC cluster integrity, limiting the ability to detect CTC clusters [124]. A prospective randomized phase II trial assessed the predictive value of CTC clusters in TNBC patients, using the CellSearch® system as the enrichment method. This study showed that the continued detection of CTC clusters over time was associated with a shorter survival [116]. This is in agreement with the results of a previous study performed in mBC patients, in which the detection of CTC clusters using a microfluidic device was found to be associated with a worse prognosis [45]. Therefore, CTC clusters have prognostic value in mBC regardless of the enrichment method used to detect them.

Other studies performed in mBC patients not only confirmed the prognostic value of CTC clusters but also highlighted that CTC cluster enumeration can provide additional prognostic value to CTC enumeration alone [125–128]. Additionally, a correlation between the size of CTC clusters and patient outcome has been established. It has been demonstrated that those patients with 3-cell CTC clusters had lower overall survival than the patients with 2-cell CTC clusters [125]. Another recent study related the expression of the stem marker CD44 in CTC clusters with patient survival, observing that the patients whose CTC clusters were CD44$^+$ had shorter overall survival than the patients with CD44$^-$ CTC clusters [109].

Despite the limited number of publications that study the CTC cluster prognostic capacity, current data supports that CTC cluster enumeration provides independent prognostic information, additional to the CTC enumeration alone. Although improvements in the detection of CTC clusters should be made, CTC cluster detection and study could potentially improve BC prognosis, leading towards a more personalised medicine.

The 'Circulating Tumour Cells Clusters (CTC clusters)' section was extracted and adapted from our previously published book chapter:

Relevance of CTC Clusters in Breast Cancer Metastasis. As part of the book: Circulating Tumor Cells in Breast Cancer Metastatic Disease.

Roberto Piñeiro, Inés Martínez-Pena, Rafael López-López.

Author affiliations:

Roberto Piñeiro[1], Inés Martínez-Pena[1], Rafael López-López[2,3].

[1]Roche-Chus Joint Unit, Translational Medical Oncology Group (Oncomet), Health Research Institute of Santiago de Compostela. Santiago de Compostela, Spain.

[2]Translational Medical Oncology Group (Oncomet), Health Research Institute of Santiago de Compostela, University Clinical Hospital of Santiago de Compostela (CHUS/SERGAS). Santiago de Compostela, Spain.

[3]Centro de Investigación Biomédica en Red Cáncer (CIBERONC). Madrid, Spain.

INÉS MARTÍNEZ PENA

Contribution to this work: I, Inés Martínez Pena, was involved in writing and reviewing the manuscript, as well as in the elaboration of the scientific drawing of this chapter.

OBJECTIVES

OBJECTIVES

The main objective of this thesis is to conduct a comparative study between individual CTCs and CTC clusters in order to go in depth in the biology of CTC clusters, to determine the biological traits of CTC clusters that confer their metastatic potential and that will help us to understand the role of CTC clusters in tumour progression. To accomplish this purpose, we have proposed different specific objectives:

1. To isolate CTCs, and specially CTC clusters from liquid biopsy samples from mBC patients.
   - To evaluate the potential of CTC clusters regarding their clinical prognostic utility in a cohort of mBC patients.
   - To evaluate the feasibility of CTC cluster isolation in the Parsortix^{TM} from different types of liquid biopsy samples (peripheral blood samples, DLA products).

2. To establish *in vitro* experimental models of CTCs and CTC clusters by using human BC cell lines and to functionally and molecularly characterise them.
   - To evaluate the metastatic potential of the CTC and CTC cluster models by performing *in vitro* assays that mimic different steps of the metastatic cascade.
   - To assess the metastatic potential of the models in preclinical *in vivo* models (*Danio rerio*, *Mus musculus*).
   - To compare gene expression signatures between individual CTCs and CTC clusters.

3. To use a mouse model of mBC as a source for CTCs and CTC clusters to deepen into their biology
   - To generate an *ex vivo* CTC derived cell line and to functionally and molecularly characterize it as

a potential and more physiological realistic model of CTCs.

- To identify genetic differences between individual CTCs and CTC clusters that could explain their differential contribution to tumour spread.

MATERIALS AND METHODS

MATERIALS AND METHODS

1. METASTATIC BREAST CANCER PATIENT SAMPLES

1.1. Peripheral blood samples from metastatic BC patients

Peripheral blood samples from patients with mBC were provided by the Oncology Department of the University Hospital of Santiago de Compostela (CHUS). Before collecting the samples, patients were properly informed regarding the conditions of the biological sample cession by signing the corresponding informed consent (protocol code: RLL-BL-2015_01), which was previously approved by the Clinical Research Ethics Committee (CEIC) of Galicia (code 2013/462 or 2015/772). Each sample consists of approximately 7.5 mL of peripheral whole blood. Samples were used for CTC and CTC cluster detection by using two different approaches for CTC enrichment. CellSearh® (Menarini Silicon Biosystems) is an immune-dependent device that isolates CTCs and CTC clusters based on their expression of EpCAM. CellSearh® is the only system approved by the FDA for clinical use. On the other hand, Parsortix™ (Angle plc) is an immune-independent system that is based on the differential physical properties of tumour cells for CTC and CTC cluster isolation. More specifically, Parsortix™ isolates CTCs and CTC clusters based on their larger size, and lower deformability, in comparison with the blood cells. For CTC and CTC cluster isolation in the Parsortix™ system, a specific protocol was used for maximising the integrity of CTC clusters and avoiding CTC cluster disaggregation. Thus, this protocol is characterised by a lower pressure, and a lower flow rate, in comparison with standard separation conditions (50 mbar and 99 mbar, respectively).

Peripheral blood samples used for CTC and CTC cluster enumeration and correlation with patient outcome were collected in CellSave Preservative tubes (Menarini Silicon Biosystems) and

processed within the following 96 hours. Blood samples were collected at baseline (before starting the first line of systemic therapy, or before starting a new line of therapy), as well as at follow up, during the course of the treatment. Follow up samples were exclusively collected from those patients who were about to start the first line of systemic therapy. The follow up period varied from 3 to 5 weeks after the beginning of the therapy. Blood samples were processed at the Liquid Biopsy Analysis Unit (Health Research Institute of Santiago de Compostela, IDIS), using the CellSearch® System (Menarini Silicon Biosystems), and the CellSearch® Epithelial Circulating Tumour Cell Kit. Hence, biological samples were enriched for CTCs and CTC clusters using metallic beads coated with an antibody that specifically recognises EpCAM antigen. Moreover, after enrichment cells were stained with fluorescent antibodies that recognise CK 8, 18, and 19. Fluorescent antibody against the WBC marker CD45, and the double-stranded DNA staining DAPI were also used for cell staining. Thus, CTCs were identified as $EpCAM^+/CKs^+/DAPI^+/CD45^-$. CTC clusters were recognised as groups of ≥ 2 cells $EpCAM^+/CKs^+/DAPI^+/CD45^-$, with intact cytoplasm membranes and non-overlapping nuclei. CTC identification was manually evaluated by two trained technicians.

Peripheral blood samples were also used to optimise a workflow for immune-independent CTC and CTC cluster isolation. These samples were collected into Vacuntainer K2 EDTA Blood Collection tubes (BD), when the blood extraction and the sample processing were performed within the same day, or into CellSave Preservative tubes (Menarini Silicon Biosystems), when the collection day and the processing day were different. Blood samples were processed in the size-based isolation system Parsortix^{TM} (Angle plc) to proceed with single CTC and CTC cluster isolation. After separation, tumour cell identification was done by immunological characterisation using a cocktail of fluorescent conjugated antibodies against the epithelial markers EpCAM, the Epidermal Growth Factor Receptor (EGFR), and E-cadherin, as well as an antibody against CD45 for the identification of leukocytes (table 3). Moreover, NucBlue^{TM} Live ReadyProbes^{TM} Reagent (Molecular Probes),

was used for nuclear staining. Therefore, it was considered as CTCs those cells with the following phenotype: EpCAM/EGFR/E-cadherin$^+$, NucBlue$^+$, CD45$^-$.

Table 3. Antibodies used for CTC identification in the Parsortix^{TM} system.

Epitope	Antibody	Proportion
EpCAM	Anti-EpCAM Alexa Fluor® 488 (Biolegend)	1:100
EGFR	Anti-EGFR Alexa Fluor® 488 (Biolegend)	1:100
E-cadherin	Anti-E-cadherin Alexa Fluor® 488 (Biolegend)	1:100
CD45	Monoclonal anti-CD45 PE (Exbio)	1:100

1.2. Diagnostic Leukapheresis (DLA) samples

DLA is a density-based method that isolates cells due to continuous centrifugation of large volumes of blood. DLA is a standard procedure used in the clinic for the isolation of mononuclear cells. However, it has been proposed that CTCs could be co-isolated with the MNCs during DLA centrifugation as they have similar density. Given its origin, DLA samples can increase the likelihood of CTC and CTC cluster detection, in comparison with a 7.5mL of peripheral blood.

DLA products obtained from mBC patients were provided by the research group of Prof. Dr. rer. nat. Hans Neubauer (University Hospital, Heinrich Heine University, Düsseldorf). DLA products were obtained after suitably signing the corresponding material transfer agreement (MTA) for sample cession by the Heinrich-Heine University of Düsseldorf (Germany) (see appendix '*Material Transfer Agreement for the use of Diagnostic Leukapheresis products*'). In particular, it was used DLA samples negative for the presence of CTCs when analysed by CellSearch® system (Menarini Silicon Biosystems), and that were named as DLA$_{CTC-}$.

DLA$_{CTC-}$ were stored in liquid nitrogen upon arrival. DLA$_{CTC-}$ products were thawed quickly in the water bath (37 °C)

while opening the cryotube regularly to avoid cell damage due to overpressure. Then, DLA$_{CTC-}$ products were transferred to a 50 mL tube (Falcon) and diluted with cold 1x EDTA in PBS (100mM) until a total volume of 15 mL. Samples were kept in ice, and they were filtered twice using a 100 µm cell strainer (Fisher Scientific).

DLA$_{CTC-}$ samples were used to perform spiking tests in which a known number of fluorescently labelled BC cells as either single CTCs or as CTC clusters were added to the sample. Particularly, a 20 µL drop containing between 100 and 300 units, that is single cells and clusters, were added to each sample. Before spiking, tumour cells were stained with the nuclear staining NucBlue^{TM} Live ReadyProbes^{TM} Reagent (Molecular Probes), following the manufacturer's instructions. Afterwards, spiked samples were processed in the Parsortix^{TM} system (Angle plc) to optimise a workflow for DLA sample processing. After finishing the separation step with the corresponding protocol, captured CTCs and CTC clusters were counted based on enhanced Green Fluorescent Protein (eGFP) expression to calculate the isolation efficiency. Afterwards, isolated CTCs were harvested, and the CTCs that remained inside the cassette after harvesting were counted, to determine the predicted recovery efficiency.

DLA$_{CTC-}$ samples were used to perform the following assays:

1.2.1. Standard separation protocol vs. Cluster separation protocol

DLA products are characterised by an elevated cellularity that could potentially block the Parsortix^{TM} system during the separation step. Thus, two different separation protocols were evaluated for CTC detection in the Parsortix^{TM} system to select the most suitable protocol to work with DLA samples.

The same DLA$_{CTC-}$ sample was divided into two different subsamples after the dilution step with 1x EDTA-PBS. On the one hand, one of the DLA$_{CTC-}$ subsamples was processed in the Parsortix^{TM} system using the standard separation protocol (figure

4, tube 1). On the other hand, the second DLA$_{CTC-}$ subsample was processed with a protocol specifically designed to maximise CTC cluster capture and to avoid CTC cluster disaggregation (figure 4, tube 2). The cluster separation protocol has lower flow rate and lower pressure (50 mbar) than the standard separation protocol (90 mbar).
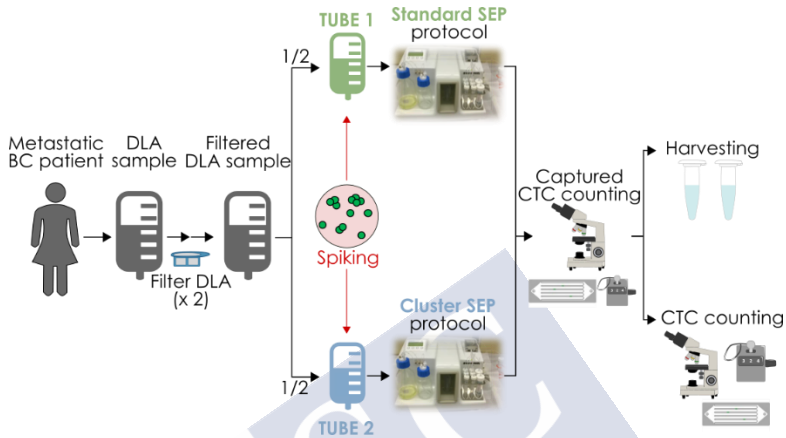

Figure 4. Graphical representation of the experimental workflow followed for the optimization of the Parsortix™ separation protocol using DLA samples. Each DLA sample was divided into two halves. Each half was spiked with a known number of tumour cells as both, individual CTCs and CTC clusters. Afterwards, CTCs were isolated in the Parsortix™ using a standard separation protocol (tube 1), or a cluster separation protocol (tube 2). Isolated CTCs and CTC clusters were counted and harvested. Finally, the CTCs that remained inside the cassette were counted and a predicted recovery was determined.

1.2.2. Filtered: Before vs. After tumour cell spiking

In this assay, the same DLA$_{CTC-}$ product was divided into two different tubes after the dilution with 1x EDTA-PBS (100 mM). One of the DLA$_{CTC-}$ subsamples was passed through the 100 µm cell strainer twice before spiking (figure 5, tube 1), while the other DLA$_{CTC-}$ subsample was filtered after performing the spiking step (figure 5, tube 2).

Later on, samples were processed in the Parsortix™ system using the standard separation protocol, and CTC counting was

performed after separation, as well as, after the harvesting step to determine the isolation and the predicted recovery efficiencies.
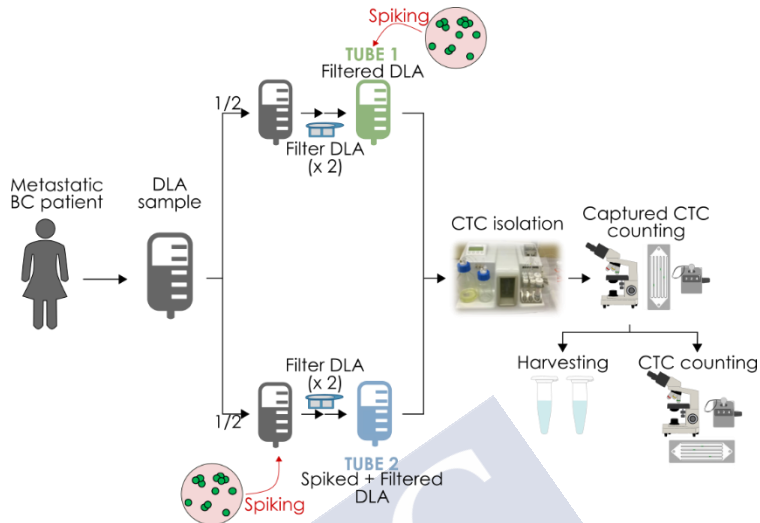

Figure 5. Graphical representation of the experimental workflow followed for evaluating the potential effect of the filtration of DLA samples over the CTC cluster enrichment. Each DLA sample was divided into two halves. Each half was spiked with a known number of tumour cells as both, individual CTCs and CTC clusters after filtration (tube 1) or before filtration (tube 2). CTCs and CTC clusters were isolated in the Parsortix™ using the standard separation protocol. Isolated CTCs and CTC clusters were counted and harvested. The CTCs that remained inside the cassette were counted and a predicted recovery was determined.

2. GENERATION OF INDIVIDUAL CTC AND CTC CLUSTER *IN VITRO* MODELS

2.1. Cell lines

Two different human BC cell lines were used to obtain *in vitro* models of single CTCs and CTC clusters: the TNBC cell line MDA-MB-231 and, the luminal A cell line MCF-7. Both cell lines were fluorescently labelled by expressing eGFP, and they also express the luciferase gene (Luc). MDA-MB-231$^{eGFP\ Luc}$ cell line was purchased at Tebu-bio (Spain), and MCF-7$^{eGFP\ Luc}$ line was purchased at GeneCopoeia, Inc. (USA). Cells were cultured

with DMEM High Glucose (Biowest), supplemented with 10% of Fetal Bovine Serum (FBS) (Corning), and 1% of Penicillin/Streptomycin (P/S) (Lonza). Cells were maintained *in vitro* at 37°C, and 5% CO_2.

2.2. Generation of single CTC and CTC cluster *in vitro* models

Two different protocols were used to generate single and cluster cell suspensions, according to the experimental necessities.

On the first protocol, cell cultures were kept overnight under low-attachment and serum-deprivation conditions at a density of $1x10^6$ cells/mL to allow cell aggregation. Afterwards, cell aggregates were subjected to a differential mechanical disaggregation to generate an individual cell suspension that mimics single CTCs, or small cell groups that represents CTC clusters (figure 6, a).

In the second protocol, cells were maintained under adherent standard conditions with DMEM High Glucose supplemented with 10% FBS and 1% P/S, at 37 °C, 5% CO_2. Adherent cell monolayers with 80% of confluence were differentially detached using different concentrations of trypsin-EDTA (10X) (Lonza). More specifically, trypsin 1X was used for the generation of the cluster population, while trypsin 10X was used in the case of single cell population (figure 6, b).

Before downstream assays, cell suspensions were counted to make sure that the same number of cells was added to the experiments. A 10 µL fraction of each population was mixed with 10 µL of 0.4% Trypan Blue solution (Sigma-Aldrich), and cells were counted using a Neubauer chamber (VWR International Eurolab, S.L.).
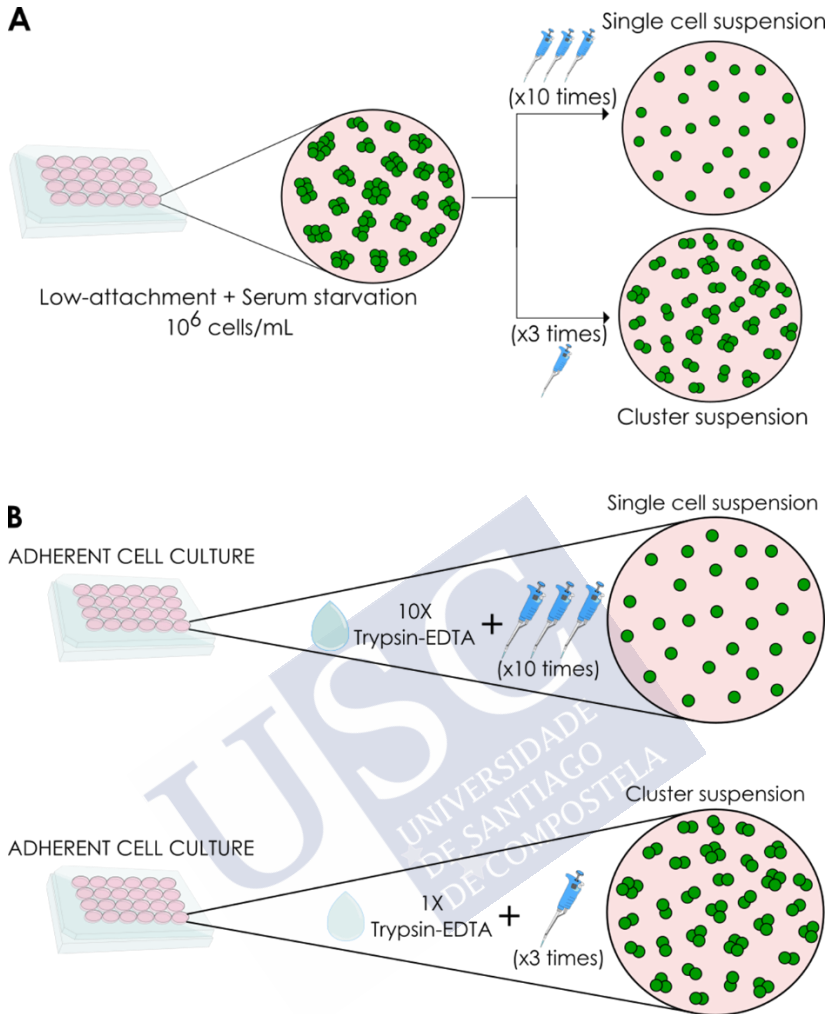
Figure 6. Protocols followed for the generation of *in vitro* single CTC and CTC cluster models. Cells were kept overnight under low-attachment and serum-deprivation conditions at a density of 1x10⁶ cells/mL to allow cell aggregation. Differential mechanical disaggregation was used to obtain a single cell suspension (single CTCs) or cellular aggregates (CTC clusters) (a). Alternatively, cells were cultured under standard conditions. For single CTC and CTC cluster generation, different concentrations of trypsin and differential mechanical disaggregation were used (b).

3. *IN VITRO* ASSAYS

3.1. Proliferation / Cell viability assays

In order to perform proliferation assays, 5×10^3 cells/well were seeded in a 96-well plate (Corning), in a final volume of 100 µL per well. The viability of cell culture was measured periodically to determine the growth rate of the cell population over time. Quantification of cell viability was assessed using two different methods, the resazurin-based solution AlamarBlue™ Cell Vialibity Reagent, and the luciferase activity.

AlamarBlue™ Cell Viability Reagent (ThermoFisher Scientific) is a resazurin-based redox-sensitive solution that uses the reducing power of the cell culture as an index of viability and cell health. AlamarBlue™ was added to each well at a final proportion of 10% of the total volume. Cell cultures were incubated with AlamarBlue™ for 2.5 hours, at 37°C and 5% CO_2. After incubation, the fluorescence signal was measured in a FLUOstar OPTIMA Microplate Reader (BMG Labtech.) using the software FLUOstar OPTIMA (BMG Labtech.).

For proliferation assays using the luciferase activity quantification, cells were cultured in a black 96-well plate (Corning), and D-luciferin (PerkinElmer) was added at a final concentration of 150 µg/mL immediately before quantification. Luciferase activity was determined by measuring the bioluminescent signal derived from the reaction catalysed by the enzyme. Bioluminiscence measurement was performed in the system EnVision Multilabel Plate Reader (PerkinElmer), and the software Wallac EnVision Manager 1.12 (PerkinElmer).

3.2. Transwell migration assays

With the purpose of assessing the migration ability of the *in vitro* single CTC and CTC cluster models, 6.5 mm Transwell™ with 8.0 µm pore polycarbonate membrane inserts (Corning) were used. For this assay, each Transwell™ insert was coated with 1:10 diluted Growth Reduced Factor (GRF) Matrigel® (Coning) to ensure the active migration of tumour cells through

the Traswell^{TM} membrane. A total of 5×10^4 cells/insert were seeded on the Transwell^{TM} either as single CTCs or CTC clusters, with a 10% of FBS (Corning) gradient (figure 7, a). Duplicates of each CTC *in vitro* model were seeded in each experiment. Cells were incubated at 37°C, 5% CO_2, during 8 hours for MDA-MB-231$^{eGFP\ Luc}$ CTC *in vitro* models, or during 24 hours for MCF-7$^{eGFP\ Luc}$ derived CTC *in vitro* models. Afterwards, the cell suspension was removed, the Transwells^{TM} washed twice with Phosphate-Buffered Saline (PBS) (Biowest), and non-migrated cells were removed by scraping the upper side of the Transwell insert with a cotton swab. Migrated cells were fixed with a solution of paraformaldehyde (PFA) (Thermo Scientific) 3%, and glutaraldehyde 1% (Sigma) in PBS, for 30 minutes at Room Temperature (RT). After fixation, inserts were washed twice with PBS and migrated cells were counted using the fluorescence microscope Leica DMi8 (Leica Microsystems), with a 10x objective. Fourteen different representative fields were counted per insert (figure 7, b).
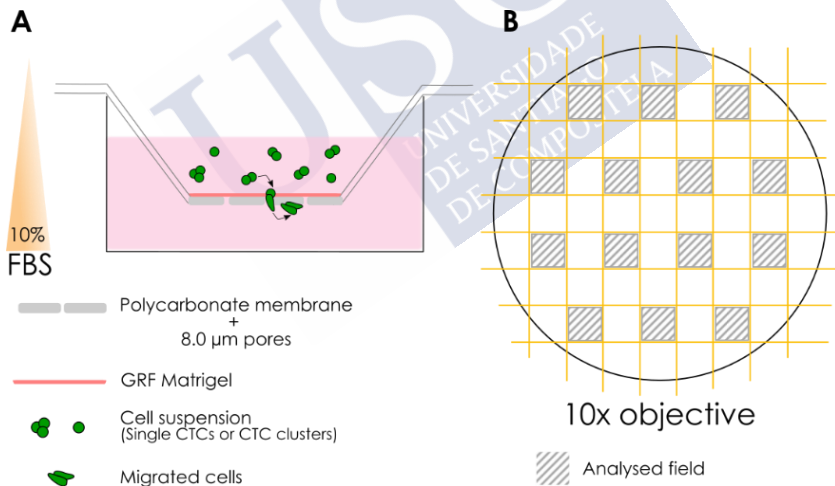

Figure 7. Transwell migration assay. Graphical representation of Transwell migration assay (a). Representation of the analysed fields. Each square with grey parallel diagonal lines represents an analysed field (b).

3.3. Transswell invasion assay

To test invasion, BioCoat Growth Factor Reduced Matrigel Invasion chamber with 8.0 µm PET membrane (Corning®) were used. A total of $5x10^4$ tumour cells were seeded under serum starvation conditions, as single CTCs or CTC clusters (figure 8). Duplicates of each CTC *in vitro* model were seeded in each experiment. The assay was incubated at 37°C, 5% CO_2 for 24 hours in the case of MDA-MB-231[eGFP Luc], and for 48 hours in the case of MCF-7 [eGFP Luc] cell line. Afterwards, non-invaded tumour cells were removed by scraping the upper side of the Transswell insert with a cotton swab and invaded tumour cells were fixed with 3% PFA (Thermo Scientific) and 1% of glutaraldehyde (Sigma) for 20 minutes. Finally, tumour cells that invaded the Matrigel and reach de transswell membrane were counted by microscopy. Seventy different Fields of View (FOV) per well were counted.
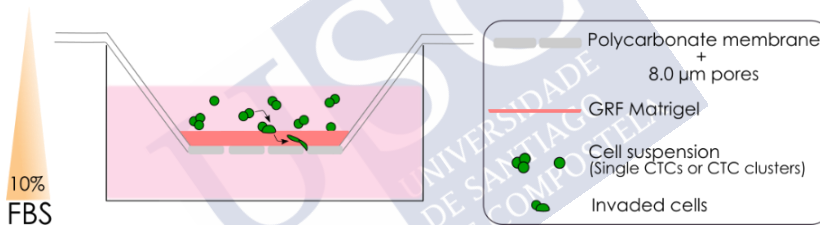
Figure 8. Graphical representation of the Transswell invasion assay.

3.4. Fluid Shear Stress (FSS) assays

For FSS assays, cell suspensions with a concentration of $5x10^5$ cells/mL were loaded into a 5 mL Norm-Ject syringe (Henke Sass Wolf). The syringes were placed into the multichannel programmable NE-1600 syringe pump (New Era Pump Systems, Inc.), whose parameters were set to mimic the flow conditions of the bloodstream (table 4). Moreover, each syringe was linked to a Microlance Hypodermic 30 gauge (G) ½'' needle (BD), in order to generate shear stress.

Table 4. Parameters of the fluid shear stress assay.

Syringe diameter (mm)	12.5
Flow rate (mL/min)	4.25
Pressure (dyn/cm^2)	1840
Volume (mL)	4

It was considered one passage (P) when the entire content of the syringes passed through the 30G needle, which was collected in a 15 mL clear polypropylene (PP) centrifuge tube (Corning). A total number of 10P were performed in each FSS assay, and a 2-minute break between consecutive passages was allowed to permit cell resting (figure 9).
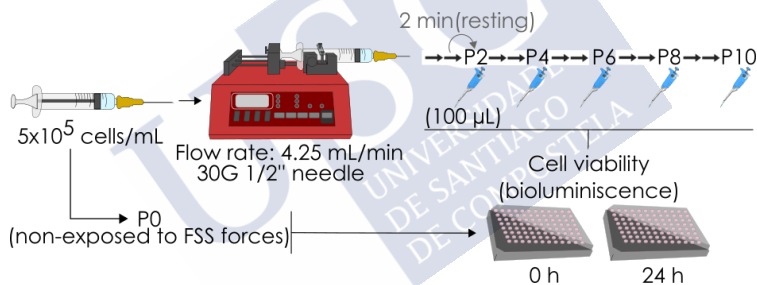
Figure 9. Workflow of the fluid shear stress (FSS) assay. Cell suspensions were loaded into a 5 mL syringe that was linked to a 30G needle, as well as to a syringe pump. It was considered a passage (P) when the cell suspension completely flowed through the needle. Each FSS assay consisted of 10P with a 2-minute break between subsequent passages. Cell viability was assessed by bioluminescence after finishing the assay (0 h), as well as 24 hours later (24 h).

Furthermore, 100 μL subsamples were taken before the first passage (P0), as well as after the even passages (P2, P4, P6, P8, and P10) and they were collected into a 96-well Flat Clear Bottom Black Polystyrene TC-treated Microplates (Corning) in order to evaluate cell viability right after finishing the assay (0 h). Besides, subsamples after even passages were collected to put in culture

for 24 hours and measure the cell viability at the 24 hour timepoint (figure 9). Cell viability was assessed across the passages at 0 and 24 hours after the FSS assay by the bioluminescent signal emitted by the luciferase enzyme when added XenoLight D-Luciferin - K+ Salt Bioluminescent Substrate (PerkinElmer) as a substrate for the reaction. Bioluminiscence was measured using the system EnVision Multilabel Plate Reader (PerkinElmer), and the software Wallac EnVision Manager 1.12 (PerkinElmer).

3.5. Endothelial adhesion assays

For endothelial adhesion assays EA.hy926 cells were seeded at a density of $5x10^4$ cells per well, in 96-well plates (Corning) previously coated with gelatine. Endothelial cells were allowed to grow to generate a cell monolayer, by incubating them at 37 °C, 5% CO_2, for 24-48 hours. Once the monolayer of endothelial cells was formed, $2.5x10^4$ tumour cells per well were added under serum deprivation conditions. A minimum of 4 replicates of each CTC *in vitro* model were seeded in each experiment (figure 10). Tumour cells were incubated together with the endothelial monolayer for 45 minutes, at 37°C, and 5% CO_2, to enable the interaction between both cell types. Non-attached tumour cells were washed away with PBS. Finally, tumour cells adhered to the endothelial monolayer were counted by microscopy, using the fluorescence microscope Leica DMi8 (Leica Microsystems). The whole surface area of each replicate was counted.
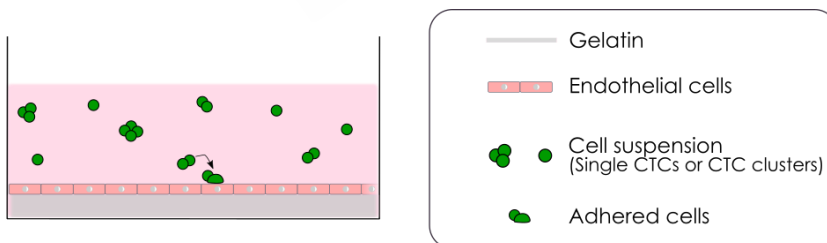

Figure 10. Overview of the endothelial adhesion assay.

3.6. Soft agar colony formation assays

Soft agar colony formation ability was assessed using 6-well and 12-well plates (Corning). Firstly, a lower layer of 0.5% agarose (Fisher Bioreagents) in the corresponding cell culture media (DMEM High Glucose, or Mammocult™, depending on the cell line) was generated. After the gelling of the bottom layer, $1x10^4$ cells/well (6-well plate), or $5x10^3$ cells/well (12-well plate) with 0.3% agarose in cell culture media were seeded. Duplicates of each CTC *in vitro* model were seeded in each experiment performed in 6-well plate, and three replicates of each condition were seeded in the case of 12-well plate assays.

Cells were incubated at 37°C, 5% CO_2 for 2-4 weeks, depending on the subtype of the cell lines used in each assay. Afterwards, the assays were stopped by incubating the wells with 0.005% crystal violet (VWR Chemicals) in methanol (Sigma-Aldrich) for 1 hour, at RT. Excess of staining was removed by washing the cells with PBS (Biowest) several times. Finally, the number of colonies was determined by microscopy using the fluorescence microscope Leica DMi8 (Leica Microsystems). In the CTC cluster models derived from BC cell lines only macroscopic colonies were counted, while for mCTC colony formation assays a ≥70 µm diameter cut-off was established for determining the number of colonies.

4. *IN VIVO* ASSAYS

4.1. Zebrafish (*Danio rerio*) embryo xenografts

Wild-type zebrafish (ZF) embryos were obtained by natural mating of adult zebrafish. Twenty-four hours post-fertilization (hpf) embryos were maintained at 28 °C, in 0.003% w/v N-phenylthiourea (PTU) (Sigma) in E3 medium, to inhibit melanogenesis and obtain transparent embryos. ZF xenografts were generated using 48 hpf embryos. All procedures were performed according to current legislation (RD53/2013). In the RD53/2013, *Danio rerio* is included in the appendix I, as a

species for experimental procedures that should be protected. However, this law only applies to embryo forms of mammals in the final third of their development, but does not apply to non-mammal embryos, such as zebrafish embryos.

Firstly, 48 hpf embryos were dechorionated and anaesthetized with 0.003% w/v tricaine methanesulfonate (MS-222) (Sigma) diluted in E3 medium with PTU. Once the embryos were immobilized by the anaesthetic, they were put in a 1.5% agarose-coated 100 mm ∅ Petri dish (VWR) to perform the microinjection of cells. Immediately preceding the microinjection, MDA-MB-231$^{eGFP\ Luc}$ and MCF-7$^{eGFP\ Luc}$ cell lines were used to generate *in vitro* models of single and cluster cell suspension, by differential disaggregation (see above '*Generation of single CTC and CTC cluster in vitro models*'). A million cells were washed once with PBS and resuspended in 6 μL of 2% polyvinylpyrrolidone (PVP) (Sigma), which is a polymer that prevents glass capillaries from blocking. Cells were maintained on ice during microinjection.

Secondly, borosilicate glass capillaries (1mm O.D. ×0.58mm) (Harvard Apparatus) were used to make needles using the PC-10 puller (Narishige International). Cell suspensions were loaded into the needles and microinjection was performed by injecting approximately 300 cells per embryo. Microinjection was done by using an IM 300 microinjector (Narishige International), set with an outpout pressure of 68.95 kPa, and 0.03 ms of injection time, as well as the Nikon SMZ800 Stereo Zoom Microscope (Nikon). ZF microinjection was performed in the perivitelline space, next to the convergence of the Duct of Cuvier (DoC) (figure 11, a).

After microinjection, ZF xenografts were examined for the presence of a fluorescent cell mass at the injection site, and the absence of tumour cells in the circulation was verified. Embryos showing tumour cells in circulation after injection were discarded and not considered for analysis. ZF xenografts were kept at 34 °C, a consensus temperature between ZF embryo development and human tumour cell viability. ZF xenografts were monitored at 0

hours post injection (hpi), 24 hpi, 48 hpi and/or 72 hpi by using the inverted fluorescent microscope Leica DMi8 (Leica Microsystems). ZF xenograft monitoring was performed to evaluate cell dissemination, and cell survival. Moreover, caudal dissemination patterns were studied in each type of ZF xenograft, dividing the tail into three different regions, according to the vasculature: dorsal (includes dorsal longitudinal anastomotic vessel), ventral (involves the caudal vein, the posterior cardinal vein, and the dorsal aorta), and lateral (intersegmental vessels) (figure 11, b).

Finally, ZF xenografts were euthanised at 120 hpf by tricain overdose, and visual verification of the absence of heart rate was used as the method for death confirmation. For molecular analysis of disseminated cells in the individual CTC and CTC cluster ZF xenografts, tails were physically separated from the rest of the body after the euthanasia for tumour cell isolation. Tails of each xenograft population were pooled in a 1.5 mL tube (Axygen). Each pool of tails was washed twice with PBS (Biowest), and later enzymatically and mechanically digested by using 100 µL of 0.25% Trypsin-EDTA (Lonza) in DMEM High Glucose (Biowest) at 37 °C, and by up-and-down pipetting for 10-15 minutes. After tissue digestion, trypsin was neutralised by 400 µL of DMEM High Glucose (with 10% FBS), and the digested product was centrifuged at 700 g, for 5 minutes. Cells were washed with PBS and pelleted for performing RNA extraction.

ZF experiments were excluded when the survival rate was less than 70%. A minimum of 30 embryos per group were included in each assay.
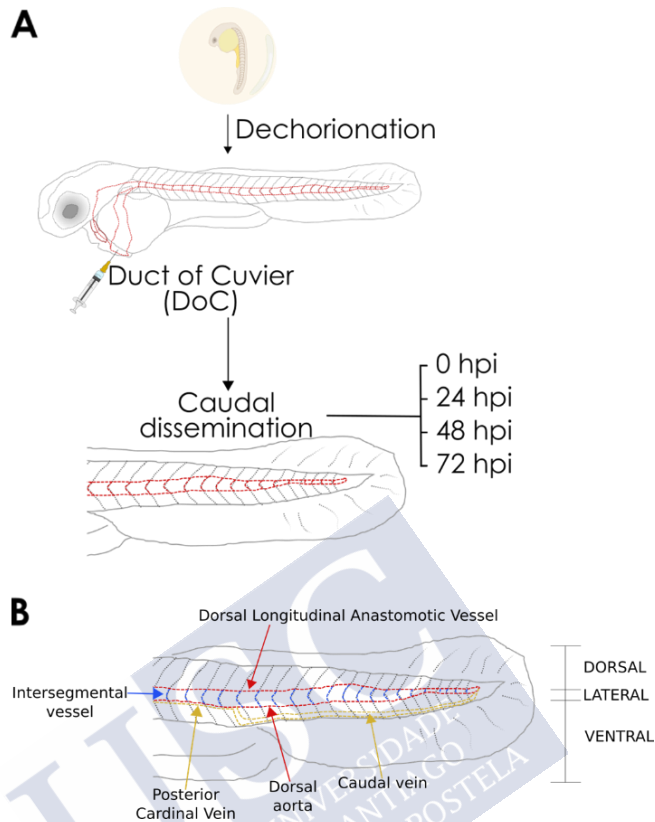
Figure 11. Zebrafish (ZF) embryo xenografts. Representation of the workflow followed for the generation of ZF xenografts. Firstly, embryo dechorionation was performed. Secondly, cells as either individual or as CTC clusters were injected in the perivitelline space, next to the convergence of the Duct of Cuvier (DoC). Finally, cell dissemination to the caudal region was assessed at 0, 24, 48 and/or 72 hours post-injection (hpi) (a). Tail division according to caudal vasculature. The caudal region was divided into three different sections, based on the vasculature: dorsal, lateral, and ventral (b).

4.2. Mouse (*Mus musculus*) experiments

Murine experimental procedures were approved by the Animal Experimentation Ethical Committee of the University of Santiago de Compostela (CEEA) (project: 15010/2019/002), according to the RD 53/2013 (see appendix '*Favourable report Animal Experimentation Ethical Committee of the University of Santiago de Compostela (CEEA)*', '*Animal experimentation*

training certificate. B function (euthanasia)', and '*Animal experimentation training certificate. C function (animal procedures)*'). Mice were obtained from Barcelona Biomedical Research Park (PRBB, Barcelona), or Charles River and hosted in the animal facilities of the Centre for Research in Molecular Medicine and Chronic Diseases (CIMUS, University of Santiago de Compostela) (REGA: ES150780275701), and in the Experimental Biomedicine Centre (CEBEGA, Santiago de Compostela) (REGA: ES150780292901), where they were provided with food and water *ad libitum*, according to the guidelines of the recipient centre.

4.2.1. Mouse lung colonization assay

To assess the lung colonization ability and tumour-initiating capacity of the *in vitro* models of single cells and clusters, suspensions of 5×10^5 cells in 100µL of PBS of the cell line MDA-MB-231$^{eGFP \ Luc}$ were injected into the lateral tail vein of SCID BEIGE mice, either as single cells, or clusters. Tumour cell injection was performed in mice immobilized using a mouse restrainer device, and 1 mL syringes linked to 27 gauge needles (BD Medical).

Then, colonization and tumour-initiating abilities were evaluated by quantifying the pulmonary metastatic incidence over time. Mice were monitored immediately after injection (1 h), as well as every 3-4 days, for 18 days. Lung tumour incidence was tracked by the bioluminescence emitted by the luciferase activity, when 15 mg/mL XenoLight D-Luciferin - K+ Salt Bioluminescent (PerkinElmer) were intra-peritoneally injected as a substrate at a final concentration of 10 µL/g of body weight.

At the endpoint, mice were euthanized by CO_2 inhalation, and cervical dislocation was used to confirm the death. Afterwards, lungs were perfused by tracheal infusion with 1% penicillin-streptomycin (Lonza) in PBS. Lungs were surgically removed and fixated in formol (VWR) to perform later histopathological analysis (figure 12). Histological analyses were performed in collaboration with the Translational Molecular

Pathology research group of the Vall d'Hebron Institute of Research (VHIR) (Barcelona).
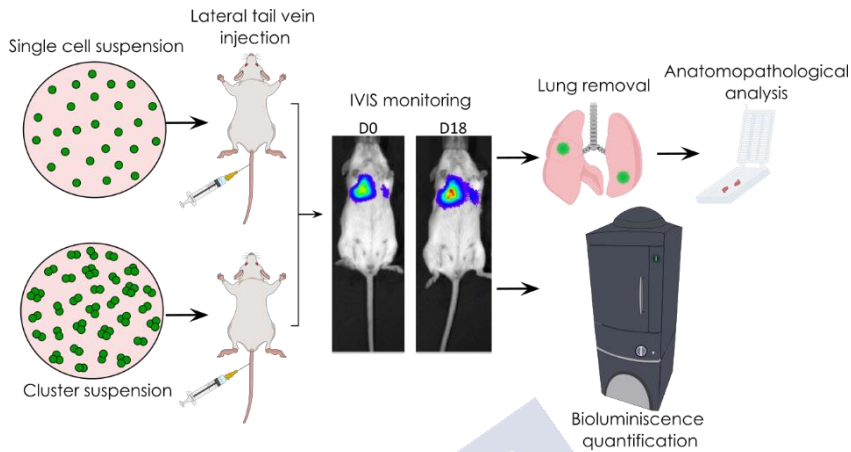
Figure 12. *In vivo* lung colonization assay. Tumour cell suspensions were injected in the lateral vein of the tail as either single CTCs or as CTC clusters. Metastatic incidence in the lungs was tracked over time by animal *in vivo* imaging. Finally, lungs were surgically removed for anatomopathological analysis.

4.2.2. Mouse orthotopic BC xenografts

Tumour cells from the MDA-MB-231$^{eGFP\ Luc}$ cell line were used for this model. A total of 2×10^6 cells resuspended in 60 µL of a mixture 1:3 of DMEM and Matrigel Growth Factor Reduced (Corning) were injected into the mammary fat pad of 8-week-old SCID BEIGE mice, using BD MicroFine^{TM} 0.33 mm (29G) x 12.7 mm insulin syringes of 0.5 mL (BD Medical).

Tumour progression was monitored weekly, during 9-10 weeks using the IVIS Spectrum *in vivo* imaging system (PerkinElmer), and the Living Image Software (PerkinElmer). Animal monitoring was performed measuring by bioluminescence, by intra-peritoneally injecting 15 mg/mL of XenoLight D-Luciferin - K+ Salt Bioluminescent (PerkinElmer), to generate a final concentration of 10 µL/g of body weight.

Once primary tumour (PT) and lung metastases were developed, mice were euthanised by inhalation of CO_2, and

cervical dislocation was used as the death confirmation method. Afterwards, blood samples were collected by cardiac puncture using insulin syringes and BD Vacuntainer K2 EDTA Blood Collection tubes (BD), or CellSave Preservative tubes (Menarini Silicon Biosystems). Murine blood samples were processed in the epitope-independent CTC isolation system Parsortix™ (Angle plc). Isolated CTC were fixed with PFA 2% for 30 minutes, washed with PBS, and pelleted in a final volume of 100 µL. Harvested cells from Parsortix™ were later individualised by the DEPArray™ system (Menarini Silicon Biosystems). These cells were used to perform genomic analysis by Whole Exome Sequencing (WES).

