Heterogeneity and Plasticity in Metastatic Breast Cancer

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

BAPTISTE HAMELIN

Basel, 2022

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultat

auf Antrag von

Prof. Dr. Mohamed Bentires-Alj

Prof. Dr. Primo Schär

Prof. Dr. Berend Snijder

Basel, den 24.05.2022

Prof. Dr. Marcel Mayor
Dekan

Summary

Breast cancer is the most common cancer amongst women. As for most solid tumours, breast cancer-related deaths are the result of metastasis, a succession of complex steps leading to the dissemination of cancer cells and colonization of distant organs. The heterogeneity and plasticity of cancer cells facilitate the selection of clones that withstand the harsh conditions encountered during their journey as circulating tumour cells and invading a foreign microenvironment. While the main steps of the metastatic cascade have been described, the precise cellular and molecular mechanisms underlying each of its steps remain poorly understood.

In my PhD studies, I aimed to identify metastasis drivers in breast cancer using *in vivo* models of metastasis and RNA-Seq. This overarching goal materialised into two different projects.

The first part of my work leveraged bulk RNA-sequencing of matched primary tumours and metastases in different sites to identify molecular processes potentially involved in metastasis. Our initial findings identified glucocorticoid signalling as promoting lung metastasis. NR3C1, also known as the glucocorticoid receptor (GR), mediates the effect of stress hormones. We showed that upon stimulation, GR controls processes involved in steps of the metastatic cascade. Mechanistically this control translates into an up-regulation of the kinase ROR1, a WNT5A receptor, modulating the Wnt and Hippo pathways. We showed how GR and ROR1 ablation decrease metastasis. Our studies highlight the importance of glucocorticoid signalling in triple-negative breast cancer and call for caution when using activators of GR such as Dexamethasone to combat the side effects of chemotherapy.

In the second part of my work, we compared matched primary tumours and lung metastases of different breast cancer PDX models using single-cell RNA-seq. This technique allowed us to assess potential mechanisms driving metastases at the level of single cells, probing biological processes occurring in specific subpopulations of cancer cells. Our data highlight the significant inter and intra-patient heterogeneity of breast cancer. We found that cell populations in primary tumours and lung macrometastases are transcriptionally similar, but that three main phenotypes are making up lesions in both sites. Each of these phenotypes appears to be engaged in the

epithelial-to-mesenchymal transition (EMT), yet at varying levels, as indicated by canonical EMT and proliferation markers. We show that the top differentially expressed genes differentiating each of these phenotypes are controlled via partial EMT transcription factors (TFs). This finding argues for the importance of partial EMT in the metastatic process and sheds further light on the processes underlying the metastatic cascade.

Table of Contents

1. Introduction

1.1 Breast cancer

1.1.1 Breast cancer epidemiology

Breast cancer has the highest incidence worldwide, accounting for 11.7% of the 19.3 million new cancer cases reported globally in 2020(Sung *et al.*, 2021). It is the leading cause of cancer-related mortality for women, with close to 700.000 annual deaths. The incidence of breast cancer keeps increasing globally and is notably higher in high-income countries, making it a concern for healthcare systems worldwide.

Breast cancer has a good prognosis, assuming an early diagnosis is made and appropriate treatment is available. Unfortunately, the heterogeneity of this disease, which includes several subtypes, means that not all breast cancer patients can benefit from an equally efficacious therapeutic regimen. Resistance to the standard of care and progression of the disease is to be expected of the most aggressive subtypes.

1.1.2 Breast cancer classification

Breast cancer is an umbrella denomination for a diverse group of neoplasms sharing little other than their tissue of origin. The different subtypes of breast cancers, whose names vary depending on the classification used, are diverse in morphologies, molecular phenotypes, treatments, and outcomes. Additionally, significant inter-patient and intra-tumour heterogeneity have been described, complicating patient stratification and therapeutic solutions.

Over time several classifications of breast cancer subtypes emerged following our understanding of the disease and the technologies available. Breast cancer may be classified according to histological features, expression of a set of specific molecular biomarkers or via gene expression profiling.

Breast tumour histology has been extensively studied and led to the description of 19 histological subtypes(WHO Classification of Tumours Editorial Board, 2019). Arising from the rich architecture of the mammary gland, the two main histological subtypes are ductal and lobular carcinomas(Harbeck *et al.*, 2019). The structure of origin and invasiveness of the lesion has significant implications for the staging of the disease and therapeutic choices.

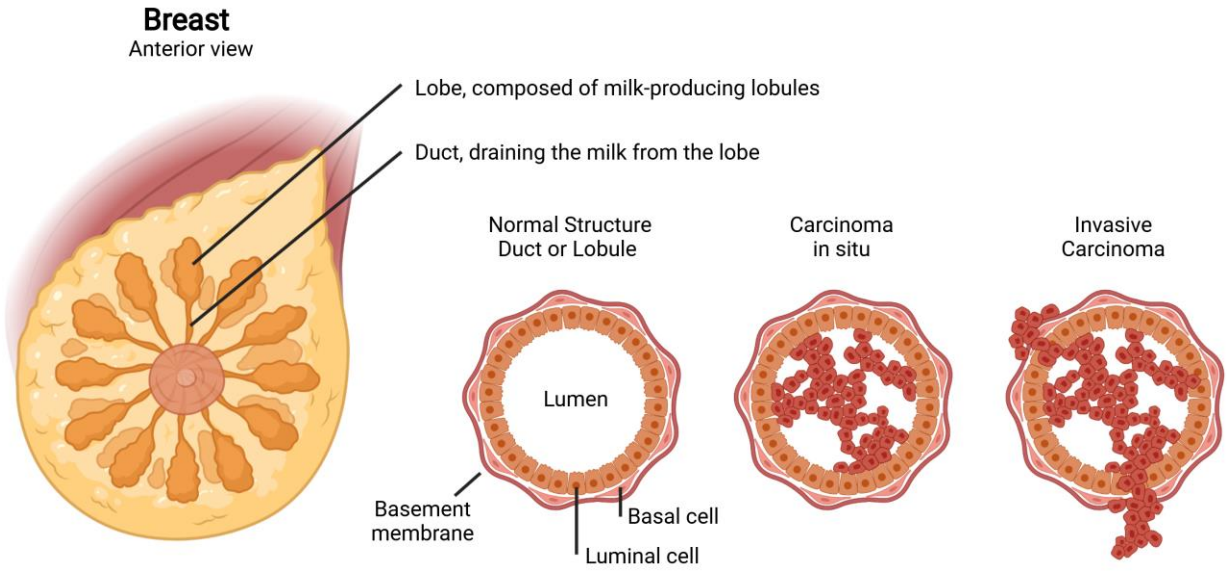


Figure 1.1 | Breast architecture and breast cancer structures of origin. The breast comprises two main elements, the milk-producing lobules, drained by ducts. As an epithelial structure, luminal cells are surrounded by a layer of basal cells and a basement membrane. When a cell acquires oncogenic traits, it starts to proliferate inside the structure, forming a restricted lesion contained by the basement membrane (Lobular or Ductal Carcinoma in situ). As the cancer cells continue to proliferate, they start the invasion process by breaching the basement membrane and spreading in the surrounding tissue (Invasive Lobular of Ductal Carcinoma).

With increasing biological insights about the disease, the characterisation of biomarkers via immunochemistry allowed to classify breast cancers according to the expression of key proteins. The Estrogen Receptor (ER), the Progesterone Receptor (PR), the human epidermal growth factor receptor 2 (HER2), and the proliferation marker Ki67 became the basis for the molecular classification of breast cancers.

Gene-profiling technologies such as microarrays and next-generation sequencing then allowed finer and unbiased molecular classification, resulting in so-called intrinsic subtypes(Perou *et al.*, 2000; Parker *et al.*, 2009).

The classification currently used in the clinic relies on these different systems to establish clinically relevant surrogate intrinsic subtypes. This classification reflects the dependencies, prognosis, and treatment options for each subtype of breast cancer:

- Luminal A

Strong co-expression of ER and PR but with low Ki67 and no HER2. Typically, low grade. Referred to as hormone-dependent, they respond to endocrine therapy targeting ER.

- Luminal B

ER-positive but at a lower level, with optional PR expression. Higher Ki67 and grade than Luminal A tumours. Still responsive to endocrine therapy, but to a lower degree. Some Luminal B may express HER2, sharing some characteristics of the HER2-enriched category.

Luminal tumours make the bulk of breast cancer cases (approx. 75%) and are regarded as cancers with a good prognosis, benefitting from an extensive range of therapies available.

- HER2-enriched

Typically, ER and PR-negative but over-expressing HER2 and with high Ki67. High grade. Their reliance on HER2 signalling allows the use of targeted therapies against this receptor.

Once known for its bad prognosis due to its aggressiveness, a better understanding of the driver of this subtype allowed the emergence of one of the first targeted therapies (Herceptin, an anti-HER2 antibody). It drastically improved the outcome of the patients affected by this subtype(Siddhartha Mukherjee, 2010).

- Triple-Negative

Characterised by their non-expression of ER, PR, nor HER2 associated with high Ki67. Regarded as the most aggressive subtype, with the poorest prognosis. The lack of understanding of the mechanisms driving this subtype limits first line therapeutic options to chemotherapy.