Moreover, primary tumours and lungs were also preserved in formol (VWR) and washed with ethanol (VWR) for performing anatomopathological (AP) examination (figure 13).

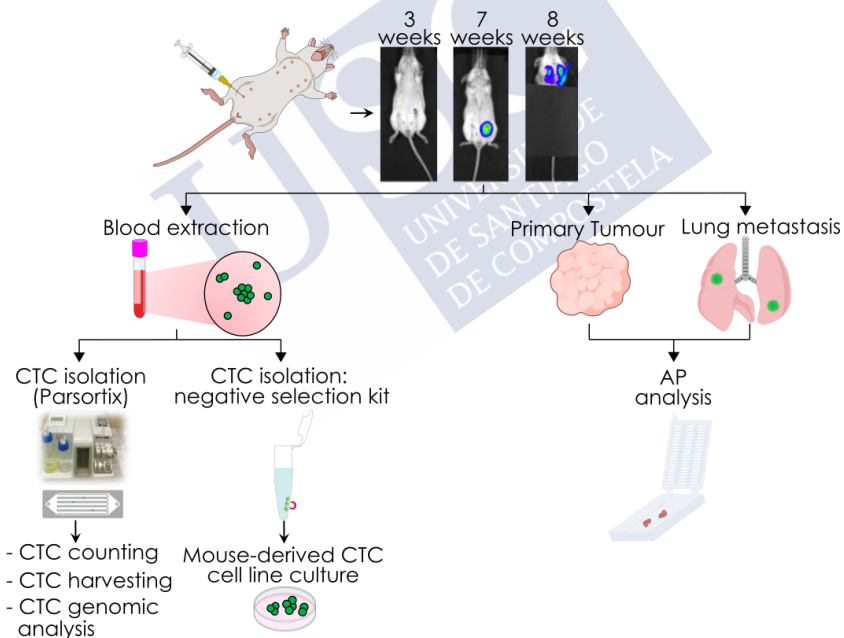**Figure 13. Orthotopic BC model workflow. MDA-MB-231^{eGFP Luc} cells were injected into the mammary fat pad of immunocompromised SCID BEIGE female mice. Tumour progression was monitored over time. Primary tumour (PT) and lungs were surgically removed for later gene expression and anathomopathological (AP) analyses. Besides, blood samples were collected for CTC isolation in the Parsortix™ system. Isolated CTCs and**

CTC clusters were used for genomic analyses, and for the obtention of an *ex vivo* CTC derived cell line.

4.3. mCTC cell line generation

Moreover, a cell line was generated from CTCs isolated from the blood of one MDA-MB-231[eGFP Luc] orthotopically injected mouse. For this, blood was processed using the negative selection kit EasySep[TM] Mouse T Cell Isolation Kit (StemCell Technologies) for CTC isolation. Isolated cells were initially cultured in a low-attachment 96-well plate (Corning), using 200 µL Mammocult[TM] (StemCell Technologies), supplemented with hydrocortisone (0.48 µg/mL), heparin (4 µg/mL), bFGF (20 ng/mL), EGF (20 ng/mL), B27 supplement (4% v/v), progesterone (0.4 µg/mL), β-estradiol (0.4 µg/mL), ultraGRO[TM] (5% v/v), and P/S (1% v/v). Isolated cells were cultured under hypoxic conditions, at 37 °C for a week. Cell culture media was renewed every two days. Afterwards, cells were cultured under low-attachment normoxic conditions (5% CO_2), at 37°C, and with supplemented Mammocult[TM]. These cells grew indefinitely generating cellular aggregates and establishing a cell line called mCTC.

5. MOLECULAR CHARACTERISATION

5.1. Gene expression analysis

To perform RNA extraction from the digested tissue of ZF tails, RNeasy Micro kit (Qiagen) was used, according to the manufacturer's instruction.

RNA quantification was done in the Nanodrop One[C] Microvolume UV-Vis Spectrophotometer (FisherScientific). Additionally, RNA sample quality and integrity was evaluated using the 4200 TapeStation System (Agilent Technologies) and the TapeStation software (Agilent Technologies), in the case of RNA samples whose expression was studied by RNAsequencing.

Equal amounts of RNA were used for generating the cDNA by reverse transcription PCR (RT-PCR). cDNA was obtained by using the enzyme Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT), 200 U/μL (ThermoFisher Scientific), following the manufacturer's protocol. In biological samples with limited availability of material that generated low RNA recoveries, cDNA was generated by the enzyme SuperScript III Reverse Transcriptase (SSIII RT), 200 U/μL (ThermoFisher Scientific). Afterwards, cDNA samples were preamplified with the Taqman Preamp Master Mix (ThermoFisher Scientific) and a pool of Taqman probes. This preamplification step was performed with 14 reaction cycles. cDNA samples were analysed in a LightCycler 480 II (Roche Diagnostics) (table 5), with Taqman Gene Expression Master Mix (Applied Biosystems, ThermoFisher Scientific), and Taqman Gene Expression Assays (Applied Biosystems, ThermoFisher Scientific) (table 6). cDNA obtained from M-MLV RT was diluted 1:20 in Braun water (Braun) for qPCR analysis, while cDNA generated by SSIII RT was diluted 1:5.

qPCR amplification was designed for 45 amplification cycles, with the following steps:

Table 5. Conditions set for the RT-qPCR.

Step		Temperature	Time
Pre-amplification		95 °C	10 minutes
Amplification **(x 45 cycles)**	**Denaturation**	95 °C	10 seconds
	Annealing	60 °C	30 seconds
	Extension	72 °C	1 second
Cooling		40 °C	30 seconds

For data analysis, gene expression levels were determined relative to the average expression of two different housekeeping genes. Specifically, Glyceraldehyde-3-Phosphate Dehydrogenase (*GAPDH*), and Beta-2-Microglobulin (*β2M*) were used as housekeeping genes.

For disseminated cells in the ZF xenografts, the following genes were analysed:

Table 6. List of genes analysed in the disseminated cells of ZF xenografts injected with either, individual CTCs or CTC clusters.

Gene	Taqman assay	Function
B2M	Hs00187842_m1	Housekeeping gene
BAX	Hs00180269_m1	Apoptosis
CCND1	Hs00765553_m1	Cell cycle regulation
CD44	Hs01075861_m1	Stem
CDK4	Hs01565683_g1	Cell cycle regulation/Proliferation
CTNNB1	Hs00355049_m1	Cell adhesion/Gene transcription
E2F4	Hs00608098_m1	Cell cycle regulation/Proliferation
GAPDH	Hs99999905_m1	Housekeeping gene
ITGA6 (CD49f)	Hs01041011_m1	Stem
JUP	Hs00158408_m1	Cell junction
MKI67	Hs01032443_m1	Proliferation
PLAU	Hs01547054_m1	Urokinase
VIM	Hs00958116_m1	EMT

5.2. RNA-sequencing (RNA-seq) análisis

For RNA-seq analysis, RNA samples extracted from paired cell lines derived from murine BC primary tumour and lung metastases were prepared for sequencing at the Wellcome Centre for Human Genetics (WHG), Oxford University. Samples were analysed using the Illumina HiSeq4000 Paired-End system and the GRCh37.EBVB95-8wt.ERCC, as the reference genome. Afterwards, RNA-seq data were analysed by the Bioinformatics Services of the Salamanca University (USAL). Data analysis was focused on the identification of differentially expressed genes between paired cell lines, as well as on the study of the functional enrichment of the genes with differential expression. On the one hand, differential expression analysis was performed using the DeSeq2 statistical package from R software (3.4.3 version). On the other hand, functional analyses were done by using the USAL Bioinformatics Service databases, which were downloaded from National Centre for Biotechnology Information (NCBI).

RNA samples of the mCTC cell line and the '231ULA' population were sent for sequencing to Macrogen Inc.. Samples were sequenced in the Illumina NovaSeq 6000 System, 150 paired-end reads (2x 150 bp), 100 M total reads. The expression profile was determined and expressed as FPKM (Fragment per Kilobase of transcript per Million mapped reads), RPKM (Reads Per Kilobase of transcript per Million mapped reads), and TPM (Transcripts Per Kilobase Million). Differentially expressed genes were analysed using DeseqR, using the cut-off: FC (fold change) ≥ 2, and p value < 0.05. Graphical representation of the differently expressed genes was performed with the FunRich (Functional Enrichment Analysis tool) software.

5.3. Genomic analysis (WES, whole exome sequencing)

Genomic studies from single CTCs and CTC clusters isolated from murine BC blood samples were performed in collaboration with the Phylogenomics research group of the Biomedical Research Centre (CINBIO), in the University of Vigo, headed by Prof. David Posada González. Briefly, CTCs and CTC clusters

isolated from blood samples of tumour-bearing mice were individualised by DEPArray™ (Menarini Silicon Biosystems). The whole exome of the individualised single CTCs (x7) and CTC clusters (x7) was amplified with the Ampli1™ WGS Kit (Menarini Silicon Biosystems) and the integrity of the DNA checked using the Ampli1™ QC Kit (Menarini Silicon Biosystems). DNA samples were sent to Oxford Genomics Centre (University of Oxford) for sequencing in an Illumina HiSeq4000 150 Paired-End ant the GRCh37.EBVB95-8wt.ERCC genome was used as the reference genome.

6. DATA ANALYSIS

For the study of CTC and CTC clusters in our cohort of BC patients, progression-free survival (PFS), and overall survival (OS) were the endpoints studied. PFS corresponded to the period of time between sample collection and the diagnosis of metastasis or death, while OS was the time between blood sample collection and the date of death. Moreover, those patients without an endpoint during follow up were eliminated from the study. Survival studies were done with the Kaplan-Meier method, and the log-rank test was used for statistical comparison between groups. Hazard ratios (HR) were used to determine correlations between the number of CTC and CTC clusters with PFS and OS, and 95% confidence intervals (CI) were obtained with univariate and multivariate Cox proportional hazard models. Multivariate Cox studies were adjusted by age, ECOG (Eastern Cooperative Oncology Group) status, breast cancer subtype, number of metastatic sites, location of metastases, and treatments. Moreover, CTC characteristics, such as CTC cluster presence, were compared at different timepoints and across different breast cancer subtypes by applying the Pearson Chi-squared analysis. Data from this cohort were analysed using the free software 'R' version 3.4.4, and the software GraphPad Prism 6.01 version (GraphPad Software).

Microscopy images were obtained with the LAS X software (Leica Microsystems) and analysed using the free software

ImageJ (NIH Image). Additionally, numerical data as the result of quantifications and measurements, their statistical analysis, as well as their graphical representation, were performed with the software GraphPad Prism 6.01 version (GraphPad Software).

Data from *in vitro* assays were statistically analysed using the Mann-Whitney test (non-parametric test for unpaired samples), except for FSS experiments, in which multiple t-test analysis was performed. ZF *in vivo* experiments were analysed by the Wilcoxon test (non-parametric test for paired samples) when comparing different timepoints within the same group, and by the Mann-Whitney test for comparisons between different groups. Besides, Mann-Whitney test was also applied for statistical analysis of murine breast cancer models.

Finally, molecular data obtained by qPCR was statistically analysed by Mann-Whitney test for comparison between groups.
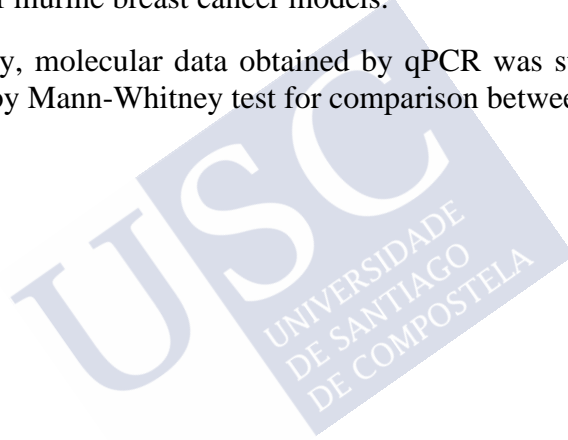
RESULTS

RESULTS

1. CTC CLUSTER DETECTION IN SAMPLES DERIVED FROM METASTATIC BREAST CANCER PATIENTS

1.1. CTC and CTC cluster enumeration in peripheral blood samples

Peripheral blood samples (or liquid biopsy samples) were collected to evaluate the relevance of CTC and CTC cluster enumeration as a prognostic factor of the outcome of a cohort of unselected mBC patients.

1.1.1. Cohort description: patient characteristics

A total number of 54 unselected female mBC patients were included in the cohort. Within this cohort, 27 patients (50.9%) were defined as having Hormone Receptor positive and Human Epidermal Growth Factor Receptor-2 negative (HR$^+$/HER2$^-$) tumours; 11 patients (20.8%) were determined as having HER2$^+$ tumours; and 15 (28.3%) were defined as having triple-negative tumours (HR$^-$/HER2$^-$). Moreover, 1 patient showed undetermined levels of HER2 histological staining. The follow up period from baseline for alive patients varied from 22 to 603 days, and the median follow up time was 197 days. Besides, 40 patients (74.1%) had visceral metastases, while 14 patients (25.9%) showed non-visceral metastases. Visceral metastases included lung, liver, peritoneal and/or pleural location of metastatic lesions. Non-visceral metastases referred to lymph node and/or bone involvement.

Eastern Cooperative Oncology Group (ECOG) performance status was also determined at baseline for this cohort of patients. ECOG is a standardised method to measure the impact of the disease in the daily routine of patients. According to the ECOG performance status, 18 patients (33.3%) were fully active, and able to continue performing all the activities without restriction (grade 0); 31 patients (57.4%) showed restrictions on the physical activity of high intensity due to the disease

but were able to perform light or sedentary activities (grade 1); while 5 patients (9.3%) were able to do all selfcare activities but were unable to do work activities, and were up and about more than 50% of waking hours (grade 2).

Moreover, NHG (Nottingham Histological Grade) was used to determine the characteristics of primary tumour, based on tubular/glandular formation, tumour cell mitotic activity, and tumour cell nuclear pleomorphism. NHG scale showed that 4 patients (7.4%) had a well differentiated primary tumour (grade I); 26 patients (48.1%) showed a moderately differentiated tumour, while 15 patients (27.8%) had a poorly differentiated primary tumour. Besides, 9 patients (16.7%) had an unknown primary tumour NHG status. Table 7 describe in detail patient characteristics.

Table 7. Patient characteristics (n = 54).

Variables	Total, n (%)
Mean age (years)	58.5
< 65 years	34 (63.0)
≥ 65 years	20 (37.0)
Tumour stage	
IV	54 (100.0)
Baseline ECOG	
0	18 (33.3)
1	31 (57.4)
2	5 (9.3)
Primary Tumour NHG	
I	4 (7.4)
II	26 (48.1)
III	15 (27.8)

131

Unknown	9 (16.7)
Breast cancer subtypes	
HR$^+$/HER2$^-$	27 (50.9)
HER2$^+$	11 (20.8)
HR-HER2$^-$ (TNBC)	15 (25.8)
Unknown	1 (1.8)
Site of metastases *	
Visceral	40 (74.1)
Non-visceral	14 (25.9)
Number of metastatic sites	
< 3	29 (53.7)
≥ 3	25 (46.3)
First line of systemic therapy (n = 44)	
Chemotherapy	23 (52.3)
Hormonal therapy	15 (34.1)
Target therapy **	6 (13.6)
Other lines (n = 10)	
Chemotherapy	8 (80.0)
Hormonal therapy	1 (10.0)
Target therapy **	1 (10.0)

Abbreviations: ECOG Eastern Cooperative Oncology Group; NHG Nottingham Histologic Grade; HR Hormone Receptor; HER2 Human Epidermal growth factor Receptor-2; TNBC Triple-Negative Breast Cancer.
*Visceral metastases include lung, liver, peritoneal and/or pleural locations. Non-visceral metastases refer to lymph node and/or bone involvement. **In combination with chemotherapy

1.1.2. CTC and CTC cluster enumeration and correlation with clinicopathological variables

A total number of 96 peripheral blood samples were used for individual CTC and CTC cluster detection. These liquid biopsy samples derived from 54 different patients that were recruited at baseline, 38 of whom had follow up sample collection. Baseline samples were collected before starting first line of systemic therapy, or before starting a new line of therapy (figure 14). The FDA-approved CellSearch® system allowed the detection and enumeration of CTCs and CTC clusters in these peripheral blood samples.
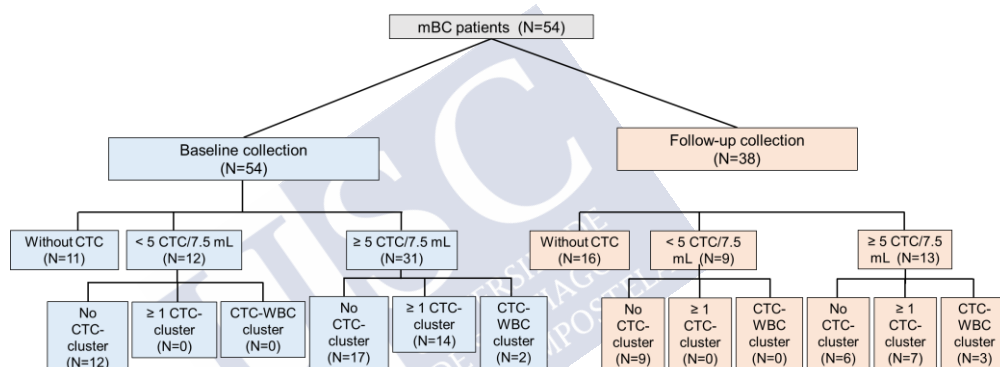

Figure 14. CONSORT flow diagram of the cohort of metastatic breast cancer (mBC) patients. Patients are stratified according to the presence or absence of CTCs and CTC clusters at baseline and during follow-up. N refers to the number of patients.

At baseline, 43 patients (79.7%) showed presence of CTCs, while 11 (20.4%) had no detectable CTCs in their blood. Besides, 12 (22.2%) of those patients positive for CTCs at baseline, showed less than 5 CTCs per 7.5 mL of blood; while 31 of them (57.4%) had 5 or more CTCs, and 14 patients (25.9%) showed at least one CTC cluster (a group of cells made up of ≥ 2 CTCs).

During follow up, 16 (42.1%) patients showed no detectable CTCs, and 22 (57.9%) of them were positive for CTCs. Moreover, 13 (34.2%)

of the CTC positive patients had ≥ 5 CTC/7.5 mL, and 7 (18.4%) of them had at least one CTC cluster in their blood (table 8).

Table 8. Analysis of CTC and CTC clusters in the study cohort.

Baseline	n = 54	Percent (%)	Mean	Min-Max
CTCs	43	79.2	183.1	1-1970
CTCs ≥5	31	57.4	253.1	5-1970
CTC cluster ≥1	14	25.9	4.28	1-14
Follow up	n = 38			
CTCs	22	57.9	95.6	1-969
CTCs ≥5	13	34.2	160.6	5-969
CTC cluster ≥1	7	18.4	12.71	1-74

It should be highlighted that CTC clusters were exclusively detected in those patients with ≥ 5 CTCs per 7.5 mL of blood in both, at baseline (14/31), and during follow up (7/13). The presence of ≥ 1 CTC cluster was associated with the number of individual CTCs detected (figure 15). In addition, it was found a higher frequency of patients with ≥ 5 CTCs and CTC clusters in the HR⁺/HER2⁻ BC subtype. Nevertheless, no significant correlation was found between CTC and CTC cluster detection and BC subtype at baseline, neither during follow up (table 9). Besides, it was not found a correlation between the distribution of CTCs and CTC clusters regarding the number and location of metastatic lesions.

Hence, CTC clusters are a subpopulation of low frequency within CTCs in mBC patients. The presence of CTC clusters is more likely in patients with high individual CTC shedding at baseline and/or during therapy.

a Baseline

Baseline	CTC-clusters=No	CTC-clusters=Yes
< 5 CTC	23	0
≥ 5 CTC	17	14

b Follow up

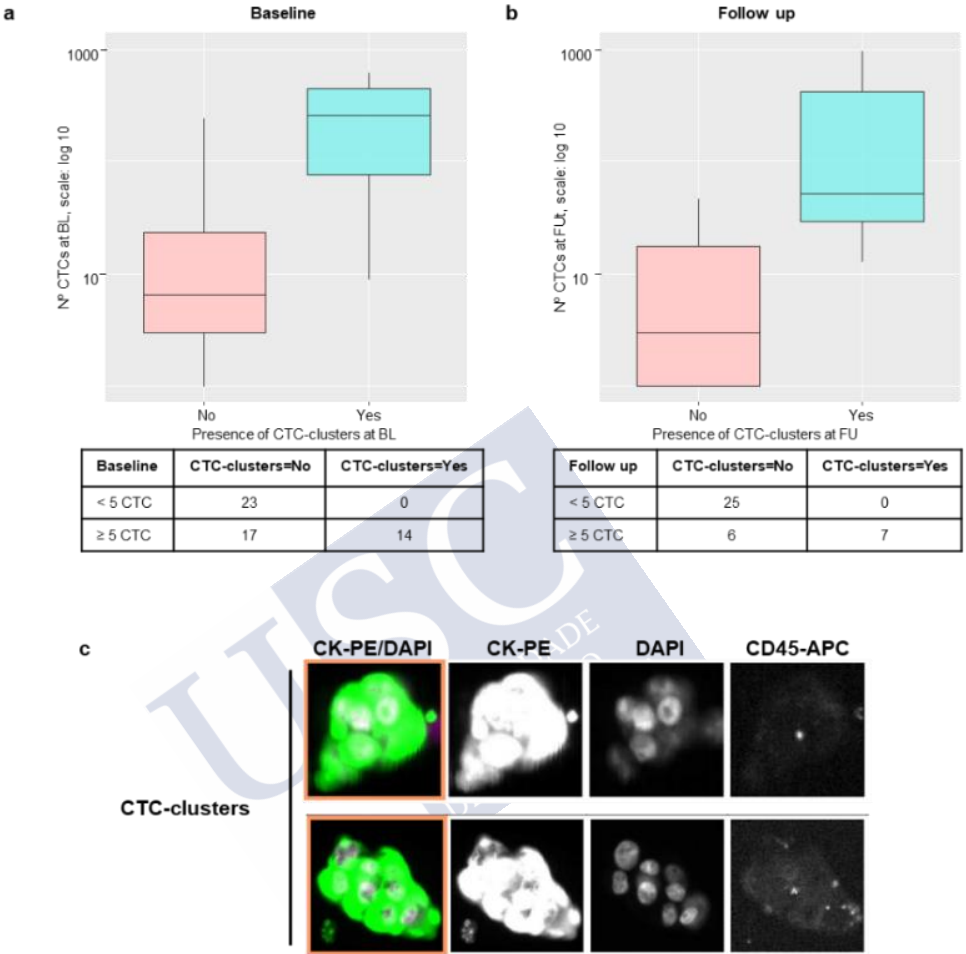| Follow up	CTC-clusters=No	CTC-clusters=Yes
< 5 CTC	25	0
≥ 5 CTC	6	7

Figure 15. Association between the number of CTCs and the presence or absence of CTC clusters. Box plots of the association between the number of CTCs and the presence/absence of CTC clusters at baseline (a), and during follow up (b). Representative images of CTC clusters isolated from the peripheral blood samples of mBC patients using the CellSearch® system (c).

Table 9. Enumeration of CTCs and CTC clusters distributed according to BC subtype.

Variable	Total, n (%)	BC subtype (%)				p value
		HR$^+$/HER2$^-$	HER2$^+$	HR$^-$/HER2$^-$	Undet.	
Baseline		n = 27	n = 11	n = 15	n = 1	
CTCs/7.5 mL						0.861
< 5 CTCs	23 (100.0)	12 (52.2)	4 (17.4)	7 (30.4)	0 (0)	
≥ 5 CTCs	31 (100.0)	15 (48.4)	7 (22.6)	8 (25.8)	1 (3.2)	
CTC clusters/7.5 mL						0.668
No	40 (100.0)	21 (52.5)	7 (17.5)	11 (27.5)	1 (2.5)	
Yes	14 (100.0)	6 (42.9)	4 (28.6)	4 (28.6)	0 (0)	
Follow up		n = 19	n = 9	n = 9	n = 1	
CTCs/7.5 mL						0.583
< 5 CTCs	25 (100.0)	11 (44.0)	7 (18.4)	6 (24.0)	1 (4.0)	
≥ 5 CTCs	13 (100.0)	8 (61.5)	2 (15.4)	3 (23.1)	0 (0)	
CTC clusters/7.5 mL						0.184
No	31 (100.0)	15 (48.4)	9 (29.0)	6 (19.4)	1 (3.2)	
Yes	7 (100.0)	4 (57.1)	0 (0)	3 (42.9)	0 (0)	

Abbreviations: CTCs Circulating Tumour Cells; HR Hormone Receptors; HER2 Human Epidermal Growth Factor Receptor 2; Under. Undetermined

1.1.3. CTC and CTC clusters as predictor factors of patient outcome

Regarding the prognostic value of CTCs and CTC clusters in relation to PFS and OS, it was observed that patients with ≥ 5 CTCs/7.5 mL of blood at baseline (31/54) had a tendency to show a lower PFS ($P_{log-rank}$ = 0.059), and a significantly shorter OS ($P_{log-rank}$ = 0.017)

(figure 16, a, b). Moreover, Cox regression analysis showed that there was a higher risk of death in those patients with ≥ 5 CTCs/7.5 mL (HR$_{OS}$ = 3.15; 95% CI: 1.16-5.55; p = 0.024) but not an increased risk of progression (HR$_{PFS}$ = 2.11; 95% CI: 0.95-4.69; p = 0.065), when compared with patients with < 5 CTCs/7.5 mL of blood (table 10). Nevertheless, both observations became statistically significant when the analysis was adjusted for other clinicopathological variables (table 11).

On the other hand, the CTC cluster-based analysis showed that those patients with ≥ 1 CTC cluster/7.5 mL of blood had a significantly higher risk of progression (HR$_{PFS}$ = 3.95; 95% CI: 1.80-8.68; p = 0.0006) and death (HR$_{OS}$ = 4.23; 95% CI: 1.8-10.1; p = 0.0009) (figure 16, c, d). In fact, those patients showed a shorter OS and PFS (OS and PFS P$_{log-rank}$ < 0.001). Besides, the significance of these observations was maintained when the analysis was adjusted for other clinicopathological parameters (table 11).

During follow up, those patients with ≥ 5 CTCs/7.5 mL blood (13/38) did not show an increased risk of progression (HR$_{PFS}$ = 2.3; 95% CI: 0.76-6.7; p = 0.15), neither a shorter PFS (P$_{log-rank}$ > 0.05). However, they showed a higher risk of death (HR$_{OS}$ = 5.3; 95% CI: 1.4-21; p = 0.017), and a shorter OS (P$_{log-rank}$ < 0.05), in comparison with patients who had < 5 CTCs/7.5 mL blood (table 10 and figure 16, e, f). These results were not significant when adjusted for other clinicopathological factors (table 11).

The follow up analysis based on CTC cluster enumeration showed no differences in the PFS and OS between patients with or without CTC clusters in their blood samples. However, it was observed a tendency by which patients with ≥ 1 CTC cluster/7.5 mL blood had a higher risk of death (HR$_{OS}$ = 3.0) (table 10 and figure 16, g, h).
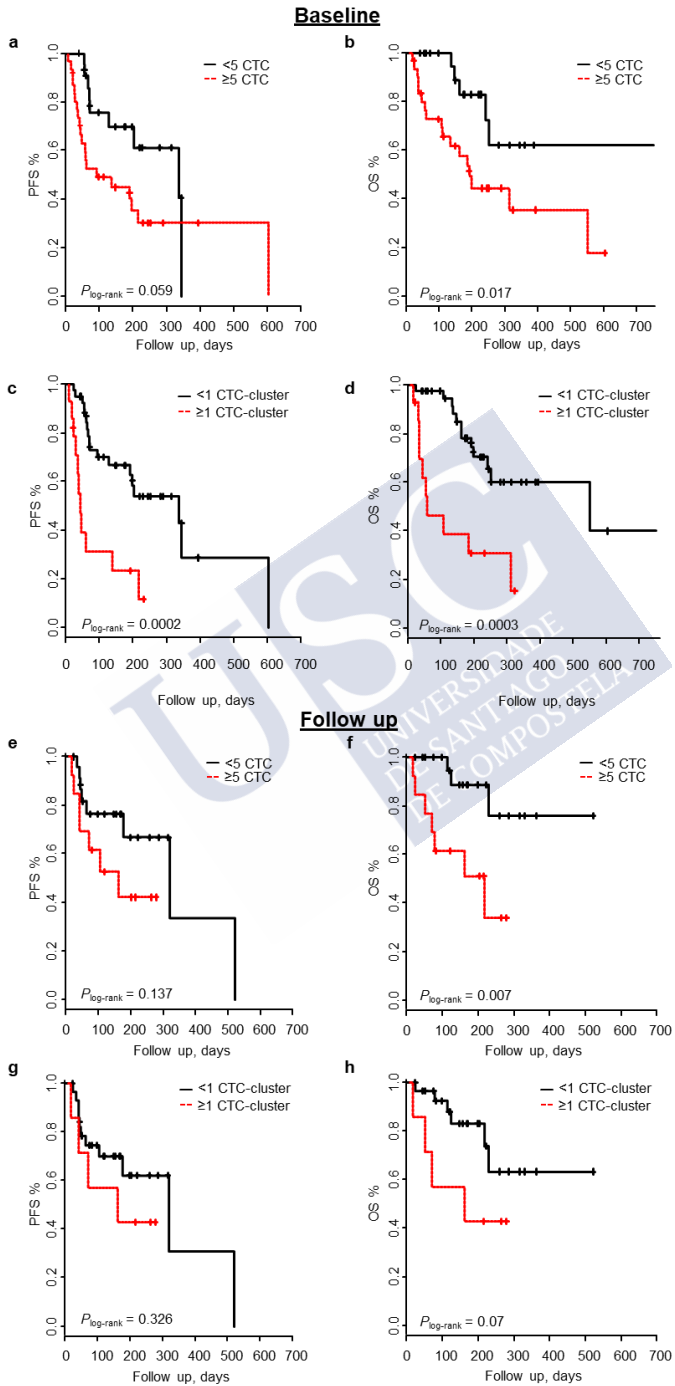
Figure 16. Correlation between CTC and CTC clusters and patient survival at baseline and during follow up. Kaplan-Meier curves show progression-free survival (PFS) and overall survival (OS) based on CTC counts (≥ 5CTCs/7.5mL) at baseline (a, b), and CTC cluster enumeration (≥ 1CTC cluster/7.5mL) (c, d). During follow up, PFS and OS are depicted based on CTC counts (e, f), and CTC cluster enumeration (g, h).

Table 10. Association between CTC and CTC cluster detection and PFS and OS of mBC patients.

Variable	Total	Events, n (%)	HR (95% CI)	p value	HR (95% CI) *	p value *
Baseline						
Associated with PFS						
CTCs/7.5 mL						
< 5 CTCs	23	11 (47.8)	1.00	0.0649	1.00	0.022
≥ 5 CTCs	31	20 (64.5)	2.11 (0.95-4.70)		2.82 (1.15-6.87)	
CTC clusters/7.5 mL						
No	40	18 (45.0)	1.00	0.0006	1.00	0.005
Yes	14	11 (78.6)	3.95 (1.80-8.68)		4.46 (1.55-12.8)	
Associated with OS						
CTCs/7.5 mL						
< 5 CTCs	23	6 (26.1)	1.00	0.024	1.00	0.027
≥ 5 CTCs	31	17 (54.8)	3.15 (1.16-8.55)		3.33 (1.14-9.73)	
CTC clusters/7.5 mL						
No	40	13 (32.5)	1.00	0.0009	1.00	0.004
Yes	14	10 (71.4)	4.23 (1.8-10.1)		6.55 (1.78-23.8)	
Follow up						
Associated with PFS						

CTCs/7.5 mL

< 5 CTCs	23	8 (32.0)	1.00	0.15	1.00	0.29
≥ 5 CTCs	15	7 (53.8)	2.3 (0.76-6.7)		2.39 (0.47-12.04)	

CTC clusters/7.5 mL

No	31	11 (35.5)	1.00	0.33	1.00	0.023
Yes	7	4 (57.1)	1.8 (0.55-5.9)		5.05 (1.24-20.52)	

<u>Associated with OS</u>

CTCs/7.5 mL

< 5 CTCs	25	3 (12.0)	1.00	0.017	1.00	0.996
≥ 5 CTCs	13	7 (53.8)	5.3 (1.4-21)		9.6×10^{17} (0.0-Inf)	

CTC clusters/7.5 mL

No	31	6 (19.4)	1.00	0.09	1.00	0.006
Yes	7	4 (57.1)	3 (0.83-11)		29.74 (2.55-345.8)	

Abbreviations: CTC Circulating Tumour Cell; PFS Progression-Free Survival; OS Overall Survival; HR Hazard Ratio; CI Confidence Interval; Inf. Infinity

* Adjusted for age, ECOG, subtype, number of metastatic sites, location of metastases, and treatments.

Table 11. Multivariate Cox regression analysis of prognostic variables included in the clinicopathological model.

Variable	PFS		OS	
	HR (95% CI)	p value	HR (95% CI)	p value *
Age at baseline				
< 65	1.0	0.15	1.0	0.27
≥ 65	0.54 (0.24-1.2)		0.59 (0.23-1.5)	
ECOG				
Ordinal scale (0.1.2)	0.73 (0.4-1.3)	0.36	0.82 (0.4-1.6)	0.56
Number of metastatic sites				
< 3	1.0	0.18	1.0	0.12
≥ 3	1.7 (0.8-3.5)		2.0 (0.58-4.7)	
Site of metastasis				
Non-visceral	1.0	0.067	1.0	0.026
Visceral	2.5 (0.94-6.6)		5.4 (1.2-23)	
BC subtype				
HR$^+$/HER2$^-$	1.0		1.0	
HER2$^+$	1.38 (0.48-3.9)	0.547	0.85 (0.25-3.5)	0.94
HR$^-$/HER2$^-$	1.63 (0.7-3.8)	0.25	1.9 (0.7-4.7)	0.16
Treatment				
Chemotherapy	1.0		1.0	
Hormone	0.13 (0.03-0.57)	0.006	0.10 (0.01-0.79)	0.028
Targeted	1.41 (0.48-4.16)	0.53	1.14 (0.3-3.9)	0.829

Abbreviations: HR Hazard Ratio, and Hormone Receptor; PFS Progression-Free Survival; OS Overall Survival; CI Confidence interval; ECOG Eastern

Cooperative Oncology Group; HER2 Human Epidermal Growth Factor
Receptor 2
*Adjusted for age, ECOG, BC subtype, number of metastatic lesions, site
of metastasis, and treatment.

In addition, it was assessed the combined effect of CTC and CTC cluster enumeration on the ability to predict patient outcome in our cohort. For this purpose, it was established three different risk groups: < 5 CTCs with no CTC clusters (low-risk), ≥ 5 CTCs without any CTC cluster (medium-risk), and ≥ 5 CTCs with ≥ 1 CTC cluster (high-risk).

At baseline, high-risk patients showed a shorter PFS and OS (PFS and OS $P_{log-rank} < 0.01$), as well as an increased risk of progression and death ($HR_{PFS} = 4.3$; 95% CI: 1.76-10.6; $p = 0.0013$; and $HR_{OS} = 5.8$; 95% CI: 1.96-17.1; $p = 0.0014$), when compared with low-risk patients. However, the medium-risk group did not show a higher risk of progression ($HR_{PFS} = 1.24$; 95% CI: 0.47-3.24; $p = 0.66$), neither of mortality ($HR_{OS} = 1.8$; 95% CI: 0.59-5.98; $p = 0.28$), in comparison with the low-risk group of patients. Besides, CTC cluster enumeration maintained its prognostic value in the multivariate analysis, after adjusting the analysis for different clinicopathological variables. This included the site of metastasis, which was the only significant prognostic marker for this cohort of patients (table 11).

Medium-risk and high-risk patients had a significantly shorter OS during follow up, compared with low-risk patients ($HR_{OS} = 5.71$; 95% CI: 1.11-29.42, $p = 0.037$, and $HR_{OS} = 5.0$; 95% CI: 1.11-22.40; $p = 0.035$, respectively) (table 12 and figure 17). Nevertheless, the significance previously observed was not maintained in the multivariate analysis in any of the above groups.
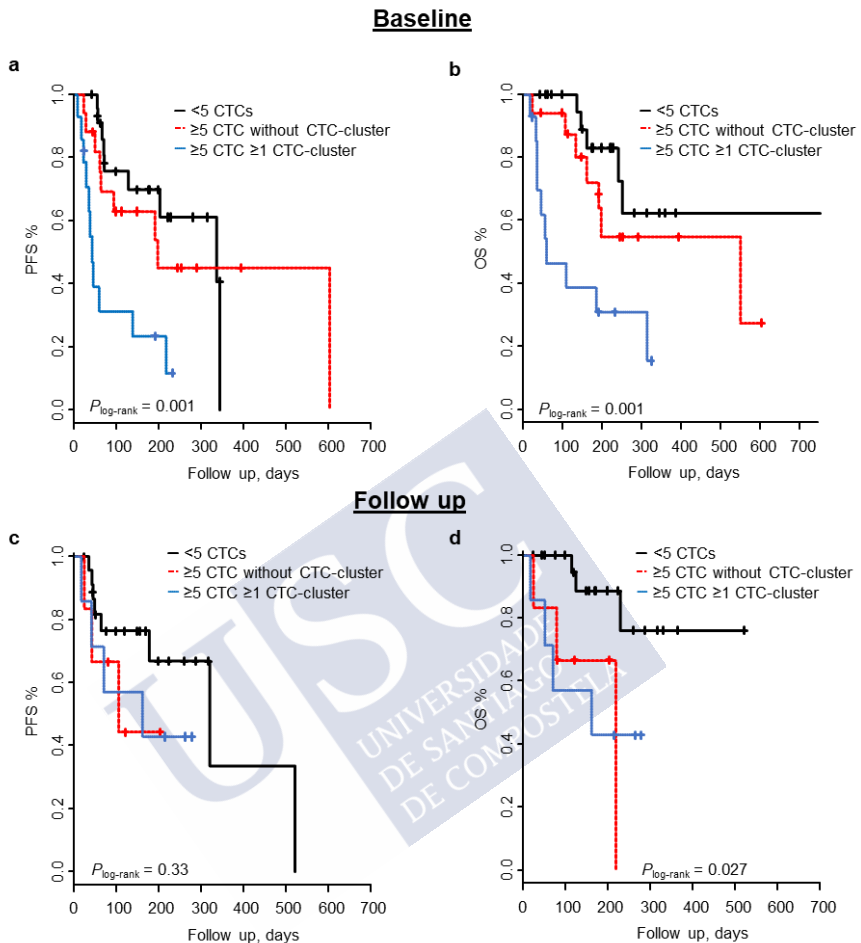
Figure 17. Progression-free survival (PFS) and overall survival (OS) in three different risk groups established based on CTC and CTC cluster enumeration. Patients with < 5 CTCs/7.5 mL were considered as low-risk, patients with ≥ 5CTCs/7.5 mL but no CTC clusters were classified as medium-risk, and patients with ≥ 5CTCs/7.5 mL and at least 1 CTC cluster corresponded to the high-risk group. Kaplan-Meier curves show PFS and OS of these groups at baseline (a, b), and during follow up (c, d).

Table 12. Study of the joint effect of CTCs and CTC clusters with patient survival (PFS and OS).

Variable	Total	Events, n (%)	HR (95% CI)	p value	HR (95% CI)*	p value*
Joint effect of CTC and CTC clusters: Baseline						
Associated with PFS						
< 5 CTCs and NO CTC clusters	23	9 (39.13)	1.00		1.00	
≥ 5 CTCs and NO CTC clusters	17	9 (52.84)	1.24 (0.47-3.24)	0.66	1.74 (0.57-5.30)	0.32
≥ 5 CTCs, ≥ 1 CTC cluster	14	11 (78.57)	4.34 (1.76-10.6)	0.0013	5.16 (1.68-15.8)	0.0041
Associated with OS						
< 5 CTCs and NO CTC clusters	23	6 (26.1)	1.00		1.00	
≥ 5 CTCs and NO CTC clusters	17	7 (41.18)	1.88 (0.59-5.98)	0.28	1.84 (0.50-6.82)	0.36
≥ 5 CTCs, ≥1 CTC cluster	14	10 (71.43)	5.79 (1.96-17.1)	0.0014	7.79 (1.93-31.4)	0.0038
Joint effect of CTC and CTC clusters: Follow up						
Associated with PFS						
< 5 CTCs and NO CTC clusters	25	8 (32.0)	1.00		1.00	
≥ 5 CTCs and NO	6	3 (50.0)	2.3 (0.57-9.27)	0.24	0.61 (0.08-4.53)	0.635

CTC clusters						
≥ 5 CTCs, ≥ 1 CTC cluster	7	4 (57.1)	2.2 (0.62-7.84)	0.22	4.45 (1.01-19.59)	0.0484
Associated with OS						
< 5 CTCs and NO CTC clusters	25	3 (12.0)	1.00		1.00	
≥ 5 CTCs and NO CTC clusters	6	3 (50.0)	5.71 (1.11-29.4)	0.037	1.2×10^{18} (0.0-Inf.)	0.996
≥ 5 CTCs, ≥ 1 CTC cluster	7	4 (57.1)	5.00 (1.11-22.4)	0.0353	2.4×10^{18} (0.0-Inf.)	0.996

Abbreviations: CTC Circulating Tumour Cell; PFS Progression-Free Survival; OS Overall Survival; HR Hazard Ratio; CI Confidence Interval; Inf. Infinity.
*Adjusted for age, ECOG, BC subtype, number of metastatic sites, metastases location, and treatment.

Furthermore, mBC liquid biopsy samples were used to study the clinical prognostic value of CTC cluster enumeration over time. To this end, the continued presence of CTC clusters throughout time was studied, regarding survival variables, PFS and OS. In our cohort and among the patients with ≥1CTC (n = 43, only 2 patients (4.4%) had detectable CTC clusters at two different timepoints, 16 (35.6%) had CTC clusters in one timepoint, and 27 (60.0%) only had individual CTCs (figure 18, a). In terms of PFS, it was observed a shorter PFS time in patients with presence of CTC clusters in their blood, in comparison with those who only had individual CTCs ($P_{log-rank} = 0.017$) (figure 18, b). Nevertheless, it was only observed a significantly increased risk of progression in those patients with CTC clusters at one timepoint (HR_{PFS} = 2.8; 95% CI: 1.22-6.34; p = 0.014). Moreover, those patients with detection of CTC clusters in one or two different timepoints had a significantly lower OS ($P_{log-rank} = 0.007$) (figure 18, c), and an increased risk of death (HR_{OS} = 3.3; 95% CI: 1.29-8.6; p = 0.012 for one timepoint, and HR_{OS} = 6.5; 95% CI: 1.31-32.7; p = 0.021).
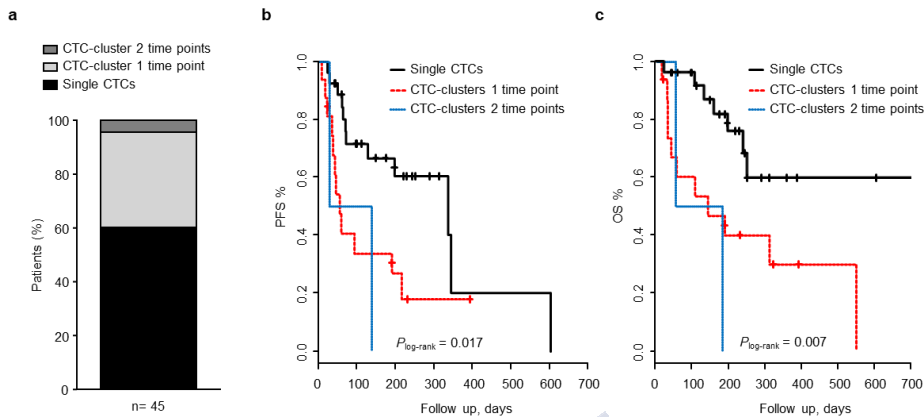
Figure 18. Longitudinal analysis of CTCs and CTC clusters. Percentage of patients with individual CTCs only (dark bar), with CTC clusters at 1 timepoint (light grey) and with CTC clusters at 2 timepoints (dark grey) (a). Kaplan-Meier curves of progression-free survival (PFS) and overall survival (OS), according to the presence of single CTCs only, CTC clusters at 1 timepoint or CTC clusters at 2 timepoints (b, c).