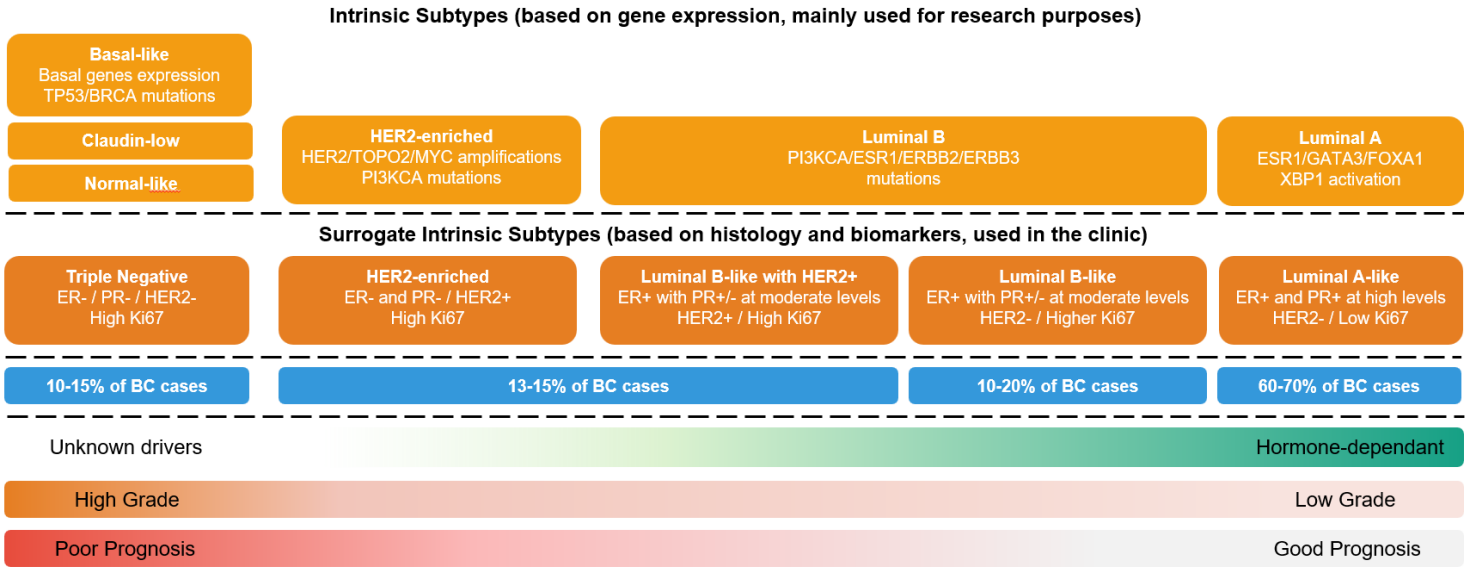


Figure 1.2 | Overview of breast cancer subtypes. Intrinsic and intrinsic surrogate subtypes, with their main characteristics (genomic alterations, biomarkers expression), frequencies, hormone-dependency, grade, and prognosis.

1.1.3 Treatments

Breast cancer treatment is based on the surrogate subtypes and clinical parameters such as the age, menopausal status of the patient, and tumour burden. The size of the tumour and lymph node status inform the type of surgery best suited to remove the primary tumour. After decades of controversy(Siddhartha Mukherjee, 2010) breast-conserving surgery has been shown to be beneficial in terms of survival for the patients. The alternative, mastectomy, is still indicated for

the most advanced tumours(Cardoso *et al.*, 2019). Breast surgery may be preceded by neo-adjuvant therapy to reduce tumour size ahead of surgery.

After surgery, systemic adjuvant treatment intended to treat potential metastases is administered, based on the surrogate subtype classification and their molecular drivers. Luminal subtypes are treated with endocrine therapies targeting ER. A whole range of therapies based on different mechanisms targets ER and its signalling. Selective estrogen receptor modulators (SERMs) and degraders (SERDs) such as Tamoxifen and Fulvestrant, respectively, are small molecule inhibitors. Tamoxifen is a competitive inhibitor of ER, preventing estrogen to bind and trigger downstream signalling. Fulvestrant induces ER degradation upon binding. Postmenopausal patients may be prescribed aromatase inhibitors in various settings. These drugs prevent estrogen production by the adipose tissue, slowing the proliferation of hormone-dependent luminal breast cancers. This subtype is usually not treated with chemotherapy, except for the most advanced (high tumour burden) or aggressive (unusually high Ki67) cases.

HER2-overexpressing breast cancers are treated with Herceptin, a monoclonal antibody directed against the product of the ERBB2 gene commonly amplified in this subtype. Different regimens of chemotherapy are usually associated with this treatment.

Finally, TNBCs are treated solely with first-line chemotherapy due to the lack of identified drivers to target. This partially explains their poor prognosis, which HER-2 positive patients once shared until the advent of Herceptin. Fortunately for TNBC patients new therapeutic solutions are being trialled.

New therapies have been emerging for all subtypes in recent years. Luminal breast cancer patients may now benefit from CDK4/6 inhibitors targeting kinases involved in the cell cycle or PI3Kα inhibitors targeting the phosphoinositide 3 kinase, the product of one of the most mutated gene (PI3KCA) in breast cancers. Advanced TNBC patients with BRCA1/2 mutations are eligible to PARP inhibitors targeting DNA repair enzymes, and recent trials have shown promising results for PARP inhibitors as first line treatment of chemotherapy naïve TNBCs(Barchiesi *et al.*, 2021; Eikesdal *et al.*, 2021). A range of antibody-drug conjugates (ADCs) are also in development, with a Trop2-targeted antibody linked to a topoisomerase being recently granted accelerated approval

by the FDA for metastatic TNBC treatment(Bardia *et al.*, 2019; Bianchini *et al.*, 2022). Immunotherapy, such as PD-1/PD-L1 checkpoint inhibitors, have shown great success in other cancer types over recent years. One such treatment, Atezolizumab, has been trialled with encouraging yet mixed results on patients with PD-L1 positive TBNCs(Tarantino *et al.*, 2022).

With increasing knowledge about breast cancer and its microenvironments and new innovative treatments such as immunotherapies or ADCs, the management of metastatic breast cancer will hopefully improve in the coming years(Bianchini *et al.*, 2022).

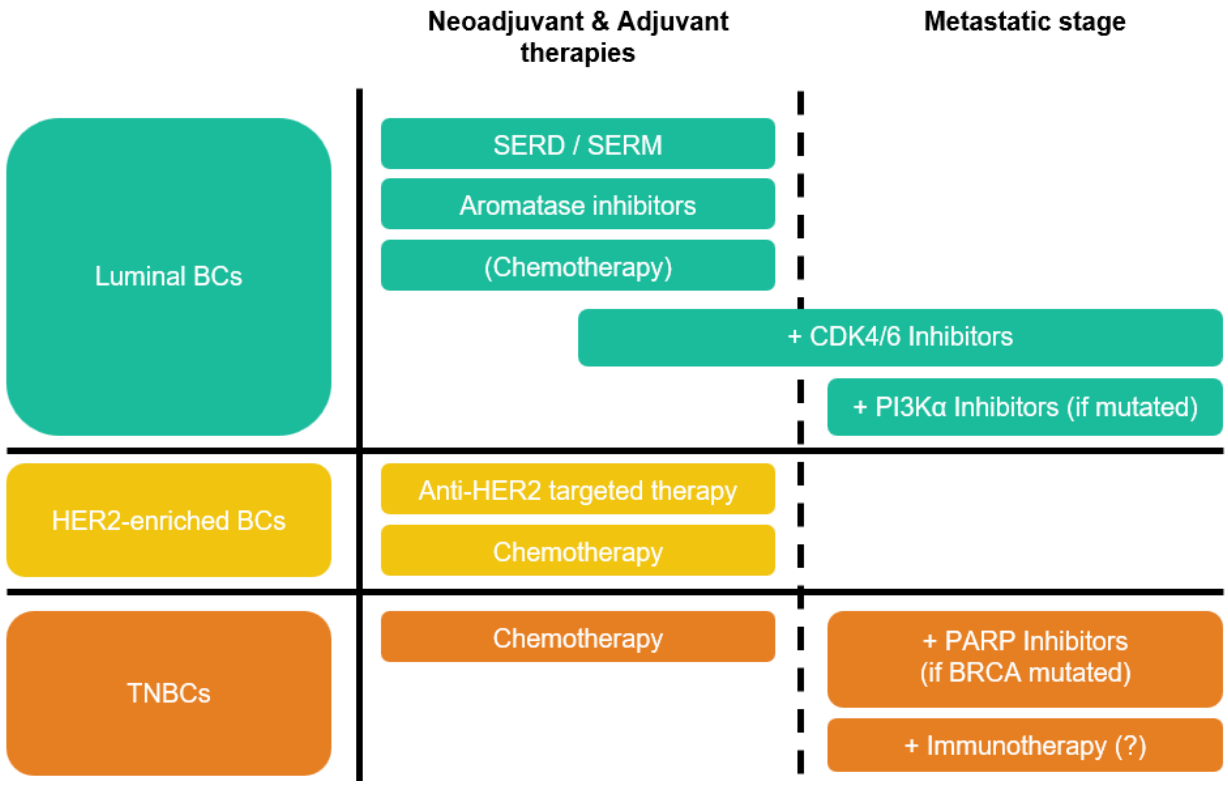


Figure 1.3 | Overview of treatment options for breast cancers. While Luminal and HER2-enriched subtypes benefit from targeted therapies in the neo-adjuvant and adjuvant settings, TNBCs are solely treated with chemotherapy in first line. Treatment of metastatic disease benefit from additional drugs but remains incurable. Metastatic patients usually also receive a variant of the first line therapy, to counter the emergence of resistance and fully exploit the potential of a drug class (example: switch between ER-targeting drugs or use of docetaxel instead of paclitaxel).

1.2 The Metastatic Cascade

In developed countries breast cancer primary tumours are generally well managed. The combination of early detection achieved through regular screening of at-risk demographics and access to proper treatments allow for early surgical removal of the tumour, usually associated with neoadjuvant and/or adjuvant treatments. As for most solid tumour type, cancer-associated mortality results from metastasis (Gupta and Qin, 2003; Dillekås, Rogers and Straume, 2019). Patients commonly develop metastases in the lungs, liver, brain, and bones(Jin and Mu, 2015) as a result of a dissemination process spanning over years(Hosseini et al., 2016).

The metastatic cascade is a complex and inefficient succession of specific events(Massagué and Obenauf, 2016). Cancer cells leave the primary tumour by invading their surrounding micro-environment, intravasate into the encountered vasculature and disseminate to a distant organ as circulating tumour cells in the bloodstream. Once at destination they extravasate into a foreign micro-environment where they may stay dormant for years(Sosa, Bragado and Aguirre-Ghiso, 2014) before eventually colonizing the organ by forming micro- then macrometastases(Baumann, Auf der Maur and Bentires-Alj, 2022). At each of these steps cancer cells display different properties enabling them to disseminate, survive cellular stress, evade the immune system, and then eventually proliferate. The plastic properties of cancer cells will allow a minority of them to form metastases.

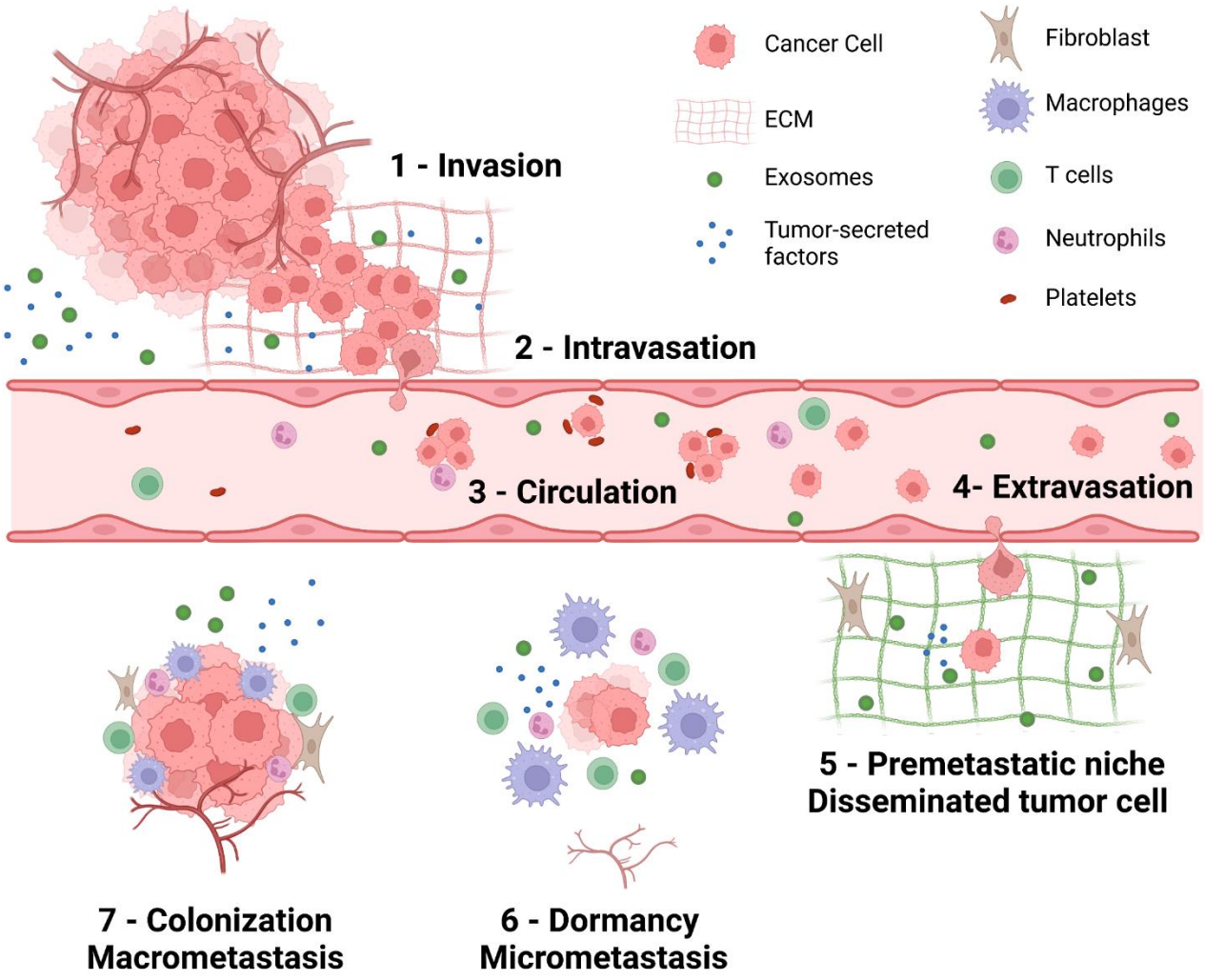


Figure 1.4 | Overview of the metastatic cascade. Steps a cancer cell go through during its dissemination to a distant site. Cancer cells start by invading (1) the surroundings of the primary lesion, eventually encountering vasculature. Upon intravasation (2) the cancer cells are taken away as single or clusters of circulating tumour cells (3), assisted by other cell types. Once arrived at a distant site, they extravasate into the host organ (4), which may have been turned into a favourable premetastatic niche by tumour-secreted exosomes (5). Depending on the predispositions of the disseminated cancer cells and their interactions with the micro-environment, micro- then macrometastases may arise (5-6).

1.2.1 Local invasion and intravasation

As primary tumours grow, cancer cells progressively invade the surrounding microenvironment, rupturing the basement membrane containing the initial neoplastic lesion. This invasion is led by cancer cells undergoing EMT, which confers them motile properties through cytoskeletal rearrangements and extracellular matrix degradation. Primary tumours also recruit their microenvironment which favour angiogenesis via secretion of angiogenic factors. This process, also termed angiogenic switch(Gupta and Qin, 2003), establishes a vasculature dedicated to supplying the primary tumour with oxygen and nutrients. Eventually invasive cancer cells will encounter one of these blood vessels and penetrate it. This step is referred to as intravasation and consist in the cancer cell breaching the endothelium and detaching into the circulation(Fares *et al.*, 2020). This process is the first of a stressful series of events, as the cells experience significant mechanical stress that can damage the nuclear membrane, leading to DNA damage or even cell death(Denais *et al.*, 2016).

1.2.2 Circulation to distant organs

Once in the vasculature, cancer cells are referred to as circulating tumour cells (CTCs). CTCs may travel alone or in clusters. These clusters have been shown to increase metastases formation compared to single CTCs (Aceto *et al.*, 2014), and may explain the heterogeneity of metastases founded by polyclonal clusters. CTCs clusters have also been shown include other cell types, such as neutrophils (Szczerba *et al.*, 2019) and platelets(Gay and Felding-Habermann, 2011). These cell-cell interactions further increase the heterogeneity and metastatic potential of the clusters by protecting them from various threats via immune suppression and structural support(Fares *et al.*, 2020). CTCs are also supported by chemo- and cytokines alleviating cell death and oxidative stress generated by fluid shear stress (Fares *et al.*, 2020; Park, Brown and Kim, 2020). Despite many factors protecting CTCs, their journey in the circulation leads to high attrition and is a limiting step in the metastatic cascade.

Additionally, CTCs may also circulate via the lymphatic system. Lymph node metastases are frequent and used for diagnostic and staging purposes. Upon tumour removal, surgeons may excise the tumour-draining lymph nodes to assess metastatic dissemination in a patient and inform treatment.

1.2.3 Arrest and Extravasation

CTC circulation is arrested in a distant organ, generally due to the reduction in size of micro-capillaries. This blockade may lead to capillary rupture seeing cancer cells passively penetrating the organ. The active process, called extravasation, consists in CTCs once again exploiting the migratory properties conferred by EMT and going through the endothelium, invading the surrounding tissue. It should be noted that only 0.01% of CTCs successfully infiltrate secondary organs(Massagué and Obenauf, 2016).

A much-discussed topic is the organotropism of cancer cells to specific distant organs(Baumann, Auf der Maur and Bentires-Alj, 2022). Part of the tropism towards frequent metastatic sites such as the bones and the liver is explained by the high permeability of the capillaries in these organs (Fouad and Aanei, 2017). However, the main factor in cancer cells being able to form metastasis in a distant organ lies in the next step of the cascade, the colonization step, which sees disseminated tumour cells proliferate or not in the microenvironment where they are extravasated.

1.2.4 Colonization

The colonization step is the final step of the metastatic cascade. It leads to the establishment of micro- then macrometastases in a distant organ, progressively impairing its functions and leading to the death of the patient.