Therefore, these results denoted that CTC cluster enumeration at baseline is an independent prognostic marker, that provides additional value to CTC enumeration alone in this cohort of mBC patients in those patients with high CTC detection. Besides, the maintained presence of CTC clusters in the peripheral blood of mBC is associated with a worse disease outcome, as they are associated with a shorter survival.

This work was extracted/adapted from our previous published work:

Analysis of a Real-World Cohort of Metastatic Breast Cancer Patients Shows Circulating Tumor Cell Clusters (CTC-clusters) as Predictors of Patient Outcomes.

Costa, C.; Muinelo-Romay, L.; Cebey-López, V.; Pereira-Veiga, T.; Martínez-Pena, I.; Abreu, M.; Abalo, A.; Lago-Lestón, R.M.; Abuín, C.; Palacios, P.; Cueva, J.; Piñeiro, R.; López-López, R.

Cancers 2020, 12(5), 1111. ISSN 2072-6694

Author affiliation:

Clotilde Costa Nogueira[1,2,‡] , Laura Muinelo Romay[2,3,‡], Víctor Cebey López[4], Thais Pereira Veiga[1], Inés Martínez Pena[1], Manuel Abreu[3], Alicia Abalo[3], Ramón M. Lago Lestón[3], Carmen Abuín[1], Patricia Palacios[4], Juan Cueva[4], Roberto Piñeiro[1,*], and Rafael López López[1,2,3,4,*]

[1]Roche-Chus Joint Unit, Translational Medical Oncology Group, Oncomet, Health Research Institute of Santiago de Compostela (IDIS), Travesía da Choupana s/n, 15706 Santiago de Compostela, Spain

[2]CIBERONC, Centro de Investigación Biomédica en Red Cáncer, 28029 Madrid, Spain

[3]Liquid Biopsy Analysis Unit, Translational Medical Oncology Group, Health Research Institute of Santiago de Santiago de Compostela (IDIS), Travesía da Choupana s/n, 15706 Santiago de Compostela, Spain

[4]Department of Oncology, Complexo Hospitalario Universitario de Santiago de Compostela (SERGAS), 15706 Santiago de Compostela, Spain

[‡]These authors contributed equally

[*]These authors jointly supervised this work

Contribution to this work: I, Inés Martínez Pena, was involved in data analysis and reviewing of the manuscript.

1.2. Optimization of Parsortix™ immuno-independent detection of CTCs and CTC clusters in BC liquid biopsy samples

1.2.1. Metastatic BC peripheral blood samples

As previously demonstrated, the FDA-approved CellSearch® system has highlighted the importance of CTC and CTC cluster enumeration for predicting mBC patient outcome. However, this system presents some limitations as it does not allow the recovery of CTCs and CTC clusters in a viable state, thus limiting the analysis that could be performed with the enriched CTCs and CTC clusters, and it does not allow the optimization of protocols for the isolation of CTC clusters. Given the limitations of this immune-dependent CTC isolation system, peripheral blood samples from mBC patients were also used to optimise a workflow for CTC and especially CTC cluster isolation in an immune-independent manner, using the Parsortix™ system. Parsortix™ is a microfluidic device that isolates CTCs based on their larger size and lower deformability, in comparison with blood cells. Parsortix™ allows the isolation of CTC clusters, as it was previously demonstrated [129] (figure 19).
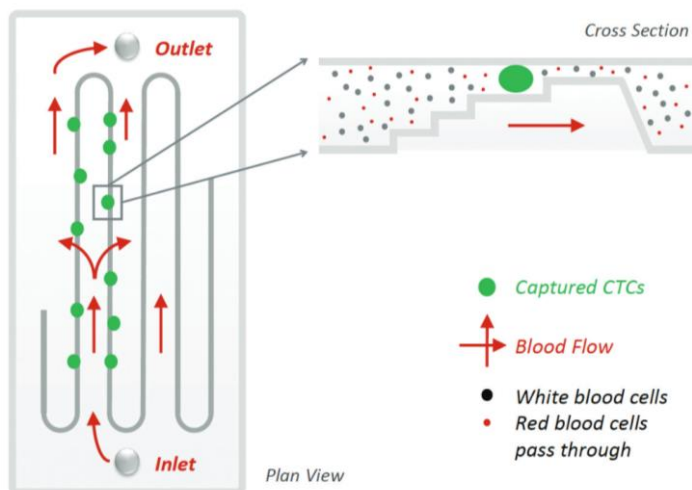
Figure 19. Overview of a Parsortix™ cassette. The sample flows through serpentine channels. These channels have a stair-like silhouette that finishes in a 6.5 µm gap in which CTCs are capture, while the rest of blood cells flow through. Red arrows indicate the direction of the flow inside the cassette. Image used with permission of Angle plc.

Firstly, we tried to isolate, and detect CTCs and especially CTC clusters from mBC patient liquid biopsy samples (n = 3) (table 13). For this purpose, a Parsortix™ separation protocol (named 'cluster separation protocol') with low pressure and reduced flow was used to minimise CTC cluster disaggregation during CTC capture. Secondly, a previously optimised live cell immunostaining was used to identify the captured CTCs and CTC clusters in the Parsortix™. This in-cassette immunostaining included a conjugated cocktail of antibodies against the epithelial markers Epithelial Cell Adhesion Molecule (EpCAM), Epidermal Growth Factor Receptor (EGFR), and E-cadherin (E-cad). Besides, an antibody against the leukocyte marker CD45, as well as the double-stranded DNA nuclear dye NucBlue™ Live ReadyProbes™ Reagent was used. Thus, CTCs and CTC clusters were defined as isolated cells with the following phenotype: EpCAM/EGFR/E-cad$^+$, NucBlue$^+$ and CD45$^-$. Moreover, a CTC cluster was defined by the presence of a group made up of ≥ 2 tumour cells. Hence, this workflow allowed the identification of both, putative CTCs and CTC clusters as a proof-of-concept experiment in a small number of patients (table 13 and figure 20).

Table 13. mBC peripheral blood samples used for CTC and CTC cluster enumeration in the Parsortix™ system.

	BC subtype	Parsortix	
		CTCs	CTC clusters
UM176MV1	HER2$^+$	57	60
UM196MV1	HR$^+$/HER2$^-$	500	179
UM88MV3	HR$^+$/HER2$^-$	9	7

Abbreviations: BC Breast Cancer; CTC Circulating Tumour Cell; HER2 Human Epidermal growth factor Receptor 2; HR Hormone Receptor
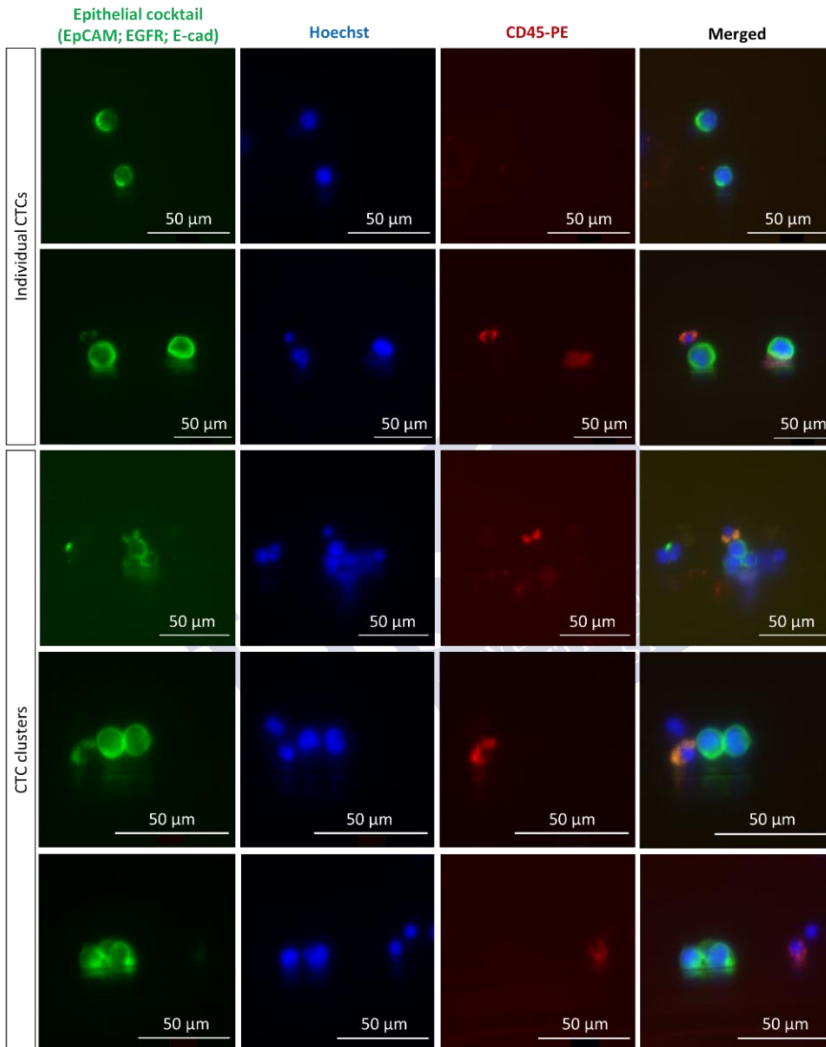
Figure 20. Representative images of single CTCs and CTC clusters isolated from the peripheral blood samples of mBC patients. Scale bars: 50 μm.

Thus, it was possible to establish a workflow that allows CTC cluster isolation from peripheral blood samples of mBC patients. Unlike

the immune-based CTC capture, CTC isolation based on their differential physical properties allowed the recovery of CTCs in a viable state as they are positive for NucBlue Live Cell Stain, potentially extending the amount of downstream analysis. It also may increase the likelihood of detecting a more heterogeneous CTC population, although for this, antibodies against other epitopes should be used.

1.2.2. Metastatic BC Diagnostic Leukapheresis (DLA) products

Despite the utility of peripheral blood samples for CTC and CTC cluster isolation, there is still a limitation regarding the low amount of biological material obtained for downstream analysis. This slows down the advancement of our knowledge in the biology of CTCs and CTC clusters. Therefore, alternative liquid biopsy samples, such as Diagnostic Leukapheresis (DLA) products, are being used to maximise the capture of CTCs. DLA is a standard density-based methodology used in the clinic for the isolation of mononuclear cells (MNCs) by continuous centrifugation of large volumes of blood (figure 21). It was recently proposed that CTCs could be co-isolated together with the MNCs during this centrifugation, as they have a similar density to MNCs. Hence, in comparison with peripheral blood samples, DLA products would increase the likelihood of CTC detection, as they come from the processing of large volumes of blood.
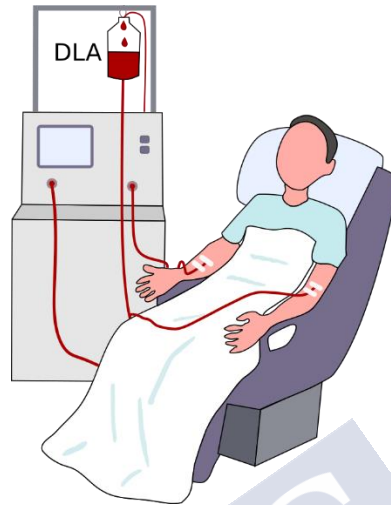
Figure 21. Graphical representation of a patient during DLA.

In collaboration with the group of Translational Research in Gynaecology of the Heinrich-Heine University (Düsseldorf, Germany), headed by Prof. Dr. rer. nat. Hans Neubauer, we worked with DLA products from mBC patients. DLA samples have a cellularity of $5x10^7$ MNCs/mL, while peripheral blood samples show a cellularity of 0.5-$3x10^6$ cells/mL of blood. Due to the high cellularity of DLA samples, a suitable optimization of the workflow is essential for the successful detection of CTCs and CTC clusters in DLA products. DLA products were used to optimise a workflow for DLA sample processing, and for testing the feasibility of CTC cluster isolation as there was no previous data on the presence of CTC clusters in these samples. We used mBC DLA samples which resulted negative for CTCs in the CellSearch® system, and that we called DLA$_{CTC-}$. These samples were initially utilised to select the most suitable CTC isolation protocol for DLA separation in the Parsortix™ system. To this end, each DLA$_{CTC-}$ sample was divided into two halves, and a known number of MDA-MB-231$^{eGFP\ Luc}$ or MCF-7 $^{eGFP\ Luc}$ cells, either as individual cells (to mimic single CTCs) or as cellular aggregates (CTC clusters), were spiked into each of them. Two different separation protocols were compared: a standard protocol (pressure: 99 mbar), and a separation protocol specifically designed to preserve CTC cluster physical integrity

('cluster separation protocol', previously mentioned. Pressure: 50 mbar). Afterwards, tumour cells were captured in the Parsortix^{TM} system, and individual CTCs (figure 22) and CTC cluster (figure 23) were counted after separation, and after the harvesting steps. Moreover, it was determined a prediction of the harvesting recovery efficiency.
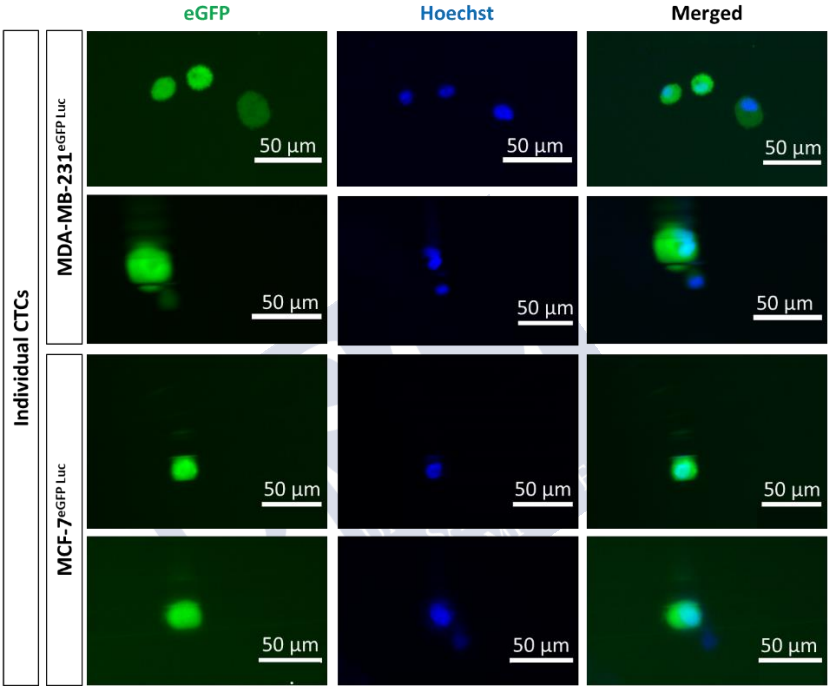

Figure 22. Representative images of MDA-MB-231$^{eGFP Luc}$ and MCF-7$^{eGFP Luc}$ individual cells isolated from DLA spiking samples. Scale bars: 50 μm.
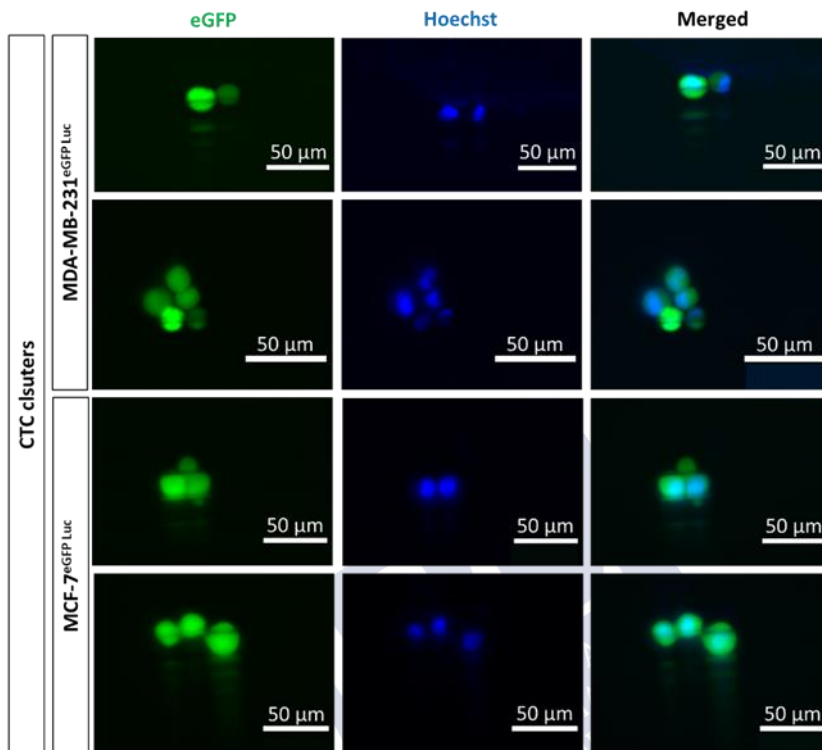
Figure 23. Representative images of MDA-MB-231[eGFP Luc] and MCF-7[eGFP Luc] CTC clusters isolated from DLA spiking samples. Scale bars: 50 μm.

MDA-MB-231[eGFP Luc] spiking showed an individual CTC isolation efficiency of 69.44% (75/108) with standard protocol, and 89.58% (86/96) for 'cluster protocol'. After harvesting, 32.41% (35/108) of individual cells remained inside the Parsortix[TM] cassette with the standard protocol, while 65.63% (63/96) of individual cells remained in the cassette after using the 'cluster protocol'. Hence, the predicted recovery for individual CTC detection was 37.04% (40/108), and 23.96% (23/96), respectively (table 14 and figure 24). Regarding CTC cluster isolation, the standard protocol captured 30.77% (24/78) of the initially spiked clusters, while 61.90% (52/84) of CTC clusters were captured with the 'cluster separation protocol'. However, the recovery efficiency was better in the case of the standard protocol, as only 6.41% (5/78) of clusters remained in the cassette after harvesting (predicted

recovery of 24.36%, or 19/78), while 42.86% (36/84) of CTC clusters stayed inside the cassette after the harvesting step with the 'cluster protocol' (predicted recovery of 19.05%, that is 16/84) (figure 24 and table 14).
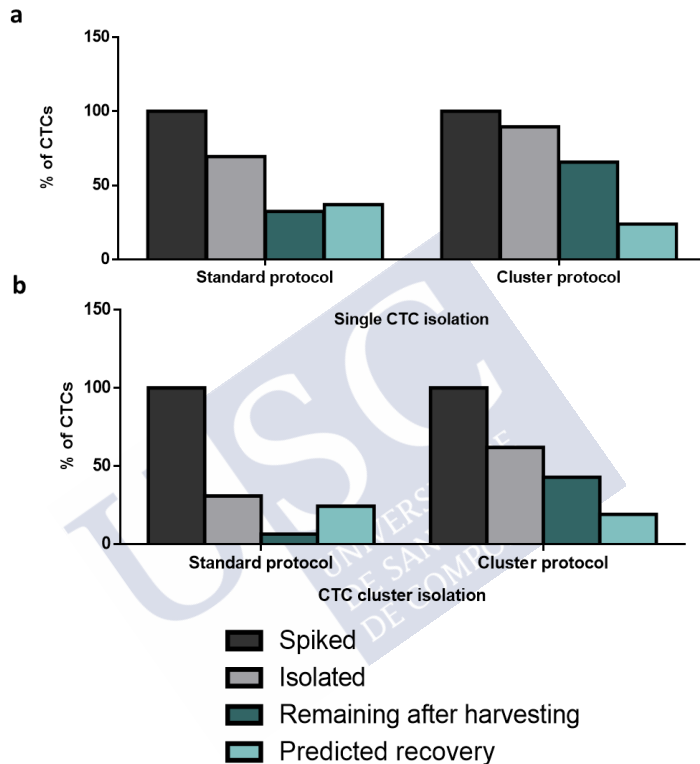

Figure 24. Comparison between standard and cluster separation protocols for MDA-MB-231[eGFP Luc] single CTC (a) and CTC cluster (b) detection in DLA samples.

Table 14. Total number of cells and isolation and recovery efficiencies of MDA-MB-231[eGFP Luc] single CTC and CTC cluster detection with standard and cluster separation protocols. Percentages are expressed relative to the initial number of spiked cells.

	MDA-MB-231[eGFP Luc]			
	Standard Protocol		Cluster protocol	
	Single CTCs	CTC cluster	Single CTCs	CTC cluster
Spiked	108 (100%)	78 (100%)	96 (100%)	84 (100%)
Isolated	75 (69.44%)	24 (30.77%)	86 (89.58%)	52 (61.90%)
Remaining after harvesting	35 (32.41%)	5 (6.41%)	63 (65.63%)	36 (42.86%)
Predicted recovery	40 (37.04%)	19 (24.36%)	23 (23.96%)	16 (19.05%)

DLA$_{CTC-}$ samples spiked with MCF-7[eGFP Luc] cell line showed that 91.85% (124/135) of spiked individual cells were captured with the standard protocol, while only 39.67% (48/121) of the cells were isolated using the 'cluster protocol'. Nevertheless, there were notable differences in terms of recovery efficiencies between both protocols. Standard protocol showed that 71.85% (97/135) of the cells were trapped in the cassette after the recovery (predicted recovery of 20.00%, that is 27/135). With 'cluster protocol', 11.57% (14/121) of cells were present inside the cassette after the recovery step (predicted recovery of 28.10%, or 34/121) (table 15 and figure 25). In the case of epithelial-like CTC cluster isolation, 77.03% (57/74) of CTC clusters were captured with the standard protocol, while only 23.26% (20/86) of the initially spiked CTC clusters were isolated with the 'cluster protocol'. In terms of recovery efficiency, standard protocol seemed to be more efficient than 'cluster protocol'. With standard separation protocol 25.68% (19/74) of CTC clusters remained in the cassette after harvesting, having a predicted recovery of 51.35% (38/74). On the contrary, 8.14% (7/86) of CTC clusters stayed in the cassette after

recovery with the 'cluster protocol' and showed a predicted recovery of 15.12% (13/86) relative to the initial number of spiked CTC clusters (figure 25 and table 15).
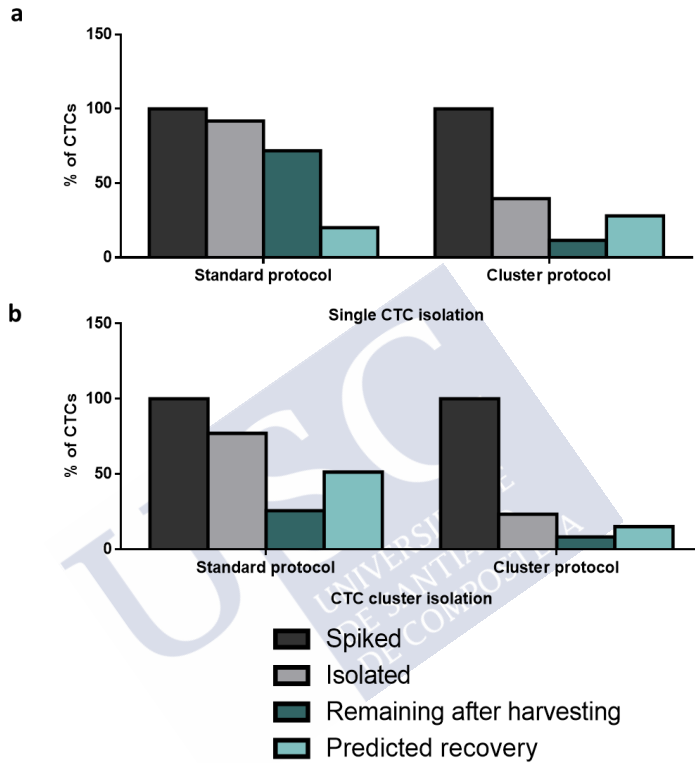

Figure 25. Comparison between standard and cluster separation protocols for MCF-7$^{eGFP\ Luc}$ single CTC (a) and CTC cluster (b) detection in DLA samples.

Table 15. Total number of cells and isolation and recovery efficiencies of MCF-7$^{eGFP\ Luc}$ single CTC and CTC cluster detection with standard and cluster separation protocols. Percentages are expressed relative to the initial number of spiked cells.

| | MCF-7$^{eGFP\ Luc}$ | | | |
| | Standard Protocol | | Cluster protocol | |
	Single CTCs	CTC cluster	Single CTCs	CTC cluster
Spiked	135 (100%)	74 (100%)	121 (100%)	86 (100%)
Isolated	124 (91.85%)	57 (77.03%)	48 (39.67%)	20 (23.26%)
Remaining after harvesting	97 (71.85%)	19 (25.68%)	14 (11.57%)	7 (8.14%)
Predicted recovery	27 (20.00%)	38 (51.35%)	34 (28.10%)	13 (15.12%)

Taken into consideration these preliminary results, it was concluded that the 'cluster separation protocol' was less suitable for DLA samples, as it generated both, lower CTC isolation and lower CTC recovery than the standard protocol. Besides, the high cellularity of DLA samples increased the risk of Parsortix^{TM} blocking during CTC separation due to the formation of cell aggregates. This together with the reduced flow rate of the 'cluster separation protocol' allowed the blockage of the system more easily and limited the efficiency of the process. Hence, the standard protocol was chosen as the most suitable for CTC and CTC cluster isolation from DLA samples.

Likewise, given the high cellularity of DLA products, sample filtration using a 100μm cell strainer before CTC isolation in the Parsortix^{TM} system is an essential step. This avoids the presence of large cell aggregates that otherwise would block the Parsortix^{TM} system easily, leading to a decrease in the efficiency of the process. However, sample filtration could also potentially increase the risk of CTC cluster

disaggregation, thus partially limiting the efficiency of CTC cluster isolation. Hence, we evaluated the potential effect of filtration in CTC and CTC cluster detection. For this, each DLA$_{CTC-}$ sample was split into two halves. A known number of MDA-MB-231$^{eGFP\,Luc}$ cells were added to each half as both, either as individual cells and as small cellular groups. One of the halves was filtered with a 100μm cell strainer before performing the spiking, and the other one was filtered after the spiking, thus exposing tumour cell clusters to potential disaggregation by the mechanical effect of filtration. After separation, the isolation efficiency of individual CTCs from samples filtered before spiking was 53.46% (154.5/289), while for samples filtered after the spiking the isolation was 29.30% (143/488). After the harvesting step, 14.88% (43/289) of the individual cells remained inside the cassette (predicted recovery of 38.58%, that is 111.5/289 cells) in the 'filtration before spiking' sample. Similarly, 10.14% (49.5/488) of the isolated individual CTCs remained in the cassette after harvesting (predicted recovery of 19.19%, or 93.5/488 cells) in the 'filtration after spiking' samples (figure 26 and table 16).

Regarding CTC cluster detection, there was no significant difference in the isolation efficiency between both conditions, tumour cell exposure and non-exposure to filtration. The percentage of CTC cluster isolation relative to the initial number of spiked cells in the 'filtration before spiking' samples was 23.48% (35/149), and in the 'filtration after spiking' samples was 19.01% (34.5/181.5). After the recovery step, only 6.04% (9/149) of CTC clusters remained in the cassette in the 'filtration before spiking' sample, and its predicted recovery was 17.44% (26/149). Likewise, the percentage of CTC clusters that remained inside the cassette after harvesting in the 'filtration after spiking' was 5.79% (10.5/181.5), and its predicted recovery was 13.22% (24/181.5) (figure 26 and table 16).
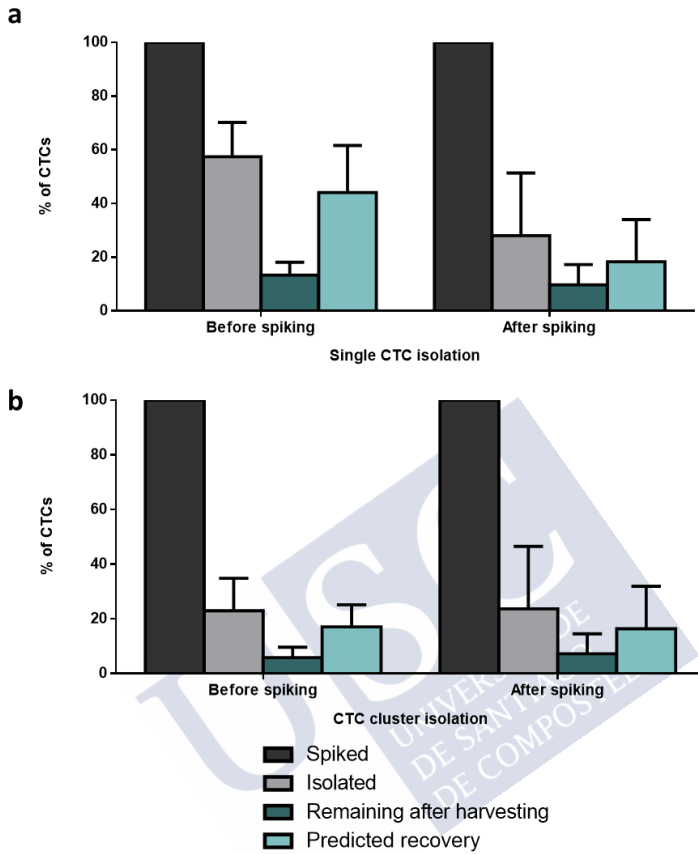
Figure 26. Bar graphs representing the isolation and recovery efficiencies of CTCs (a) and CTC clusters (b) spiked in DLA samples non-exposed ('before spiking') and exposed ('after spiking') to filtration previous Parsortix™ separation.

Table 16. Potential effect of DLA pre-filtration over MDA-MB-231^{eGFP Luc} CTC and CTC cluster isolation and recovery efficiencies. In the 'before spiking' sample tumour cells were not exposed to filtration, while in the 'after spiking' filtration was applied after the spiking of tumour cells.

	MDA-MB-231$^{eGFP\ Luc}$ & DLA filtration (n=2)			
	Before spiking		**After spiking**	
	Single CTCs	**CTC cluster**	**Single CTCs**	**CTC cluster**
Spiked	289 (100%)	149 (100%)	488 (100%)	181.5 (100%)
Isolated	154.5 (53.46%)	35 (23.48%)	143 (29.30%)	34.5 (19.01%)
Remaining after harvesting	43 (14.88%)	9 (6.04%)	49.5 (10.14%)	10.5 (5.79%)
Predicted recovery	111.5 (38.58%)	26 (17.44%)	93.5 (19.16%)	24 (13.22%)

Therefore, the optimisation of the key steps of the workflow allowed us to demonstrate that it is feasible to detect and isolate CTCs and CTC clusters from DLA samples. Thus, DLA products can be a suitable type of sample for maximising the isolation of CTCs and especially CTC clusters.

2. CTC CLUSTER MODELS HAVE A DIFFERENTIAL METASTATIC BEHAVIOUR THAN INDIVIDUAL CTCS

As previously mentioned, CTC clusters correspond to a minority subpopulation within CTCs, as they only represent about 1-30% of all CTCs. The low frequency of CTC clusters in the peripheral blood of BC patients enormously restricts the advances in the study of the biology of CTC clusters. This generates the need of developing alternative tools to overcome this limitation, such as the generation of

experimental models of CTC clusters. The *in vitro* models of CTC clusters should recapitulate the physiology of those CTC clusters isolated from mBC patients and they could have a great potential to promote a deeper understanding of the biology of CTC clusters.

Two human BC cell lines, MDA-MB-231[eGFP Luc] and MCF-7 [eGFP Luc], were used to obtain *in vitro* models of single CTCs and CTC clusters As previously mentioned, two different protocols were used to generate single and cluster models (see '*Generation of single CTC and CTC cluster in vitro models*', on the Materials and Methods section). Briefly, cells were cultured overnight under low-attachment and serum deprivation conditions to allow cell aggregation. Afterwards, cell aggregates were differentially disaggregated to obtain an individual cell suspension that mimicked individual CTCs, and cell groups that represented CTC clusters (figure 27, a). Alternatively, CTC models were generated by culturing cells under adherent standard conditions and using different concentrations of trypsin-EDTA to differentially detach cells. The generation of single CTC models was performed using a more concentrated trypsin, while the obtaining of CTC cluster models was done with a more diluted trypsin (figure 27, b).
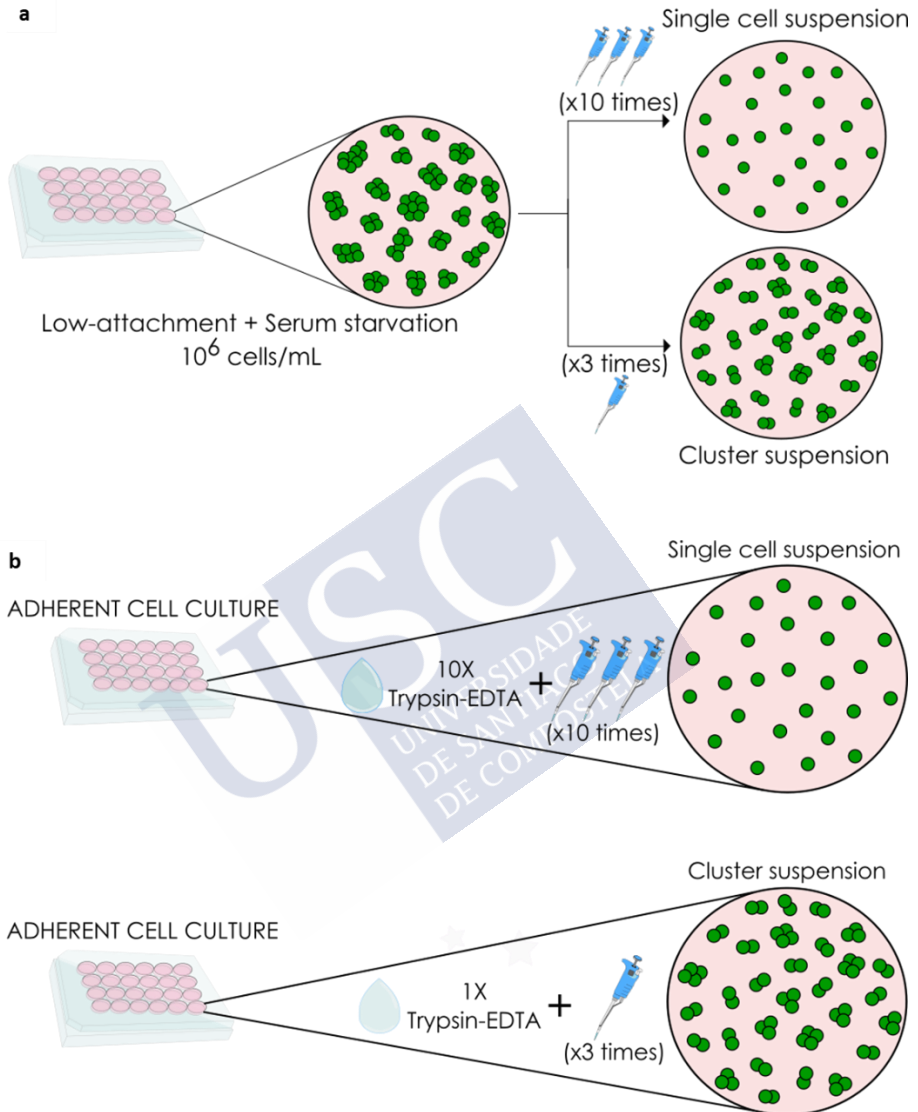
Figure 27. Protocols followed for the generation of *in vitro* single CTC and CTC cluster models. Cells were kept overnight under low-attachment and serum-deprivation conditions at a density of 1x10^6 cells/mL to allow cell aggregation. Differential mechanical disaggregation was used to obtain single cell suspension (single CTCs) or cellular aggregates (CTC clusters) (a). Cells were cultured under standard conditions. For single CTC and CTC cluster generation, different

concentrations of trypsin and differential mechanical disaggregation were used (b).

2.1. Functional characterization shows differences between individual CTC and CTC cluster experimental models

Different *in vitro* and *in vivo* assays were performed to assess functional characteristics of individual CTC and CTC cluster *in vitro* models, as well as their differences regarding the metastatic potential of each model.

Each *in vitro* assay represents a different step within the metastatic cascade and evaluates a phenotypic trait that is typically associated with metastatic potential. Individual CTC and CTC cluster models were generated using two different human BC cell lines, MDA-MB-231$^{\text{eGFP Luc}}$ and MCF-7$^{\text{eGFP Luc}}$, representing a different BC molecular subtype and a different phenotypic profile of CTC. Hence, the MDA-MB-231$^{\text{eGFP Luc}}$ cell line is representative of TNBC, and it is characterised by having a mesenchymal-like, more stem phenotype. On the contrary, the MCF-7$^{\text{eGFP Luc}}$ is a luminal A BC cell line that shows a well differentiated epithelial phenotype. These different *in vitro* models were generated with the purpose of having a wider heterogeneity and mimicking the biological reality as accurately as possible.

2.1.1. CTC clusters show a higher migration and invasion capacity than individual CTCs

Transwell migration assays were performed to evaluate the ability of active migration of the CTC cluster model in comparison with individual CTC model.

The migration experiments performed with MDA-MB-231$^{\text{eGFP Luc}}$ showed that the CTC cluster population had a significantly higher transwell migration ability compared to the single CTC model. We observed that after 8 hours of incubation, the number of migrated cells in the CTC cluster suspension was almost twice as high as in the single CTC population ($p = 0.0008$) (figure 28, a).

On the other hand, MCF-7$^{eGFP\ Luc}$ cell line showed negligible migration ability after 8 hours of incubation (data not shown), as it is expected due to their epithelial phenotypic traits. Hence, it was necessary to increase the incubation period to see possible differences between cell populations in terms of active migration. After 24 hours, the CTC cluster population showed a tendency to have a increase in the number of migrated cells compared to single CTCs (1.62 fold change). However, this observation did not reach statistical significance (p = 0.4331) (figure 28, b).
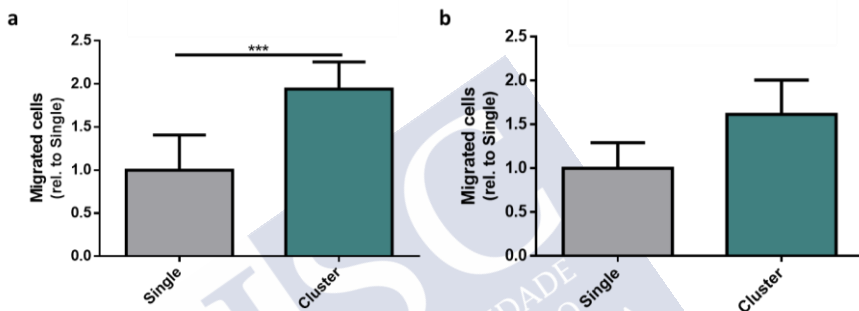

Figure 28. Graph bar showing the Transwell migration of single CTC and CTC cluster models derived from the MDA-MB-231$^{eGFP\ Luc}$ cell line (n = 6) (a), and from the MCF-7$^{eGFP\ Luc}$ cell line (n = 4) (b). Data are expressed as the fold change of migrated cells relative to the average number of migrated cells observed in the single CTC population. *** p < 0.001.

Matrigel-coated transwell invasion assays were used to evaluate the behaviour of the models regarding their ability to perform alterations in an extracellular matrix, that allow them to invade through the ECM and reach the pores of the Transwell membrane.

CTC cluster model of MDA-MB-231$^{eGFP\ Luc}$ cells showed a higher ability to invade when compared to individual CTC population. In fact, after 24 hours of incubation, CTC cluster showed an average of 30.1 ± 21.7 invaded cells per field of view (FOV), while single CTC population only showed 22.4 ± 18.0 invaded cells per FOV (p = 0.0255) (figure 29, a).

Regarding MCF-7$^{eGFP\ Luc}$ CTC models, no significant differences were observed in the number of invaded cells between individual CTCs and CTC clusters. This observation is not unexpected, given the epithelial traits of this cell line. Hence, we did not observe any difference between single and cluster populations in terms of *in vitro* invasion capacity (0.50±0.34 and 0.53±0.26, respectively) (figure 29, b).
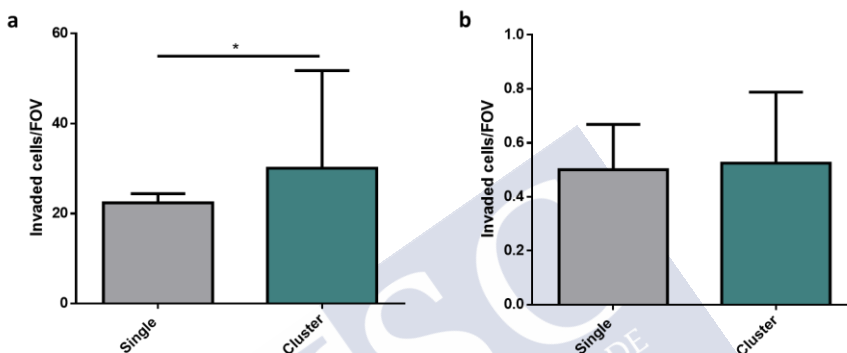

Figure 29. Transwell invasion assay of single CTC and CTC cluster models generated from MDA-MB-231$^{eGFP\ Luc}$ cell line (n = 2) (a), and from the MCF-7$^{eGFP\ Luc}$ cell line (n = 4) (2). Data are expressed as the average of invaded cells per Field Of View (FOV). * p < 0.05.

Therefore, we observed differences between individual CTC and CTC cluster *in vitro* models in terms of migration and invasion capacity. This differential behaviour showed that the MDA-MB-231^{eGFPLuc} CTC cluster model had a higher capacity to migrate and invade *in vitro*, which are abilities required in the initial steps of the metastatic cascade. On the contrary, the epithelial characteristics of the MCF-7^{eGFPLuc} models showed subtle differences in terms of migration but no differences were observed in the invasion capacity.

2.1.2. CTC clusters show a lower ability to adhere to the endothelium but a higher capacity to form colonies

Endothelial adhesion tests were made to determine the ability of tumour cells, as both single CTCs and CTC clusters, to interact with

a monolayer of endothelial cells. This interaction is required when CTCs intravasate and extravasate from the bloodstream, as a previous step of colonizing a new niche. Endothelial adhesion assays were done using the cell line EA.hy926 to generate the monolayer of endothelial cells. After the generation of the endothelial monolayer, tumour cells as either, individual CTCs or CTC clusters, were seeded over the monolayer under serum deprivation conditions. Tumour cells were allowed to interact with the endothelial cell for 45 minutes. Afterwards, non-attached tumour cells were washed away, and adhered cells were counted.

Although there was a wide biological variability, MDA-MB-231[eGFP Luc] CTC cluster population had a significantly lower ability to adhere to an endothelial monolayer *in vitro*, compared to individual CTCs (p = 0.006) (figure 30, a). More specifically, CTC clusters showed an average of 115.1±26.8 adhered cells, while individual CTCs had 145.5±34.7 adhered cells, meaning that CTC clusters showed 20.95% fewer adhered cells than single CTC population.