As with the other steps of the cascade, cancer cells face arduous conditions in the foreign microenvironment they try to colonize. They may fail to do so, going through apoptosis due to the

stressful conditions they are exposed to (and potentially not adapted to). They may also fail to evade the immune system in their new host organ, which may be more surveyed by the immune system than the primary tumour was. Disseminated tumour cells might also not have the required support (vascularization, nutrients, oxygen) necessary for proliferation.

In recent years studies have shown that disseminated cancer cell may also go dormant instead of dying or starting to proliferate into a macrometastases. This dormant state, quiescent and reversible state is referred to as cellular dormancy. Another type is organ dormancy in which there is a balance between slow proliferation and cell death due to stress or the immune surveillance(Sosa, Bragado and Aguirre-Ghiso, 2014; Correia *et al.*, 2021; Baumann, Auf der Maur and Bentires-Alj, 2022). Combined with early dissemination, dormancy is considered as a major therapeutic hurdle. Cancer cells may disseminate early, before primary tumour detection and treatment, and stay dormant at a distant site for years if not decades. Hardly detectable, they are also potentially well suited to resist systemic therapies targeting cycling cancer cells such as chemotherapy. Their awakening, still poorly understood, may result from a favourable balance of different signalling pathways, immune evasion as less immunogenic clones are selected over time, decreased immune surveillance (ageing), or inflammatory conditions leading to increased proliferation(Albrengues *et al.*, 2018; Correia *et al.*, 2021). Fortunately, cancer cell dormancy also represents a therapeutic opportunity, as drugs maintaining cancer cells dormant (potentially by modulating their microenvironment) could slow down progression.

The fact that cancer cell colonization of a distant site strongly depends upon their adaptation to the conditions in this site, led to the seed and soil hypothesis(Langley and Fidler, 2011). The emergence of metastases (and by extension metastatic patterns) results from favourable interactions between seeds (cancer cells) and soil (the organ microenvironment). This concept has been illustrated by studies showing that a premetastatic niche may be primed for the seeds by factors secreted by the primary tumour, such as exosomes(Fares *et al.*, 2020). These soluble factors prepare a favourable soil, the premetastatic niche (Hanahan and Weinberg, 2011), by coercing several microenvironmental compartments. Fibroblasts, hepatocytes, endocytes, and even astrocytes have been shown to be hijacked by cancer cells to support subsequent cancer cell homing and colonization (Baumann, Auf der Maur and Bentires-Alj, 2022).

Assuming cancer cells find suitable conditions or adapt to them, the stemness of CSCs and plasticity conferred by the MET allow them to display proliferative properties. A few DTCs will proliferate into a micrometastasis, which then repeats the angiogenic process described in the primary tumour and fuel the development of the lesion. Micrometastases will then grow into macrometastases as they colonize the host organ.

Cancer cells found in metastases are often therapy-resistant since they evaded the cytotoxic treatments given upon primary tumour detection and removal. The heterogeneity of metastatic cancer cells, that disseminated to different organs and adapted to different microenvironments contributes further to the difficulty of treating metastatic disease.

1.3 Cancer properties relevant to metastasis

1.3.1 Hallmarks of cancer

In a series of landmark publications D. Hanahan and B.Weinberg (Hanahan and Weinberg, 2000, 2011; Hanahan, 2022) defined a set of biological capabilities acquired during the development of cancer. These key aspects reflect the ability of cancer to proliferate uncontrollably, invade its surroundings, evade the immune system, and recruit other cell types which fuel their growth. More recent publications expanded the hallmarks and added enabling characteristics as our knowledge of cancer progressed. Taken together, they represent the phenotype of a given cancer cell to resist death, proliferate, invade, and eventually form metastases depending on the mutations it carries, the epigenetic programs governing its gene expression and the external stimuli it receives.

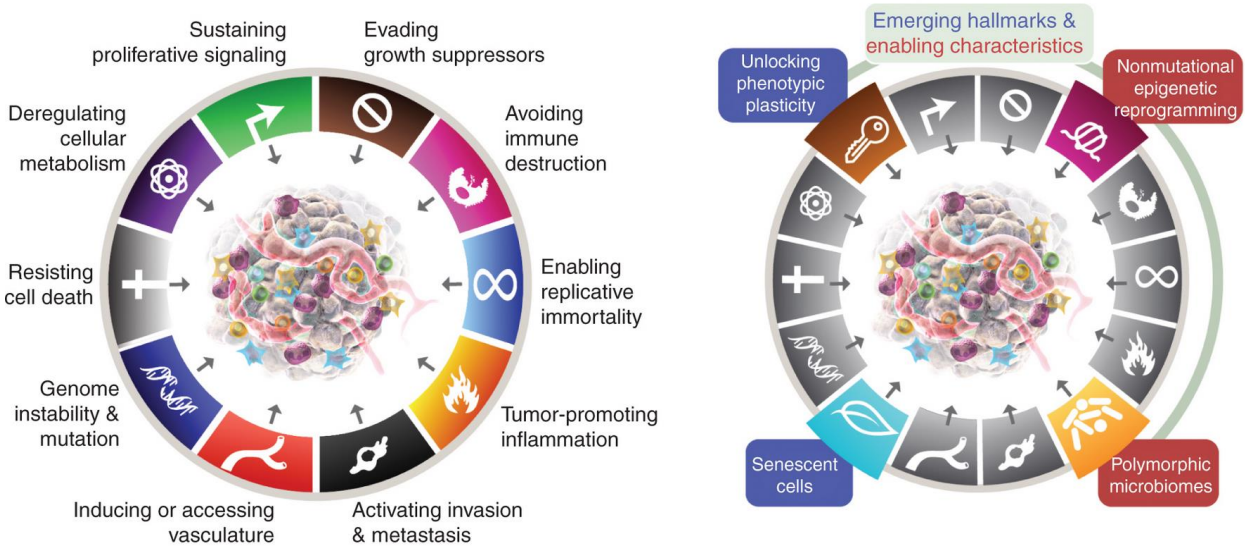


Figure 1.5 | The hallmarks of cancer. (Hanahan, 2022) The most up-to-date hallmarks, supplemented with enabling characteristics. The influence of epigenetics and extrinsic factors (immune system, microbiome) were recognized as increasingly important to cancer over the years.

1.3.2 Heterogeneity

Breast cancer heterogeneity is multifaceted and a significant hurdle for therapy. Heterogeneity can manifest at a patient's level (intra-patient and intra-tumour heterogeneity) or between patients (inter-patient heterogeneity).

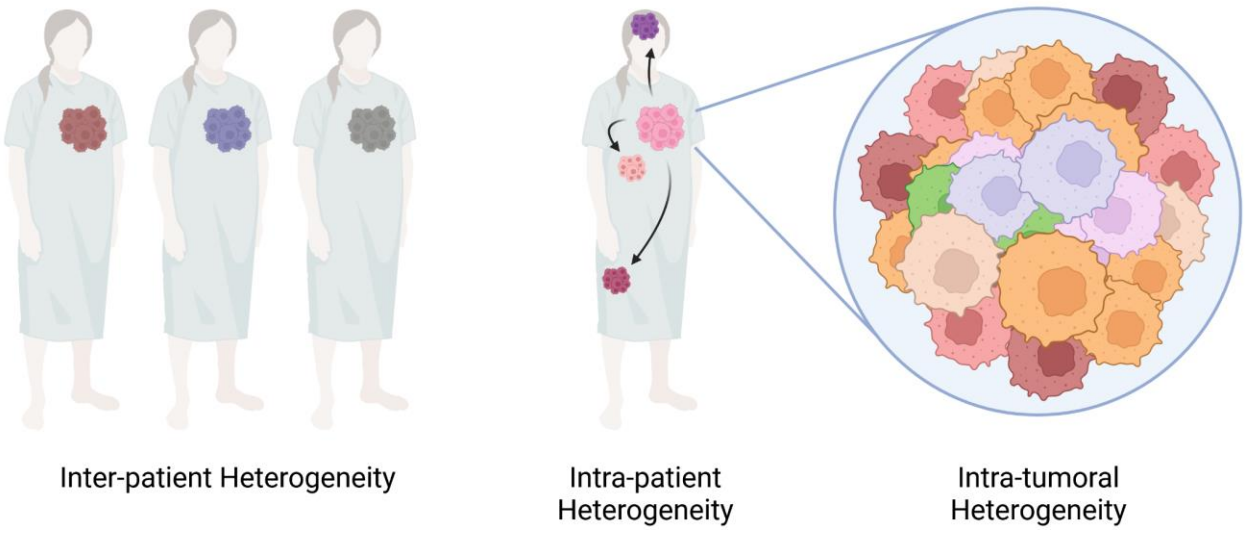


Figure 1.6 | Heterogeneity of breast cancers. Breast cancers vary at different levels, complicating characterization, stratification of patients and ultimately treatment. For example, different patients suffering from a given BC subtype will exhibit different mutations (inter-patient variability). Metastases will adapt to the distant site they disseminated to, adapting their gene expression and metabolism (intra-patient heterogeneity). A given tumour or metastasis is comprised of cancer cell clones sharing a common ancestor but displaying minor genetic and epigenetic alterations leading to different phenotypes (intra-tumoral heterogeneity).

Breast cancer cells display molecular and phenotypic diversity. This results from genetic and non-genetic factors. Cancer being a genetic disease, genomic alterations have been extensively studied. Tumours arise from the accumulation of mutations over many cell cycles. Over time driver (starting and benefiting tumorigenesis) and passengers (mutations without advantages for the cancer cell) mutations emerge and accumulate. They progressively confer hallmarks capabilities to a normal cell, transforming it into a cancer cell. A Darwinian-like selection process leads to the emergence of subpopulations of cancer cells with similar yet different

mutational backgrounds(Black and McGranahan, 2021). These alterations lead to clones with different proliferative, invasive, and metastatic potentials and may enable resistance to targeted therapies. While all these cells share the hallmarks of cancer, their slight differences contribute to the heterogeneity of the disease. Clonal diversity is especially problematic for diagnosis as a single biopsy may not represent the rest of the tumour or metastatic sites(Koren and Bentires-Alj, 2015). This is of particular importance if a targeted therapy aimed at a specific mutation which might not be present in all clones is administered to a patient.

The diverse mutational landscape of breast cancers is also challenging in terms of driver mutations. While breast cancer drivers have been identified (TP53, PI3KCA, MYC, PTEN, CCND1, for example), their mutations remain relatively rare (Pereira *et al*, 2016). This results in patients with very heterogeneous mutational profiles, where the cumulative effects of multiple low penetrance mutations drive the disease (rather than select high penetrance, strong effect mutations). The identification of breast cancer driver mutations has nonetheless opened the door to targeted therapies aimed at PI3K signalling and the DNA repair mechanism (in the case of BRCA1/2 mutations).

In recent years, the tumour and metastases clonal dynamics have been extensively studied, suggesting that genomic alterations are not the main driver of the metastatic process. Metastases-private mutations were seldom found, and actionable mutations were identical in primary tumours and metastases(Hu *et al.*, 2020),(van de Haar *et al.*, 2021).

Non-genetic heterogeneity in breast cancer is also significant and increasingly regarded for its importance in disease progression and therapy. One of the first aspects to consider is epigenetic heterogeneity. Epigenetic regulation of gene expression is one of the main drivers behind phenotypic diversity. It allows cells sharing a single genome to differentiate into different cell types and fulfil different functions in an organism. Such an important mechanism and its roles in cancer are to be considered. DNA and histones modifications (such as methylation or acetylation) are prime examples of epigenetic modifications occurring during cancer. They modulate chromatin conformation and DNA accessibility, mediating gene activation or repression, thereby contributing to phenotypic diversity(Harbeck *et al.*, 2019).

Finally, the microenvironment surrounding cancer cells is also a contributor to cancer heterogeneity. During tumorigenesis, cancer cells recruit and thrive among stromal (such as cancer-associates fibroblasts (CAF)) and immune cells making up a significant part of the tumour. The recruitment of these cells to the lesion lead to a disorganization of the normal tissue architecture and abnormal conditions. The extracellular matrix (ECM) is remodelled by CAFs, and immune cells may kill cancer cells, or lead to inflammatory conditions benefiting them. Disruption of normal vasculature disturbs nutrients and oxygen supply, leading to hypoxic regions in the tumour. The immune compartment, initially keeping in check the tumour, may eventually support its growth via the production of specific cytokines or attraction of immune-suppressive immune cell types (Fares *et al.*, 2020). In other cases, immune cells may be entirely excluded from the tumour, leading to "immune cold" lesions(Yao, Li and Wang, 2022). The microenvironment, supporting or attacking cancer cells, contributes to the overall heterogeneity.

All these factors, intrinsic and extrinsic to cancer cells, result in significant heterogeneity taking the form of phenotypic diversity.

1.3.3 Plasticity

Another key trait of cancer cells is their phenotypic plasticity. This concept illustrates the ability of cancer cells to reversibly switch between different cellular states, each with distinct properties(Jehanno *et al.*, 2022). This switch is vital as it has been shown that cancer cells need different properties (i.e., proliferation, motility) at different stages of the disease.

While the precise nature and characteristics of the cellular states in question are still debated, they concern at least two main mechanisms.

1.3.3.1 Stemness

The first is the differentiation state of the cells and stemness properties of so-called stem cells. In normal tissues such as epitheliums, tissue homeostasis is maintained by stem cells able to divide into more differentiated cell types. These stem cells can exist in different states, regenerate the tissue, or remain quiescent if the tissue renewal is slow or not needed. Studies have suggested the existence of cancer stem cells (CSCs), characterized by their ability to self-renew to give rise to a new tumour in serial limiting dilution experiments(Clarke et al., 2006;Nguyen et al., 2012). This suggests that not all cancer cells have stemness potential, but maybe be in a more differentiated state (non-CSC) preventing them from proliferating. Similar to quiescent stem cells, cancer cells have been shown to exist in a dormant state after disseminating into a distant organ(Risson et al., 2020;Correia et al., 2021). These cells would then awaken and give rise to metastases, demonstrating their stemness potential. However, the nature of CSCs is still debated. While cell surface markers defining CSC populations have been described they have proven highly variable depending on the setting (cancer type and subtype). Additionally, cell populations expressing none of these markers have also been shown to have tumour-initiating potential in different studies(Tang *et al.*, 2012).

1.3.3.2 EMT

The second mechanism comprising different cellular states that are exploited by cancer is the epithelial-to-mesenchymal transition (EMT, and conversely MET). This process is heavily involved in embryogenesis and wound healing. It is a complex cellular program resulting in the epithelial and mesenchymal phenotypes. During embryonic development or wound healing, cells reorganize to form new structures or repair injured ones. Genes that are part of the EMT program encode for proteins involved in cell-to-cell interactions (tight/adherens/gap junctions, desmosomes), cytoskeleton and ECM-related elements. When epithelial cells undergo EMT, epithelial genes are repressed, and mesenchymal ones are activated. A newly mesenchymal cell hereby acquires motility as cell-to-cell contact molecules are down-regulated, the cytoskeleton

reorganized for migration (formation of filopodia, invadopodia) and the ECM sensed via integrins and reorganized via matrix metalloproteases (MMPs)(Lamouille, Xu and Derynck, 2014). Once the cell homes in a distant organ it can revert to the epithelial state, re-establishing contact with surrounding cells. While vital for the development of structures in the embryo, or repair of an injury, it is easy to see how EMT may be of similar importance for cancer and the metastatic process. Thereby its role has been extensively studied in this context and found to be a major source of phenotypic plasticity and stemness for cancer cells (Polyak and Weinberg, 2009; Pastushenko and Blanpain, 2019; Lüönd, Tiede and Christofori, 2021). EMT and consecutive MET have been shown to be necessary for metastatic seeding(Ocaña *et al.*, 2012; Wei *et al.*, 2015). Induction of EMT leads to increased CTCs(Ye *et al.*, 2015; Xu *et al.*, 2017). These findings highlight the importance of EMT for dissemination. But cells with a mesenchymal state show little stemness and are not able to proliferate. The reverse process, MET, allows the reacquisition of epithelial traits and the proliferation of cancer cells, effectively seeding metastases in a distant organ. A better understanding of EMT, the TFs responsible for it, and markers associated with its two extreme states lead to the concept of partial EMT. Cancer cell populations co-expressing both epithelial and mesenchymal markers were found by different studies (Strauss *et al.*, 2009; Yu *et al.*, 2013). These hybrid states, laying on a continuum between the epithelial and mesenchymal, were found to be associated with stemness, aggressiveness, and poor outcome(Grosse-Wilde *et al.*, 2015; Smith and Bhowmick, 2016; Yamashita *et al.*, 2018).

The intricacies of partial EMT in cancer cells or the existence of malignant stem cells remain to be fully understood. However, the plastic properties of cancer cells, demonstrated in the light of these concepts, are tangible and represent major mechanisms in the progression of the disease.