Similarly, MCF-7[eGFP Luc] CTC clusters showed a significantly lower ability to adhere to the endothelial monolayer than single CTCs (p = 0.0011), despite having a greater ability than MDA-MB-231[eGFP Luc] to adhere to the endothelium *in vitro* due to the epithelial traits of this cell line. CTC cluster population had an average of 580.9±162.4 adhered cells, while individual CTCs showed an average of 873.7±206.4 adhered cells, indicating that CTC clusters had 33.51% fewer adhered cells than the single CTC model (figure 30, b).
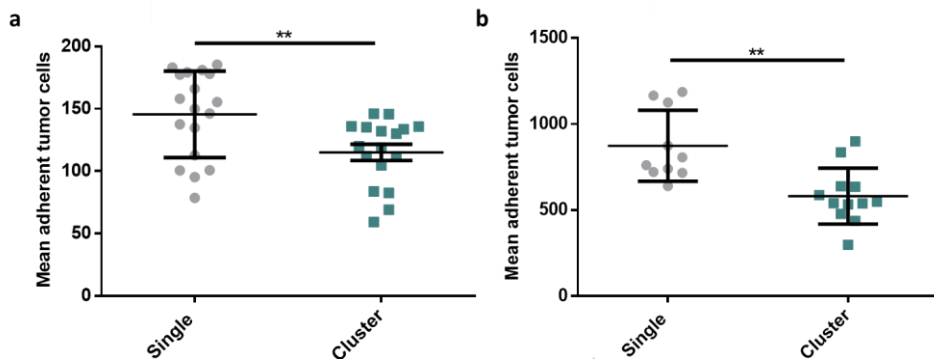
Figure 30. MDA-MB-231[eGFP Luc] **(n = 4) (a) and MCF-7**[eGFP Luc] **(n = 2) (b) individual CTC and CTC cluster adhesion to the endothelium. Data are expressed as the average of adhered tumour cells per well. ** p < 0.01.**

Additionally, the ability of the CTC models to seed tumour lesions was tested. Soft agar colony formation assay is used to assess the anchorage-independent growth capacity of a cell population, as well as their ability of self-maintenance and self-renewal under the absence of a solid ECM. Hence, with this experimental procedure we evaluated the potential of the single CTC and the CTC cluster models to grow in an anchorage-independent manner, which is necessary to generate micrometastases.

Clusters of MDA-MB-231[eGFP Luc] showed a higher capacity to form colonies than the single cell population. On average, CTC clusters showed 9.67 ± 6.8 colonies, while single CTCs had only 1.67 ± 1.4 colonies. Hence, CTC clusters had 5.8 more capacity for colony generation than the single cell population. However, this difference did not reach statistical significance ($p = 0.200$) (figure 31, a).

MCF-7[eGFP Luc] CTC clusters also showed a tendency towards a higher ability to generate colonies, when compared to single CTCs. This tendency showed that on average CTC clusters had 19.67 ± 8.7 colonies, while single CTCs only had 4.17 ± 0.17 colonies, meaning that CTC clusters had a colony formation ability 4.7 times higher than single

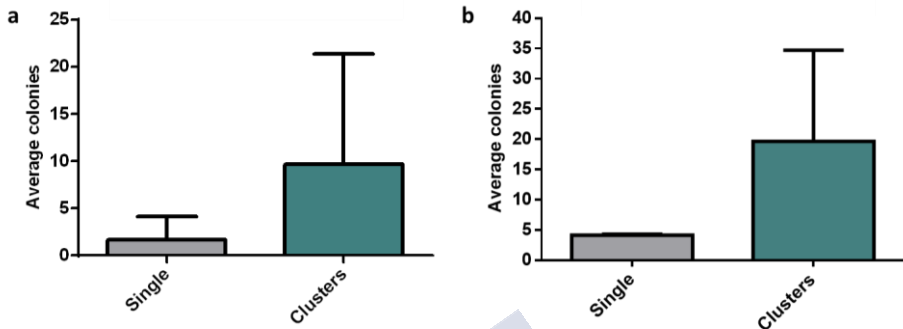CTCs. Nevertheless, the observed tendency did not reach statistical significance (p = 0.100) (figure 31, b).

Figure 31. Soft agar colony formation ability of single CTC and CTC cluster models generated from MDA-MB-231$^{eGFP\ Luc}$ (n = 3) (a), and MCF-7$^{eGFP\ Luc}$ (n = 3) (b) cell lines. Data are expressed as the average of colonies per well.

As a whole, these data point to a lower ability of CTC cluster model to interact with endothelial cells than single CTCs. However, this observation should be taken with caution as this experimental approach was performed under static conditions that did not represent the dynamic environment in which these interactions take place. Furthermore, these results also demonstrate that the CTC cluster model has a higher ability to grow in the absence of anchorage, which is an essential feature to generate micrometastasis.

2.1.3. CTC clusters disseminate less within the circulation of the zebrafish embryo than individual CTCs

We next evaluated the potential ability of CTCs and CTC clusters to disseminate in a biologically complex system, using the zebrafish (ZF) embryos as a model to study BC progression. Their extra uterine development, the absence of an adaptive immune system, and their transparency that allow image-based monitoring very easily make ZF embryos an ideal model to study tumour progression.

Firstly, the ability of the human BC cell lines, MDA-MB-231$^{eGFP\ Luc}$ and MCF-7$^{eGFP\ Luc}$, to survive under the maintenance conditions of

the xenografted ZF embryos was tested. The maintenance conditions of ZF after injecting the tumour cells included a temperature of 34 °C, which is a suboptimal temperature for human cell growth. Despite the reduction in their proliferation capacity over time, it was observed that both BC cell lines survived and grew under these conditions (figure 32).
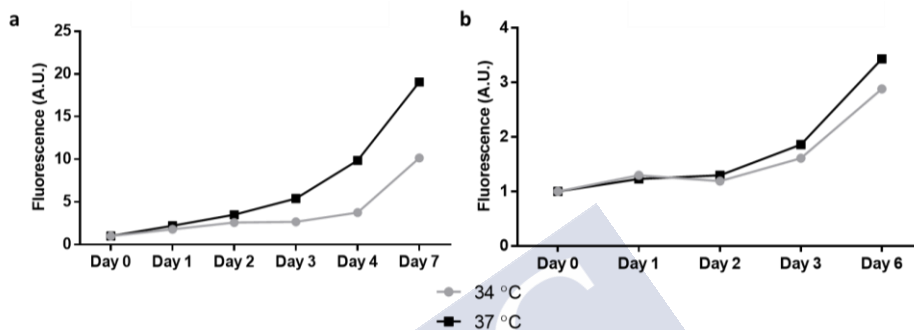

Figure 32. Proliferation of MDA-MB-231[eGFP Luc] **(a), and MCF-7**[eGFP Luc] **(b) cell lines at 34˚C (light grey), and 37˚C (dark grey).**

Afterwards, we proceeded to generate the ZF xenografts to study the ability of the CTC models to migrate and disseminate *in vivo*. For this purpose, equal numbers of single CTC and CTC cluster cells were injected near the convergence with the Duct of Cuvier and the pericardial space, and cell dissemination was quantified by the fluorescent signal detected in the tail of the ZF xenografts over time. ZF xenografts with disseminated cells at the time of injection were discarded (passive dissemination), as we were seeking for an active component during cell dissemination.

ZF xenografts injected with single and CTC cluster models derived from MCF-7[eGFP Luc] cell line showed that MCF-7[eGFP Luc] cells barely survive over time inside the embryo's body. Moreover, low tumour cell dissemination was observed with MCF-7[eGFP Luc] derived CTC models when injected close to the convergence of the Duct of Cuvier. There was a low percentage of ZF xenografts with dissemination to the tail at 24 hours post-injection (hpi) (30.6%) that was even lower at 72 hpi (3.4%), confirming that the disseminated cells did not survive over time

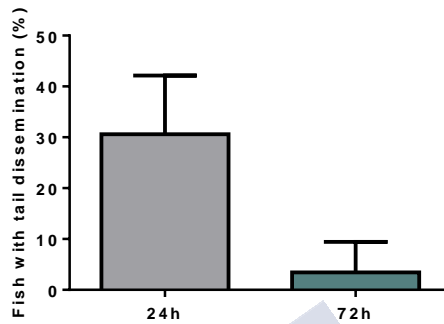in the embryo's body (figure 33). Consequently, MCF-7[eGFP Luc] cell line was discarded for later analysis.

Figure 33. ZF xenograft generated with MCF-7[eGFP Luc] cell line. Percentage of ZF embryos with disseminated cells in the tail at 24 hpi and 72 hpi.

Unlike MCF-7[eGFP Luc] models, MDA-MB-231[eGFP Luc] ZF xenografts showed that this cell line not only survived inside the embryo's body over time but also disseminated to the tail, potentially allowing us to observe differences between single CTC and CTC cluster models (figure 34). Thus, MDA-MB-231[eGFP Luc] was chosen to perform later *in vivo* experiments.
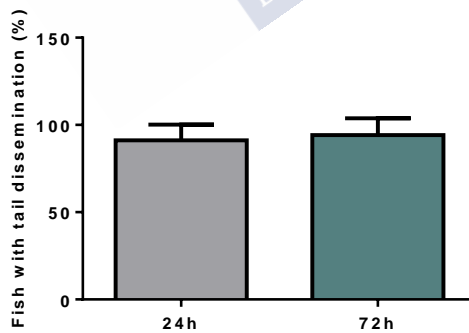

Figure 34. ZF xenograft generated with MDA-MB-231[eGFP Luc] cell line. Percentage of ZF embryos with disseminated cells in the tail at 24 hpi and 72 hpi.

Equal numbers of single and clustered MDA-MB-231[eGFP Luc] cells were injected in the embryos and tumours of equal size generated for both cell populations (figure 35).
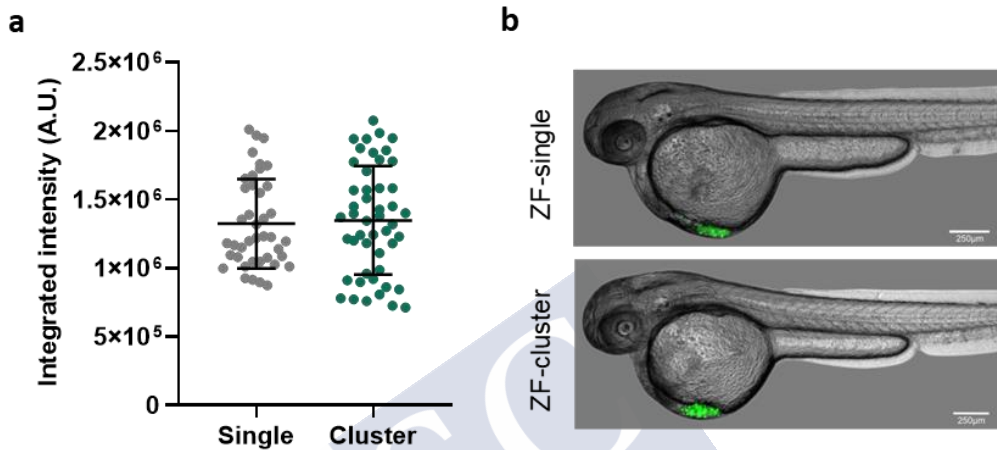

Figure 35. Quantification based on the fluorescence signal of the tumours generated in ZF injected with single CTCs or with CTC clusters at the site of injection (a). Representative images of ZF injected with single CTC and CTC cluster models (b). The injection was performed into the perivitelline space next to the convergence of the Duct of Cuvier (DoC). Scale bars: 250 μm.

ZF embryos injected with MDA-MB-231[eGFP Luc] single CTCs had a higher fluorescent signal in the tail region at 24 and 72 hours post-injection (hpi), compared to the fluorescent signal observed in the ZF injected with CTC clusters at the same timepoints. On average, the intensity of the fluorescent (A.U.) signal at 24 hpi was $1.998 \times 10^6 \pm 1.309 \times 10^6$ for single CTC xenografts, while for CTC cluster xenograft was $1.079 \times 10^6 \pm 8.98 \times 10^5$ ($p < 0.01$). At 72 hpi, the fluorescence intensity for single CTC and CTC cluster xenografts was $1.651 \times 10^6 \pm 1.315 \times 10^6$ and $1.037 \times 10^6 \pm 9.396 \times 10^5$ ($p > 0.05$), respectively. Taken into consideration that ZF with tumour cell dissemination right after injection were excluded from the analysis, these results indicated that single CTCs had a higher capacity to abandon the injection site than CTC clusters. Besides, it was observed

a significantly decrease in the fluorescence intensity between 24 hpi and 72 hpi in the ZF injected with single CTCs, while in ZF injected with CTC clusters the fluorescence intensity remained the same. This result suggests a better tumour cell survival in the ZF injected with CTC clusters than the ZF xenografts generated with individual CTCs (figure 36).
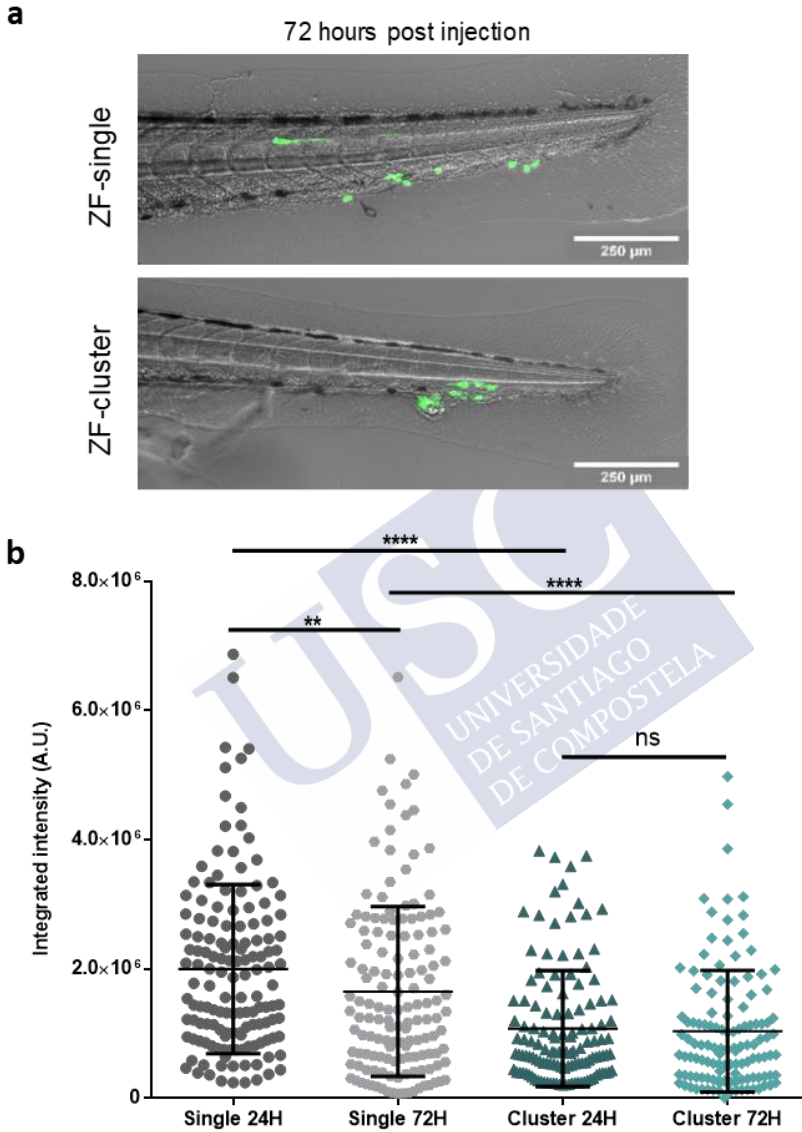
Figure 36. Representative images of cell dissemination to the tail of ZF injected with the MDA- MB-231[eGFP Luc] single CTC or CTC cluster models. Scale bars: 250µm (a). Fluorescence intensity of disseminated cells to the tail at 24 and 72 hours post-injection in ZF xenograft injected with single CTCs or with CTC clusters. (n = 6 independent experiments). ns not significant, ** p < 0.01, **** p < 0.0001.

Additionally, we examined the ability of CTCs and CTC clusters to move within the tail of the ZF embryo once disseminated (figure 37, a). For this, we assessed the dissemination patterns of the cells based on the division of the tail into three different areas, according to the vasculature distribution: dorsal (dorsal longitudinal anastomotic vessel), ventral (area of caudal vein, posterior cardinal vein, and dorsal aorta), and lateral (intersegmental vessels) (see '*Zebrafish (Danio rerio) embryo xenografts*', on the Materials and Methods section). The percentage of ZF with dorsal dissemination was slightly higher in the xenografts injected with single CTCs (63.2%), compared to the ZF injected with CTC clusters (37.9%) (p = 0.200) (figure 37, b). Similar levels of ventral dissemination were observed between both types of xenografts, as 94.7% of single cell ZF had ventral dissemination, while 98.6% of CTC cluster ZF had disseminated cells in this region (p = 0.700) (figure 37, c). Regarding the dissemination to the ventral area, it should be taken into consideration that CTCs are dragged by the direction of the blood flow in the ZF from the site of injection along the dorsal aorta to the cardinal vein, where CTCs are arrested at the caudal plexus. Moreover, the proportion of ZF with lateral dissemination also tended to be higher in the single CTC ZF xenograft (37.0%) than the CTC cluster ZF xenografts (15.0%) (p = 0.400) (figure 37, d).
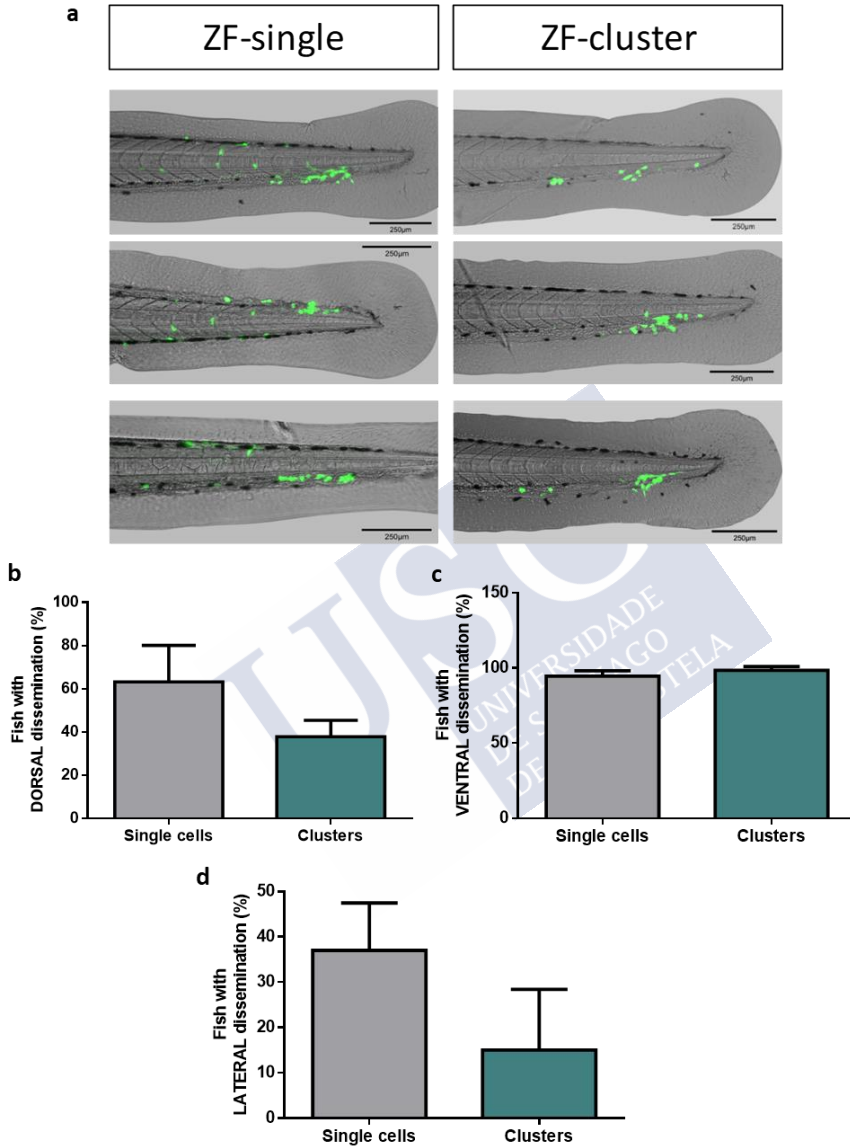
Figure 37. Representative images of the dissemination patterns of ZF injected with single CTCs (left) or with CTC clusters (right). Scale bars: 250 μm (a). Percentage of ZF xenografts generated with single CTCs or with CTC clusters with dorsal (b), ventral (c), and lateral (d) dissemination in the tail (n = 6).

In a more detailed analysis, the number of foci established by the disseminated cells in the three different regions of the caudal area was also determined. ZF xenografts generated with single CTCs showed a significantly higher number of foci in dorsal location than CTC cluster ZF xenografts (2.2±0.7, and 1.2±0.3, respectively, $p < 0.05$) (figure 38, a). As expected, no difference was observed between ZF xenograft populations in the number of ventral foci. Single CTC ZF xenografts showed 4.6±0.3 ventral foci, while ZF injected with CTC cluster had 4.2 ± 0.2 ventral foci (figure 38, b). Finally, ZF injected with single CTCs showed a significantly higher number of lateral foci than CTC cluster ZF xenografts (0.9±0.3 and 0.3±0.08, respectively, $p < 0.05$) (figure 38, c).
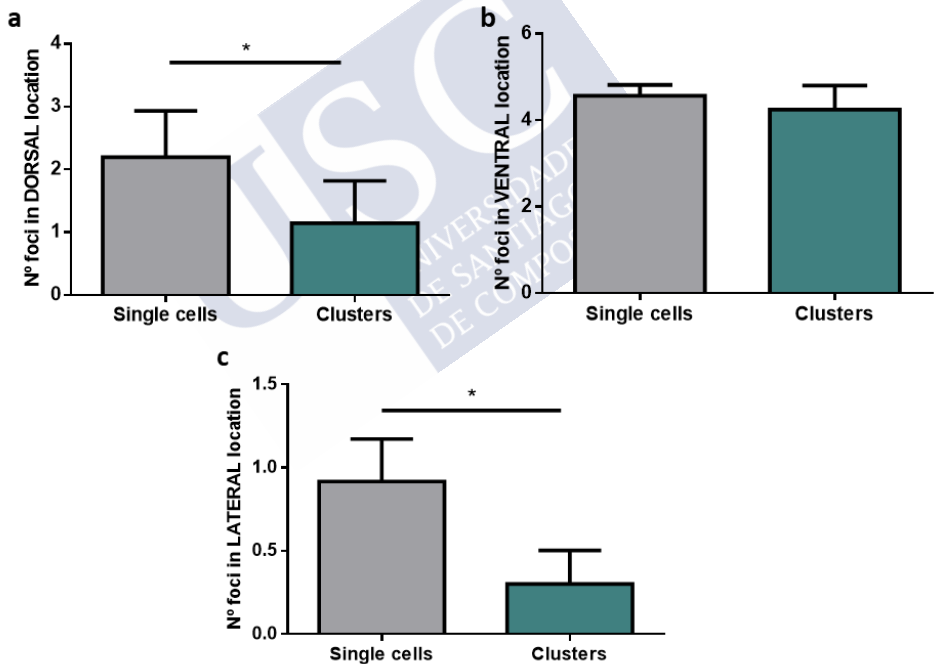

Figure 38. Number of tumour foci in the tail with dorsal (a), ventral (b), and lateral (c) location in the ZF injected with single CTCs or with CTC clusters (n = 6 independent experiments). * p < 0.05.

The restricted ability of the MDA-MB-231$^{eGFP\ Luc}$ CTC clusters to abandon the injection site and disseminate through the circulation of the ZF embryo is explained by their larger size, which physically restricts their movement. This can be the reason for the low frequency of CTC clusters in the circulation, in line with the observations performed in other animal models (i.e mouse) and in BC patients. However, the lower dissemination of CTC clusters also seems to suggest a lower metastatic potential of CTC clusters.

2.1.4. CTC clusters survive better in the circulation of the zebrafish embryo xenografts and have a proliferative advantage in comparison with single CTCs

After intravasation, CTCs should survive within the bloodstream. However, during this critical step of metastasis, CTCs are exposed to mechanical stress due to the presence of fluid shear forces. To evaluate how fluid shear stress (FSS) affects both individual CTCs, and CTC clusters, we carried out *in vitro* FSS assays. Cell suspensions either as individual CTCs, or as CTC clusters were loaded into a syringe, which was linked to a 30G needle, and to a syringe pump to simulate the shear stress forces of the bloodstream (see '*Fluid Shear Stress (FSS) assays*', on Materials and Methods section). It was considered a passage (P) when the entire content of the syringe passed through the needle. A total of 10 passages were performed in each assay, and a 2-minute resting time was allowed between subsequent passages. Under these conditions, FSS assays allowed us to observe a differential impact of shear stress forces on the viability of the *in vitro* models, based on the bioluminescent signal. Bioluminescence was measured after finishing the assay (0 h timepoint). After each even FSS passage aliquots of cells were also put in culture to measure the bioluminescence signal 24 hours later (24 h timepoint).

MDA-MB-231$^{eGFP\ Luc}$ CTC cluster model exhibited significantly higher viability than single CTC population, at both timepoints, 0 and 24 hours (figure 39, a, b). Although both populations experienced a decrease in survival throughout FSS cycles at 0-hour timepoint, CTC cluster population showed a smaller reduction in survival than single

CTCs, which suggests that the immediate negative impact of FSS is smaller in the CTC cluster population than in the single CTC cell suspension. Specifically, CTC cluster population showed survival of 47.32% at P4, while individual CTC showed a survival of 16.48%. The lower impact of FSS over the survival of CTC clusters was also observed in prolonged exposure to FSS, as CTC clusters had a survival of 28.68% at P6 that decreased to 20.62% at P8. However, single CTC survival was 9.27% and 5.13% for the same timepoints, respectively (figure 39, c).

Differences between both single CTCs and CTC clusters were also noticeable at the timepoint of 24 hours. The 24-hour timepoint not only allowed the evaluation of the survival but also showed the ability of cells to overcome the mechanical stress and even to potentially proliferate in culture. CTC clusters exposed to early FSS cycles (P2-P4) showed sustained viability showing a survival of 79.42% at P4, while intermediate FSS passages (P6-P8) generated a progressive reduction in their viability reaching a survival percentage of 45.30% at P8, when compared to the control (P0). On the other hand, single CTC population exposed to early FSS cycles (until P4) experienced a large reduction of viability, based on the drop of the bioluminescence signal. Individual CTC population had a great reduction in the viability as early as P4 (16.2% cell survival), and the viability decreased to 5.7% at P8. The final viability registered after 10 cycles of FSS (P10), at the 24-hour timepoint, was 13.15% for CTC cluster population, while single CTC population showed a final viability of 4.12% when compared to the viability of the non-FSS exposed population (P0) (figure 39, d).

Additionally, the comparison of the bioluminescence signal at 24h timepoint relative to the 0h timepoint of the different FSS passages showed that CTC clusters have a higher ability to recover from FSS than individual CTCs (figure 39, e, f). In fact, it was observed that CTC clusters exposed to early FSS passages (P4) had a bioluminescence fold change of 1.32 relative to the corresponding value at 0h, meaning that there were 32% more viable cells in the culture. On the contrary, single CTCs showed a fold change of 0.83 at P4, which indicates a decrease in the cell viability and the total number of cells. Similarly, CTC

clusters had a fold change of 2.02 and 1.53 at prolonged FSS passages (P6 and P8, respectively). This indicates that despite the negative effect of the shear stress forces there was cell proliferation in the CTC cluster population after several FSS passages. However, single CTC population did not proliferate between 0 and 24 h timepoints, as it is showed by the fold change of 0.85 at P6, and of 1.0 at P8. This indicates that the number of cells at 24h is similar, or even slightly lower, to the one observed at 0h, thus reflecting an absence of proliferation between 0 and 24 h timepoints in the single CTC population.
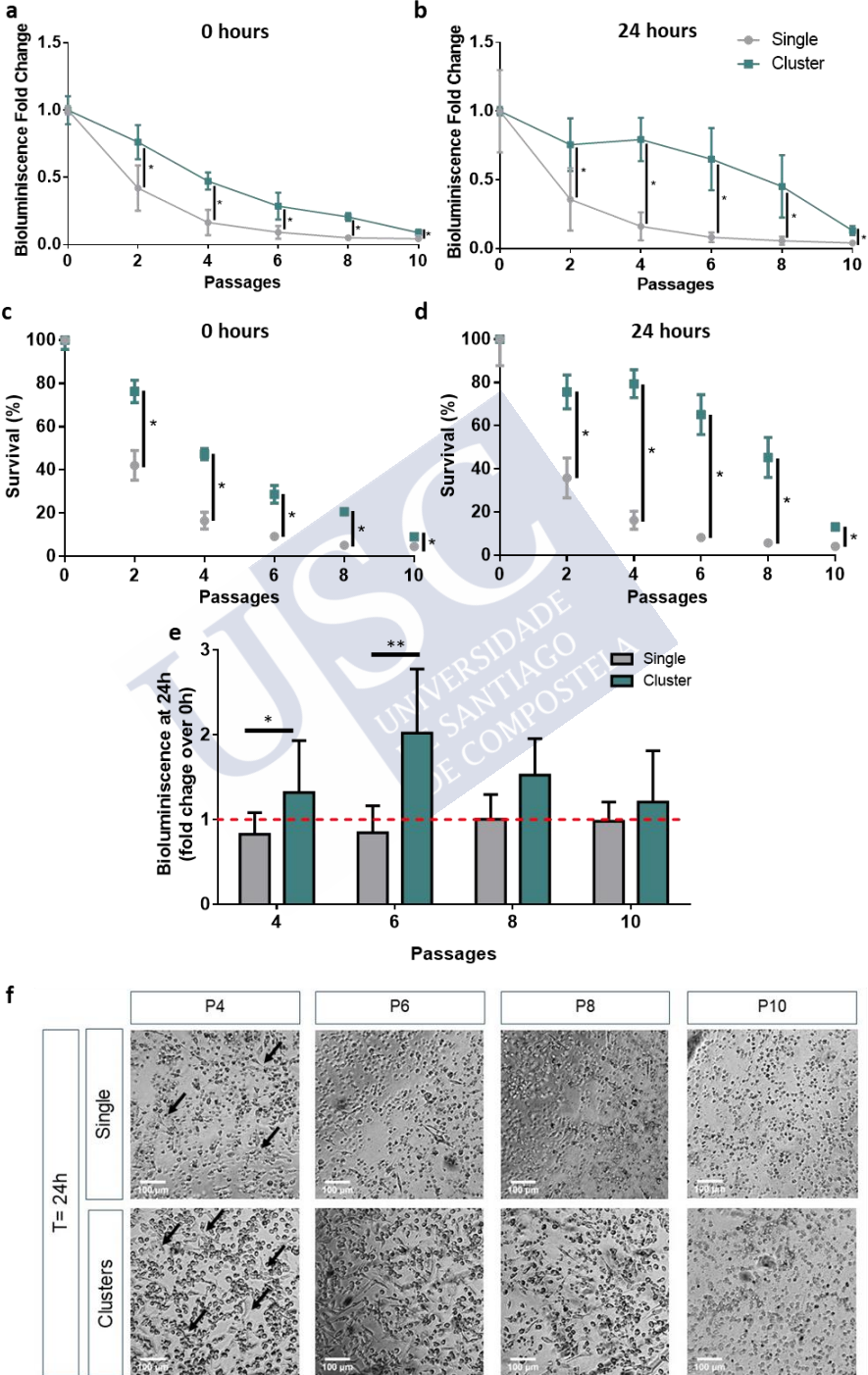
Figure 39. Fluid Shear Stress (FSS) assay (n = 3). Cell survival of single CTCs and CTC clusters subjected to FSS cycles at 0 hours (a), and 24 hours later (b). Data are expressed as the bioluminescence fold change relative to the average bioluminescence measured for the non-FSS exposed control (P0). Percentage of cell survival of single and CTC clusters after the exposure to FSS at 0 hours (c) and 24 hours (d). * p < 0.05. Data are expressed relative to the survival observed in the non-FSS exposed control (P0). Bioluminescence fold change of 24 hours relative to the average of bioluminescence of the corresponding passage (P) at 0 hours (e). * p < 0.05, ** p < 0.01. Representative images of single CTCs and CTC cluster cell suspension 24 hours after exposure to different passages (P) of FSS (P4, P6, P8 and P10). Black arrows indicate tumour cells adhered to the culture surface. Scale bars: 100 μm (f).

Therefore, the differences in cell survival observed at 0-hour and 24-hour timepoints showed that CTC cluster population could have a higher capacity to overcome the negative effect of mechanical stress during circulation than the individual CTC model, and a higher ability to recover from the FSS exposure.

The results of the *in vitro* FSS assay and cell survival are in line with the variations observed in the fluorescent signal over time between ZF injected with single CTCs and CTC clusters, in which ZF injected with single CTCs decreased their fluorescence while ZF injected with CTC clusters maintained their fluorescence over time (figure 40). To analyse a possible proliferation of CTC clusters within the ZF, the caudal fluorescent signal of tumour cells in each ZF xenograft was individually analysed along time.

We observed that the fluorescence intensity was reduced in most of the single CTC ZF xenografts from 24 to 48 or 72 hpi. In fact, in 48.3% of single cell ZF xenografts the fluorescence intensity was reduced between 24 and 48 hpi, while in the CTC cluster ZF xenografts this percentage was 43.5%. At 72 hpi, the fluorescence intensity decreased in 66.2% of the single CTC ZF xenograft respect to 24 hpi, while only 52.6% of the ZF injected with CTC cluster showed a decreased fluorescent signal. This decrease in the intensity of the fluorescent signal evidenced a reduction in the number of cells present in the tail between 24 hpi and 48 or 72 hpi.

In addition, the percentage of ZF injected with single CTCs that maintained the fluorescence levels between 24 hpi and 48 hpi was 14.2%, while in ZF injected with CTC cluster this percentage was 11.9%. At 72 hpi, 17.6% of the ZF injected with single CTCs had similar fluorescence levels to 24 hpi, meanwhile this percentage was 8.0% in the case of CTC cluster ZF xenografts. Besides, 37.5% of the xenografts injected with single CTCs showed an increase in the fluorescence in the tail at 48 hpi, while in the ZF injected with CTC clusters this percentage reached 44.7%. Importantly, the fluorescent intensity increased in the 39.4% of ZF injected with CTC clusters from 24 to 72 hpi, while the percentage was significantly lower (p = 0.0286) in the single CTC ZF xenografts (16.3%). The increase in the fluorescence over time observed in these fish may be explained by a higher number of cells in tail, probably due to cell proliferation of the disseminated cells.

Therefore, at 72 hpi we observed a significantly higher percentage of ZF xenografts in which the fluorescence intensity increased among the embryos injected with CTC cluster, compared to those injected with single CTCs, suggesting that CTC clusters not only survive better in the tail of the ZF than single CTCs, but also have a proliferative advantage once disseminated. These results also showed a non-significant tendency in the CTC cluster ZF xenograft population to have a lower percentage of embryos with reduced fluorescence intensity than single CTC xenografts (p = 0.1143) (figure 40).
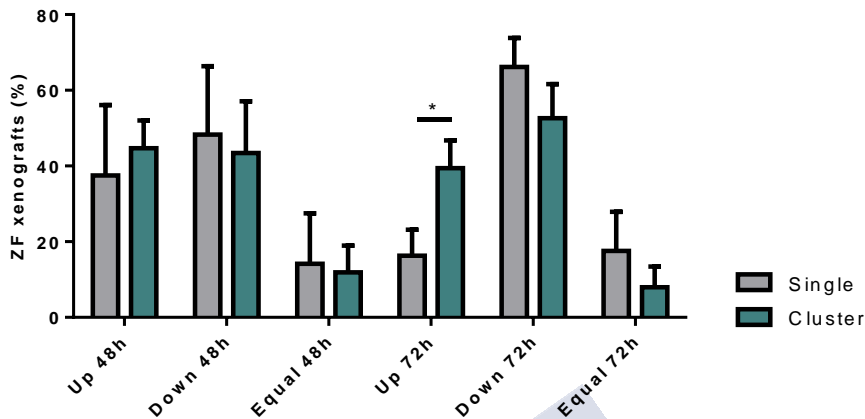
Figure 40. Percentage of ZF xenografts generated with single CTCs or with CTC clusters whose fluorescence intensity increased (up), decreased (down) or was maintained (equal) between 24 and 48 or 72 hours post-injection (n = 6 independent experiments). * p < 0.05.

Therefore, although CTC clusters disseminate to the tail of ZF embryos in fewer numbers than single CTCs, they have an enhanced survival within the circulation and a proliferative advantage, as shown by the FSS assays and ZF experiments. This greater survival of CTC clusters during circulation and their ability to proliferate after being exposed to FSS can potentially allow them to generate larger metastatic lesions, as these are essential features to seed successful metastases, and could partially explain their higher metastatic potential. Furthermore, these results also highlight the great potential of the zebrafish as an animal model to unravel the biology of CTCs, and especially of CTC clusters.

2.1.5. Gene expression profiling supports the higher metastatic potential of the CTC cluster model

Gene expression analyses were performed using RT-qPCR to find genes differentially expressed that could potentially explain the differential behaviour observed between single CTC and CTC clusters models in the ZF. Gene expression assays were performed on pools of

disseminated cells isolated from ZF xenografts of four different experiments at 48 hpi. To this end, 48 hpi zebrafish tails were physically separated from the rest of the embryo body, mechanically and enzymatically digested to isolate the disseminated tumour cells and the RNA extracted. A panel of genes related to proliferation (i.e., MKI67), cell cycle regulation (i.e., CDK4, E2F4), cell survival (i.e., PLAU), and stemness (i.e., CD44, ITGA6) was used for this analysis and the expression levels were determined relative to the average expression of two different housekeeping genes (figure 41, a).
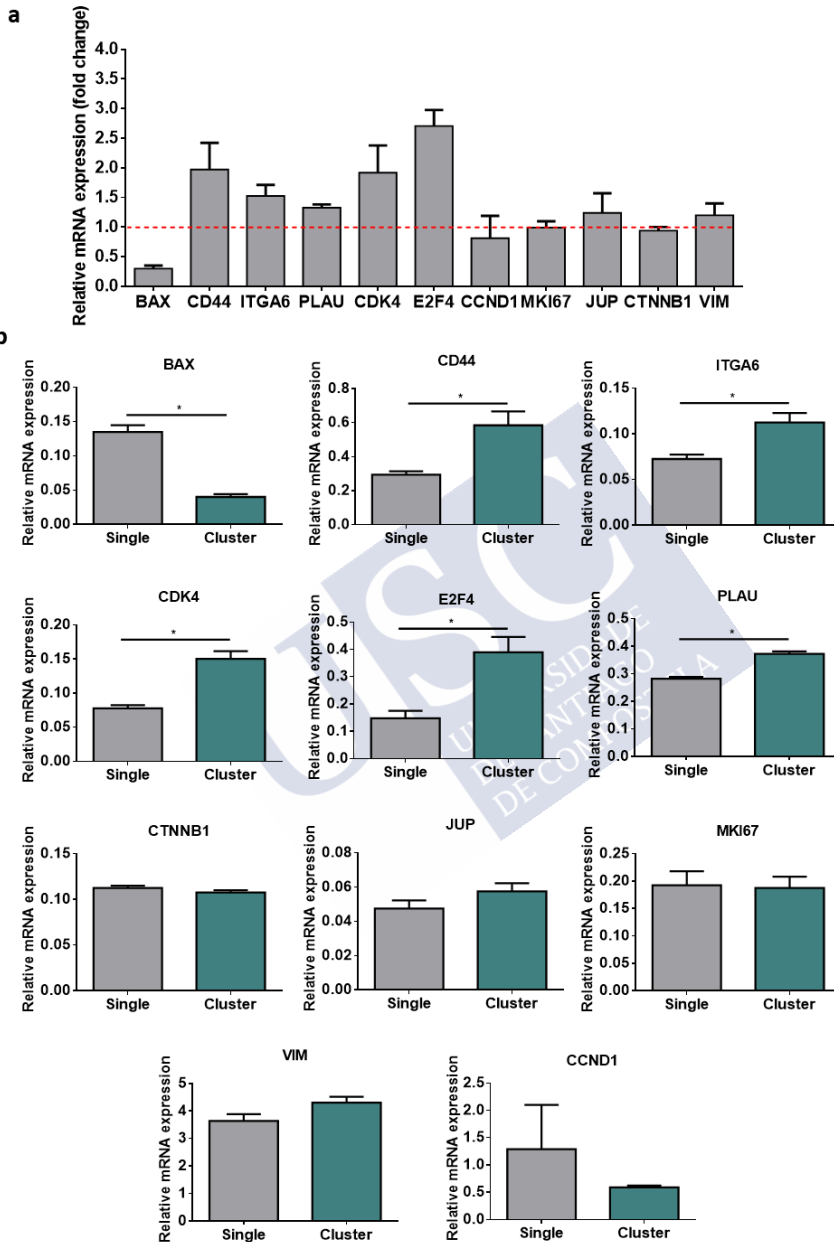
Figure 41. Gene expression analysis of the disseminated cells of ZF injected with CTC clusters relative to the expression of ZF injected with single CTCs (a). Gene expression levels in both cell populations expressed as $2^{-\Delta Ct}$ (b). Data are

normalised by the average expression of the housekeeping genes GAPDH (glyceraldehyde-3-phosphate dehydrogenase) and ß2M (beta-2-microglobulin). * p < 0.05.

Gene expression data showed that the disseminated cells of the ZF injected with the MDA-MB-231$^{\text{eGFP Luc}}$ CTC cluster model had a lower expression of the pro-apoptotic gene BAX when compared to single CTC xenografts (p = 0.0286). Moreover, we found a tendency towards a higher expression of the stem cell-like related gene CD44, and a significantly higher expression of ITGA6 in the CTC cluster ZF xenografts (p = 0.0286; p = 0.0286, respectively). Urokinase-type plasminogen activator gene PLAU was found to be more expressed in the CTC cluster xenografts (p = 0.0286). It was also observed that those ZF injected with CTC clusters had a significantly higher expression of genes related to cell cycle regulation, such as CDK4, E2F4, and CCND1 (p = 0.0286; p = 0.0286, p = 0.2857, respectively). Nevertheless, no significant differences were observed in the expression of the proliferation marker MKI67 gene between single CTCs and CTC cluster ZF disseminated cells (p = 0.9714). Likewise, no differences were found in the expression of the cell junction related gene plakoglobin (JUP), the cell adhesion related gene β-catenin (CTNNB1) and the EMT related gene VIM (p = 0.3143; p = 0.5714; p = 0.0857) (figure 41, b).

These findings are in line with the *in vitro* and *in vivo* functional differences observed between the CTC models and support the suitability of the CTC cluster model to deepen the study of the biology of CTC clusters.

2.1.6. CTC cluster has a higher colonization capacity in mice than individual CTCs

To seek if the functional differences are maintained in a more complex biological system than the ZF, a mouse lung colonization assay was done. This assay was performed to test the capacity of CTC models to survive within the bloodstream, as well as their ability to colonize and proliferate in a new niche, such as the pulmonary tissue,

for the development of metastatic lesions. Hence, equal numbers of single CTC or CTC cluster suspensions were directly injected into the bloodstream, through the lateral vein of the tail of immunodeficient SCID BEIGE mice.

Despite the biological variability observed in this assay, we could observe that during the initial days upon injection a large proportion of cells of both populations died at the lungs, based on a noticeable drop in the bioluminescence counts. However, the decrease in the bioluminescence signal from day 0 (taken 1 hour post-injection) to day 3 was less pronounced in the group of mice injected with CTC clusters than in the group injected with single CTCs (bioluminescence counts fold change 0.3 ± 0.2 vs. 0.03 ± 0.05 respectively; $p = 0.031$), meaning that there was a lower incidence of cell death in this group (figure 42, a). This tendency was maintained at 6 days post-injection (dpi), where CTC cluster xenografts had a bioluminescence fold change of 0.2 ± 0.1, while the bioluminescence fold change relative to the signal emitted at 1 hour post-injection timepoint for single CTC xenografts was 0.1 ± 0.1. However, the tendency observed at 6 dpi was not statistically significant ($p = 0.5635$) (figure 42, b). These data evinced that although there was a massive cell death during the first days after injection in both populations, the proportion of dead cells was lower in the CTC cluster injected mice.