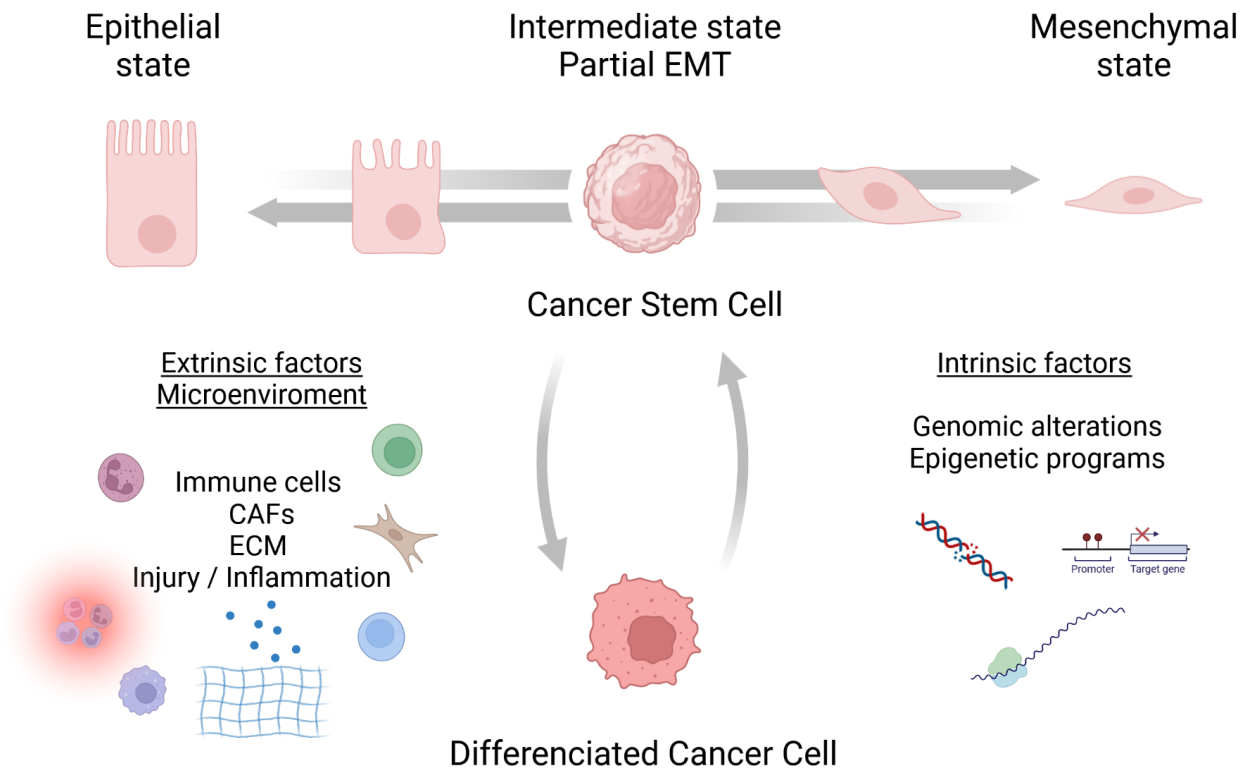


Figure 1.7 | Plasticity in cancer. The ability of cancer cells to display different phenotypes and switch between them is pivotal for cancer progression. Different cellular states are adapted to different conditions, allowing cancer cells to withstand the harsh conditions of the metastatic cascade and eventually result in metastasis. These different states are influenced by many factors, intrinsic to the cancer cells or mediated by external stimuli.

2. Rationale of the work

Breast cancer is the most frequent and lethal cancer type among women worldwide(Sung *et al.*, 2021). Primary tumour surgery and treatments allow good management of the initial lesion. However, breast cancers (TNBCs in particular) eventually progress to the metastatic stage and metastases are the main cause of cancer-related deaths. The origin of cancer is genetic, with the sequential accumulation of genomic alterations leading to the acquisition of malignant properties (Hanahan & Weinberg, 2000, 2011). While clonal dynamics studies have shown that cancer cells keep accumulating genomic alterations during the metastatic process, they have failed to clearly identify genomic drivers of progression. Instead, epigenetic mechanisms of adaptation, and overall heterogeneity and plasticity of cancer cells are suspected to be the main culprits behind metastatic disease. Studies have suggested that the switch between different cellular states/phenotypes and the pathways supporting such states, are vital for metastatic dissemination and colonization. However, the underlying cellular and molecular mechanisms enabling and supporting metastasis are still only partially understood.

To identify potential drivers of metastasis we performed unbiased, *in vivo* RNA-Seq screens comparing the gene expression of primary tumours and matched metastases of TNBC models.

In a first project, using bulk RNA-sequencing, we identified the glucocorticoid receptor (GR, product from the NR3C1 gene) as a driver of lung metastases in TNBC.

For a second project, we performed single-cell RNA-Sequencing in a similar setting. The deeper insights permitted by single-cell sequencing allowed us to characterize single cell states and compare cancer cells against each other, in primary tumours and metastases.

Specifically, we addressed the following points:

- What is the effect of glucocorticoid signalling in breast cancer metastasis?
- What is the single-cell landscape of metastatic breast cancer?

3. Results Part I: The effect of the Glucocorticoid receptors in breast cancer metastasis

Research Article - Published in *Nature*

Glucocorticoids promote breast cancer metastasis.

Milan M. S. Obradović[1,2,7], Baptiste Hamelin[1], Nenad Manevski[3,8], Joana Pinto Couto[1,2,9], Atul Sethi[1,2,4], Marie-May Coissieux[1,2], Simone Münst[5], Ryoko Okamoto[1,2], Hubertus Kohler[2], Alexander Schmidt[6] & Mohamed Bentires-Alj[1,2*]

[1] Department of Biomedicine, Department of Surgery, University Hospital Basel, University of Basel, Basel, Switzerland.

[2] Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland.

[3] Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, University of Helsinki, Helsinki, Finland.

[4] Swiss Institute of Bioinformatics, Basel, Switzerland.

[5] Institute of Pathology, University Hospital Basel, University of Basel, Basel, Switzerland.

[6] Proteomics Core Facility, Biozentrum, University of Basel, Basel, Switzerland.

[7] Present address: Wellmera AG, Basel, Switzerland. [8]Present address: UCB Celltech, Development Sciences, Slough, UK. [9]Present address: Novartis Institutes for BioMedical Research, Basel, Switzerland.

*e-mail: m.bentires-alj@unibas.ch

3.1 Abstract

Diversity within or between tumours and metastases, known as intra-patient tumour heterogeneity, that develops during disease progression is a serious hurdle for therapy (Almendro, Marusyk and Polyak, 2013; Polzer and Klein, 2013; Koren and Bentires-Alj, 2015). Metastasis is the fatal hallmark of cancer and the mechanisms of colonization, the most complex step in the metastatic cascade(Hanahan and Robert A. Weinberg, 2011) (Weinberg, 2014), remain ill-defined. A clearer understanding of the cellular and molecular processes underlying both intra-patient tumour heterogeneity and metastasis is crucial for the success of personalized cancer therapy. In our study, transcriptional profiling of tumours and matched metastases revealed cancer site-specific phenotypes and increased glucocorticoid receptor (GR) activity in distant metastases. GR mediates the effects of stress hormones and synthetic derivatives used widely in the clinic as anti-inflammatory and immunosuppressive agents. We show that increase in stress hormones during breast cancer (BC) progression results in GR activation at distant metastatic sites, increased colonization and, ultimately, reduced survival. Transcriptomics, proteomics and phosphoproteomics studies implicated GR in the activation of multiple processes in metastasis and in the increased expression of kinase ROR1, which correlate with shorter patient survival. Ablation of ROR1 reduced metastatic outgrowth and prolonged survival in preclinical models. Our results suggest that GR activation increases heterogeneity and metastasis, which thus calls for cautious use of glucocorticoids in the treatment of BC patients with cancer-related complications.

3.2 Increase in GR activation in distant metastases

Intra-patient tumour heterogeneity describes a poorly understood phenomenon during malignant progression by which cancer cells and patients themselves undergo genetic and epigenetic as well as hormonal and immunological changes (Marusyk, Almendro and Polyak, 2012; Almendro, Marusyk and Polyak, 2013; Meacham and Morrison, 2013; Koren and Bentires-Alj, 2015; Chaffer et al., 2016)(Meacham and Morrison, 2013). Phenotypic changes in cancer cells are a consequence of selection and adaptation that may result in cancer growth at distant sites years after primary tumour diagnosis and removal(Vanharanta and Massagué, 2013),(Obenauf and Massagué, 2015). Intra-patient tumour heterogeneity is an obstacle to treatment, spawning divergence in diagnostic markers between primary tumours and matched metastases that may lead to inadequate treatment(Lindström *et al.*, 2012). Our understanding of such global phenotypic changes arising at distant metastatic sites is still fragmentary(Alizadeh *et al.*, 2015).

To explore heterogeneity between tumours and distant metastases in clinically relevant models, we implanted 17 primary-derived xenografts (PDX) or cell lines into mammary glands of NOD-*scid IL2rγnull* (NSG) immunodeficient mice, resected the primary tumours, and monitored metastasis. Metastases were found in lung, liver, spleen, and ovaries, and as circulating tumour cells (CTC) (Fig. 3.1a, Extended Data Fig. 3.1a-e and Supplementary Table 1). To compare matched tumours and metastases, we isolated cancer cells by fluorescence-activated cell sorting (FACS) based on expression of GFP in the MDA-MB 231 model or of the human specific marker CD298(Lawson *et al.*, 2015) in PDX models (Extended Data Fig. 3.1b,e), and performed global transcriptional profiling. Principal component analysis (PCA) revealed cancer cell clusters mainly according to the site of the metastases (Fig. 3.1b and Extended Data Fig. 3.1f, g-m). The cellular processes that differed most frequently between tumours and matched metastases were hypoxia, various metabolic processes and mTOR signalling (Extended Data Fig. 3.1g-i and Supplementary Table 2).

Using an integrated system for motif activity response analysis (ISMARA)(Balwierz *et al.*, 2014), we modelled transcription factor activity in these samples and in publicly available datasets(Hosseini et al., 2016). We found a recurrent increase in glucocorticoid receptor (GR)

activity in metastases (Extended Data Fig. 3.2a-c). Ingenuity Pathway Analysis (IPA) showed correspondence between differential GR expression in tumours and matched metastases and the expression profiles evoked by glucocorticoids, such as dexamethasone (DEX) and triamcinolone acetonide (Fig. 3.1d and Extended Data Fig. 3.1g-i). In addition, distant lung metastases showed GR nuclear localization, a further indication of GR activation in metastases (Extended Data Fig. 3.2d). The levels of the stress hormones cortisol and corticosterone were higher in the plasma of mice with metastases (M1 animals) than in the healthy controls or mice with tumours but no metastases (M0) (Fig. 3.1e, f). However, genes involved in stress hormone biosynthesis were not expressed in cells isolated from tumours or metastases (Extended Data Fig. 3.2e-h), suggesting that the increases in the hormone were not generated by the cancer cells themselves. Adrenocorticotropic hormone (ACTH), which promotes the production and release of cortisol and corticosterone, was also higher in the plasma of mice with metastases (Fig. 3.1g). Increased cortisol and corticosterone levels were associated with increased tumour volume (Extended Data Fig. 3.2i-k). Taken together, our data reveal enhanced GR activity in BC metastases, most likely due to an increase in glucocorticoids during BC progression.

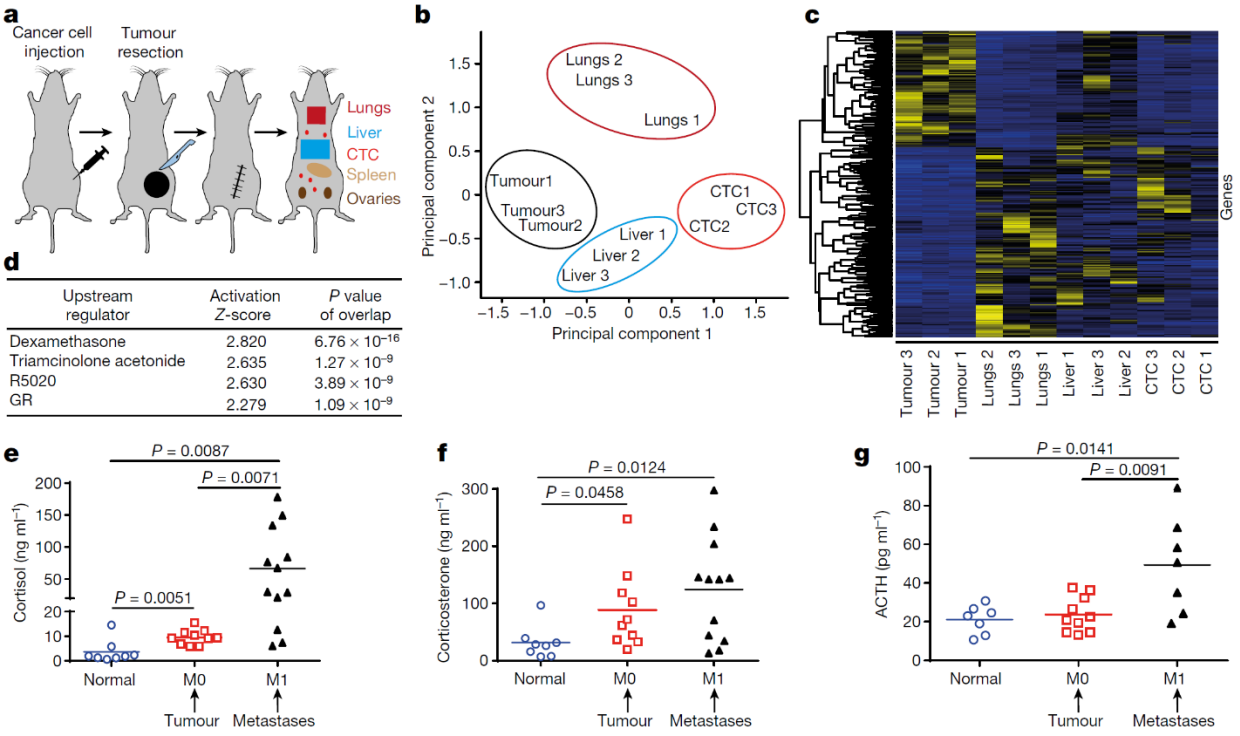


Figure 3.1 - Increase in GR activation in distant metastases. a, Experimental design of the study. PDX models or MDA-MB 231 cells were transplanted into mammary fat-pads of female NSG mice. Cells isolated from the "tumour" (tumour cells found in the mammary fat-pad after orthotopic injection/transplantation) or matched "metastases" (cancer cells found at distant sites such as lungs, liver etc. weeks after tumour resection) were sorted by FACS. b, Principal component analysis of tumours and matched metastases in the MDA-MB 231 model: clustered transcriptional profiles based on the sites of metastases. The identifiers in the PCA plot represent tumour cells isolated from the indicated sites; the numbers within the identifiers indicate a specific mouse; n=3 biological replicates (mice). c, Heatmap of genes differentially expressed in tumours, CTC, liver and lung metastases in the MDA-MB 231 model; 280 genes, yellow = upregulation, blue = downregulation, fold change≥2, FDR<0.05, n=3 biological replicates (details on statistical tests are available in the methods section). d, Ingenuity Pathway Analysis revealed GR and GR ligands as the upstream regulators of the lung metastatic phenotype, n=3 biological replicates, Fisher's Exact Test. Increased levels of e, cortisol; f, corticosterone (n=8 Normal, n=10 M0, n=12 M1 animals) and g, ACTH (n=7 Normal, n=10 M0, n=7 M1 animals) in plasma of animals prior to tumour resection (M0) and of animals with distant metastases (M1) compared to age-matched tumour-free animals (Normal). Means and single data points are represented, *n* represents biological replicates (mice), two-tailed Student's *t*-test.

3.3 GR activation escalates metastatic colonization and reduces survival

Stress hormone levels are also higher in BC patients with metastases than in age-matched healthy women or patients without metastases(Van Der Pompe, Antoni and Heijnen, 1996), and abnormal or flattened cortisol rhythms correspond with shorter survival of patients with advanced BC(Sephton, 2000). We sought to examine the cell autonomous effect of glucocorticoids on metastasis. Given the observed GR activation profile in BC metastases, we hypothesized that GR activation in cancer cells promotes metastatic colonization in lungs of experimental animals. *In vitro* GR activation of MDA-MB 231 cells by DEX resulted in the anticipated expression of GR target genes (GR signature), which was reversible upon DEX withdrawal (Extended Data Fig. 3.3a, b). We used an experimental metastases assay to test colonization capacity in which immunodeficient and immunocompetent mice were injected *i.v.* with GR-activated or control MDA-MB 231 or 4T1 cancer cells, respectively. Notably, GR activation with a clinically relevant DEX dose(Pang *et al.*, 2006) increased the colonization capacity of MDA-MB 231 and 4T1 cells (Fig. 3.2a,b and Extended Data Fig 3.3c). Next, we downregulated GR in MDA-MB 231 cells

using short hairpin RNAs targeting GR (shGR1, shGR2) (Extended Data Figure 3.3d,e). DEX treatment had no effect on the expression of GR targets in shGR cells (Extended Data Fig. 3.3f). To test whether the observed colonization is evoked by GR activation, DEX-treated shGR and control (shCTRL) MDA-MB 231 cells were *i.v.* inoculated into NSG mice. DEX resulted in metastatic escalation in shCTRL but not in shGR cells lacking GR (Extended Data Fig. 3.3g, h). The data suggest that the glucocorticoid-mediated escalation of metastasis is GR dependent. Interestingly, treatment of MDA-MB 231 cells with the GR antagonist mifepristone decreased metastatic colonization relative to vehicle-treated and GR-activated cells (Extended Data Fig. 3.3i). However, this did not substantially prolong overall survival when compared to vehicle treatment, possibly due to the shortness of the mifepristone treatment as against sustained GR downregulation by shRNA (Extended Data Fig. 3.3j). Taken together, the enhanced growth of cancer cells in lungs upon GR activation and the decreased colonization capacity of GR down-regulated cells suggest that GR activation enhances the colonization step of metastases via a cancer cell-autonomous mechanism.

Glucocorticoids are widely used in the treatment of metastatic BC to combat side-effects of chemotherapy and to treat symptoms related to advanced cancer. To address the effects of GR activation on the treatment of BC metastasis, we applied paclitaxel to animals injected *i.v.* with 4T1, shCTRL or shGR MDA-MB 231 cells. In the 4T1 model, DEX- and paclitaxel-treated Balb/c animals displayed shorter overall survival than animals that received vehicle and paclitaxel (Extended Data Figure 3.4a). In the MDA-MB 231 model, paclitaxel treatment reduced metastatic progression and survival in shGR1 and shGR2 MDA-MB 231, regardless of DEX application. Remarkably, GR activation in shCTRL-injected animals offset the paclitaxel effects, escalating metastases and decreasing survival compared to shCTRL non-DEX-treated animals (Extended Data Fig. 3.4b-d).

The experimental metastasis assay recapitulates only the last steps of the metastatic cascade(Saxena and Christofori, 2013). Therefore, we assessed the effects of DEX administration after resection of orthotopically implanted tumours and found that it precipitated metastasis and death of MDA-MB 231, PDX1 and 4T1-bearing animals (Fig. 3.2c, Extended Data Fig. 3.5a, b).

The data thus show that DEX enhanced metastases and reduced overall survival of the animals in both the experimental and the orthotopic metastasis assays.

Next, we asked whether GR signalling also affects primary tumour growth. GR downregulation had no effect on tumour volume (Extended Data Fig. 3.5c) but increased cancer cell motility as shown by two-photon intravital microscopy (Supplementary Data Videos 1-3). Moreover, GR knockdown enhanced the number of tumour cells in the circulation and in the lungs (Extended Data Fig. 3.5d, e). Yet, this increased dissemination did not translate into a decrease in overall survival of the animals (Extended Data Fig. 3.5f) presumably because of the lower colonization capacity of cells lacking GR (Extended Data Fig. 3.3f, g). Of note, orthotopic injection of DEX- or vehicle-treated cells resulted in a slight increase in tumour volume in shCTRL but not in the shGR MDA-MB 231 model (Extended Data Fig. 3.5g).