The monitoring over time and the determination of the metastatic incidence in the lungs by bioluminescence quantification permitted us to observe that mice injected with CTC cluster population showed a higher bioluminescence signal in the lungs over time when compared to mice injected with the single cell suspension (figure 42, c).
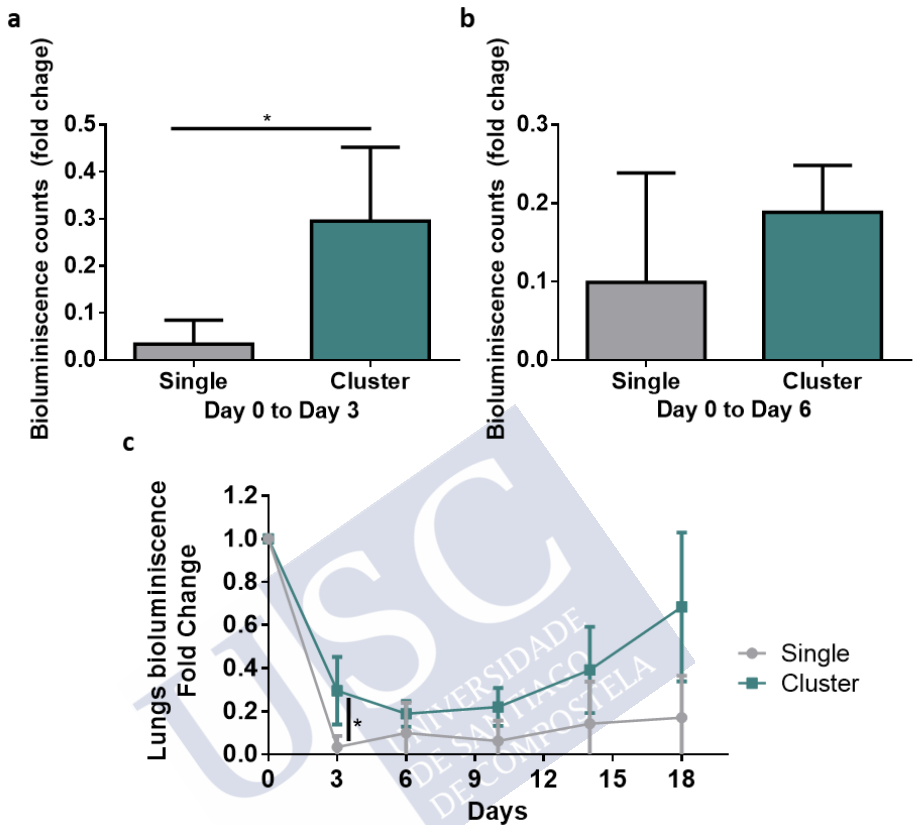
Figure 42. Bioluminescence counts fold change between day 0 (1 hour after injection) and day 3 after injection (a). Bioluminescence counts fold change between day 0 and day 6 after injection. Bioluminescent signal emitted over time by the tumour cells that colonized the lungs (n = 4 mice per group) (c). * p < 0.05.

Moreover, histopathological analysis of lungs also showed a clear tendency to a higher metastatic incidence in those mice injected with the CTC cluster suspension. On average, the number of metastatic foci identified in the single CTC mouse group was 99.0±105, while the number of foci detected in mice injected with CTC clusters was 321.3±101 (p = 0.0635) (figure 43).
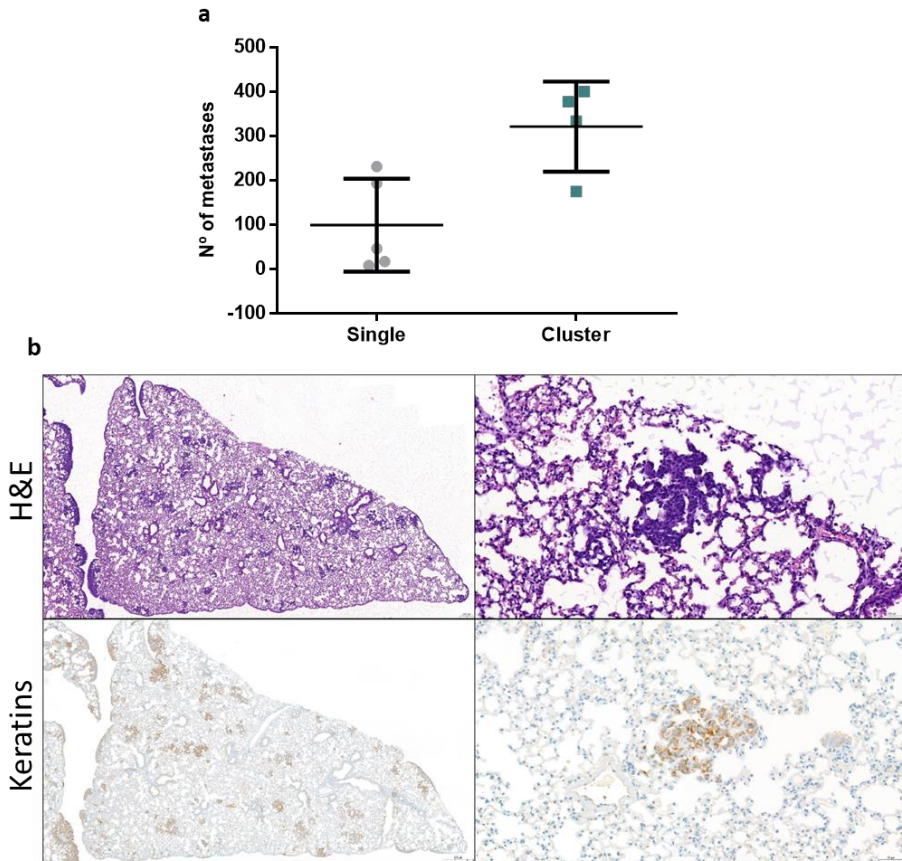
Figure 43. Quantification of the number of metastatic foci observed in the lungs of mice injected with either, individual CTCs or CTC clusters (a). Representative sections of pulmonary tissue of a mouse injected with CTC clusters with metastasis stained with Hematoxylin-eosin (H&E, upper panels), and with antibody against keratins (bottom panels). Note the brown staining in tumour cell foci and nodules. Left images are shown with low magnification (5x), while right images are at high magnification (20x) (b).

The quantification was in line with the final bioluminescence signal of surgically removed lungs. On average, the endpoint bioluminescent signal of the lungs was $4.03 \times 10^8 \pm 4.57 \times 10^8$ for the single CTC mouse group, while for the CTC cluster mouse group the endpoint bioluminescence was $1.03 \times 10^9 \pm 1.57 \times 10^8$ (figure 44).

a

Single – D18
(R6 – R7)

Cluster – D18
(R21 – R23)

b

Lungs (R6)

Lungs (R21)

Lungs (R7)
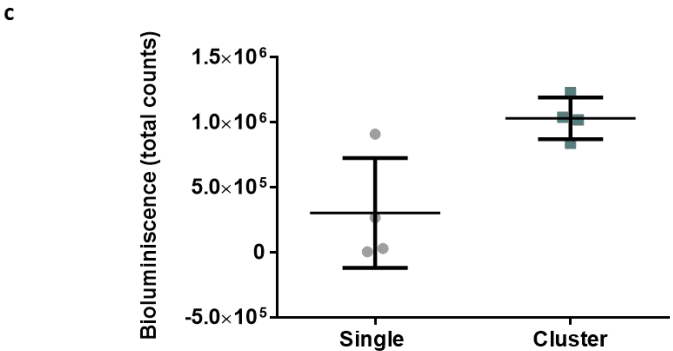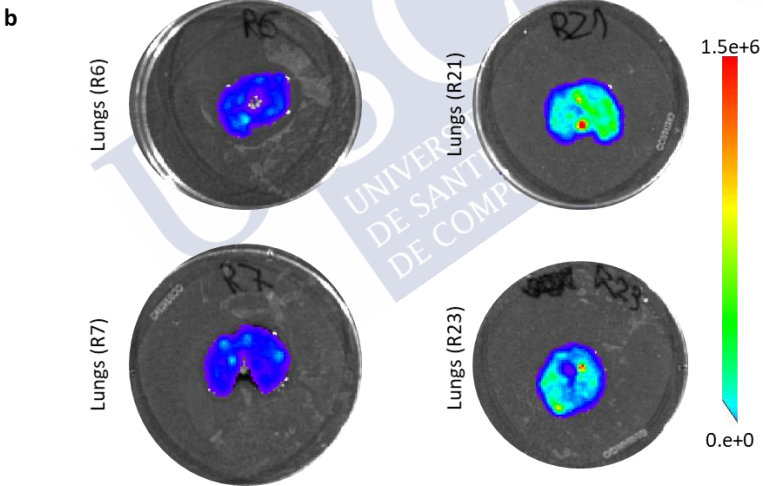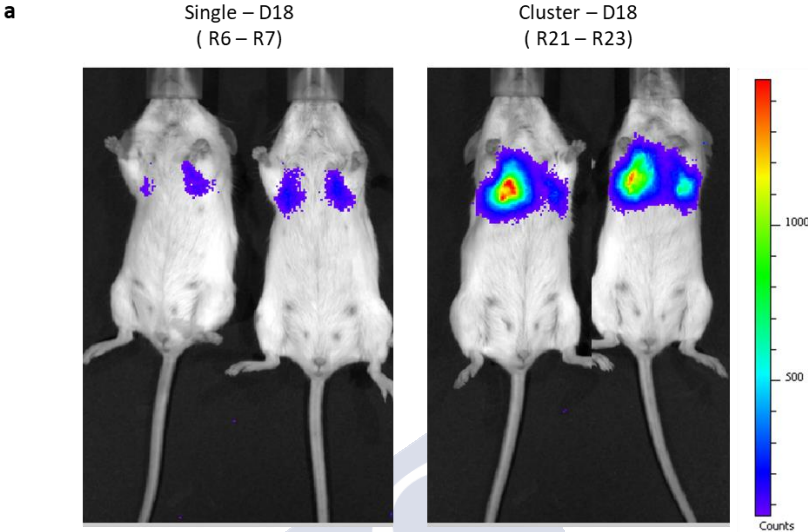
Lungs (R23)

1.5e+6

0.e+0

c

Figure 44. Mouse lung colonization assay. Representative images of the pulmonary metastatic incidence in mice injected with either single CTCs or CTC clusters. R6 and R7 mice were injected with single CTCs (left) whereas R21 and R23 were injected with CTC clusters (right) (a). Representative images of the metastatic incidence in the lungs of mice injected with single CTCs or with CTC clusters after being surgically removed (b). Quantification of the bioluminescence signal (total counts) measured in the surgically removed lungs of mice injected with single CTCs or CTC-clusters at the endpoint (c).

These results demonstrate that the CTC clusters have a higher ability to colonize a new niche and generate new tumour lesions than the single CTCs in mice.

Taken together, functional *in vitro* and *in vivo* characterisation showed the existence of functional differences between the *in vitro* models of individual CTCs and CTC clusters. These differences were in concordance with the gene expression analyses and suggest that the *in vitro* CTC cluster model has a higher metastatic potential than the single CTC model. Therefore, the CTC cluster model can be a useful tool that contributes to overcoming the limitation of material, due to the low frequency of CTC clusters in patient samples, to study the biology of CTC clusters in a wider manner. Besides, these results also outline the great potential of the zebrafish as a model to study the metastatic process, as well as the role of CTCs and CTC clusters during tumour spread.

3. USE OF A METASTATIC ORTHOTOPIC BC MURINE MODEL AS A SOURCE FOR CTCS AND CTC CLUSTERS

In addition to the development of CTC cluster models from BC cell lines, a metastatic murine orthotopic BC model established with the MDA-MB-231$^{eGFP\ Luc}$ cells was used to test the feasibility of CTC cluster detection and isolation from blood samples. This would allow us to use it as a potential source of CTCs, and especially CTC clusters. MDA-MB-231$^{eGFP\ Luc}$ cells were injected at the mammary fat pad of immunodeficient SCID BEIGE female mice.

The bioluminescent signal associated with the development of the primary tumour in the mammary fat pad was detected between weeks 3 and 4 after injection. Afterwards, tumour development was assessed until the endpoint, at week 8 after injection. During this period of time, it was observed a progressive tumour growth (figure 45, a).

Pulmonary metastasis formation was also tracked over time. Generally, the metastatic bioluminescent signal was detectable at week 6 after injection. Over the weeks, bioluminescence associated with lung metastases was progressively increasing until the endpoint (figure 45, b).
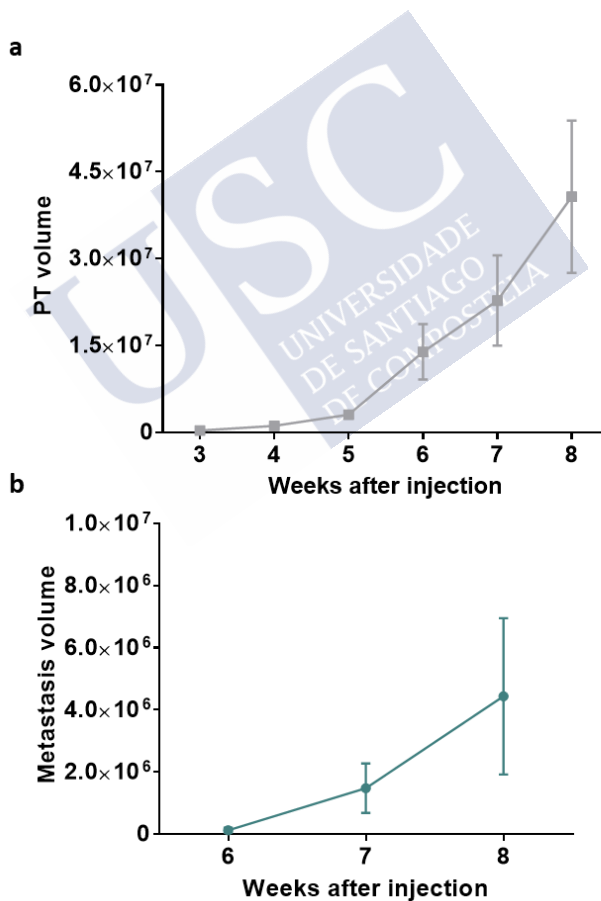
Figure 45. Monitoring of tumour progression in the orthotopic BC murine model based on the quantification of the bioluminescence signal. Primary tumour growth over time (a). Lung metastasis development over time (b). Data are expressed as total bioluminescence counts (n = 6 mice per group).

Blood samples were collected at the endpoint by cardiac puncture and used for CTC and CTC cluster isolation with the immune-independent system Parsortix[TM]. Individual CTC and CTC cluster enumeration was performed after CTC isolation based on the presence of eGFP positive cells within the separation cassette. Importantly, it was possible to detect and isolate CTCs and CTC clusters from the blood and determine a ratio of individual CTCs:CTC clusters. This ratio was 1:0.127 in this murine BC model, meaning that CTC cluster population represents 12.7% of all CTCs. Moreover, a positive correlation was found between the number of individual CTCs per 100 µL of blood, and the number of CTC clusters per 100 µL of blood ($R^2 = 0.9267$). Thus, the higher is the individual CTC frequency of detection, the more CTC clusters are expected to be found in a blood sample (figure 46).
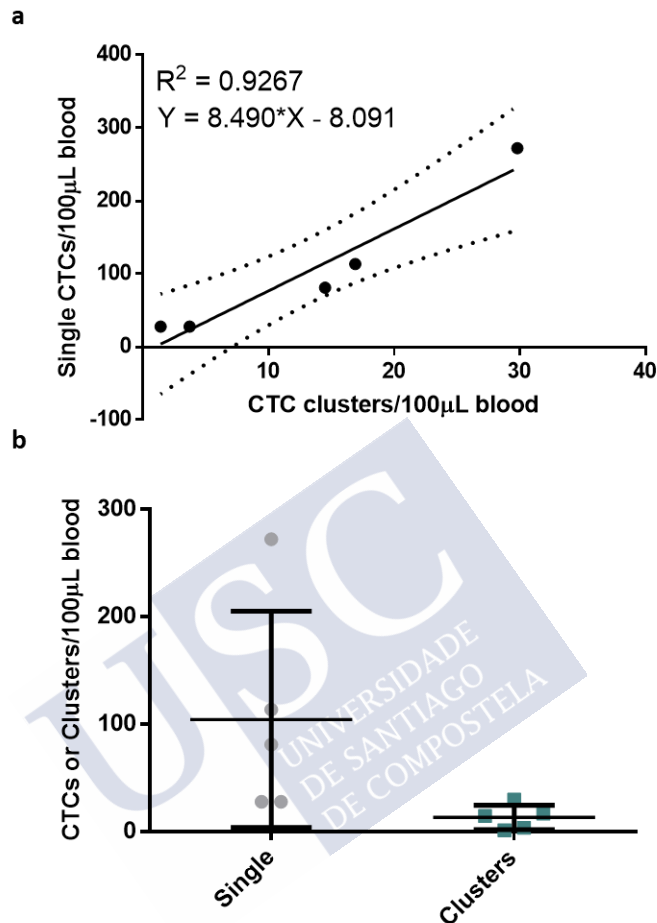
Figure 46. Analysis of the CTCs and CTC clusters isolated from the blood samples of the orthotopic BC murine model. Tumour cells were isolated in the Parsortix[TM] system. Correlation between the number of individual CTCs/100 μL blood and the number of CTC clusters/100 μL blood (a). Number of individual CTCs and CTC clusters detected per 100 μL of blood. Each dot represents an individual mouse (b).

3.1. Establishment of an *ex vivo* CTC cluster surrogate model

Given that CTCs and CTC clusters were detected in blood samples from the metastatic MDA-MB-231[eGFP Luc] mouse model, in collaboration with the Nano-oncology research line of the Roche-

CHUS Joint Unit, we isolated CTCs in a viable state and put them in culture with the aim of establishing a CTC-derived cell line *in vitro*. These cells grew *in vitro* as cellular aggregates under low-attachment conditions, and these properties were maintained after cryopreservation periods. The cell line established was as mCTC (*murine model CTC derived cell line*) (figure 47, a). Hence, the mCTC cell line was preliminarily characterised, as it could potentially represent a more realistic CTC cluster model, from a physiological point of view.

Firstly, taking into consideration the origins of this cell line, mCTC resistance to fluid shear stress (FSS) was studied in comparison with the MDA-MB-231$^{eGFP\ Luc}$ CTC cluster model. At 0-hour timepoint, it was observed that both populations decreased their survival at P2, showing a survival of 65.95% in the case of MDA-MB-231$^{eGFP\ Luc}$ CTC cluster model, and 71.40% in the mCTC cell line. Nevertheless, mCTC cell line showed sustained survival between short and intermediate exposure to shear stress (P2-P6), having a survival of 73.05% after 6 cycles of shear stress (P6), and a final survival of 55.90% after 10 cycles of shear stress (P10). In contrast, MDA-MB-231$^{eGFP\ Luc}$ CTC clusters progressively decreased their survival reaching a 25.88% at P6, and a final survival of 9.28% at P10. This suggests that mCTC cell line has a higher ability to resist the immediate negative impact of FSS over cell survival than the MDA-MB-231$^{eGFP\ Luc}$ CTC cluster model (figure 47, b).

Differences between both CTC cluster models were also observed at the 24-hour timepoint, after their maintenance in cell culture. MDA-MB-231$^{eGFP\ Luc}$ derived CTC cluster population showed a cell survival of 64.88% after P2, while mCTC cell line had a survival of 89.22% after 2 FSS cycles (P2). Moreover, MDA-MB-231$^{eGFP\ Luc}$ CTC clusters showed survival of 32.04% after intermediate exposure to shear stress, at P8, and a final survival of 10.54% after 10 FSS cycles (P10). On the contrary, mCTC experienced a progressive increase in its survival after P2 and throughout the following FSS cycles. Thus, mCTC cell line not only showed a higher survival capacity than MDA-MB-231$^{eGFP\ Luc}$ CTC cluster model but also greater resilience to FSS, as it proliferated after a long exposure to FSS (survival at P10) (figure 47, c).
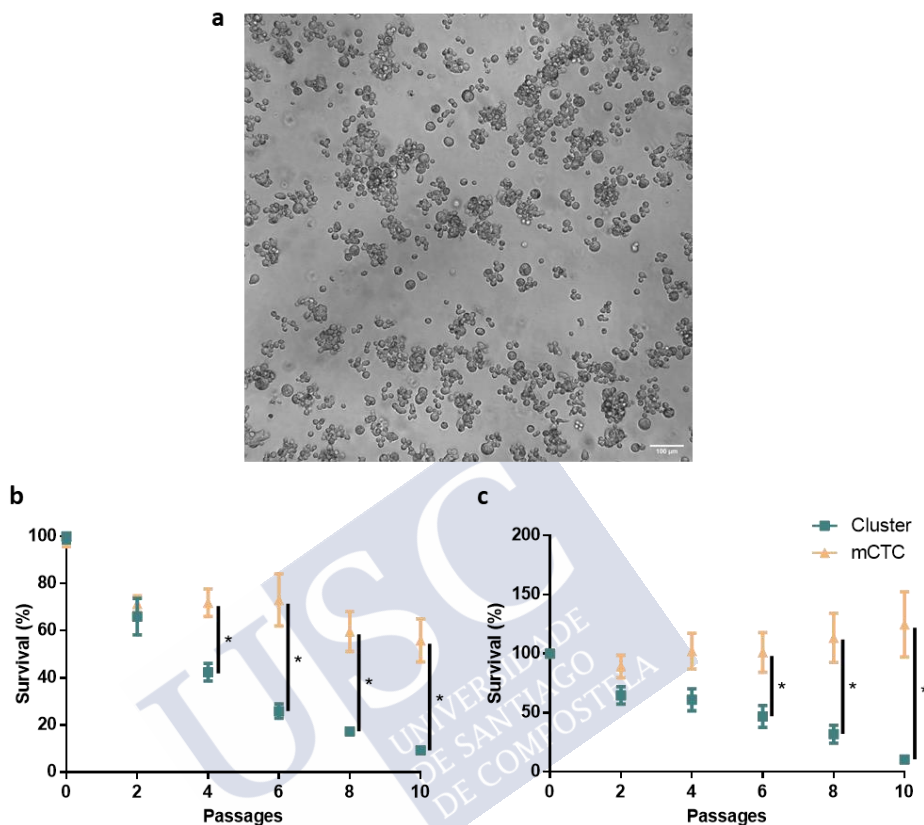
Figure 47. Evaluating the effect of the FSS over the mCTC cell line.
Representative image of the mCTC cell line growing in culture. Scale bar: 100 µm (a). Cell survival at 0 h (b) and at 24 h (c) (n = 3). * $p < 0.05$.

Furthermore, *in vivo* characterization using zebrafish (ZF) embryo xenografts was carried out to test the behaviour of the mCTC cell line when injected into the perivitelline space, close to the Duct of Cuvier (DoC). The dissemination ability of the mCTC cell line was compared with the dissemination of the parental MDA-MB-231[eGFP Luc] cell line. ZF xenograft generation permitted to observe that mCTC cell line showed a significantly higher fluorescent signal at the fish tails, meaning higher cell dissemination to the caudal region than the MDA-MB-231[eGFP Luc] cell line at both, 24 hpi ($2.17 \times 10^7 \pm 2.55 \times 10^7$ vs.

$2.75 \times 10^6 \pm 2.99 \times 10^6$, respectively; $p < 0.001$) and 72 hpi ($2.73 \times 10^7 \pm 3.49 \times 10^7$ vs. $1.05 \times 10^6 \pm 1.45 \times 10^6$). Besides, MDA-MB-231[eGFP Luc] cell line showed a significant decrease in the fluorescence intensity at 72 hpi when compared to 24 hpi, while mCTC cell line had no significant difference in the fluorescence intensity between 24 and 72 hpi. This indicates that disseminated mCTC cells survive better in the tail than the MDA-MB-231[eGFP Luc] population (figure 48).
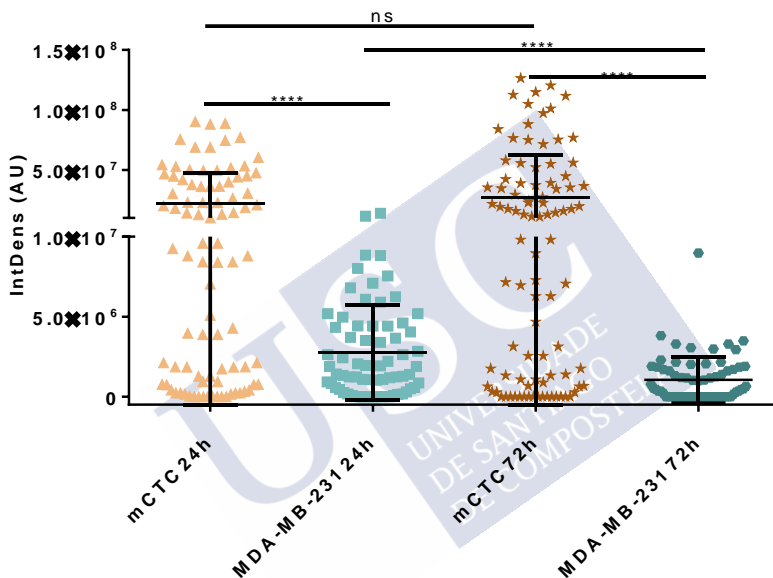
Figure 48. Studying the ability of mCTC to disseminate and survive in the zebrafish xenografts in comparison with the MDA-MB-231[eGFP Luc] cell line (n = 6 independent experiments). ns not significant, **** p < 0.0001.

Additionally, a third cell population, called '231 ULA' was included in the assay. This population corresponded to the MDA-MB-231[eGFP Luc] cell line maintained *in vitro* under the same conditions as the mCTC cell line, which are under low-attachment conditions and cultured with Mammocult[TM] media. In this way, the potential bias derived from the differential cell culture conditions between mCTC and MDA-MB-231[eGFP Luc] was eliminated. The mCTC cell line showed significantly higher dissemination to the tail than the '231 ULA'

population at 24 and 72 hpi, which is denoted by the higher fluorescence intensity of the ZF injected with the mCTC cell line. Besides, mCTC showed a sustained fluorescence intensity between 24 and 72 hpi, while '231 ULA' population showed a significant decrease in the fluorescence intensity at 72 hpi, when compared with 24 hpi. This means that disseminated mCTC cells survive better in the tail than the '231ULA' population (figure 49).
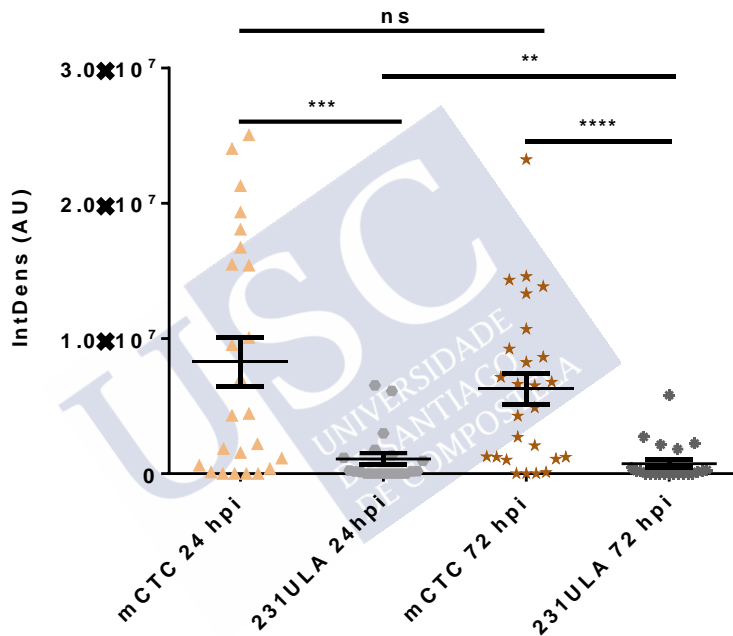
Figure 49. Comparing the dissemination of the mCTC cell line with the '231ULA' population in the zebrafish xenografts (n = 3 independent experiments). ns not significant, ** p < 0.01, *** p < 0.001, **** p < 0.0001.

Hence, the functional differences observed between mCTC and MDA-MB-231[eGFP Luc] cell lines in terms of *in vitro* and *in vivo* resistance to FSS cannot be explained by the differential conditions of the *in vitro* cell culture maintenance.

Secondly, the anchorage-independent growth capacity, as well as the self-maintenance and self-renewal abilities of the mCTC cell line were preliminarily studied by performing soft agar colony formation assays. It was observed that the mCTC cell line had a significantly higher ability to generate colonies *in vitro* than the parental MDA-MB-231$^{eGFP\ Luc}$ cell line, which based on the ≥ 70 µm cut-off diameter established for the quantification did not give rise to any colony (2.87 ± 0.67 vs. 0.0 ± 0.0). However, no significant differences were observed between the mCTC cell line, and the population '231 ULA' (3.37 $\pm$ 0.70). This can be explained by the use of supplemented Mammocult^{TM} media for both cell populations. Mammocult^{TM} specifically selects those cell clones with stem-like characteristics, such as self-maintenance and self-renewal capacities, which are the cells able to grow in an anchorage-independent manner and generate colonies (figure 50).
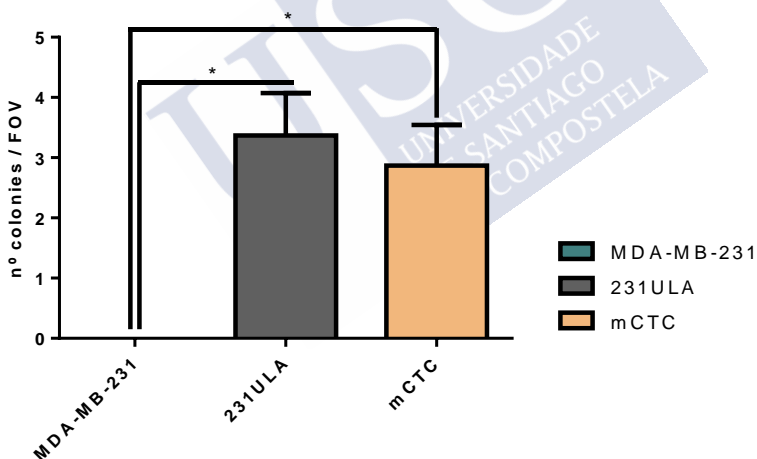
Figure 50. Assessing the ability of the mCTC cell line to form colonies *in vitro*. mCTC cell line was compared to the MDA-MB-231eGFP Luc cell line, as well as with the '231ULA' population (n = 2). * p < 0.05.

In addition, a preliminary molecular characterization of the mCTC cell line was performed. A transcriptomic analysis performed by RNAseq revealed 1696 genes differently expressed between mCTC and

'231ULA', with a fold change ≥2, and a p value < 0.05. More specifically, it has been found that 984 genes were significantly up-regulated in the '231ULA' cell line, while 712 genes were found to be significantly up-regulated in the mCTC cell line. These genes could potentially support the preliminary functional differences observed. However, further validation is required in order to test this hypothesis (figure 51).
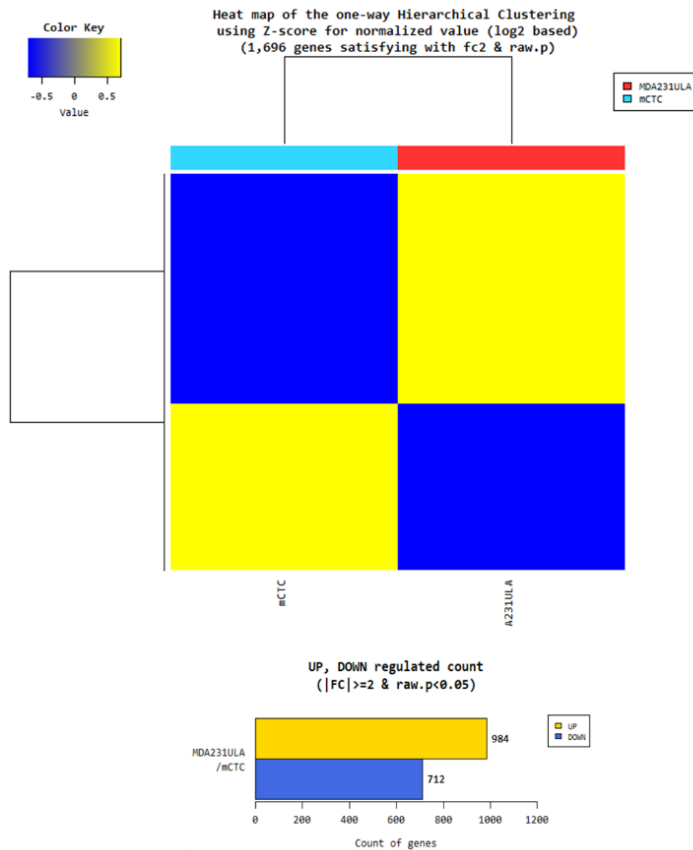
Figure 51. Transcriptomic analysis of the mCTC and the '231ULA' population by RNAseq. Genes were considered to be differentially expressed when the fold change (FC) ≥ 2, and a p value < 0.05.

RNAseq analysis evinced differences in the gene expression profiles of mCTC and 231ULA concerning different molecular functions. Hence, mCTC cell line had 2.99% of its up-regulated genes related to SMAD binding, while 231ULA only showed 0.09% of its up-regulated genes related to this molecular function. Receptor binding is another molecular function that shows differences between both cell lines. mCTC cell line showed a higher percentage of genes than 231ULA involved in this molecular function (15.86%, and 5.60%, respectively) (figure 52).
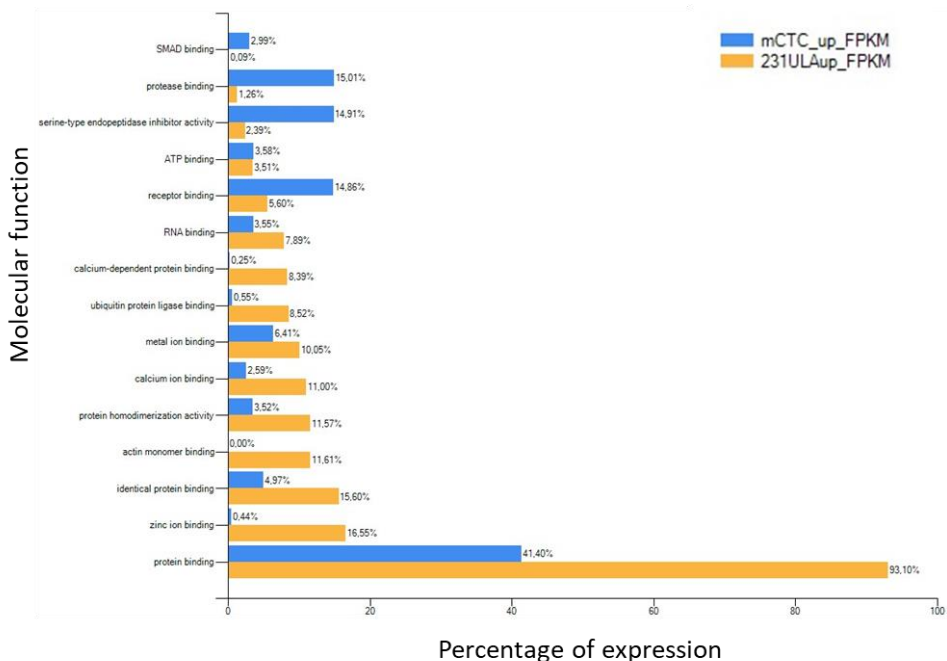
Figure 52. Molecular function of the differentially expressed genes between the mCTC cell line and the 231ULA population.
Image obtained with FunRich Functional Enrichment Analysis Tool.

These preliminary data showed that mCTC have functional differences compared to the *in vitro* CTC cluster model derived from the MDA-MB-231$^{eGFP\ Luc}$ cell line. mCTC could potentially be a CTC cluster model more realistic from a physiological point of view.

However, further functional, and molecular characterization should be performed to test this hypothesis.

3.2. Molecular analysis to study single CTC and CTC cluster genomic heterogeneity

Preclinical data indicate that CTC clusters seem to bear a higher metastatic potential than single CTCs, and clinical studies show that the presence of CTC clusters in the peripheral blood of BC patients is associated with a worse outcome. This data is in agreement with our own research work [130] (see '*CTC and CTC clusters as predictor factors of patient outcome*' in chapter 1). Thus, individual CTCs and CTC clusters could have specific genomic changes that can contribute to unveil their differential metastatic potential. Moreover, CTC clusters can be oligoclonal/polyclonal entities, meaning that CTC clusters may originate from different tumour cell clones that cooperate to generate polyclonal metastases in BC. Therefore, CTC clusters could be genetically more heterogeneous than single CTCs. Nevertheless, there is little knowledge regarding the genetic features that make CTCs and CTC clusters metastatic. To test if single CTCs and CTC clusters have differential genomic traits a proof-of-concept experiment was designed based on the development of genomic studies (DNAseq) at a single cell level. This work was developed in collaboration with the Phylogenomics research group of the Biomedical Research Centre (CINBIO), in the University of Vigo, headed by Prof. David Posada González. Prof. David Posada, an expert in molecular evolutionary biology and the analysis of NGS (Next-Generation Sequencing) data on cancer.

CTCs and CTC clusters were isolated from blood samples of murine orthotopic BC xenografts with the Parsortix^{TM} system. More specifically, single CTCs and CTC clusters were enriched from a pool of blood from three different tumour-bearing mice. After recovering the CTCs and CTC clusters isolated, the cells were individualised by DEPArray^{TM} system. The whole genome of individualised CTCs and CTC clusters was amplified, checked for its integrity, and sequenced

(WES, whole exome sequencing) at a single-cell and single-cluster level (figure 53 and 54).
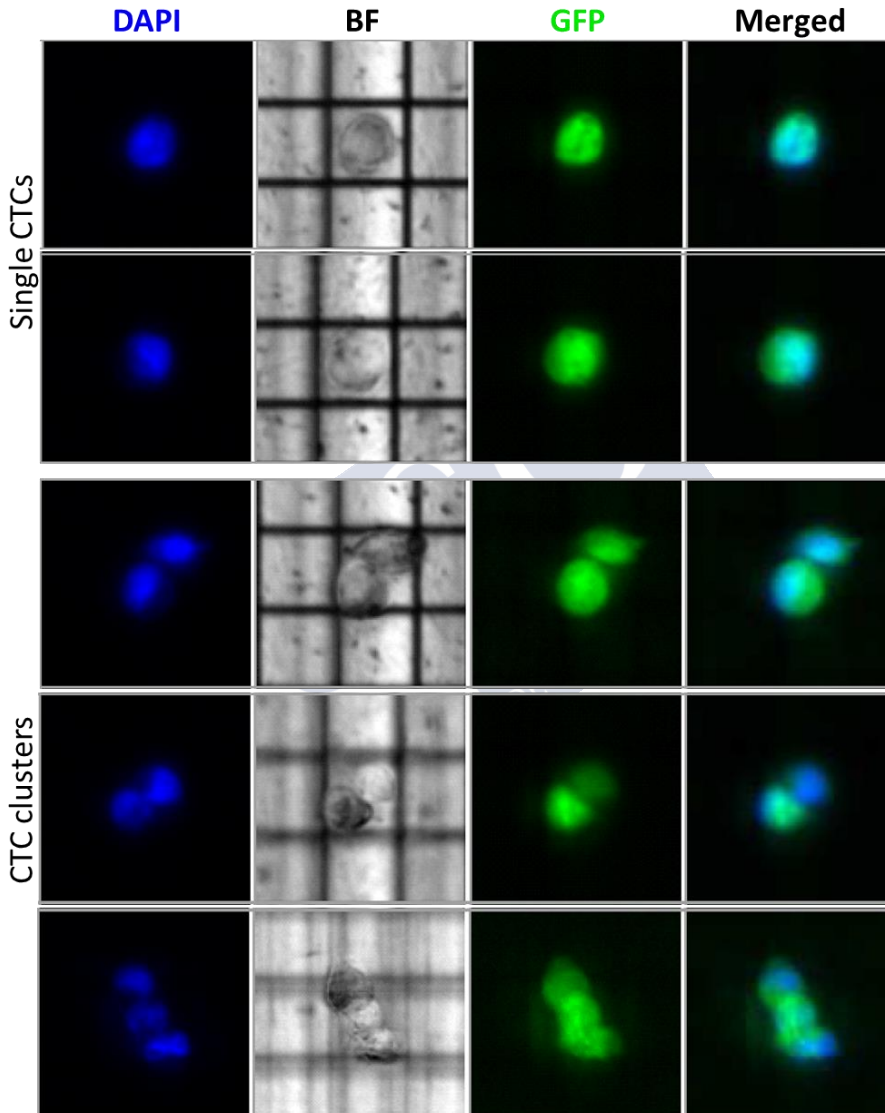
Figure 53. Representative examples of single CTCs and CTC clusters isolated from blood samples of tumour-bearing mice and individualised by the DEPArray™.

Single-cell sequencing was performed on 7 single CTCs and 7 CTC clusters as the initial biological material. However, 2 CTC clusters were discarded due to the low quality of the sequencing data (figure 54, red arrow heads). Hence, the final analysis was performed with 7 single CTCs and 5 CTC clusters.
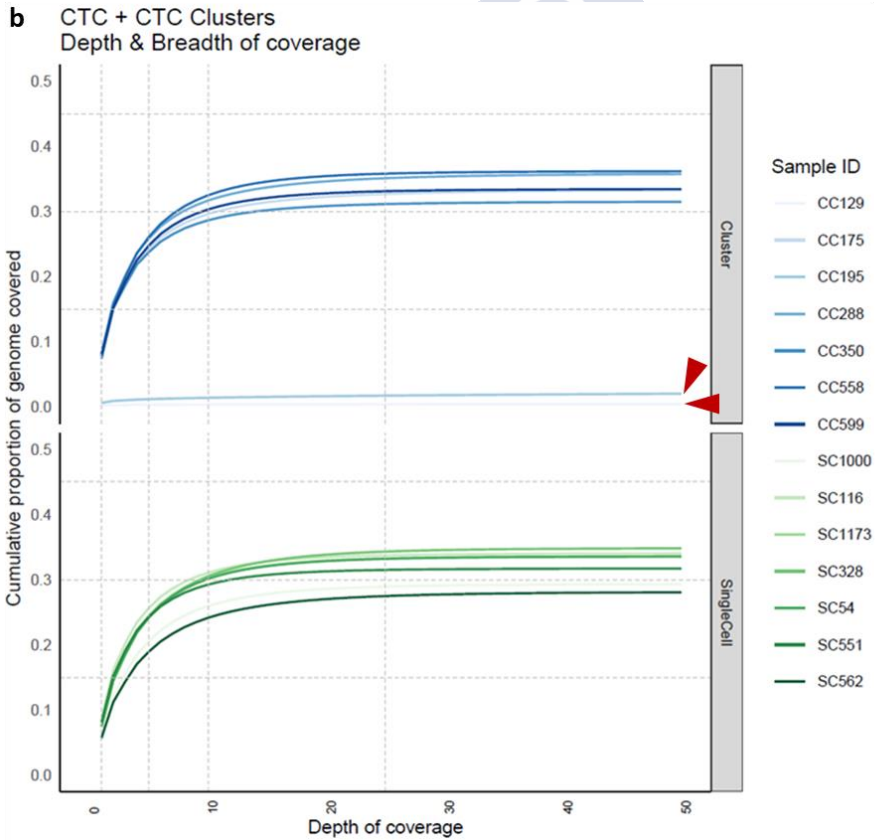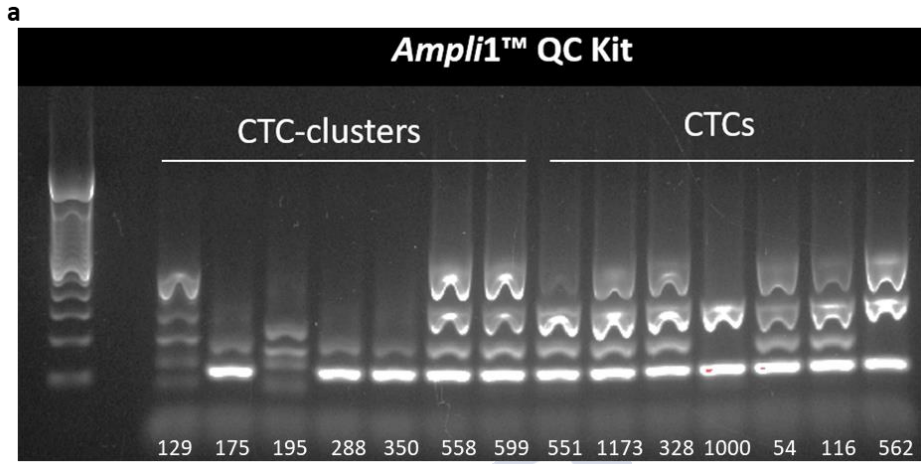
Figure 54. Evaluating the quality of the genetic material extracted from single CTCs and CTC cluster isolated from blood samples of tumour-bearing mice. Whole genome amplification (WGA) DNA quality control performed with the Ampli1 QC kit (a). DNA sequencing depth and breadth coverage of single CTCs and CTC clusters. Red arrow heads indicate low coverage of the genome in two CTC clusters (b).

Bioinformatic analysis of sequencing data showed that CTC clusters had a higher number of Single Nucleotide Variants (SNVs) than individual CTCs. As expected, some of the detected SNVs were common to single CTCs and CTC clusters (figure 55).
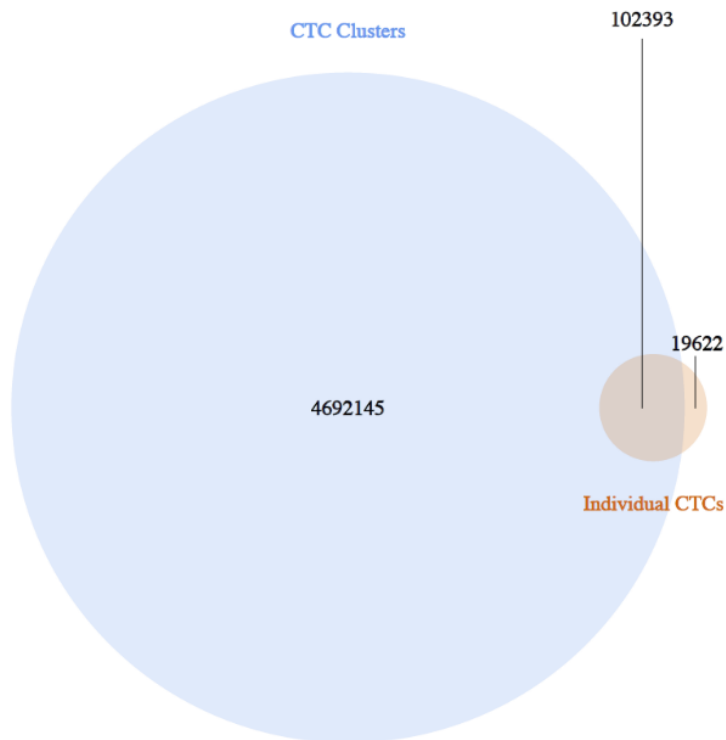
Figure 55. Venn diagram of SNV calls. Blue circle corresponds to the total number of SNVs identified in CTC clusters and the orange circle represents the number of SNVs detected in individual CTCs.

Moreover, SNVs in CTC clusters were classified as 'potential germline mutations', meaning probably present in the xenografted cell line, and 'potential somatic mutations', probably developed during tumour establishment in the mouse, based on the Variant Allele Fraction (VAF). Mutations with a VAF range between 0.6 and 1 were considered potential germline mutations, while SNVs with a VAF range between 0.05 and 0.6 were considered potential somatic mutations. The classification of the CTC cluster SNVs overlapping with the single CTC SNVs allowed us to determine that most of the SNVs in single CTCs might be potential germline mutations, showing a VAF ranging from 0.6-1 (figure 56).

a
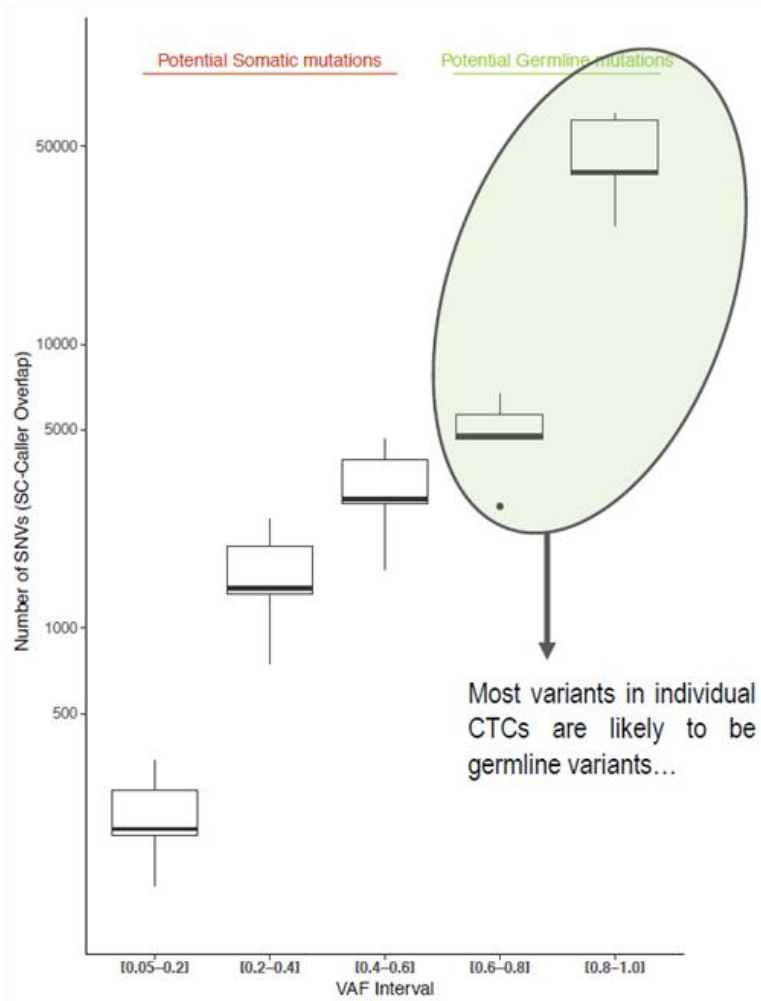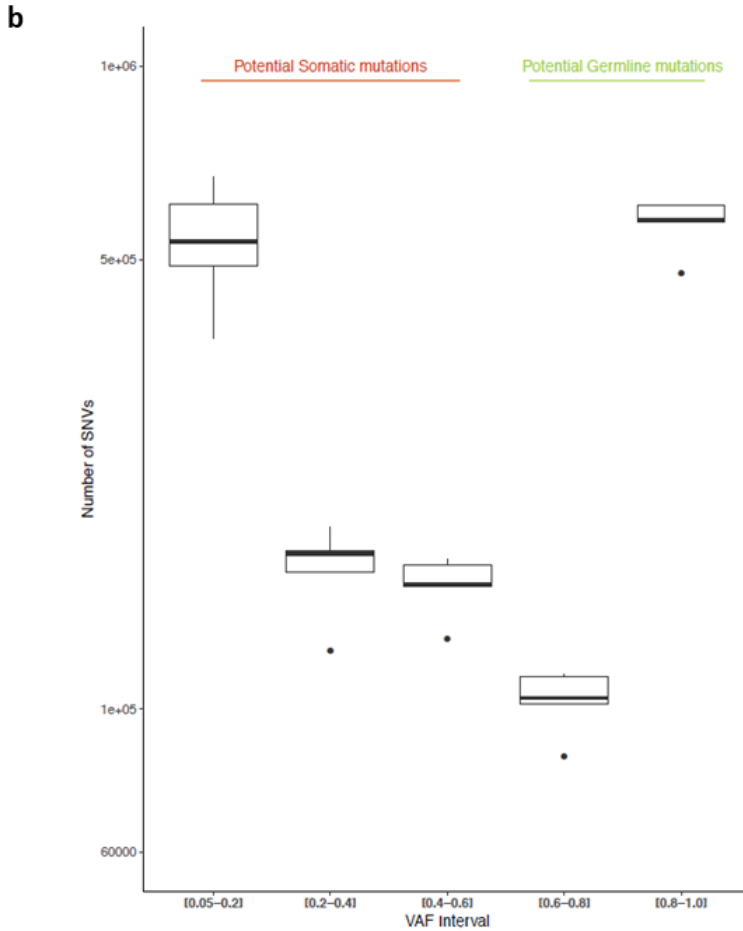
Figure 56. Number of SNVs (Single Nucleotide Variants) per VAF (Variant Allele Fraction). Number of SNVs per VAF in single CTCs. Note that most of the mutations in the single CTCs are 'potential germline mutations' (a). Number of SNVs per VAF in CTC clusters (b). Variants with a VAF ranging from 0.05-0.6 are classified as 'potential somatic mutations', while those with a VAF ranging from 0.6-1 are considered as 'potential germline mutations'.

Furthermore, VAFs of 'potential somatic mutations' were used to investigate the genetic variability present in each CTC cluster in comparison to the genetic variability estimated for all 7 single CTCs, based on the estimation of the expected heterozygosity (He). A higher He was found in each CTC cluster than in all single CTCs, meaning that CTC clusters have higher genetic variability than individual CTCs (figure 57).
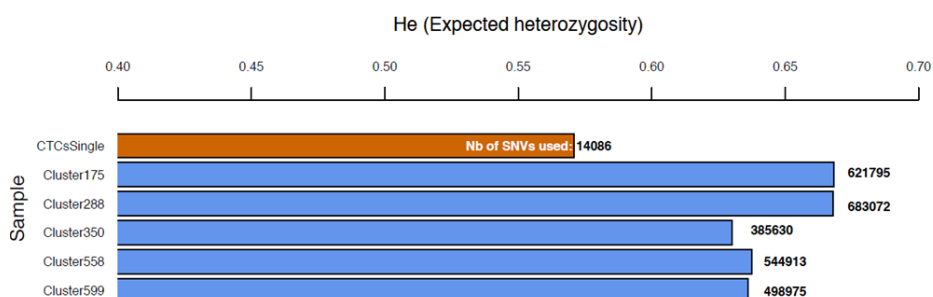
Figure 57. Expected Heterozygosity (He) of the single CTC population (orange bar), and of 5 different CTC clusters (blue bars). For each sample, He was estimated using the VAFs of 'potentially somatic mutations'.

Tile plots allowed us to visualise the mutational profile of each of the 7 single CTCs and the 5 CTC clusters (figure 58). We observed that some mutations were present in the majority of the individual CTCs analysed, e.g., AHNAK (11:62292882), or FUS (16:31195279). However, some mutations were present in a few of single CTCs, such as SVEP1 (9:113308525), or SPTAN1 (9:131375701), which were only in two out of the seven single CTCs analysed (figure 58, a). Similarly, CTC clusters showed some common mutations, e.g., TRIO, MGA, or SVEP1, while other SNV mutations were only observed in some of the CTC clusters analysed (3-4/5). However, a relevant number of mutations are only present in few CTC clusters (1-2/5) (figure 58, b).

In particular 595 mutations were exclusively found in CTC clusters, 4 in single CTCs and 9 mutations were common to both single CTCs and CTC clusters (figure 59. See also tables S1 and S2, in the *Supplementary Material* section).
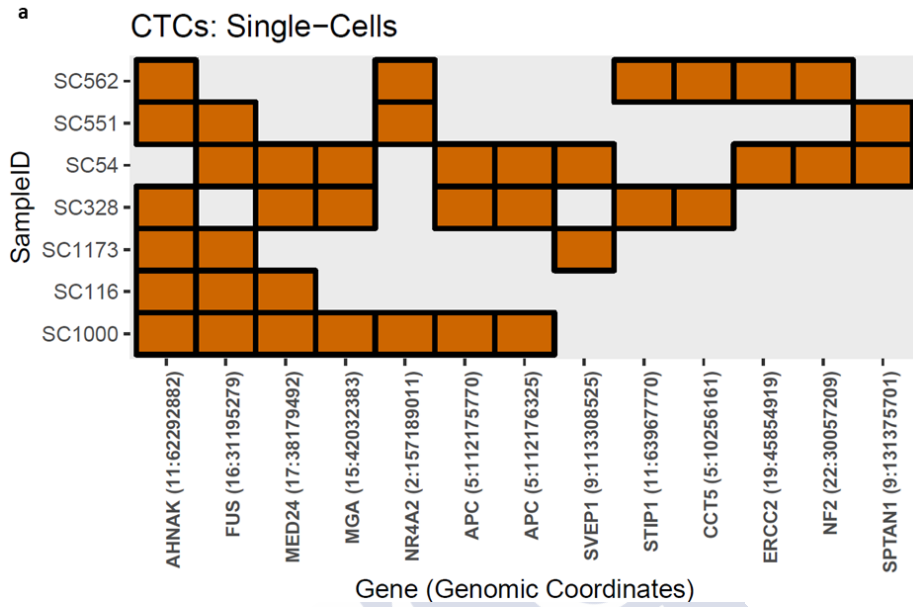
b CTCs: CTC−Clusters
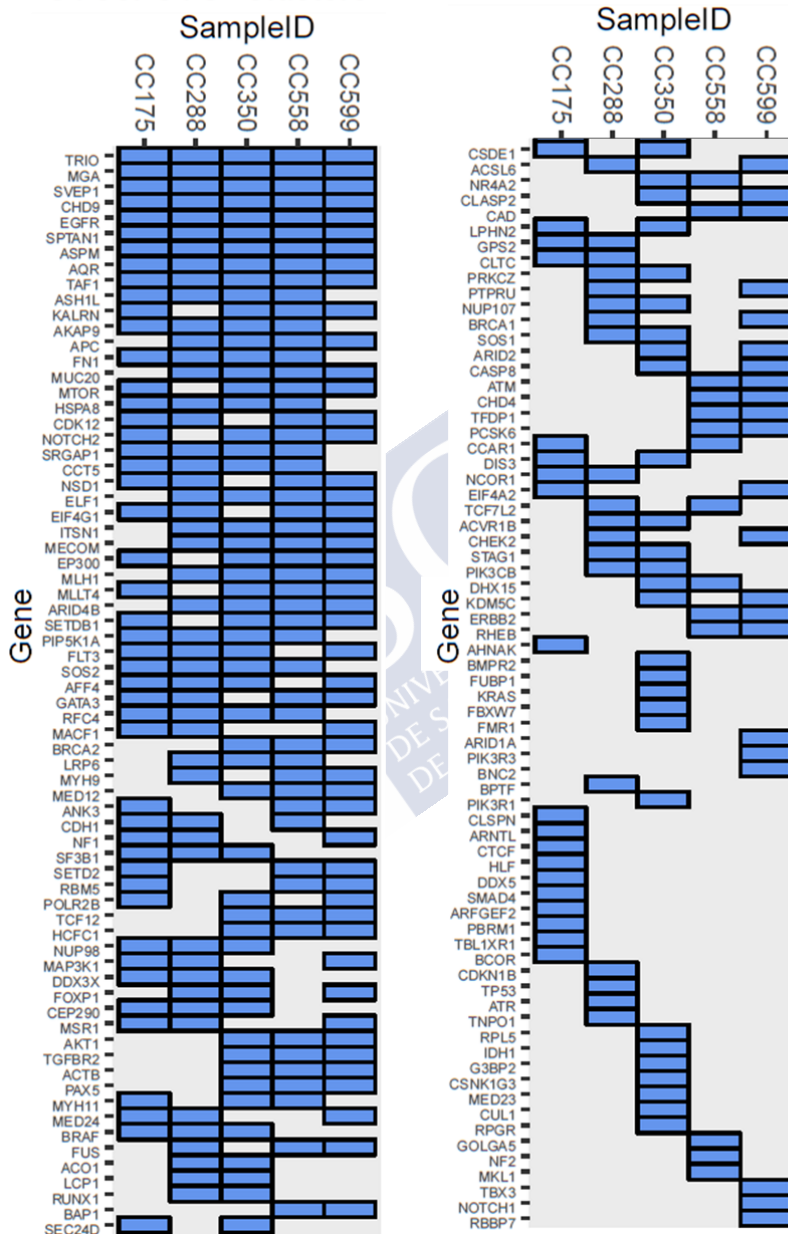
Figure 58. Exome-wide genomic profile of single CTCs and CTC clusters isolated from mice. Tile plot depicting absence/presence mutational profiles of seven CTCs isolated form mice. Somatic mutations sorted according to their overall frequency. Gene names are displayed at the bottom of the plot together with the corresponding genomic coordinates. Each row represents a single CTC (a). Tile plot depicting absence/presence mutational profile of five CTC clusters isolated from mice. Somatic mutations sorted according to their frequency across all clusters. Gene names are displayed at the left of the plot. Each column represents a single CTC cluster (b).
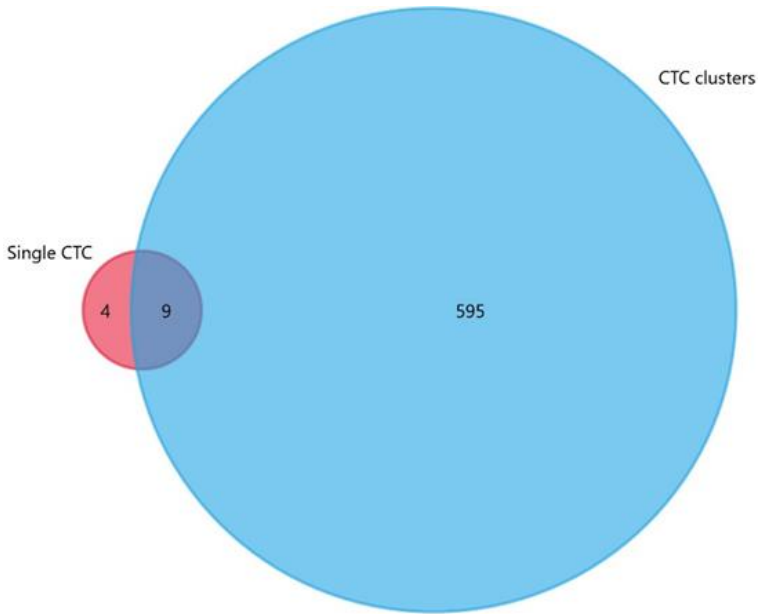
Figure 59. Venn diagram depicting common mutations to single CTCs and CTC clusters and mutations exclusively found in each cell group.

Therefore, preliminary results extracted from the genomic data of CTCs and CTC clusters isolated from tumour-bearing mice showed that CTC clusters not only have a higher genomic variability in comparison with single CTCs, but also a large set of specific mutations that were not present in single CTCs. According to this preliminary data, CTC clusters could potentially be genetically more complex entities than

single CTCs. However, further studies are required to unravel the genomic traits of the single CTCs and the CTCs clusters responsible for seeding metastasis.

DISCUSSION

DISCUSSION

CTC clusters can be detected in the peripheral blood of BC patients, among other cancer types [116,131–133]. Although CTC clusters are associated with a higher metastatic potential compared with individual CTCs [45,108,134], their role during metastasis is still in the process of being elucidated. Different mechanisms that partially explain the role of CTC clusters in tumour dissemination are now being discovered. The presence of strong cell-cell junctions within CTCs in the clusters, the expression of stem cell markers, or changes in the DNA methylation profile are some of the mechanisms proposed to explain their greater metastatic potential [45,109,117,131,135]. Given the increasing relevance of CTC clusters in tumour spread, it is not unexpected that CTC cluster enumeration by CellSearch® provides valuable additional clinical information of patient outcome, in comparison with CTC enumeration alone. It was previously reported that longitudinal study of CTC clusters added prognostic value to CTC enumeration alone in mBC patients, and that the presence of CTC clusters of larger size increased the risk of death [126–128,136]. However, the majority of the currently available data in this regard are generated from well-selected homogeneous cohorts of patients. On the contrary, we studied a small cohort of unselected mBC patients to evaluate the potential clinical value of CTC cluster enumeration. In our cohort, luminal BC subtypes (HR^+/$HER2^-$) were under-represented (50.9% in comparison with 70% in the general population), while TNBC was over-represented (28.3% of patients, compared with the 15% of the general population) [137]. Despite the reduced size of our cohort, some results that corroborated previous findings were observed. Thus, the presence of a high number of CTCs (≥5CTCs/7.5mL) correlated with a shorter OS at baseline. However, we did not observe a significant association between PFS and the previously established cut-off of ≥5 CTC/7.5 mL of blood [121] at baseline, neither during follow up. This can be explained by the heterogeneity of our cohort. Besides, CTC cluster enumeration at baseline showed that the presence of CTC

clusters correlated with a shorter PFS and OS. This corroborated the results observed in previous works, demonstrating the prognostic value of CTC clusters at baseline in mBC [126,128,136]. However, follow up data showed that CTC enumeration correlated with a worse prognosis, but CTC cluster detection did not correlate with clinical outcome, unlike the findings reported in previous studies [128,136,138]. This discrepancy can be explained by the heterogeneity included in our cohort of patients, as well as the limited number of patients with follow up samples (38/54). The joint analysis, in which patients were classified into risk groups according to the number of CTCs (<5 or ≥5 CTCs/7.5mL) and the presence or absence of CTC clusters (<1 or ≥1 CTCs/7.5mL) showed that CTC cluster enumeration at baseline constituted an independent prognostic factor predicting patient outcome. However, a study performed by Paoletti and colleagues, showed that the presence of CTC clusters at baseline in patients with high CTC counts (≥5 CTCs/7.5mL) is not a prognostic factor and that the prognostic value depends on the number of CTCs/7.5mL, suggesting that CTC clusters do not play a major role in the outcome of MBC patients starting first line chemotherapy [139]. Moreover, the sustained presence of CTC clusters over time was associated with shorter survival. Although our data are in agreement with previous reports [45,125], given the size of our cohort, this result should be evaluated in a larger cohort.

As a whole, and despite the size of the cohort studied, our data demonstrated that CTC cluster enumeration is an independent prognostic marker at baseline, that provides additional clinical information to CTC count alone. Hence, data obtained from our cohort of patients suggested that CTC cluster monitoring throughout time could potentially increase their prognostic value. However, this should be evaluated in a larger cohort of patients to confirm our observations, as we only studied the presence of CTC clusters at two different timepoints in a reduced number of patients.

Given the clinical relevance of CTC cluster analysis, maximising CTC cluster detection and capture is essential to

better understand tumour progression. However, most clinical data regarding the enumeration of CTC clusters are generated using the CellSearch® system for CTC detection. This is because CellSearch® is the only CTC isolation system approved by the FDA for clinical use in breast, prostate, and colorectal cancer [140]. Despite its demonstrated clinical value, CellSearch® isolates CTCs and CTC clusters based on the use of magnetic beads functionalized with antibodies against the epithelial marker EpCAM. Hence, it exclusively isolates the epithelial-like CTCs but it might miss the mesenchymal-like subset of CTCs, which have a lower or no expression of EpCAM [78]. Due to this limitation, several alternative CTC separation systems were developed. The majority of these CTC capture systems are based on the differential physical properties of CTCs and CTC clusters, in comparison with blood cells [78,140]. Parsortix™ is a microfluidic device that isolates CTCs based on their larger size and their lower deformability than blood cells [97,98]. Therefore, Parsortix™ can potentially detect not only the epithelial-like CTCs but also the CTCs with a more mesenchymal-like (or stem-like) phenotypic traits. Unlike the CellSearch®, Parsortix™ also allows the recovery of the isolated CTCs in a viable state for downstream analysis, such as CTC *in vitro* culture or transcriptomic analysis [78,141]. Moreover, the Parsortix™ system might allow the capture of a more heterogeneous CTC population, which would provide a more complete landscape of the metastatic process. The optimisation of the CTC enrichment parameters during CTC isolation in the Parsortix™ system, and the development of a live cell staining allowed us to capture and identify both, individual CTCs and CTC clusters in peripheral blood samples from mBC patients. Our optimisation was specifically aimed to potentially increase the likelihood of CTC cluster identification.

Despite the clinical value of CTC enumeration in peripheral blood samples for predicting patient outcome, the study of CTCs at a molecular level can potentially provide additional and valuable information about the tumour status. Currently, the study of the tumour genomic mutations is performed on tumour biopsy (or solid biopsy). In this context, the molecular analysis of CTCs and CTC clusters, as tumour surrogates, is a non-invasive

alternative to assess tumour progression, or response to therapy [142]. However, the low frequency of CTCs and especially CTC clusters present in 7.5 mL of blood is one of their main limitations of liquid biopsy, as it narrows the variety of downstream molecular analysis that can be performed, restricting the improvement of the knowledge of the biology CTCs and CTC clusters [92,94]. To overcome this limitation, alternative pre-enrichment steps, such as Diagnostic Leukapheresis (DLA) are being used for CTC isolation. DLA samples come from the continuous centrifugation of large volumes of blood, and it was originally used to isolate MNCs as a therapy for leukaemia. Epithelial cells have a similar density to MNCs, thus CTCs can be co-isolated by this procedure [92–94]. Therefore, DLA samples can partially solve the problem of low CTC frequency of the peripheral blood samples (7.5 mL of blood). It was estimated that the use of DLA products for CTC detection in the CellSearch® system generated a 0-32 increased yield of CTC detection in comparison with 7.5 mL of peripheral blood [95]. In collaboration with the group of Translational Research in Gynaecology of the Heinrich-Heine University (Düsseldorf, Germany), we used DLA products from mBC patients to test the feasibility of isolating CTC clusters. The optimisation of a workflow for DLA sample processing with the Parsortix™ using spiking tests not only allowed us to detect individual CTCs but also to isolate CTC clusters. These results showed the great potential of DLA products for the isolation of CTC clusters. Although additional analysis should be performed, our data suggest that DLA can be a useful tool to increase CTC cluster isolation in BC. In this sense, it would be recommendable to compare the ability to detect CTC clusters in DLA products and peripheral blood samples when the cellularity of both samples is equivalent.

CTC clusters seem to have a higher metastatic potential than single CTCs. Studies performed in animal models showed that CTC clusters are the main responsible for tumour spread and metastasis formation [23,45,104,108]. Hence, unravelling the phenotypic traits that make CTC clusters more metastatic is essential for a better understanding of tumour dissemination. However, the insufficient number of CTC clusters isolated from

peripheral blood samples, the technical limitations existing regarding the identification and isolation of CTCs, and the challenges of *ex vivo* CTC culture created the need of developing alternative models of CTCs and CTC clusters to allow the in-depth study of their metastatic potential, as well as their role within the metastatic spread of BC.

Consequently, we worked on the development of experimental CTC cluster models directly derived from two well-characterised BC cell lines to shed light on their metastatic traits using a set of different *in vitro* and *in vivo* experimental approaches. On the one hand, the *in vitro* characterisation intended to mimic different steps of the metastatic cascade, as well as to evaluate different phenotypic characteristics linked to the metastatic potential. On the other hand, *in vivo* characterisation was performed to assess the metastatic competence of our CTC models in a more realistic environment, such as a living organism.

Transwell migration assays mimic the cell migration phenomena required in the initial steps of the metastatic cascade. Migration assays showed that the CTC clusters had a significantly higher ability to migrate in response to FBS gradient than individual CTCs in both, MDA-MB-231$^{eGFP\ Luc}$ and MCF-7$^{eGFP\ Luc}$ derived models. These results might initially seem counterintuitive as a higher ability to migrate would be expected in the individual CTC population based on their lower physical restrictions, which would allow it to go through the pores of the transwell membrane more easily. However, it was reported that CTC clusters can approach each other and aggregate generating bigger clusters, which increases their velocity of migration [143]. This could explain our observations regarding the differential migration potential of CTC clusters. Additionally, invasion assays mimic the initial steps of the metastatic cascade, in which malignant cells should break the basement membrane, and migrate to invade other adjacent tissues during local invasion, or even, during early steps of intravasation. The MDA-MB-231$^{eGFP\ Luc}$ CTC cluster model had a significantly higher capacity to invade an ECM *in vitro* than the single CTC model. However,

MCF-7$^{eGFP\ Luc}$ CTC models did not show differences in the invasion ability between single CTCs and CTC clusters, probably due to the low levels of invaded cells detected for both models. The negligible invasion capacity detected was not unexpected given the epithelial nature of the MCF-7$^{eGFP\ Luc}$ cells, characterised by a low migration and invasion ability [144].

During intravasation and extravasation, tumour cells should interact and adhere to the endothelium. The *in vitro* endothelial adhesion assays showed that CTC clusters adhered significantly less to the endothelial monolayer than single CTCs. These results were observed in both, MDA-MB-231$^{eGFP\ Luc}$ and MCF-7$^{eGFP\ Luc}$ CTC cluster models. However, the static conditions of this assay should be taken into consideration, as did not favour the generation of transient weak interactions between tumour cells and the endothelial monolayer. These interactions are considered an essential step for tumour transmigration throughout the endothelial layer during extravasation, promoting cell rolling, similar to the cell rolling of WBCs [51,145]. Hence, the lack of a dynamic environment promoting interactions between both cell types and the absence of tumour cell rolling could potentially explain the lower adhesion ability of CTC clusters observed. *In vivo* studies provide dynamic conditions for studying the interactions between CTCs and endothelial cells. Thus, studies in ZF embryos demonstrated that CTC clusters could potentially have a higher ability to adhere to endothelium due to a physical arrest in the blood vessels of the ZF [146]. Hence, further experimental approaches should be performed to test the endothelial adhesion capacity and extravasation ability of our CTC cluster models.

Besides, the self-maintenance and self-renewal capacities of the CTC models were studied by soft agar colony formation assays, as these are traits that would favour successful metastatic seeding. The higher ability of colony formation observed in the CTC cluster models of MDA-MB-231$^{eGFP\ Luc}$ and MCF-7^{eGFP} cells compared to individual CTC models goes in line with previous publications [147] and could be explained by the higher resistance to apoptosis of clustered cells. The survival during this

step of the metastasis is driven by key genetic programs [56,108,148,149]. For instance, a work performed by Zhang and colleagues demonstrated that cell aggregation induced the tyrosine-phosphorylation of PECAM-1 and Pyk2, which promoted anchorage-independent growth and resistance to anoikis [147].

ZF experiments with MDA-MB-231[eGFP Luc] cell line showed higher dissemination of single CTCs than CTC clusters at 24 hpi. The dissemination of CTC clusters can be physically restricted by their larger size, which may explain their lower frequency in circulation. This was supported by the different dissemination patterns observed in single CTC and CTC cluster ZF xenografts. Almost all of the ZF injected with single CTCs showed the majority of cells arrested in the ventral area of the caudal region (area of caudal vein and dorsal aorta). This could be explained by the reduced flow in this area, which would allow the interaction between tumour cells and the endothelial cells of the vessel [146]. Arrested cells were less frequently found at the dorsal (dorsal longitudinal anastomotic vessel) and lateral (intersegmental vessels). Likewise, ZF injected with CTC clusters showed disseminated cells preferentially arrested in the ventral region. However, their presence at dorsal and lateral locations was occasionally found and at a lower frequency than in single CTC ZF xenografts. CTC clusters reached dorsal and lateral locations probably due to their ability to reorganise their shape generating a chain-like cluster able to transverse capillary blood vessels, as was previously reported [150]. However, the lower percentage of ZF and the lower frequency of CTC cluster disseminated in the dorsal and lateral regions compared to those injected with single CTCs, confirmed that CTC cluster dissemination is physically restricted by the size of the cluster, which favours their physical occlusion into the capillaries of low diameter. Hence, this would result in fast CTC cluster entrapment and a corresponding lower circulation half-life than individual CTCs, as it was previously published [45].

Despite the lower dissemination observed in the ZF injected with CTC clusters than those injected with single CTCs, most of

the ZF injected with single CTCs showed a reduced fluorescence intensity at 72 hpi compared to 24 hpi, indicating that a great proportion of the disseminated cells did not survive in the caudal region. However, ZF xenografts injected with CTC clusters showed a sustained fluorescence throughout the time, having similar fluorescence intensities at both timepoints, 24 hpi and 72 hpi, suggesting that CTC clusters survive better into the bloodstream of the ZF embryo than single CTCs. The survival advantage of CTC clusters exposed to the environmental stress of the circulation was also observed *in vitro* by the FSS assays. Survival into the bloodstream is a key step during tumour dissemination for developing successful metastases. The mechanical stress experienced by CTCs into the bloodstream due to shear forces is thought to be one of the main causes of tumour cell death during hematogenous dissemination, along with immune system response, and loss of ECM adhesion [45,151]. Consequently, the significantly lower mortality and higher proliferation observed in the MDA-MB-231[eGFP Luc] CTC cluster model when exposed to FSS *in vitro* agreed with the observations performed in the ZF individual analysis. The ZF analysis showed a significant proportion of ZF injected with CTC clusters in which increased the fluorescence intensity throughout time, alike ZF injected with single CTCs. These results indicate that CTC clusters not only have a greater survival but also a proliferative advantage over single CTCs in the ZF. The higher FSS resistance detected in the MDA-MB-231[eGFP Luc] CTC clusters could be explained by the presence of strong cell-cell junctions, which prevent them from *anoikis* [107,152]. Moreover, some tumour cells within the clusters may be less exposed to the FSS, which could reduce apoptosis related to mechanical stress and contribute to cell survival under these conditions [153]. Furthermore, it was reported that CTC clusters can modify their shape by selective cleavage of cell junctions to go through capillary vessels [150], which can also contribute to limiting the negative impact of FSS and explain the survival advantage observed in the CTC cluster population.

Furthermore, gene expression analyses showed a differential expression profile between disseminated cells from single CTC

and CTC cluster ZF xenografts that supported the results obtained from functional analyses. Disseminated cells of the CTC cluster ZF xenografts had a lower expression of the pro-apoptotic gene BAX and an up-regulation of the pro-survival gene PLAU [154,155], when compared with the disseminated cells of the ZF xenografts injected with single CTCs. Urokinase-type plasminogen activator gene (PLAU) encodes a proteinase involved in the hydrolyzation and remodelling of the ECM, as well as in the activation of growth factors, which promotes cell proliferation, migration, and invasion [156]. Thus, PLAU up-regulation was previously described to be associated with tumour progression and with a poorer outcome in BC, among other cancer types [157–160]. Nevertheless, it was also shown that urokinase activity has anti-metastatic effects due to the dissociation of CTC clusters [161]. Although further studies are required to unravel the role of PLAU gene in tumour progression, our results point towards a positive role of the urokinase activity during metastasis. This observation denoted that CTC clusters could have a survival advantage when injected into the circulation of the ZF embryos.

In addition, disseminated cells from CTC cluster ZF xenografts had a higher expression of stem related genes (CD44, ITGA6) and cell cycle regulation genes (CDK4, E2F4), in comparison with the disseminated cells of single CTC xenografts. CD44 is a common marker of stem cells in BC that regulates cell adhesion, migration, survival, invasion and EMT [162]. The up-regulation of CD44 in the CTC cluster population is in line with previous reports, demonstrating that CD44 induces tumour cell clustering through the formation of homophylic CD44-CD44 interactions, thus promoting tumorigenesis and polyclonal metastases formation [109]. ITGA6 (also known as CD49f) is another gene related to BC stem cells that encodes an integral membrane protein involved in the interaction with the ECM [163]. ITGA6 is also associated with cell invasion and tumour initiation capacity [163–165]. Moreover, high expression levels of ITGA6 were reported to correlate with shorter PFS and OS in BC patients [166]. Furthermore, the development of murine PDXs using samples derived from ER-negative BC patients allowed the identification of a CD44$^+$/CD49f ^{high}/CD133/2 ^{high} cell subset with higher

tumour initiating capacity [167]. Importantly, CTC clusters from mBC patients had a higher stem-like profile than individual CTCs [109,117]. In fact, it was demonstrated that CD49f-based isolation improves CTC detection in different BC subtypes [168].

The analysis of the fluorescence variation over time performed in each ZF xenograft individually allowed us to observe an increase in the fluorescent intensity over time in the disseminated cells of ZF injected with CTC clusters, probably due to the growth of the disseminated cells. This observation agreed with the gene expression observed in the cell cycle related genes CDK4 and E2F4. Cyclin-dependent-kinase 4 gene (CDK4) correlates with a worse OS and PFS in BC, and it is usually overexpressed in TNBC [169]. Treatment with CDK4/6 inhibitors block proliferation and foster apoptosis [169]. Hence, the higher expression of CDK4 observed in the CTC cluster ZF xenograft could be promoting cell cycle progression and cell proliferation. Besides, CDK4 seems to regulate the expression of stem cell surface markers, such as CD44 and CD24 in TNBC. CDK4 inhibitors reduced the proportion of the CD44$^+$/CD24$^-$ subset of cells and prevented cancer cells from self-renewal [170]. Moreover, E2F4 was also up-regulated in the CTC cluster ZF xenografts. E2F4 was described as both, tumour suppressor [171,172], and as an oncogene [173]. Furthermore, it was highlighted that the E2F4 activity, rather than E2F4 expression, is a negative prognostic factor for BC survival [174]. Although our results of gene expression did not allow us to assess E2F4 activity, they support the role of E2F4 as a factor promoting proliferation in CTC clusters.

Taken together, these data suggest an enhanced survival capacity of CTC clusters during circulation in the bloodstream derived from cell aggregation, in concordance with previous studies [45,56,108,124]. Moreover, these results are supported by clinical observations in CTCs where the incidence of apoptosis among CTCs within the clusters is much lower than the observed in single CTCs isolated from the peripheral blood of patients with BC and small-cell lung cancer [116,131]. The comparison with the *in vivo* ZF assays and the clinical data currently available cannot be performed since we did not quantify the incidence of apoptosis in

our models. However, it brings to light the relevance of this animal model in oncology, especially to study tumour progression and metastasis formation in a short period of time.

Although our data showed the great potential of ZF as an animal model to study tumour progression, murine models of cancer continue to be the gold standard for *in vivo* modelling of tumour spread. For this reason, we also performed an *in vivo* functional characterization of our CTC cluster model by using the mouse as an animal model.

The injection of MDA-MB-231[eGFP Luc] single CTC and CTC cluster models into the lateral tail vein allowed to evaluate their survival capacity within the bloodstream, as well as their colonization and tumour-initiating abilities. This experimental approach showed that CTC clusters have a higher ability to survive in the lungs, and colonize a new anatomical location than single CTCs, giving rise to larger metastatic lesions. Despite the initial massive cell death observed in both populations, the monitoring of the tumour bioluminescent signal in the lungs over time allowed us to observe that CTC clusters showed a milder decrease in the bioluminescence than single CTC xenografts. This suggests that CTC clusters suffered a lower incidence of cell death than single CTCs during the early steps of lung colonization. These data reproduced a former work in which CTC clusters injected into the bloodstream of mice were more resistant to apoptosis during colonization of distant sites than single CTCs [45]. This advantage of CTC clusters allowed them to grow faster generating greater metastases than single CTCs and was in line with the results derived from the ZF xenografts.

The murine orthotopic BC model was utilised as a potential source of CTC and CTC-cluster obtaining. We were able to detect and isolate CTCs as both, single CTCs and clusters. Additionally, we estimated that CTC clusters represented approximately 12.7% of all the detected events. Although this may be an underestimation of the real number of CTCs and CTC clusters, it was reported that CTC clusters are in a minority within the CTC population, representing 1-30 % of all CTCs [104]. Hence, this orthotopic model of BC recapitulated human BC progression

properly, and it is a suitable model for functional and molecular studying the role of CTCs and CTC clusters in tumour spread. CTCs and CTC cluster quantification allowed us to observe a direct correlation between CTCs and CTC clusters. It implies that the more single CTCs are detected, the more CTC clusters are expected to be found, as it was also observed in the clinical setting.

In collaboration with the Nano-oncology research line of the Roche-CHUS Joint Unit, we established the mCTC cell line, which grows *in vitro* forming cellular aggregates in suspension resembling CTC clusters. Thus, the mCTC cell line could potentially be a surrogate model of CTC clusters, as culturing and characterization of CTCs may be a suitable tool to perform drug screening or to study epigenomic changes in the expression profile of CTCs [117,175]. *In vitro* mCTC characterisation allowed us to observe that mCTC had a significantly higher resistance to FSS than the CTC cluster model directly derived from MDA-MB-231[eGFP Luc]. Moreover, it was observed that mCTC cell line had a higher colony formation ability than the parental cell line MDA-MB-231[eGFP Luc]. However, these preliminary results could be explained by the differential culture conditions of both cell lines. Consequently, we included a third cell population, called '231ULA', derived from the culture of the MDA-MB-231[eGFP Luc] parental cell line under the same conditions as mCTC. The 231 ULA cells showed a similar capacity to give rise to colonies than mCTC cells. Nevertheless, *in vivo* ZF xenografts showed significantly higher dissemination and survival ability of the mCTC cell line than both, parental MDA-MB-231[eGFP Luc] and 231ULA. Besides, RNAseq data also showed potential differences in the expression profile between mCTC and the 231ULA. Interestingly, among the differently expressed gene the ADTRP gene (androgen dependent TFPI regulating protein) was found to be overexpressed in the mCTC cell line (FC: 16.55). This fatty acid metabolism related gene was also found to be overexpressed in a RNAseq previously performed in the laboratory (data not published) in which we compared four paired cell lines derived from the primary tumour (PT) and pulmonary metastases (L) of the BC orthotopic mouse model. Specifically, it

was observed that ADTRP was overexpressed in the metastasis-derived cell lines (see figure S1, in the *Supplementary Material* section). Thus, ADTRP could potentially be a suitable candidate gene for later studies. However, further analyses are required to validate these results and to check these observations. Although further studies are required, preliminary characterisation of mCTC suggested that this CTC-derived cell line may have a higher metastatic potential than the *in vitro* CTC cluster models generated from MDA-MB-231$^{\text{eGFP Luc}}$. These observations could derive from the origin of this cell line, where blood shear stress could have selected the more resistant clones to this environment, thus selecting those clones with higher metastatic potential and with a greater probability of seeding a metastatic lesion.

Altogether, preliminary results indicate that mCTC cells may represent a more physiologically realistic model of CTC clusters and therefore a useful tool to study the biology of CTC clusters more accurately. Evaluating the colonization ability by performing a lung colonization assay in mice, assessing the tumorigenic capacity of the mCTC cell line *in vivo*, or studying drug resistance would be interesting experimental approaches to fully test the functional behaviour of the mCTC cells. Besides, studying the mutational heterogeneity of the mCTC may show the molecular traits underlying the functional observations. In this way, we would have a complete landscape at both, functional and molecular level, of the mCTC cells.

Additionally, CTCs and CTC clusters isolated from tumour-bearing mice were genomically studied, in collaboration with the Phylogenomics research group of the Biomedical Research Centre (CINBIO), at the University of Vigo. The individual genomic sequencing of 7 CTCs and 5 CTC clusters showed that CTC clusters carried a large set of mutations that were not present in single CTCs, which suggests a greater genomic variability than single CTCs. According to this, we hypothesise that CTC clusters may be genetically more heterogeneous than single CTCs, probably due to the presence of multiple tumour cell clones within the cluster, and they might bear specific genomic changes that may increase the chance of succeeding in metastatic colonization.

Proving this hypothesis may have important implications in our understanding of BC progression, particularly in the context of the polyclonal metastases seeding observed in BC [45,108,176,177]. Nevertheless, additional studies are required to decipher the genomic characteristics of those CTCs and CTC clusters responsible for seeding metastases, as well as to determine the contribution to metastases formation.

Taken together, *in vitro* and *in vivo* assays permitted to functionally characterise the single CTC and CTC cluster models. They also allowed us to observe differences between the *in vitro* single CTC and CTC cluster models, showing a higher metastatic potential associated with CTC clusters. This is in line with the current knowledge regarding the differential contribution of CTC clusters to the metastatic process. Hence, the CTC cluster model could potentially be a suitable surrogate for studying the biology of CTC clusters and their role within metastasis. Besides, the mCTC cell line is a promising tool for a more realistic study of the biology of CTC clusters. However, further studies are required to confirm its suitability as a potential CTC cluster surrogate. Finally, preliminary genomic analyses of CTCs and CTC clusters showed genomic differences between single CTC and CTC clusters, demonstrating that genomic profiling of CTCs can be of relevance to understand the metastatic seeding. As a whole, the models developed in this project contribute to overcoming the limitation of the low frequency of CTC clusters in the blood of BC patients and are useful tools to study the role of CTC clusters within the tumour spread.

CONCLUSIONS

CONCLUSIONS

1. CTC cluster enumeration constitutes an independent prognostic factor that adds clinical value to CTC enumeration alone at baseline. Moreover, the sustained presence of CTC clusters in the peripheral blood samples is associated with shorter patient survival, hence correlating with a worse disease outcome.

2. It was possible to isolate and detect individual CTCs and CTC clusters from peripheral blood samples derived from metastatic breast cancer patients by using an immune-independent CTC separation workflow. This allowed the recovery of CTCs in a viable state that can potentially increase the variety of downstream analyses.

3. Preliminary optimisation of a workflow with DLA products showed that it was feasible to isolate CTC clusters Hence, DLA products can increase the likelihood of CTC and CTC cluster isolation.

4. A CTC cluster model from MDA-MB-231$^{eGFP\ Luc}$ TNBC cell line was developed. This model showed differential functional traits *in vitro* and *in vivo* in comparison with the individual CTC model, such as higher migration, invasion, colony formation, resistance to fluid shear stress, greater survival and proliferation within the ZF circulation and higher ability to colonize the lungs of mice.

5. The significantly higher expression of CD44, ITGA6, CDK4, E2F4, PLAU, and the significantly lower expression of BAX gene points to a greater survival and a proliferative advantage of CTC clusters *in vivo*, in comparison with individual CTCs.

6. Preliminary characterization showed that the mCTC cell line has functional *in vitro* and *in vivo* differences that suggest its higher metastatic potential. Although further analyses are needed,

mCTC cell line could potentially be a more realistic CTC cluster model, from a physiological point of view.

7. Preliminary genomic analyses showed that CTC clusters have a higher genetic heterogeneity than single CTCs, as well as a set of specific mutations. This heterogeneity can be explained by the possible oligoclonal/polyclonal origin of CTC clusters. However, further studies are required to test this hypothesis.

8. Together, the experimental approaches and the CTC models developed in this thesis provide alternative biological material for studying the biology of CTC clusters and their role in the metastatic process. These tools can help overcome the limitation of the lower number of CTC clusters isolated from patient samples.

SUPPLEMENTARY MATERIAL
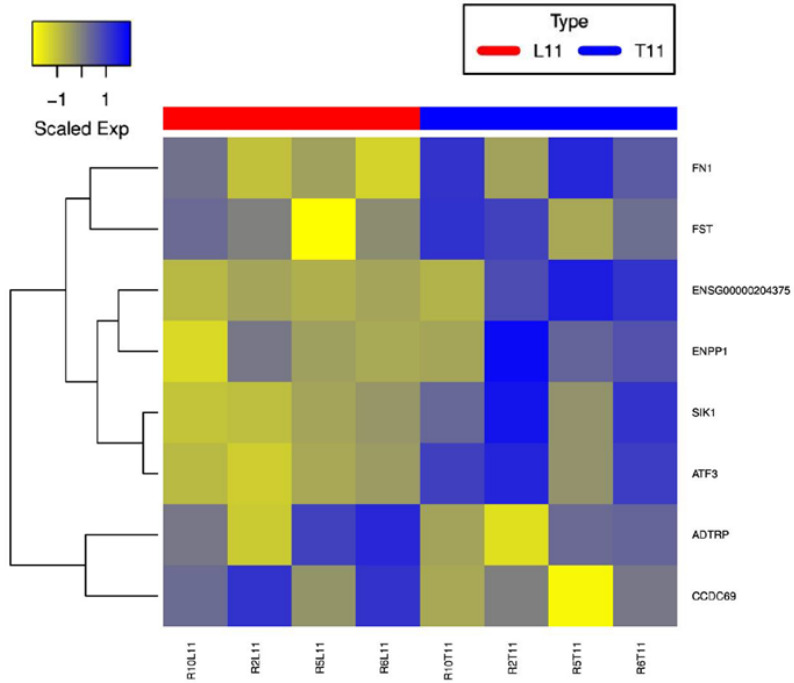
Figure S1. Differently expressed genes determined by RNAseq between the primary tumour (PT) and lung metastasis (L) paired cell lines derived from four different mice orthotopic BC models. R10T11, R2T11, R5T11 and R6T11 correspond to PT samples and R10L11, R2L11, R5L11, R6L11 are the L samples.

Table S1. Mutations found in CTC clusters.

Gene symbol	Description	Map location
MTOR	mechanistic target of rapamycin (serine/threonine kinase)	1p36.2
CLSPN	claspin	1p34.2
MACF1	microtubule-actin crosslinking factor 1	1p32-p31
CSDE1	cold shock domain containing E1, RNA-binding	1p22
NOTCH 2	notch 2	1p13-p11
SETDB1	SET domain, bifurcated 1	1q21
PIP5K1A	phosphatidylinositol-4-phosphate 5-kinase, type I, alpha	1q21.3
ASH1L	ash1 (absent, small, or homeotic)-like (Drosophila)	1q22
ASPM	asp (abnormal spindle) homolog, microcephaly associated (Drosophila)	1q31
GATA3	GATA binding protein 3	10p15
ANK3	ankyrin 3, node of Ranvier (ankyrin G)	10q21
CCAR1	cell division cycle and apoptosis regulator 1	10q21.3
NUP98	nucleoporin 98kDa	11p15.5
ARNTL	aryl hydrocarbon receptor nuclear translocator-like	11p15
AHNAK	AHNAK nucleoprotein	11q12.2
HSPA8	heat shock 70kDa protein 8	11q24.1
SRGAP1	SLIT-ROBO Rho GTPase activating protein 1	12q14.2
CEP290	centrosomal protein 290kDa	12q21.32
FLT3	fms-related tyrosine kinase 3	13q12
DIS3	DIS3 homolog, exosome endoribonuclease and 3'-5' exoribonuclease	13q22.1
SOS2	son of sevenless homolog 2 (Drosophila)	14q21
AQR	aquarius intron-binding spliceosomal factor	15q14
MGA	MGA, MAX dimerization protein	15q14
MYH11	myosin, heavy chain 11, smooth muscle	16p13.11
CHD9	chromodomain helicase DNA binding protein 9	16q12.2
CTCF	CCCTC-binding factor (zinc finger protein)	16q21-q22.3
CDH1	cadherin 1, type 1	16q22.1
GPS2	G protein pathway suppressor 2	17p13
NCOR1	nuclear receptor corepressor 1	17p11.2

NF1	neurofibromin 1	17q11.2
CDK12	cyclin-dependent kinase 12	17q12
MED24	mediator complex subunit 24	17q21.1
HLF	hepatic leukemia factor	17q22
CLTC	clathrin, heavy chain (Hc)	17q23.1
DDX5	DEAD (Asp-Glu-Ala-Asp) box helicase 5	17q21
SMAD4	SMAD family member 4	18q21.1
SF3B1	splicing factor 3b, subunit 1, 155kDa	2q33.1
FN1	fibronectin 1	2q34
ARFGEF2	ADP-ribosylation factor guanine nucleotide-exchange factor 2 (brefeldin A-inhibited)	20q13.13
EP300	E1A binding protein p300	22q13.2
SETD2	SET domain containing 2	3p21.31
RBM5	RNA binding motif protein 5	3p21.3
PBRM1	polybromo 1	3p21
KALRN	kalirin, RhoGEF kinase	3q21.2
TBL1XR1	transducin (beta)-like 1 X-linked receptor 1	3q26.32
EIF4G1	eukaryotic translation initiation factor 4 gamma, 1	3q27.1
EIF4A2	eukaryotic translation initiation factor 4A2	3q28
RFC4	replication factor C (activator 1) 4, 37kDa	3q27
POLR2B	polymerase (RNA) II (DNA directed) polypeptide B, 140kDa	4q12
SEC24D	SEC24 family member D	4q26
CCT5	chaperonin containing TCP1, subunit 5 (epsilon)	5p15.2
TRIO	trio Rho guanine nucleotide exchange factor	5p15.2
MAP3K1	mitogen-activated protein kinase kinase kinase 1, E3 ubiquitin protein ligase	5q11.2
AFF4	AF4/FMR2 family, member 4	5q31
NSD1	nuclear receptor binding SET domain protein 1	5q35
MLLT4	myeloid/lymphoid or mixed-lineage leukemia; translocated to, 4	6q27
EGFR	epidermal growth factor receptor	7p12
AKAP9	A kinase (PRKA) anchor protein 9	7q21-q22
BRAF	B-Raf proto-oncogene, serine/threonine kinase	7q34
MSR1	macrophage scavenger receptor 1	8p22
SVEP1	sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1	9q32
SPTAN1	spectrin, alpha, non-erythrocytic 1	9q34.11

BCOR	BCL6 corepressor	Xp11.4
DDX3X	DEAD (Asp-Glu-Ala-Asp) box helicase 3, X-linked	Xp11.3-p11.23
TAF1	TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa	Xq13.1
PRKCZ	protein kinase C, zeta	1p36.33-p36.2
PTPRU	protein tyrosine phosphatase, receptor type, U	1p35.3
ARID4B	AT rich interactive domain 4B (RBP1-like)	1q42.3
TCF7L2	transcription factor 7-like 2 (T-cell specific, HMG-box)	10q25.3
LRP6	low density lipoprotein receptor-related protein 6	12p13.2
CDKN1B	cyclin-dependent kinase inhibitor 1B (p27, Kip1)	12p13.1-p12
ACVR1B	activin A receptor, type IB	12q13
NUP107	nucleoporin 107kDa	12q15
ELF1	E74-like factor 1 (ets domain transcription factor)	13q13
LCP1	lymphocyte cytosolic protein 1 (L-plastin)	13q14.3
FUS	FUS RNA binding protein	16p11.2
TP53	tumor protein p53	17p13.1
BRCA1	breast cancer 1, early onset	17q21
BPTF	bromodomain PHD finger transcription factor	17q24.3
SOS1	son of sevenless homolog 1 (Drosophila)	2p21
ITSN1	intersectin 1 (SH3 domain protein)	21q22.1-q22.2
RUNX1	runt-related transcription factor 1	21q22.3
CHEK2	checkpoint kinase 2	22q12.1
MYH9	myosin, heavy chain 9, non-muscle	22q13.1
MLH1	mutL homolog 1	3p21.3
FOXP1	forkhead box P1	3p14.1
STAG1	stromal antigen 1	3q22.3
PIK3CB	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit beta	3q22.3
ATR	ATR serine/threonine kinase	3q23
MECOM	MDS1 and EVI1 complex locus	3q26.2
MUC20	mucin 20, cell surface associated	3q29
TNPO1	transportin 1	5q13.2
APC	adenomatous polyposis coli	5q21-q22
ACSL6	acyl-CoA synthetase long-chain family member 6	5q31.1
ACO1	aconitase 1, soluble	9p21.1

FUBP1	far upstream element (FUSE) binding protein 1	1p31.1
RPL5	ribosomal protein L5	1p22.1
KRAS	Kirsten rat sarcoma viral oncogene homolog	12p12.1
ARID2	AT rich interactive domain 2 (ARID, RFX-like)	12q12
BRCA2	breast cancer 2, early onset	13q12.3
AKT1	v-akt murine thymoma viral oncogene homolog 1	14q32.32
TCF12	transcription factor 12	15q21
NR4A2	nuclear receptor subfamily 4, group A, member 2	2q22-q23
CASP8	caspase 8, apoptosis-related cysteine peptidase	2q33-q34
BMPR2	bone morphogenetic protein receptor, type II (serine/threonine kinase)	2q33-q34
IDH1	isocitrate dehydrogenase 1 (NADP+), soluble	2q33.3
TGFBR2	transforming growth factor, beta receptor II (70/80kDa)	3p22
CLASP2	cytoplasmic linker associated protein 2	3p22.3
DHX15	DEAH (Asp-Glu-Ala-His) box helicase 15	4p15.3
G3BP2	GTPase activating protein (SH3 domain) binding protein 2	4q21.1
FBXW7	F-box and WD repeat domain containing 7, E3 ubiquitin protein ligase	4q31.3
PIK3R1	phosphoinositide-3-kinase, regulatory subunit 1 (alpha)	5q13.1
CSNK1G 3	casein kinase 1, gamma 3	5q23
MED23	mediator complex subunit 23	6q22.33-q24.1
ACTB	actin, beta	7p22
CUL1	cullin 1	7q36.1
PAX5	paired box 5	9p13
RPGR	retinitis pigmentosa GTPase regulator	Xp21.1
KDM5C	lysine (K)-specific demethylase 5C	Xp11.22-p11.21
MED12	mediator complex subunit 12	Xq13
FMR1	fragile X mental retardation 1	Xq27.3
HCFC1	host cell factor C1	Xq28
ATM	ATM serine/threonine kinase	11q22-q23
CHD4	chromodomain helicase DNA binding protein 4	12p13
TFDP1	transcription factor Dp-1	13q34
GOLGA 5	golgin A5	14q32.12

PCSK6	proprotein convertase subtilisin/kexin type 6	15q26.3
ERBB2	erb-b2 receptor tyrosine kinase 2	17q12
CAD	carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase	2p22-p21
NF2	neurofibromin 2 (merlin)	22q12.2
MKL1	megakaryoblastic leukemia (translocation) 1	22q13
BAP1	BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase)	3p21.1
RHEB	Ras homolog enriched in brain	7q36
ARID1A	AT rich interactive domain 1A (SWI-like)	1p35.3
PIK3R3	phosphoinositide-3-kinase, regulatory subunit 3 (gamma)	1p34.1
TBX3	T-box 3	12q24.21
BNC2	basonuclin 2	9p22.2
NOTCH 1	notch 1	9q34.3
RBBP7	retinoblastoma binding protein 7	Xp22.2

Table S2. Mutations found in single CTCs.

Gene symbol	Description	Map location
AHNAK	AHNAK nucleoprotein	11q12.2
MGA	MGA, MAX dimerization protein	15q14
FUS	FUS RNA binding protein	16p11.2
MED24	mediator complex subunit 24	17q21.1
NR4A2	nuclear receptor subfamily 4, group A, member 2	2q22-q23
APC	adenomatous polyposis coli	5q21-q22
SVEP1	sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1	9q32
STIP1	stress-induced phosphoprotein 1	11q13
CCT5	chaperonin containing TCP1, subunit 5 (epsilon)	5p15.2
ERCC2	excision repair cross-complementation group 2	19q13.3
NF2	neurofibromin 2 (merlin)	22q12.2
SPTAN1	spectrin, alpha, non-erythrocytic 1	9q34.11

APPENDIX

APPENDIX

LIST OF PUBLICATIONS:

- Piñeiro R., Martínez-Pena I., López-López R. (2020) Relevance of CTC Clusters in Breast Cancer Metastasis. In: Piñeiro R. (eds) Circulating Tumor Cells in Breast Cancer Metastatic Disease. Advances in Experimental Medicine and Biology, vol 1220. Springer, Cham. https://doi.org/10.1007/978-3-030-35805-1_7

Used with the permission of Springer Nature Switzerland AG, who gave the authors the right of using the content for educational purposes, among other uses.
Contribution to this work: I, Inés Martínez Pena, was involved in writing and reviewing the manuscript, as well as in the elaboration of the scientific drawing of this chapter.

- Costa, C.; Muinelo-Romay, L.; Cebey-López, V.; Pereira-Veiga, T.; Martínez-Pena, I.; Abreu, M.; Abalo, A.; Lago-Lestón, R.M.; Abuín, C.; Palacios, P.; Cueva, J.; Piñeiro, R.; López-López, R. Analysis of a Real-World Cohort of Metastatic Breast Cancer Patients Shows Circulating Tumor Cell Clusters (CTC-clusters) as Predictors of Patient Outcomes. Cancers 2020, 12, 1111. https://doi.org/10.3390/cancers12051111

© Cancers is an open access journal with Creative Commons Attribution Licence which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal Rank: Q1 (Oncology)
Impact Factor (IF): 6.126
5-year IF: 6.433

Contribution to this work: I, Inés Martínez Pena, was involved in data analysis and reviewing of the manuscript

FAVOURABLE REPORT (CLINICAL RESEARCH ETHICS COMMITTEE OF GALICIA, CEIC)

XUNTA DE GALICIA
CONSELLERÍA DE SANIDADE
Secretaría Xeral Técnica

Edificio Administrativo San Lázaro
15703 SANTIAGO DE COMPOSTELA
Teléfono: 881546425
ceic@sergas.es
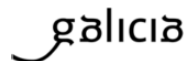
galicia

DICTAMEN DEL COMITE AUTONOMICO DE ETICA DE LA INVESTIGACION DE GALICIA

Paula M. López Vázquez, Secretaria del Comité Autonómico de Ética de la Investigación de Galicia

CERTIFICA:

Que este Comité evaluó en su reunión del día 22/12/15 :

Título: Biopsia líquida para oncología de precisión
Promotor: Rafael López López
Tipo de estudio: EPA-SP
Version: versión 30 de noviembre de 2015 y HIP/CI (cultivo, control biopsia, y biopsia líquida) de la misma fecha
Código del Promotor: RLL-BL-2015_01
Código de Registro: 2015/772

Y, tomando en consideración las siguientes cuestiones:
- La pertinencia del estudio, teniendo en cuenta el conocimiento disponible, así como los requisitos legales aplicables, y en particular la Ley 14/2007, de investigación biomédica, el Real Decreto 1716/2011, de 18 de noviembre, por el que se establecen los requisitos básicos de autorización y funcionamiento de los biobancos con fines de investigación biomédica y del tratamiento de las muestras biológicas de origen humano, y se regula el funcionamiento y organización del Registro Nacional de Biobancos para investigación biomédica, la ORDEN SAS/3470/2009, de 16 de diciembre, por la que se publican las Directrices sobre estudios Posautorización de Tipo Observacional para medicamentos de uso humano, y la Circular nº 07 / 2004, investigaciones clínicas con productos sanitarios.
- La idoneidad del protocolo en relación con los objetivos del estudio, justificación de los riesgos y molestias previsibles para el sujeto, así como los beneficios esperados.
- Los principios éticos de la Declaración de Helsinki vigente.
- Los Procedimientos Normalizados de Trabajo del CEIC de Galicia.

Emite un **INFORME FAVORABLE** para la realización del estudio por el/la investigador/a del centro:

Centros	Investigadores Principales
C.H. Universitario de Santiago	Clotilde Costa Nogueira

XUNTA DE GALICIA
CONSELLERÍA DE SANIDADE
Secretaría Xeral Técnica

Edificio Administrativo San Lázaro
15703 SANTIAGO DE COMPOSTELA
Teléfono: 881546425
ceic@sergas.es

galicia

Y HACE CONSTAR QUE:

1   El CAEIG cumple los requisitos legales vigentes (R.D 223/2004 por el que se regulan los ensayos clínicos con medicamentos, y la Ley 14/2007 de Investigación Biomédica).
2   El CAEIG tanto en su composición como en sus PNTs cumple las Normas de Buena Práctica Clínica (CPMP/ICH/135/95).
3   La composición actual del CAEIG es:

Manuel Portela Romero. (Presidente). Médico Especialista en Medicina Familiar y Comunitaria.
Irene Zarra Ferro. (Vicepresidenta). Farmacéutica de Atención Especializada.
Paula Mª López Vázquez, (Secretaria). Médico Especialista en Farmacología Clínica.
Juan Vázquez Lago (Secretario Suplente). Médico Especialista en Medicina Preventiva y Salud Pública.
Jesús Alberdi Sudupe. Médico especialista en Psiquiatría.
Rosendo Bugarín González. Médico Especialista en Medicina Familiar y Comunitaria.
Juan Casariego Rosón. Médico Especialista en Cardiología.
Xoán X. Casas Rodríguez. Médico Especialista en Medicina Familiar y Comunitaria.
Juana Mª Cruz del Río. Trabajadora Social.
Juan Fernando Cueva Bañuelos. Médico Especialista en Oncología Médica.
José Álvaro Fernández Rial. Médico Especialista en Medicina Interna.
José Luis Fernández Trisac. Médico Especialista en Pediatría.
Mª José Ferreira Díaz. Diplomada Universitaria de Enfermería
Pablo Nimo Ríos. Licenciado en Derecho. Miembro externo
Pilar Gayoso Diz. Médico Especialista en Medicina Familiar y Comunitaria.
Agustín Pía Morandeira. Farmacéutico de Atención Primaria
Salvador Pita Fernández. Médico Especialista en Medicina Familiar y Comunitaria.
Carmen Rodríguez-Tenreiro Sánchez. Licenciada en Farmacia.
Susana María Romero Yuste. Médico Especialista en Reumatología.
Mª Asunción Verdejo González. Médico Especialista en Farmacología Clínica.

En Santiago de Compostela, a 05 de enero de 2015

MATERIAL TRANSFER AGREEMENT FOR THE USE OF DIAGNOSTIC LEUKAPHERESIS PRODUCTS

MATERIAL TRANSFER AGREEMENT

THIS MATERIAL TRANSFER AGREEMENT ("Agreement"), shall commence on the date of the last signature (the "Effective Date") is entered into between:

By one side,
HEINRICH-HEINE-UNIVERSITY DUESSELDORF, acting on its behalf: University Hospital of Duesseldorf ("UHD"), an organization with a principal business address at Moorenstrasse 5, D-40225 Düsseldorf, (Germany) (hereinafter referred to as "**UHD**").
Executing entity:HANS NEUBAUER researcher of the Department of Obstetrics and Gynecology at University Hospital of Duesseldorf with address at Moorenstraße 5, 40225 Düsseldorf, (Germany), (hereinafter referred to as "**PROVIDER SCIENTIST**").

Hereinafter referred to as "**PROVIDER**"

By the other side,
FUNDACIÓN INSTITUTO DE INVESTIGACIÓN SANITARIA DE SANTIAGO DE COMPOSTELA, an organization with a principal business address at Hospital Clínico Universitario, C/ Choupana s/n, 15706 Santiago de Compostela (Galicia, Spain), (hereinafter referred to as "**FIDIS**")
and
ROBERTO PIÑEIRO CID, PHD, researcher of the Roche-Chus Joint Unit of the Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS) at Complexo Hospitalario Universitario de Santiago, with address at Travesía A Choupana, s/n, 15706 Santiago de Compostela (Galicia, Spain) (hereinafter referred to as "**RECIPIENT SCIENTIST**").
Both hereinafter referred to as "**RECIPIENT**".

MATERIAL (Description of Material to be transferred by PROVIDER to RECIPIENT; pls state in detail):

Frozen products of Diagnostic Leukapheresis obtained from breast cancer patients
(Patients have consented that material can be shipped to co-operation partners and that they can be used in research projects)

The MATERIAL includes the original MATERIAL, Progeny and Unmodified Derivatives. The MATERIAL shall not include: (a) Modifications, or (b) other substances created by the RECIPIENT through the use of the MATERIAL, which are not Modifications, Progeny or Unmodified Derivatives. As used in this Agreement, "Progeny" means an unmodified descendant from the MATERIAL, such as virus from virus, cell from cell or organism from organism; "Unmodified Derivatives" means substances created by the RECIPIENT which constitute an unmodified functional subunit or product expressed by the MATERIAL (some examples include: subclones of

unmodified cell lines, purified or fractionated subsets of the MATERIAL, proteins expressed by DNA/RNA supplied by the PROVIDER, or monoclonal antibodies secreted by a hybridoma cell line); and "Modifications" mean substances created by the RECIPIENT which contain/incorporate MATERIAL.

RESEARCH: The research to be conducted by the RECIPIENT SCIENTIST using the MATERIAL is described in Appendix A attached hereto (the "RESEARCH").

TERMS AND CONDITIONS

1. **Ownership/Use**. The MATERIAL is the property of PROVIDER. Nothing contained in this Agreement shall restrict, modify or limit any ownership rights of the PROVIDER. The MATERIAL is to be used by RECIPIENT solely for the RESEARCH to be conducted by the RECIPIENT SCIENTIST. The RECIPIENT and the RECIPIENT SCIENTIST agree that the MATERIAL will not be used in human subjects, in clinical trials, or for diagnostic purposes involving human subjects without the written consent of the PROVIDER; is to be used only at the RECIPIENT organization and only in the RECIPIENT SCIENTIST'S laboratory under the direction of the RECIPIENT SCIENTIST or others working under RECIPIENT SCIENTIST'S direct supervision; will not be used for any commercial purpose or in connection with any commercially-sponsored research and will not be transferred to anyone else (including anyone else within the RECIPIENT'S organization) without the prior written consent of the PROVIDER; will not be sequenced or otherwise analyzed in order to determine its structure or composition. RECIPIENT and RECIPIENT SCIENTIST agree to use the MATERIAL in compliance with all applicable statutes and regulations.

2. **Confidentiality**. The MATERIAL as well as all related information provided to RECIPIENT or RECIPIENT SCIENTIST shall constitute the confidential and/or proprietary information of PROVIDER (the "Confidential Information"). RECIPIENT and RECIPIENT SCIENTIST agree to maintain the confidentiality of the Confidential Information and to possess and use the Information solely for the purposes set forth in this Agreement. As used in this Agreement, "Confidential Information" shall not mean or include any information which: (a) was known to RECIPIENT or RECIPIENT SCIENTIST prior to the receipt of the MATERIAL or the Information from PROVIDER; (b) becomes known to the public without any breach of this Agreement by RECIPIENT or RECIPIENT SCIENTIST; (c) is acquired from a third party without breach of any obligation of confidentiality; (d) is developed independently by RECIPIENT or RECIPIENT SCIENTIST without reference to or reliance on any Confidential Information; or (e) is required to be disclosed pursuant to subpoena or other judicial or administrative order, or pursuant to applicable law. The obligations of confidentiality contained in this Section shall continue, as to any item of Confidential Information for a period of five (5) years from the date of receipt of such Confidential Information and shall survive the termination or expiration of this Agreement for any reason.

3. **Availability**. RECIPIENT agrees to refer to PROVIDER any requests for the MATERIAL from other scientists. To the extent supplies are available, RECIPIENT

agrees to make the MATERIAL available under a separate agreement to other scientists for teaching or not-for-profit research purposes only.

4. **Report**. RECIPIENT and RECIPIENT SCIENTIST agree to submit to PROVIDER a written report describing the data generated in the course of the RESEARCH (the "DATA") and setting for the results of the RESEARCH (together with the DATA, the "RESULTS") within six (6) months after the expiration or the termination of this Agreement, whichever is the first to occur. PROVIDER is hereby granted a non-exclusive, royalty-free, perpetual right and license to use the RESULTS for internal academic and research purposes only.

5. **Publication**. In case of joint research, research results shall be published jointly and in mutual consent. If the RESEARCH conducted solely by the RECIPIENT SCIENTIST results in publication, the RECIPIENT SCIENTIST agrees to provide appropriate acknowledgment of the source of the MATERIAL and/or to give credit to PROVIDER or PROVIDER SCIENTIST, as scientifically appropriate, for any contribution which PROVIDER or PROVIDER SCIENTIST may make to the subject matter of the publication.
In the case of a joint publication Prof. Dr. H. Neubauer and André Franke shall be listed as co-authors.
The RECIPIENT SCIENTIST shall give due consideration to the concerns of the PROVIDER when making scientific publications that relate to the subject matter of the Agreement and, as such, will send the text of the intended publication to PROVIDER before it is published to give the PROVIDER an opportunity to review and comment on it in order to avoid having any applications for IP rights jeopardised by any prior publications that would compromise the novelty thereof or reveal any trade secrets. No later than four weeks after submitting the intended publication to the PROVIDER, and to the extent the PROVIDER has not expressed any objection thereto, the RECIPIENT may publish the findings in accordance with sentence 2. Should the PROVIDER raise any objections, the Parties shall endeavour to arrive at an amicable solution to enable publication.

6. **No Warranty**. The MATERIAL is understood to be experimental in nature and may have hazardous properties. The PROVIDER MAKES NO REPRESENTATIONS AND EXTENDS NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, WITH RESPECT TO THE MATERIAL. THERE ARE NO EXPRESS OR IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE, OR THAT THE USE OF THE MATERIAL WILL NOT INFRINGE ANY PATENT, COPYRIGHT, TRADEMARK, OR OTHER PROPRIETARY RIGHTS.

7. **Liability**. Except to the extent prohibited by law, the RECIPIENT assumes all liability for damages, which may arise from its use, storage or disposal of the MATERIAL. The PROVIDER will not be liable to the RECIPIENT for any loss, claim or demand made by the RECIPIENT, or made against the RECIPIENT by any other party, due to or arising from the use of the MATERIAL by the RECIPIENT.

256

8. **Inventions**. Any patentable invention which relates to new uses of the MATERIAL or which could not have been made but for the contribution of the MATERIAL (an "Invention"), will be jointly owned by PROVIDER and RECIPIENT. Any revenues arising from any use or implementation of such Invention will be shared by the PROVIDER and the RECIPIENT, with their respective shares to be negotiated in good faith and based on the relative contribution made by the MATERIAL to the Invention.

9. **Rights and Licenses**. The RECIPIENT acknowledges that the MATERIAL is or may be the subject of a patent application. Except as provided in this Agreement, no express or implied licenses or other rights are provided to the RECIPIENT under any patents, patent applications, trade secrets or other proprietary rights of the PROVIDER, including any altered forms of the MATERIAL made by the PROVIDER. In particular, no express or implied licenses or other rights are provided to use the MATERIAL, MODIFICATIONS, or any related patents of the PROVIDER for COMMERCIAL PURPOSES. If the RECIPIENT desires to use or license the MATERIAL or MODIFICATIONS for COMMERCIAL PURPOSES, the RECIPIENT agrees, in advance of such use, to negotiate in good faith with the PROVIDER to establish the terms of a commercial license. It is understood by the RECIPIENT that the PROVIDER shall have no obligation to grant such a license to the RECIPIENT, and may grant exclusive or non-exclusive commercial licenses to others, or sell or assign all or part of the rights in the MATERIAL to any third party(ies), subject to any pre-existing rights held by others.
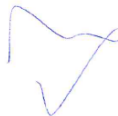
10. **Costs**. The MATERIAL is provided at no cost (if no amount is stated below) or for the optional transmittal fee stated below solely to reimburse the PROVIDER for preparation and distribution costs. .

11. **Termination**. Either PROVIDER or RECIPIENT may terminate this Agreement upon the giving of at least thirty (30) days advance written notice. If not sooner terminated, this Agreement shall expire one (1) year after the Effective Date. Upon expiration or termination of this Agreement, RECIPIENT shall, if so directed by PROVIDER, to return any unused portions of the MATERIAL or destroy the MATERIAL. Sections 2, 4, 5, 6, 7, 8 and 9 shall survive termination or expiration.

12. **No Assignment**. RECIPIENT shall have no right to assign or otherwise transfer any of its rights in this Agreement or delegate any of its duties under this Agreement, without the written consent of PROVIDER.

13. **Governing Law**. This Agreement will be governed, construed, interpreted and enforced in accordance with the laws of Germany (without regard to the laws of conflict thereof).

14. **Contacts**. All notices or reports permitted or required under this Agreement shall be in writing and shall be delivered by personal delivery, electronic mail, facsimile

transmission or by certified or registered mail, return receipt requested, and shall be deemed given upon personal delivery, five (5) days after deposit in the mail, or upon acknowledgment of receipt of electronic transmission.

Notices shall be sent to the following persons for scientific issues:
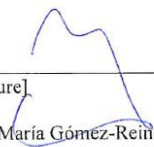
UHD PROVIDER SCIENTIST:	FIDIS RECIPIENT SCIENTIST:
Name: Prof. Dr. Hans Neubauer	Name: Roberto Piñeiro Cid, PhD
Address: Department of Obstetrics and Gynecology University Hospital of Duesseldorf Moorenstrasse 5 D-40225 Düsseldorf, Germany	Address: Cancer Modelling Lab Roche-Chus Joint Unit. Travesía da Choupana S/N 15706 Santiago de Compostela, Spain Email: roberto.pineiro.cid@sergas.es Phone: +34 981 955 602 http://www.roche-chus.es/

Notices shall be sent to the following persons for intellectual property issues:

UHD:	FIDIS:
Name: Prof. Dr. Hans Neubauer	Name: Mabel Sampedro Parada
Address: Department of Obstetrics and Gynecology University Hospital of Duesseldorf Moorenstrasse 5 D-40225 Düsseldorf, Germany	Address: Fundación Instituto de Investigación Sanitaria de Santiago de Compostela (FIDIS). Travesía da Choupana S/N. 15706 Santiago de Compostela, Spain e-mail: mabel.sampedro.parada@sergas.es Phone: +34 981 95 11 95

IN WITNESS WHEREOF, the PROVIDER and RECIPIENT have entered into this Agreement as of the Effective Date.

PROVIDER	RECIPIENT
Heinrich Heine University of Duesseldorf, acting on its behalf. University Hospital Duesseldorf	**Fundación Instituto de Investigación Sanitaria de Santiago de Compostela**
By:	By:
[Signature]	[Signature]
Name: Ekkehard Zimmer	Name: María Gómez-Reino Garrido
Title: Commercial Director	Title: Managing Director
Date: _____	Date: _____13/06/2019_____
PROVIDER SCIENTIST	As RECIPIENT SCIENTIST I acknowledge and accept the terms and provisions of this Agreement:
[Signature]	[Signature]
Name: Dr. Hans Neubauer	Name: Roberto Piñeiro Cid
Title: Prof. Dr.	Title: Researcher
Date: _____	Date: _____04/06/2019_____

259

APPENDIX A

The RESEARCH

Background: Circulating tumor cells (CTCs) can travel in the blood of breast cancer (BC) patients as single cells or cluster (CTC-clusters). CTC-clusters are small groups of 2 up to 100 cells released from the primary tumour, and they are believed to have a greater metastatic potential than individual CTCs. The presence of CTC-clusters in the blood of BC patients is associated with a decrease in metastasis-free survival and with the formation of new metastatic foci. Indeed, the presence of CTC-clusters is significantly associated with a worse prognosis, and their persistent presence upon therapy is associated with an adverse clinical outcome, adding prognostic value to CTC enumeration alone. However, CTC-clusters is an extremely rare population of cells. Thus, it has been estimated that CTC-clusters represent only a 2 to 5% of all CTCs, making their isolation and study very challenging. These evidences show the importance to screen larger volumes of blood than obtained with a regular blood draw (7.5 – 10ml) in order to increase the likelihood of detecting and isolating this particular subpopulation of CTCs and to study their contribution to BC metastasis. DLA is a clinically safe method to collect CTCs from liters of blood. On the one side it does not interfere with therapeutic regimen and on the other side it significant increases CTC detection frequency and yield. **Objective of the research**: MATERIAL (Frozen products of Diagnostic Leukapheresis obtained from breast cancer patients) will be used to:
1) Investigate the feasibility to isolate CTC-clusters
2) Characterize CTC-clusters and single CTCs
3) Compare genome and transcriptome data
4) Understand their (differential) contributions to metastasis.

Work to be done: DLA products obtained from BC patients will be processed with the Parsortix microfluidic technology (ANGLE) to capture CTCs and CTC-clusters present in the sample. DLA products will be run in the system under a modified version of the standard CTC-separation protocol specifically established to capture CTC-clusters: the sample is run under reduced pressure and flow conditions avoiding breakage or disaggregation of CTC-clusters. The protocol's performance to capture intact CTC-cluster has been verified by the Recipient Scientist using peripheral blood drawn from BC patients. After capture, CTCs and CTC-clusters retained within the system will be stained with antibodies against epithelial antigens for identification and harvested for downstream analysis by reverting the flow direction.

WP#1 Test of modified Parsortix protocol with DLA-products
1. It will be tested whether the adapted protocol for the Parsortix system is able to separate DLA products given the high cellularity of these samples. We will use cell numbers and protocols provided by PROVIDER OF MATERIAL
2. Capture of CTCs and CTC-clusters will be tested and performed.

To realize 1) and 2) the following tasks will be performed:
a) DLA products tested CTC-negative (based on CellSearch analysis, provided) will be spiked with a known number of individual cells and clusters of BC cell lines (GFP labelled) and run in the Parsortix system. The number of individual tumor

cells and clusters captured will be counted and the integrity of the clusters assessed under a fluorescent microscope (to determine capture efficiency). In addition, the captured cells will be harvested from the system and the number of individual cells and clusters recovered counted (to determine recovery efficiency).

b) DLA products tested CTC-negative (based on CellSearch analysis, provided) will be spiked with a known number of individual cells and clusters of BC cell lines (not labelled) and run in the Parsortix system. Captured cells will be stained with antibodies against epithelial antigens (EpCAM, HER2, EGFR, pan-CK) and the number of individual tumor cells and clusters captured will be counted (to determine capture efficiency). Later, captured cells will be harvested from the system and the number of individual cells and clusters recovered counted (to determine recovery efficiency upon the process of in cassette staining).

c) The next step will be to run CTC-positive DLA products (based on CellSearch analysis, provided). Cells captured within the system will be stained with antibodies against epithelial antigens, and the number and CTCs and CTC-clusters in the cassette will be counted. Later, captured cells will be harvested from the system, and the recovered populations of single CTCs and CTC-clusters will be counted again.

WP#2 Isolation and analysis of CTCs and CTC-clusters

a) CTCs and CTC-clusters will be separated or individualized either by manual micromanipulation or by the DEPArray system (Menarini Silicon Biosystems).

b) Downstream analysis will include techniques of next generation sequencing, NGS. Individualized sample will be characterized by genomic (WES and RNA-seq) studies in order to prove their tumor nature and to identify differential characteristics between single CTCs and CTC-clusters. Cell lysis, DNA extraction and amplification (WGA, whole genome amplification) will be performed using a single cell specific kit (Ampli 1™ WGA Kit), and for RNA studies the Smart-seq2 protocol will be used.

FAVOURABLE REPORT ANIMAL EXPERIMENTATION ETHICAL COMMITTEE OF THE UNIVERSITY OF SANTIAGO DE COMPOSTELA (CEEA)

VICERREITORADO DE INVESTIGACIÓN
E INNOVACIÓN
Oficina de Investigación e Tecnoloxía

Edificio CACTUS – Campus universitario sur
15782 Santiago de Compostela
Tel. 981 547 040 - Fax 981 547 077
Correo electrónico: citainfo@usc.es
http://imaisd.usc.es

Informe del Comité de Ética de Experimentación Animal (CEEA) de los centros usuarios de animales de experimentación de la USC en el Campus de Santiago

El CEEA de los centros usuarios de animales de experimentación de la USC en el Campus de Santiago, tras evaluar el Proyecto titulado "Desarrollo de xenotrasplantes derivados de Células Tumorales Circulantes (CTC) de pacientes metastásicos en modelos murinos" del que es Investigadora Responsable Dña. Clotilde Costa Nogueira, acordó con fecha 1 de Octubre de 2018 emitir

INFORME FAVORABLE

para la realización de dicho proyecto, así como los procedimientos que incluye, en las instalaciones del establecimiento usuario Animalario del CIMUS, con número de registro ES150780275701, siempre que, en cumplimiento del RD 53/2013, se obtenga la correspondiente autorización administrativa.

En Santiago, a 3 de Octubre de 2018

Fdo. El secretario

Fdo. El presidente

Responsable administrativo:	Nombre: Raúl Vieira Miguel	Vº Bº Firma y sello
	Cargo: Director de la RIAIDT de la USC	

ANIMAL EXPERIMENTATION TRAINING CERTIFICATE. B FUNCTION (EUTHANASIA)

XUNTA DE GALICIA
CONSELLERÍA DO MEDIO RURAL

Avda. do Camiño Francés, 10, Baixo
15703 Santiago de Compostela
A Coruña

AGACAL
Axencia Galega
da Calidade Alimentaria

galicia

MÓDULOS FUNDAMENTAIS OU TRONCAIS E ESPECÍFICOS CORRESPONDENTES Á CATEGORÍA "B" EUTANASIA DOS ANIMAIS – ORDE ECC/566/2015, DE 20 DE MARZO

MÓDULOS FUNDAMENTAIS OU TRONCAIS

1.- Lexislación nacional (1 hora).

2.- Ética, benestar animal e as "tres erres", nivel 1 (2 horas).

3.- Bioloxía básica e adecuada, nivel 1 (3 horas).

4.- Coidado, saúde e manexo dos animais, nivel 1 (5 horas).

5.- Recoñecemento do dolor, o sufrimento e a angustia (3 horas).

6.- Métodos incruentos de sacrificio, nivel 1 (2 horas)

MÓDULOS ESPECÍFICOS DA CATEGORÍA "B"

1.-Bioloxía básica e adecuada, nivel 2 (3 horas)

2.- Métodos incruentos de sacrificio, nivel 2 (3 horas)

....

MÓDULOS FUNDAMENTALES O TRONCALES

1.- Legislación nacional (1 hora).

2.- Ética, bienestar animal y las "tres erres", nivel 1 (2 horas).

3.- Biología básica y adecuada, nivel 1 (3 horas).

4.- Cuidado, salud y manejo de los animales, nivel 1 (5 horas).

5.- Reconocimiento del dolor, el sufrimiento y la angustia (3 horas).

6.- Métodos incruentos de sacrificio, nivel 1 (2 horas)

MÓDULOS ESPECÍFICOS DE LA CATEGORÍA "B"

1.-Biología básica y adecuada, nivel 2 (3 horas)

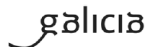*2.- Métodos incruentos de sacrificio, nivel 2 (3 horas)*

Xacobeo 2021

XUNTA DE GALICIA
CONSELLERÍA DO MEDIO RURAL

Avda. do Camiño Francés, 10, Baixo
15703 Santiago de Compostela
A Coruña

AGACAL
Axencia Galega
da Calidade Alimentaria

galicia

Certificado de capacitación en materia de protección de animais utilizados, criados ou subministrados con fins de experimentación e otros fins científicos, incluíndo a docencia conforme coa Orde ECC/566/2015 de 20 de marzo.

Certificado de capacitación en materia de protección de animales utilizados, criados o suministrados con fines de experimentación y otros fines científicos, incluyendo la docencia conforme con la Orden ECC/566/2015 de 20 de marzo.

1. IDENTIFICACIÓN		
1.1. Apelidos / *Apellidos / Surname:* **MARTÍNEZ PENA**		
1.2.Nome / *Nombre / First names:* **INÉS**		DNI /*DNI / Identity card number:* **79338843K**
1.3. Categoría/*Categoría/Category:* **"b"**	1.4. Especies/*Especies/Species:* **ROEDORES**	1.5. Válido ata/ *válido hasta/expires:* **15/09/2028**

2. Nº DO CERTIFICADO / *Nº DEL CERTIFICADO / CERTIFICATE NUMBER*
B489

3.ORGANISMO QUE EXPIDE O CERTIFICADO / *ORGANISMO QUE EXPIDE EL CERTIFICADO / BODY ISSUING THE CERTIFICATE:*

3.1.Nome e enderezo do organismo que expide o certificado / *Nombre y dirección del organismo que expide el certificado / Name and address of the body issuing the certificate:*

CONSELLERÍA DO MEDIO RURAL – XUNTA DE GALICIA
AXENCIA GALEGA DA CALIDADE ALIMENTARIA
Avda. Do Camiño Francés, 10 baixo
15781 Santiago de Compostela
A Coruña (España)

3.2.Teléfono / *Teléfono / Telephone:* 981 546 654	3.3.Fax / *Fax / Fax:* 981 546 651	3.4.Correo electrónico / *Correo electrónico / Email:* formacion.cmrm@xunta.es
3.5.Data / *Fecha / Date:* **15/01/2021**	3.6.Lugar / *Lugar / Place:* **Santiago de Compostela**	

3.7.Nome e sinatura / Nombre y firma / Name and signature
3.8.Selo / Sello / Stamp

Asdo.: Manuel López Luaces
Xefe de Area de Formación, Innovación e Investigación Agraria
(asinado dixitalmente)

Xacobeo 2021

ANIMAL EXPERIMENTATION TRAINING CERTIFICATE. C FUNCTION (ANIMAL PROCEDURES)

MÓDULOS FUNDAMENTAIS OU TRONCAIS E ESPECÍFICOS CORRESPONDENTES Á CATEGORÍA "C" REALIZACIÓN DOS PROCEDEMENTOS – ORDE ECC/566/2015, DE 20 DE MARZO

MÓDULOS FUNDAMENTAIS OU TRONCAIS

1.- Lexislación nacional (1 hora).

2.- Ética, benestar animal e as "tres erres", nivel 1 (2 horas).

3.- Bioloxía básica e adecuada, nivel 1 (3 horas).

4.- Coidado, saúde e manexo dos animais, nivel 1 (5 horas).

5.- Recoñecemento do dolor, o sufrimento e a angustia (3 horas).

6.- Métodos incruentos de sacrificio, nivel 1 (2 horas)

MÓDULOS ESPECÍFICOS DA CATEGORÍA "C"

1.- Bioloxía básica e adecuada, nivel 2 (3 horas)

2.- Procedementos minimamente invasores sen anestesia, nivel 1 (5 horas)

3.- Procedementos minimamente invasores sen anestesia, nivel 2 (10 horas)

4.- Anestesia para procedementos menores (5 horas)

5.- Anestesia avanzada para intervencións cirúrxicas ou procedementos prolongados (8 horas)

6.- Principios de cirurxía (5 horas)

....

MÓDULOS FUNDAMENTALES O TRONCALES

1.- Legislación nacional (1 hora).

2.- Ética, bienestar animal y las "tres erres", nivel 1 (2 horas).

3.- Biología básica y adecuada, nivel 1 (3 horas).

4.- Cuidado, salud y manejo de los animales, nivel 1 (5 horas).

5.- Reconocimiento del dolor, el sufrimiento y la angustia (3 horas).

6.- Métodos incruentos de sacrificio, nivel 1 (2 horas)

MÓDULOS ESPECÍFICOS DE LA CATEGORÍA "C"

1.- Biología básica y adecuada, nivel 2 (3 horas)

2.- Procedimientos mínimamente invasivos sin anestesia, nivel 1 (5 horas)

3.- Procedimientos mínimamente invasivos sin anestesia, nivel 2 (10 horas)

4.- Anestesia para procedementos menores (5 horas)

5.- Anestesia avanzada para intervenciones quirúrgicas o procedimientos prolongados (8 horas)

6.- Principios de cirugía (5 horas)

XUNTA DE GALICIA
CONSELLERÍA DO MEDIO RURAL

Avda. do Camiño Francés, 10, Baixo
15703 Santiago de Compostela
A Coruña

AGACAL
Axencia Galega
da Calidade Alimentaria

galicia

Certificado de capacitación en materia de protección de animais utilizados, criados ou subministrados con fins de experimentación e outros fins científicos, incluíndo a docencia conforme coa Orde ECC/566/2015 de 20 de marzo.

Certificado de capacitación en materia de protección de animales utilizados, criados o suministrados con fines de experimentación y otros fines científicos, incluyendo la docencia conforme con la Orden ECC/566/2015 de 20 de marzo.

1. IDENTIFICACIÓN		
1.1. Apelidos / *Apellidos / Surname:* **MARTÍNEZ PENA**		
1.2. Nome / *Nombre / First names:* **INÉS**		DNI /*DNI / Identity card number:* **79338843K**
1.3. Categoría/*Categoría/Category:* **"c"**	1.4. Especies/*Especies/Species:* **ROEDORES**	1.5. Válido ata/ *válido hasta/expires:* **15/09/2028**

2. Nº DO CERTIFICADO / *Nº DEL CERTIFICADO / CERTIFICATE NUMBER*
C523

3. ORGANISMO QUE EXPIDE O CERTIFICADO / *ORGANISMO QUE EXPIDE EL CERTIFICADO / BODY ISSUING THE CERTIFICATE:*

3.1. Nome e enderezo do organismo que expide o certificado / *Nombre y dirección del organismo que expide el certificado / Name and address of the body issuing the certificate:*

CONSELLERÍA DO MEDIO RURAL – XUNTA DE GALICIA
AXENCIA GALEGA DA CALIDADE ALIMENTARIA
Avda. Do Camiño Francés, 10 baixo
15781 Santiago de Compostela
A Coruña (España)

3.2. Teléfono / *Teléfono / Telephone:* 981 546 654	3.3. Fax / *Fax / Fax:* 981 546 651	3.4. Correo electrónico / *Correo electrónico / Email:* formacion.cmrm@xunta.es
3.5. Data / *Fecha / Date:* **15/01/2021**	3.6. Lugar / *Lugar / Place:* **Santiago de Compostela**	

3.7. Nome e sinatura / *Nombre y firma / Name and signature*

Asdo.: **Manuel López Luaces**
Xefe de Área de Formación, Innovación e Investigación Agraria
(asinado dixitalmente)

Xacobeo 2021

AGRADECIMIENTOS

AGRADECIMIENTOS

Me gustaría agradecer a mis directores Rafael López y Roberto Piñeiro por darme la oportunidad de iniciar esta increíble experiencia en el mundo de la investigación científica y por haberme guiado durante estos 4 años, en los que he aprendido tanto.

Gracias al Servicio de Oncología y a todo el grupo ONCOMET por el gran aprendizaje que me llevo de nuestras reuniones y por aportarme una visión diferente y más clínica sobre las CTC, los clústeres de CTC y de su gran potencial para entender mejor la metástasis. Ojalá pronto puedan beneficiarse mucho más de ello los pacientes con cáncer de mama. Gracias a Miguel Abal por su ayuda y por tener siempre un momento para interesarse por mi cada vez que me ve. Gracias a Jorge por las sesiones de *coaching* y de *mentoring*. Gracias también a Ali por su gran ayuda y a Ramón por su ayuda, su buen humor y por tener siempre unas palabras amables que te alegran el día.

También me gustaría darle las gracias al grupo de la Unidad Mixta tanto a los que lo integran actualmente, como a las distintas personas que han pasado por él a lo largo de estos 4 años. Gracias a Gloria, Ana y Cloti por su ayuda y sus buenos consejos que me han hecho aprender y crecer tanto. Gracias a Carmen por hacernos el día a día mucho más fácil. Gracias a Celso y Miriam por su compañerismo y a mi compañero de línea, Pablo, por echar una mano siempre que hace falta.

Gracias a nuestros colaboradores, por haber depositado tanta confianza en nosotros.

Gracias a mis estudiantes, tanto a los que habéis estado de prácticas en nuestro laboratorio, como a los que he conocido en mis prácticas docentes en las diferentes facultades, por todo lo que me habéis enseñado. He aprendido mucho sobre cómo conectar con los alumnos y ha sido muy gratificante ver cómo poco a poco todo vuestro esfuerzo fue dando sus frutos. Ha sido una experiencia muy enriquecedora, donde el aprendizaje ha sido bidireccional y en la que espero que también vosotros hayáis

disfrutado y aprendido. Gracias también a todos los investigadores y profesores con los que he dado clase, por estar siempre dispuestos a ayudarme y por haberse interesado por mi incluso después de haber terminado nuestras clases. Gracias a Miguel F. y Luis L. por la ayuda que me habéis dado y por lo bien que os habéis portado siempre conmigo, ¡no sabéis cuánto lo aprecio!

En especial me gustaría darle las gracias al 'Club de Wendoline' por los buenos momentos, tanto dentro, como fuera del laboratorio. Sin nuestras 'terapias de grupo' y vuestro apoyo no habría sido tan fácil llegar hasta aquí. Habéis enriquecido mucho mi camino y lo habéis hecho mucho más fácil. Tengo mucha suerte de poder decir que además de la experiencia, ¡me llevo muy buenas amistades! Gracias a Aida, Aitor (y a Manu) por haberme escuchado cuando fue necesario y por recordarme lo importante que es reírse hasta de las cosas menos agradables. Gracias a Carlos, Patri y Óscar por estar siempre ahí, dispuestos a escucharme y a ayudar en lo que fuese necesario, siempre con una sonrisa. Gracias a Sandra A. (y Luis), Sandra D., Marce, María y Sainza por los '¿qué tal estás?' cada vez que nos vemos. Gracias a Thais que aunque ya no estés en el grupo, te sigo teniendo igual de presente que si trabajásemos una al lado de la otra. Gracias por las charlas, por la ayuda, por los buenos momentos, por los increíbles viajes… Gracias a Carol y a Nuria que habéis sido un gran apoyo, siempre ahí para darme ánimos cuando hizo falta y también para celebrar las cosas buenas. Gracias por los vermuts de domingo, los vinos de entresemana, las cenas improvisadas, los planes de escapada y las charlas que siempre hacen que vuelva a casa con una sonrisa al final del día. Esto no se acaba aquí, ni se me olvida que tenemos un viaje pendiente juntas ;-) ¡Hemos vivido mucho junt@s y no os podéis imaginar cuánto os voy a echar de menos a todos!

Gracias al grupo 'Reflote', Javi, Jorge, Lucía, Manu, Santi, por las risas aseguradas cada vez que nos vemos.

Gracias a Lorena C., Cris V., y a Lucía A. (y a Dani) por los cafés, las cañas, las conversaciones, las risas y sobre todo, por ser tan comprensiv@s conmigo estos últimos meses.

Muchas gracias a mi entrenador Norberto G. y a mis compañeros de karate por recordarme la importancia de la constancia y por estar siempre para ayudarme y apoyarme, tanto dentro como fuera del *dojo*. Después de estos 23 años de aprendizaje juntos, más que un grupo somos una gran familia.

Muchísimas gracias a mi familia, tanto a mis padres, como a mis abuelos maternos, por vuestro apoyo incondicional y por enseñarme la importancia de pelear por aquello que me guste, aunque a veces no sea fácil. Sois un ejemplo de fortaleza, trabajo duro, constancia y lo más importante, de cómo son las buenas personas.

Finalmente, gracias a Martín por estar conmigo, apoyándome cuando lo he necesitado y por tu paciencia infinita estos últimos meses. Has sido una fuente de calma siempre que ha hecho falta y has hecho llevaderos hasta los días más amargos. He elegido el mejor compañero de viaje que se puede tener.

Espero no haberme dejado a nadie en el tintero. Si fuese así, perdonadme, ha sido un largo viaje de aprendizaje en el que me he cruzado con muchas personas. Para los que me haya podido dejar atrás en estos agradecimientos, gracias por haber enriquecido mi experiencia a lo largo de esto 4 años.

Si tuviese que resumir estos 4 años, diría que sido un viaje muy inspirador, con sus altos y sus bajos pero que me ha hecho crecer mucho, tanto a nivel profesional, como a nivel personal. Se cierra una etapa importante para mí por todo lo que ha implicado, así que ahora toca disfrutarlo y sobre todo ¡celebrarlo juntos!

REFERENCES

REFERENCES

1.   Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2020. *CA. Cancer J. Clin.* **70**, 7–30 (2020).

2.   Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* **68**, 394–424 (2018).

3.   Ahmad, A. Breast Cancer Statistics: Recent Trends. in 1–7 (2019). doi:10.1007/978-3-030-20301-6_1.

4.   Barnard, M. E., Boeke, C. E. & Tamimi, R. M. Established breast cancer risk factors and risk of intrinsic tumor subtypes. *Biochim. Biophys. Acta - Rev. Cancer* **1856**, 73–85 (2015).

5.   ROJAS, K. & STUCKEY, A. Breast Cancer Epidemiology and Risk Factors. *Clin. Obstet. Gynecol.* **59**, 651–672 (2016).

6.   Tao, Z. Q. *et al.* Breast Cancer: Epidemiology and Etiology. *Cell Biochem. Biophys.* **72**, 333–338 (2015).

7.   Eroles, P., Bosch, A., Alejandro Pérez-Fidalgo, J. & Lluch, A. Molecular biology in breast cancer: Intrinsic subtypes and signaling pathways. *Cancer Treat. Rev.* **38**, 698–707 (2012).

8.   Sonnenblick, A., Fumagalli, D., Sotiriou, C. & Piccart, M. Is the differentiation into molecular subtypes of breast cancer important for staging, local and systemic therapy, and follow up? *Cancer Treat. Rev.* **40**, 1089–1095 (2014).

9.   Prat, A. *et al.* Clinical implications of the intrinsic molecular subtypes of breast cancer. *The Breast* **24**, S26–S35 (2015).

10. Dai, X. *et al.* Breast cancer intrinsic subtype classification, clinical use and future trends. *Am. J. Cancer Res.* **5**, 2929–2943 (2015).

11. Sorlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci.* **98**, 10869–10874 (2001).

12. Gao, J. J. & Swain, S. M. Luminal A Breast Cancer and Molecular Assays: A Review. *Oncologist* **23**, 556–565 (2018).

13. Ignatiadis, M. & Sotiriou, C. Luminal breast cancer: from biology to treatment. *Nat. Rev. Clin. Oncol.* **10**, 494–506 (2013).

14. Garrido-Castro, A. C., Lin, N. U. & Polyak, K. Insights into molecular classifications of triple-negative breast cancer: Improving patient selection for treatment. *Cancer Discov.* **9**, 176–198 (2019).

15. Plichta, J. K. *et al.* Implications for Breast Cancer Restaging Based on the 8th Edition AJCC Staging Manual. *Ann. Surg.* **271**, 169–176 (2020).

16. Sobin, L. H. TNM: Evolution and relation to other prognostic factors. *Semin. Surg. Oncol.* **21**, 3–7 (2003).

17. Giuliano, A. E., Edge, S. B. & Hortobagyi, G. N. Eighth Edition of the AJCC Cancer Staging Manual: Breast Cancer. *Ann. Surg. Oncol.* **25**, 1783–1785 (2018).

18. Hortobagyi, G. N. *et al.* Breast. in *AJCC Cancer Staging Manual* (ed. Amin, M.B., Edge, S., Greene, F., Byrd, D.R., Brookland, R.K., Washington, M.K., Gershenwald, J.E., Compton, C.C., Hess, K.R., Sullivan, D.C., Jessup, J.M., Brierley, J.D., Gaspar, L.E., Schilsky, R.L., Balch, C.M., Winchester, D.P., Asare, E.A., Madera, L. R. (Eds. .) 589–636 (Springer International Publishing, 2017). doi:10.1007/978-3-319-40618-3_48.

19. Giuliano, A. E. *et al.* Breast Cancer-Major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA. Cancer J. Clin.* **67**, 290–303 (2017).

20. Massagué, J. & Obenauf, A. C. Metastatic colonization by circulating tumour cells. *Nature* **529**, 298–306 (2016).

21. Lambert, A. W., Pattabiraman, D. R. & Weinberg, R. A. Emerging Biological Principles of Metastasis. *Cell* **168**, 670–691 (2017).

22. Jemal, A., Siegel, R., Xu, J. & Ward, E. Cancer Statistics, 2010. *CA. Cancer J. Clin.* **60**, 277–300 (2010).

23. Cheung, K. J. & Ewald, A. J. A collective route to metastasis: Seeding by tumor cell clusters. *Science (80-.).* **352**, 167–169 (2016).

24. Lintz, M., Muñoz, A. & Reinhart-King, C. A. The Mechanics of Single Cell and Collective Migration of Tumor Cells. *J. Biomech. Eng.* **139**, (2017).

25. Scully, O. J., Bay, B.-H., Yip, G. & Yu, Y. Breast cancer metastasis. *Cancer Genomics Proteomics* **9**, 311–20 (2012).

26. Cheung, K. J. & Ewald, A. J. Illuminating breast cancer invasion: diverse roles for cell–cell interactions. *Curr. Opin. Cell Biol.* **30**, 99–111 (2014).

27. Yu, W., Yang, L., Li, T. & Zhang, Y. Cadherin Signaling in Cancer: Its Functions and Role as a Therapeutic Target. *Front. Oncol.* **9**, (2019).

28. Sousa, B., Pereira, J. & Paredes, J. The Crosstalk Between Cell Adhesion and Cancer

Metabolism. *Int. J. Mol. Sci.* **20**, 1933 (2019).

29. Kourtidis, A., Lu, R., Pence, L. J. & Anastasiadis, P. Z. A central role for cadherin signaling in cancer. *Exp. Cell Res.* **358**, 78–85 (2017).

30. Li, D.-M. & Feng, Y.-M. Signaling mechanism of cell adhesion molecules in breast cancer metastasis: potential therapeutic targets. *Breast Cancer Res. Treat.* **128**, 7–21 (2011).

31. Padmanaban, V. *et al.* E-cadherin is required for metastasis in multiple models of breast cancer. *Nature* **573**, 439–444 (2019).

32. Blaschuk, O. W. N-cadherin antagonists as oncology therapeutics. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140039 (2015).

33. Micalizzi, D. S., Haber, D. A. & Maheswaran, S. Cancer metastasis through the prism of epithelial-to-mesenchymal transition in circulating tumor cells. *Mol. Oncol.* **11**, 770–780 (2017).

34. Mendonsa, A. M., Na, T.-Y. & Gumbiner, B. M. E-cadherin in contact inhibition and cancer. *Oncogene* **37**, 4769–4780 (2018).

35. Mittal, V. Epithelial Mesenchymal Transition in Tumor Metastasis. *Annu. Rev. Pathol. Mech. Dis.* **13**, 395–412 (2018).

36. Yadav, L. *et al.* Matrix Metalloproteinases and Cancer - Roles in Threat and Therapy. *Asian Pacific J. Cancer Prev.* **15**, 1085–1091 (2014).

37. Das, A., Monteiro, M., Barai, A., Kumar, S. & Sen, S. MMP proteolytic activity regulates cancer invasiveness by modulating integrins. *Sci. Rep.* **7**, 14219 (2017).

38. Winograd-Katz, S. E., Fässler, R., Geiger, B. & Legate, K. R. The integrin adhesome: from genes and proteins to human disease. *Nat. Rev. Mol. Cell Biol.* **15**, 273–288 (2014).

39. Jang, I. & Beningo, K. A. Integrins, CAFs and Mechanical Forces in the Progression of Cancer. *Cancers (Basel).* **11**, 721 (2019).

40. Alday-Parejo, Stupp & Rüegg. Are Integrins Still Practicable Targets for Anti-Cancer Therapy? *Cancers (Basel).* **11**, 978 (2019).

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

42. Parvani, J. G., Gujrati, M. D., Mack, M. A., Schiemann, W. P. & Lu, Z.-R. Silencing β3 Integrin by Targeted ECO/siRNA Nanoparticles Inhibits EMT and Metastasis of Triple-Negative Breast Cancer. *Cancer Res.* **75**, 2316–2325 (2015).

43. Steeg, P. S. Targeting metastasis. *Nat. Rev. Cancer* **16**, 201–218 (2016).

44. Wu, P.-H., Opadele, A. E., Onodera, Y. & Nam, J.-M. Targeting Integrins in Cancer Nanomedicine: Applications in Cancer Diagnosis and Therapy. *Cancers (Basel).* **11**, 1783 (2019).

45. Aceto, N. *et al.* Circulating Tumor Cell Clusters Are Oligoclonal Precursors of Breast Cancer Metastasis. *Cell* **158**, 1110–1122 (2014).

46. Ullah, I. *et al.* Evolutionary history of metastatic breast cancer reveals minimal seeding from axillary lymph nodes. *J. Clin. Invest.* **128**, 1355–1370 (2018).

47. Guelte, A., Dwyer, J. & Gavard, J. Jumping the barrier: VE-cadherin, VEGF and other angiogenic modifiers in cancer. *Biol. Cell* **103**, 593–605 (2011).

48. Katt, M. E., Wong, A. D. & Searson, P. C. Dissemination from a Solid Tumor: Examining the Multiple Parallel Pathways. *Trends in Cancer* **4**, 20–37 (2018).

49. Melzer, C., von der Ohe, J. & Hass, R. Breast Carcinoma: From Initial Tumor Cell Detachment to Settlement at Secondary Sites. *Biomed Res. Int.* **2017**, 1–11 (2017).

50. Marciel, M. P. & Hoffmann, P. R. Selenoproteins and Metastasis. in 85–108 (2017). doi:10.1016/bs.acr.2017.07.008.

51. Strell, C. & Entschladen, F. Extravasation of leukocytes in comparison to tumor cells. *Cell Commun. Signal.* **6**, 10 (2008).

52. Allen, T. A. *et al.* Angiopellosis as an Alternative Mechanism of Cell Extravasation. *Stem Cells* **35**, 170–180 (2017).

53. Allen, T. A. *et al.* Circulating tumor cells exit circulation while maintaining multicellularity, augmenting metastatic potential. *J. Cell Sci.* **132**, (2019).

54. Yu, M. *et al.* Circulating Breast Tumor Cells Exhibit Dynamic Changes in Epithelial and Mesenchymal Composition. *Science (80-.).* **339**, 580–584 (2013).

55. Zhao, Q. *et al.* Interaction between circulating galectin-3 and cancer-associated MUC1 enhances tumour cell homotypic aggregation and prevents anoikis. *Mol. Cancer* **9**, 154 (2010).

56. Fabisiewicz, A. & Grzybowska, E. CTC clusters in cancer progression and metastasis. *Med. Oncol.* **34**, 12 (2017).

57. Krog, B. L. & Henry, M. D. Biomechanics of the Circulating Tumor Cell

Microenvironment. in 209–233 (2018). doi:10.1007/978-3-319-95294-9_11.

58. Corbin, E. A., Kong, F., Lim, C. T., King, W. P. & Bashir, R. Biophysical properties of human breast cancer cells measured using silicon MEMS resonators and atomic force microscopy. *Lab Chip* **15**, 839–847 (2015).

59. Pepona, M. *et al.* Investigating the Interaction Between Circulating Tumor Cells and Local Hydrodynamics via Experiment and Simulations. *Cell. Mol. Bioeng.* **13**, 527–540 (2020).

60. Marrella, A. *et al.* High blood flow shear stress values are associated with circulating tumor cells cluster disaggregation in a multi-channel microfluidic device. *PLoS One* **16**, e0245536 (2021).

61. Méhes, G., Witt, A., Kubista, E. & Ambros, P. F. Circulating Breast Cancer Cells Are Frequently Apoptotic. *Am. J. Pathol.* **159**, 17–20 (2001).

62. Kallergi, G. *et al.* Apoptotic Circulating Tumor Cells in Early and Metastatic Breast Cancer Patients. *Mol. Cancer Ther.* **12**, 1886–1895 (2013).

63. Chaffer, C. L., Thompson, E. W. & Williams, E. D. Mesenchymal to Epithelial Transition in Development and Disease. *Cells Tissues*

Organs **185**, 7–19 (2007).

64. Chin, V. L. & Lim, C. L. Epithelial-mesenchymal plasticity—engaging stemness in an interplay of phenotypes. *Stem Cell Investig.* **6**, 25–25 (2019).

65. Pantel, K. & Speicher, M. R. The biology of circulating tumor cells. *Oncogene* **35**, 1216–1224 (2016).

66. Akhtar, M., Haider, A., Rashid, S. & Al-Nabet, A. D. M. H. Paget's "Seed and Soil" Theory of Cancer Metastasis: An Idea Whose Time has Come. *Adv. Anat. Pathol.* **26**, 69–74 (2019).

67. Psaila, B. & Lyden, D. The metastatic niche: adapting the foreign soil. *Nat. Rev. Cancer* **9**, 285–293 (2009).

68. Mashouri, L. *et al.* Exosomes: composition, biogenesis, and mechanisms in cancer metastasis and drug resistance. *Mol. Cancer* **18**, 75 (2019).

69. Chinen, L. T. D. *et al.* Circulating Tumor Cells as Cancer Biomarkers in the Clinic. in 1–41 (2017). doi:10.1007/978-3-319-55947-6_1.

70. Ashworth, T. R. A case of cancer in which cells similar to those in the tumours were seen in blood after death. *Aust. Med. J.* **14**, 146–

147 (1869).

71. Haber, D. A. & Velculescu, V. E. Blood-Based Analyses of Cancer: Circulating Tumor Cells and Circulating Tumor DNA. *Cancer Discov.* **4**, 650–661 (2014).

72. Agarwal, A., Balic, M., El-Ashry, D. & Cote, R. J. Circulating Tumor Cells: Strategies for Capture, Analyses, and Propagation. *Cancer J.* **24**, 70–77 (2018).

73. Fabisiewicz, A., Szostakowska-Rodzos, M., Zaczek, A. J. & Grzybowska, E. A. Circulating Tumor Cells in Early and Advanced Breast Cancer; Biology and Prognostic Value. *Int. J. Mol. Sci.* **21**, 1671 (2020).

74. Poulet, G., Massias, J. & Taly, V. Liquid Biopsy: General Concepts. *Acta Cytol.* **63**, 449–455 (2019).

75. Agarwal, A., Balic, M., El-Ashry, D. & Cote, R. J. Circulating Tumor Cells: strategies for Capture, Analysis, and Propagation. *Cancer J.* **24**, 70–77 (2018).

76. Rushton, A. J., Nteliopoulos, G., Shaw, J. A. & Coombes, R. C. A Review of Circulating Tumour Cell Enrichment Technologies. *Cancers (Basel).* **13**, 970 (2021).

77. Genna, A. *et al.* EMT-Associated

Heterogeneity in Circulating Tumor Cells: Sticky Friends on the Road to Metastasis. *Cancers (Basel).* **12**, 1632 (2020).

78. Costa, C. & Dávila-Ibáñez, A. B. Methodology for the Isolation and Analysis of CTCs. in 45–59 (2020). doi:10.1007/978-3-030-35805-1_4.

79. Cristofanilli, M. *et al.* Circulating Tumor Cells, Disease Progression, and Survival in Metastatic Breast Cancer. *N. Engl. J. Med.* **351**, 781–791 (2004).

80. Danila, D. C. *et al.* Circulating Tumor Cell Number and Prognosis in Progressive Castration-Resistant Prostate Cancer. *Clin. Cancer Res.* **13**, 7053–7058 (2007).

81. Abalde-Cela, S., Piairo, P. & Diéguez, L. The Significance of Circulating Tumour Cells in the Clinic. *Acta Cytol.* **63**, 466–478 (2019).

82. N. Saucedo-Zeni, S. Mewes, R. Niestroj, L. Gasiorowski, D. Murawa, P. Nowaczyk, T. Tomasi, E. Weber, G.Dworacki, N.G. Morgenthaler, H. Jansen, C. Propping, K. Sterzynska, W. Dyszkiewicz, M. Zabel, M. Kiechle, U. Reuning, M. Schmitt, K. L. A novel method for the in vivo isolation of circulating tumor cells from peripheral blood of cancer patients using a functionalized and structured medical wire. *Int. J. Oncol.* (2012)

doi:10.3892/ijo.2012.1557.

83. Schneck, H. *et al.* EpCAM-Independent Enrichment of Circulating Tumor Cells in Metastatic Breast Cancer. *PLoS One* **10**, e0144535 (2015).

84. Eslami-S, Z., Cortés-Hernández, L. E. & Alix-Panabières, C. Epithelial Cell Adhesion Molecule: An Anchor to Isolate Clinically Relevant Circulating Tumor Cells. *Cells* **9**, 1836 (2020).

85. Giordano, A. & Cristofanilli, M. CTCs in Metastatic Breast Cancer. in 193–201 (2012). doi:10.1007/978-3-642-28160-0_18.

86. Bredemeier, M. *et al.* Gene Expression Signatures in Circulating Tumor Cells Correlate with Response to Therapy in Metastatic Breast Cancer. *Clin. Chem.* **63**, 1585–1593 (2017).

87. Keup, C. *et al.* RNA Profiles of Circulating Tumor Cells and Extracellular Vesicles for Therapy Stratification of Metastatic Breast Cancer Patients. *Clin. Chem.* **64**, 1054–1062 (2018).

88. Labib, M. & Kelley, S. O. Circulating tumor cell profiling for precision oncology. *Mol. Oncol.* **15**, 1622–1646 (2021).

89. Pereira-Veiga, T. *et al.* CTCs-derived

xenograft development in a triple negative breast cancer case. *Int. J. Cancer* ijc.32001 (2018) doi:10.1002/ijc.32001.

90. Ramirez, J.-M. *et al.* Prognostic Relevance of Viable Circulating Tumor Cells Detected by EPISPOT in Metastatic Breast Cancer Patients. *Clin. Chem.* **60**, 214–221 (2014).

91. Boya, M., Chu, C.-H., Liu, R., Ozkaya-Ahmadov, T. & Sarioglu, A. F. Circulating Tumor Cell Enrichment Technologies. in 25–55 (2020). doi:10.1007/978-3-030-26439-0_2.

92. Franken, A. *et al.* Label-Free Enrichment and Molecular Characterization of Viable Circulating Tumor Cells from Diagnostic Leukapheresis Products. *Clin. Chem.* **65**, 549–558 (2019).

93. Stoecklein, N. H., Fischer, J. C., Niederacher, D. & Terstappen, L. W. M. M. Challenges for CTC-based liquid biopsies: low CTC frequency and diagnostic leukapheresis as a potential solution. *Expert Rev. Mol. Diagn.* **16**, 147–164 (2016).

94. Reinhardt, F. *et al.* Diagnostic Leukapheresis Enables Reliable Transcriptomic Profiling of Single Circulating Tumor Cells to Characterize Inter-Cellular Heterogeneity in Terms of Endocrine Resistance. *Cancers (Basel).* **11**, 903 (2019).

95. Andree, K. C. *et al.* Toward a real liquid biopsy in metastatic breast and prostate cancer: Diagnostic LeukApheresis increases CTC yields in a European prospective multicenter study (CTCTrap). *Int. J. Cancer* **143**, 2584–2591 (2018).

96. Vona, G. *et al.* Isolation by Size of Epithelial Tumor Cells. *Am. J. Pathol.* **156**, 57–63 (2000).

97. Miller, M. C., Robinson, P. S., Wagner, C. & O'Shannessy, D. J. The Parsortix™ Cell Separation System—A versatile liquid biopsy platform. *Cytom. Part A* **93**, 1234–1239 (2018).

98. Cohen, E. N. *et al.* Antigen-agnostic microfluidics-based circulating tumor cell enrichment and downstream molecular characterization. *PLoS One* **15**, e0241123 (2020).

99. Koch, C. *et al.* Pre-Analytical and Analytical Variables of Label-Independent Enrichment and Automated Detection of Circulating Tumor Cells in Cancer Patients. *Cancers (Basel).* **12**, 442 (2020).

100. Hvichia, G. E. *et al.* A novel microfluidic platform for size and deformability based separation and the subsequent molecular characterization of viable circulating tumor

cells. *Int. J. Cancer* **138**, 2894–2904 (2016).

101. Alshareef, M. *et al.* Separation of tumor cells with dielectrophoresis-based microfluidic chip. *Biomicrofluidics* **7**, 011803 (2013).

102. Fabbri, F. *et al.* Detection and recovery of circulating colon cancer cells using a dielectrophoresis-based device: KRAS mutation status in pure CTCs. *Cancer Lett.* **335**, 225–231 (2013).

103. Castro-Giner, F. & Aceto, N. Tracking cancer progression: from circulating tumor cells to metastasis. *Genome Med.* **12**, 31 (2020).

104. Aceto, N. Bring along your friends: Homotypic and heterotypic circulating tumor cell clustering to accelerate metastasis. *Biomed. J.* **43**, 18–23 (2020).

105. VIRCHOW, R. As Based upon Physiological and Pathological Histology. *Nutr. Rev.* **47**, 23–25 (1989).

106. Watanabe, S. The metastasizability of tumor cells. *Cancer* **7**, 215–223 (1954).

107. Hong, Y., Fang, F. & Zhang, Q. Circulating tumor cell clusters: What we know and what we expect (Review). *Int. J. Oncol.* **49**, 2206–2216 (2016).

108. Cheung, K. J. *et al.* Polyclonal breast cancer metastases arise from collective dissemination

of keratin 14-expressing tumor cell clusters. *Proc. Natl. Acad. Sci.* **113**, E854–E863 (2016).

109. Liu, X. *et al.* Homophilic CD44 Interactions Mediate Tumor Cell Aggregation and Polyclonal Metastasis in Patient-Derived Breast Cancer Models. *Cancer Discov.* **9**, 96–113 (2019).

110. Glinsky, V. V *et al.* Intravascular metastatic cancer cell homotypic aggregation at the sites of primary attachment to the endothelium. *Cancer Res.* **63**, 3805–11 (2003).

111. Haeger, A., Krause, M., Wolf, K. & Friedl, P. Cell jamming: Collective invasion of mesenchymal tumor cells imposed by tissue confinement. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 2386–2395 (2014).

112. Küsters, B. *et al.* Micronodular transformation as a novel mechanism of VEGF-A-induced metastasis. *Oncogene* **26**, 5808–5815 (2007).

113. Fidler, I. J. & Talmadge, J. E. Evidence that intravenously derived murine pulmonary melanoma metastases can originate from the expansion of a single tumor cell. *Cancer Res.* **46**, 5167–71 (1986).

114. Thangavel, H. *et al.* A CTC-Cluster-Specific Signature Derived from OMICS Analysis of Patient-Derived Xenograft Tumors Predicts

Outcomes in Basal-Like Breast Cancer. *J. Clin. Med.* **8**, 1772 (2019).

115. Paoli, P., Giannoni, E. & Chiarugi, P. Anoikis molecular pathways and its role in cancer progression. *Biochim. Biophys. Acta - Mol. Cell Res.* **1833**, 3481–3498 (2013).

116. Paoletti, C. *et al.* Significance of Circulating Tumor Cells in Metastatic Triple-Negative Breast Cancer Patients within a Randomized, Phase II Trial: TBCRC 019. *Clin. Cancer Res.* **21**, 2771–2779 (2015).

117. Gkountela, S. *et al.* Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding. *Cell* **176**, 98-112.e14 (2019).

118. Reduzzi, C. *et al.* Circulating Tumor Cell Clusters Are Frequently Detected in Women with Early-Stage Breast Cancer. *Cancers (Basel).* **13**, 2356 (2021).

119. Krol, I. *et al.* Detection of clustered circulating tumour cells in early breast cancer. *Br. J. Cancer* **125**, 23–27 (2021).

120. Silvestri, M. *et al.* Detection of Genomically Aberrant Cells within Circulating Tumor Microemboli (CTMs) Isolated from Early-Stage Breast Cancer Patients. *Cancers (Basel).* **13**, 1409 (2021).

121. Bidard, F.-C. *et al.* Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. *Lancet Oncol.* **15**, 406–414 (2014).

122. Zhang, L. *et al.* Meta-Analysis of the Prognostic Value of Circulating Tumor Cells in Breast Cancer. *Clin. Cancer Res.* **18**, 5701–5710 (2012).

123. Smerage, J. B. *et al.* Circulating Tumor Cells and Response to Chemotherapy in Metastatic Breast Cancer: SWOG S0500. *J. Clin. Oncol.* **32**, 3483–3489 (2014).

124. Amintas, S. *et al.* Circulating Tumor Cell Clusters: United We Stand Divided We Fall. *Int. J. Mol. Sci.* **21**, 2653 (2020).

125. Wang, C. *et al.* Longitudinally collected CTCs and CTC-clusters and clinical outcomes of metastatic breast cancer. *Breast Cancer Res. Treat.* **161**, 83–94 (2017).

126. Mu, Z. *et al.* Prospective assessment of the prognostic value of circulating tumor cells and their clusters in patients with advanced-stage breast cancer. *Breast Cancer Res. Treat.* **154**, 563–571 (2015).

127. Jansson, S., Bendahl, P.-O., Larsson, A.-M., Aaltonen, K. E. & Rydén, L. Prognostic impact of circulating tumor cell apoptosis and

clusters in serial blood samples from patients with metastatic breast cancer in a prospective observational cohort. *BMC Cancer* **16**, 433 (2016).

128. Larsson, A.-M. *et al.* Longitudinal enumeration and cluster evaluation of circulating tumor cells improve prognostication for patients with newly diagnosed metastatic breast cancer in a prospective observational trial. *Breast Cancer Res.* **20**, 48 (2018).

129. Szczerba, B. M. *et al.* Neutrophils escort circulating tumour cells to enable cell cycle progression. *Nature* **566**, 553–557 (2019).

130. Costa, C. *et al.* Analysis of a Real-World Cohort of Metastatic Breast Cancer Patients Shows Circulating Tumor Cell Clusters (CTC-clusters) as Predictors of Patient Outcomes. *Cancers (Basel).* **12**, 1111 (2020).

131. Hou, J.-M. *et al.* Clinical Significance and Molecular Characteristics of Circulating Tumor Cells and Circulating Tumor Microemboli in Patients With Small-Cell Lung Cancer. *J. Clin. Oncol.* **30**, 525–532 (2012).

132. Hou, J.-M. *et al.* Circulating Tumor Cells as a Window on Metastasis Biology in Lung Cancer. *Am. J. Pathol.* **178**, 989–996 (2011).

133. Kats-Ugurlu, G. *et al.* Circulating tumour tissue fragments in patients with pulmonary metastasis of clear cell renal cell carcinoma. *J. Pathol.* **219**, 287–293 (2009).

134. Topal, B., Roskams, T., Fevery, J. & Penninckx, F. Aggregated colon cancer cells have a higher metastatic efficiency in the liver compared with nonaggregated cells: an experimental study. *J. Surg. Res.* **112**, 31–37 (2003).

135. Aceto, N., Toner, M., Maheswaran, S. & Haber, D. A. En Route to Metastasis: Circulating Tumor Cell Clusters and Epithelial-to-Mesenchymal Transition. *Trends in Cancer* **1**, 44–52 (2015).

136. Wang, C. *et al.* Longitudinally collected CTCs and CTC-clusters and clinical outcomes of metastatic breast cancer. *Breast Cancer Res. Treat.* **161**, 83–94 (2017).

137. Waks, A. G. & Winer, E. P. Breast Cancer Treatment. *JAMA* **321**, 288 (2019).

138. Paoletti, C. *et al.* Significance of Circulating Tumor Cells in Metastatic Triple-Negative Breast Cancer Patients within a Randomized, Phase II Trial: TBCRC 019. *Clin. Cancer Res.* **21**, 2771–2779 (2015).

139. Paoletti, C. *et al.* Circulating Tumor Cell Clusters in Patients with Metastatic Breast

Cancer: a SWOG S0500 Translational Medicine Study. *Clin. Cancer Res.* **25**, 6089–6097 (2019).

140. Sharma, S. *et al.* Circulating tumor cell isolation, culture, and downstream molecular analysis. *Biotechnol. Adv.* **36**, 1063–1078 (2018).

141. Xu, L. *et al.* Optimization and Evaluation of a Novel Size Based Circulating Tumor Cell Isolation System. *PLoS One* **10**, e0138032 (2015).

142. Vasseur, A., Kiavue, N., Bidard, F., Pierga, J. & Cabel, L. Clinical utility of circulating tumor cells: an update. *Mol. Oncol.* **15**, 1647–1666 (2021).

143. Puliafito, A. *et al.* Three-dimensional chemotaxis-driven aggregation of tumor cells. *Sci. Rep.* **5**, 15205 (2015).

144. Ziegler, E., Hansen, M.-T., Haase, M., Emons, G. & Gründker, C. Generation of MCF-7 cells with aggressive metastatic potential in vitro and in vivo. *Breast Cancer Res. Treat.* **148**, 269–277 (2014).

145. Rejniak, K. A. Circulating Tumor Cells: When a Solid Tumor Meets a Fluid Microenvironment. in 93–106 (2016). doi:10.1007/978-3-319-42023-3_5.

146. Follain, G. *et al.* Hemodynamic Forces Tune the Arrest, Adhesion, and Extravasation of Circulating Tumor Cells. *Dev. Cell* **45**, 33-52.e12 (2018).

147. Zhang, X., Xu, L. & Yu, Q. Cell aggregation induces phosphorylation of PECAM-1 and Pyk2 and promotes tumor cell anchorage-independent growth. *Mol. Cancer* **9**, 7 (2010).

148. Micalizzi, D. S., Maheswaran, S. & Haber, D. A. A conduit to metastasis: circulating tumor cell biology. *Genes Dev.* **31**, 1827–1840 (2017).

149. Shen, M. & Kang, Y. Stresses in the metastatic cascade: molecular mechanisms and therapeutic opportunities. *Genes Dev.* **34**, 1577–1598 (2020).

150. Au, S. H. *et al.* Clusters of circulating tumor cells traverse capillary-sized vessels. *Proc. Natl. Acad. Sci.* **113**, 4947–4952 (2016).

151. Strilic, B. & Offermanns, S. Intravascular Survival and Extravasation of Tumor Cells. *Cancer Cell* **32**, 282–293 (2017).

152. Maeshiro, M. *et al.* Colonization of distant organs by tumor cells generating circulating homotypic clusters adaptive to fluid shear stress. *Sci. Rep.* **11**, 6150 (2021).

153. Agnoletto, C. *et al.* Heterogeneity in

Circulating Tumor Cells: The Relevance of the Stem-Cell Subset. *Cancers (Basel).* **11**, 483 (2019).

154. Ma, Z., Webb, D. J., Jo, M. & Gonias, S. L. Endogenously produced urokinase-type plasminogen activator is a major determinant of the basal level of activated ERK/MAP kinase and prevents apoptosis in MDA-MB-231 breast cancer cells. *J. Cell Sci.* **114**, 3387–96 (2001).

155. Alfano, D., Iaccarino, I. & Stoppelli, M. P. Urokinase Signaling through Its Receptor Protects against Anoikis by Increasing BCL-xL Expression Levels. *J. Biol. Chem.* **281**, 17758–17767 (2006).

156. Mahmood, N., Mihalcioiu, C. & Rabbani, S. A. Multifaceted Role of the Urokinase-Type Plasminogen Activator (uPA) and Its Receptor (uPAR): Diagnostic, Prognostic, and Therapeutic Applications. *Front. Oncol.* **8**, (2018).

157. Gutierrez, L. S. *et al.* Tumor development is retarded in mice lacking the gene for urokinase-type plasminogen activator or its inhibitor, plasminogen activator inhibitor-1. *Cancer Res.* **60**, 5839–47 (2000).

158. Moquet-Torcy, G., Tolza, C., Piechaczyk, M. & Jariel-Encontre, I. Transcriptional

complexity and roles of Fra-1/AP-1 at the uPA/Plau locus in aggressive breast cancer. *Nucleic Acids Res.* **42**, 11011–11024 (2014).

159. Li, Z. *et al.* Overexpressed PLAU and its potential prognostic value in head and neck squamous cell carcinoma. *PeerJ* **9**, e10746 (2021).

160. Harbeck, N. *et al.* Clinical Utility of Urokinase-Type Plasminogen Activator and Plasminogen Activator Inhibitor—1 Determination in Primary Breast Cancer Tissue for Individualized Therapy Concepts. *Clin. Breast Cancer* **3**, 196–200 (2002).

161. Choi, J. W. *et al.* Urokinase Exerts Antimetastatic Effects by Dissociating Clusters of Circulating Tumor Cells. *Cancer Res.* **75**, 4474–4482 (2015).

162. Al-Othman, N. *et al.* Role of CD44 in breast cancer. *Breast Dis.* **39**, 1–13 (2020).

163. Hehlgans, S., Haase, M. & Cordes, N. Signalling via integrins: Implications for cell survival and anticancer strategies. *Biochim. Biophys. Acta - Rev. Cancer* **1775**, 163–180 (2007).

164. Hu, T., Zhou, R., Zhao, Y. & Wu, G. Integrin α6/Akt/Erk signaling is essential for human breast cancer resistance to radiotherapy. *Sci. Rep.* **6**, 33376 (2016).

165. Tanaka, S. *et al.* In silico analysis-based identification of the target residue of integrin α6 for metastasis inhibition of basal-like breast cancer. *Genes to Cells* **24**, 596–607 (2019).

166. Brooks, D. L. P. *et al.* ITGA6 is directly regulated by hypoxia-inducible factors and enriches for cancer stem cell activity and invasion in metastatic breast cancer models. *Mol. Cancer* **15**, 26 (2016).

167. Meyer, M. J. *et al.* CD44 pos CD49f hi CD133/2 hi Defines Xenograft-Initiating Cells in Estrogen Receptor–Negative Breast Cancer. *Cancer Res.* **70**, 4624–4633 (2010).

168. Mostert, B. *et al.* CD49f-based selection of circulating tumor cells (CTCs) improves detection across breast cancer subtypes. *Cancer Lett.* **319**, 49–55 (2012).

169. Li, T. *et al.* Ribociclib (LEE011) suppresses cell proliferation and induces apoptosis of MDA-MB-231 by inhibiting CDK4/6-cyclin D-Rb-E2F pathway. *Artif. Cells, Nanomedicine, Biotechnol.* **47**, 4001–4011 (2019).

170. Dai, M. *et al.* CDK4 regulates cancer stemness and is a novel therapeutic target for triple-negative breast cancer. *Sci. Rep.* **6**, 35383 (2016).

171. Sun, C.-C. *et al.* Comprehensive Analysis of the Expression and Prognosis for E2Fs in Human Breast Cancer. *Mol. Ther.* **27**, 1153–1165 (2019).

172. Liu, Z.-L. *et al.* Expressions and prognostic values of the E2F transcription factors in human breast carcinoma. *Cancer Manag. Res.* **Volume 10**, 3521–3532 (2018).

173. Rakha, E. A., Pinder, S. E., Paish, E. C., Robertson, J. F. & Ellis, I. O. Expression of E2F-4 in invasive breast carcinomas is associated with poor prognosis. *J. Pathol.* **203**, 754–761 (2004).

174. Khaleel, S. S., Andrews, E. H., Ung, M., DiRenzo, J. & Cheng, C. E2F4 regulatory program predicts patient survival prognosis in breast cancer. *Breast Cancer Res.* **16**, 486 (2014).

175. Yu, M. *et al.* Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. *Science (80-.).* **345**, 216–220 (2014).

176. Yates, L. R. *et al.* Genomic Evolution of Breast Cancer Metastasis and Relapse. *Cancer Cell* **32**, 169-184.e7 (2017).

177. Siegel, M. B. *et al.* Integrated RNA and DNA sequencing reveals early drivers of metastatic breast cancer. *J. Clin. Invest.* **128**, 1371–1383

(2018).