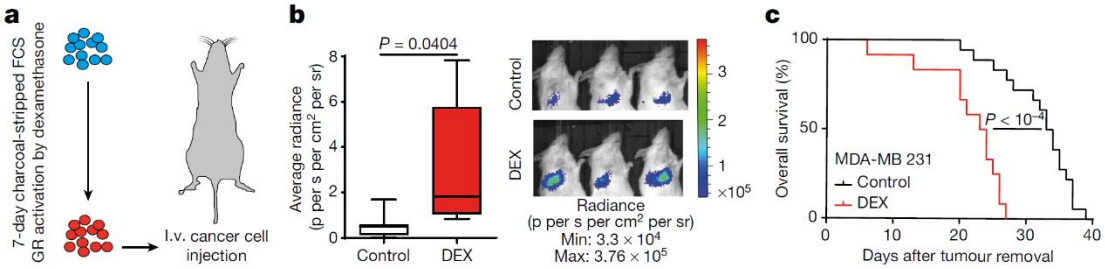
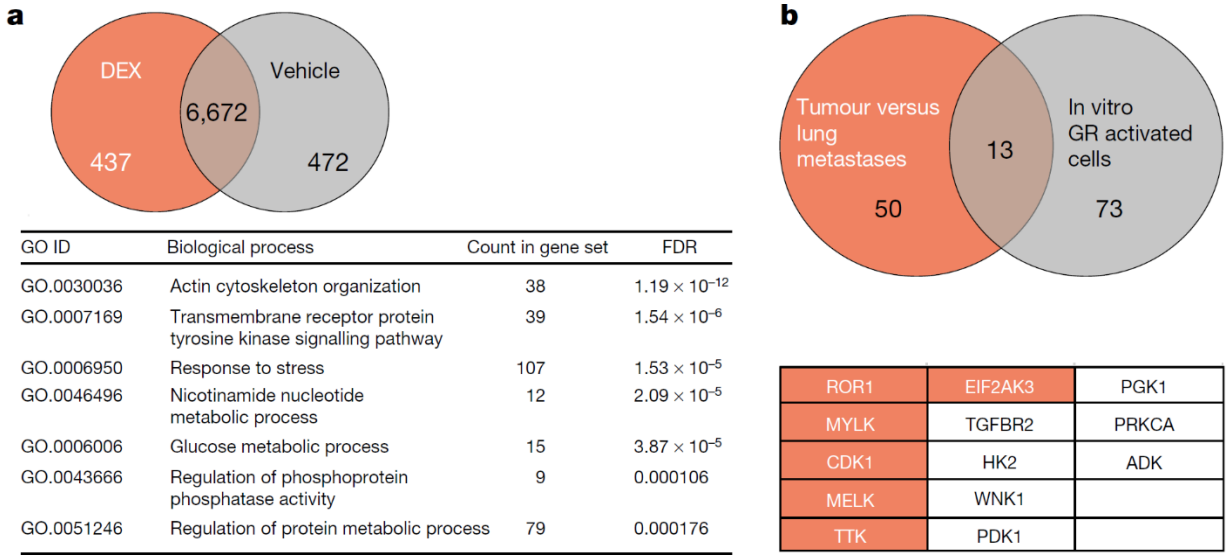


Figure 3.2. GR activation escalates metastatic colonization and reduces survival. a, Experimental design. MDA-MB 231 cells were propagated in the presence of DEX for 7 days before injection into mice. **b**, Box plot: GR activation increased lung colonization upon *i.v.* injection with MDA-MB 231 cells. Right, bioluminescence imaging two weeks after cell injection, *n*=7 mice, two-tailed Student's *t*-test. Box-plots, centre line- median, box extends from the 25th to 75th percentiles, whiskers extend to the most extreme data points. **c,** DEX administration after MDA-MB 231 tumour removal (orthotopic model) reduced survival, *n*=12 DEX and *n*=18 control (vehicle) treated mice, pooled data from 3 independent experiments, two-tailed log-rank test.

3.4 GR activation induces signalling networks and protein kinases implicated in breast cancer progression

We assessed the molecular consequences of GR activation in order to determine the mechanism causing escalated metastatic colonization in GR-activated cells. Global proteomic analysis of lysates from GR-activated cells revealed 437 up- and 472 downregulated proteins. Markers of GR activation and processes such as epithelial–mesenchymal transition (EMT), glucose and nicotinamide metabolism, cytoskeleton organization, and pathways involved in metastases (e.g., EGFR, Hippo)(Marotta et al., 2011),(Liu et al., 2018),(Li *et al.*, 2017) all increased upon GR activation. Sixty-three of the upregulated proteins were kinases, of which six were higher at the RNA level in metastases than in tumours of the MDA-MB 231 and PDX models (Fig. 3.3a,b, Extended Data Fig. 3.6a-d, Extended Data Fig. 3.7, and Supplementary Table 3). Notably, expression levels of GR-upregulated kinases in tumours of BC patients were predictive of decreased relapse-free survival (Extended Data Fig. 3.8).



GO ID	Biological process	Count in gene set	FDR
GO.0030036	Actin cytoskeleton organization	38	1.19×10^{-12}
GO.0007169	Transmembrane receptor protein tyrosine kinase signalling pathway	39	1.54×10^{-6}
GO.0006950	Response to stress	107	1.53×10^{-5}
GO.0046496	Nicotinamide nucleotide metabolic process	12	2.09×10^{-5}
GO.0006006	Glucose metabolic process	15	3.87×10^{-5}
GO.0043666	Regulation of phosphoprotein phosphatase activity	9	0.000106
GO.0051246	Regulation of protein metabolic process	79	0.000176

ROR1	EIF2AK3	PGK1
MYLK	TGFBR2	PRKCA
CDK1	HK2	ADK
MELK	WNK1	
TTK	PDK1	

Figure 3.3. GR activation induces signalling networks and protein kinases implicated in breast cancer progression. a, Proteomic analysis of DEX GR-activated and vehicle-treated MDA-MB 231 cells (upper panel) and String proteome network analysis of differentially regulated networks (lower panel) upon DEX treatment of MDA-MB 231 cells, *n*=3 "vehicle" biological replicates, *n*=4 "DEX" biological replicates, Fisher's Exact Test (details

available in the methods section). **b**, Overlap of differentially regulated protein kinases in lung metastases of the MDA-MB 231 model and *in vitro* propagated cells. List of 13 common kinases in both datasets (the 6 upregulated kinases are shown in orange).

3.5 ROR1 mediates GR-induced lung metastatic colonization

One of the kinases upregulated in transcriptome and proteome analyses was ROR1, which has been implicated previously in BC(Li *et al.*, 2017),(Cui *et al.*, 2013),(Chien *et al.*, 2016). The ROR1 signature of tumour samples showing ROR1 amplification(Lorinc Pongor *et al.*, 2015) was associated with decreased survival (Extended Data Fig. 3.9a). Transcriptome analyses of recently published BC metastasis data(Robinson *et al.*, 2017) revealed a correlation between high ROR1 levels in distant metastases and the GR activation signature (Fig. 3.4a and Extended Data Fig. 3.9b). Analysis of published BC datasets(Curtis et al., 2012),(Pereira *et al.*, 2016) indicated co-expression of GR, the GR signature and ROR1 (Extended Data Fig. 3.9c). Of note, the GR signature correlated with the claudin-low intrinsic BC subtype (Extended Data Fig. 3.9d). Furthermore, FACS analysis confirmed higher ROR1 levels in metastases of MDA-MB 231 and PDX1 than in matched primary tumours (Fig. 3.4b). DEX treatment increased ROR1 levels in MDA-MB 231 parental and shCTRL compared to vehicle-treated cells but not in shGR cells (Extended Data Fig. 3.10a). Conversely, DEX administration *in vivo* increased ROR1 levels in shCTRL but not in shGR MDA-MB 231 tumours or lung metastases (Extended Data Fig. 3.10b, c). GR activation *in vitro* enhanced the expression and abundance of the ROR1 ligand Wnt5a (Extended Data Fig. 3.10d, e). Wnt5a has been implicated in multiple processes in BC and was recently reported to be a GR target(West *et al.*, 2018). Wnt5a levels were also higher in MDA-MB 231 metastases than in primary tumours in shCTRL, shGR1 and shGR2 suggesting that it can be induced by additional means than GR activation (Extended Data Fig. 3.10f). Consistently, phosphoproteomic analysis revealed higher activities of multiple members of the Wnt and Hippo pathways in GR-activated MDA-MB 231 cells than in controls (Extended Data Fig. 3.10g and Supplementary Table 5).

Next, we addressed the effect of ROR1 knockdown (Extended Data Fig. 3.10h) on metastasis using both the experimental metastases and the orthotopic assays. Downregulation of ROR1 by two independent inducible shRNAs decreased metastasis and prolonged survival in both assays (Fig. 3.4c and d; Extended Data Fig. 3.10i). Notably, ROR1 ablation halted metastases evoked by GR activation and prevented the precipitated death of the animals (Fig. 3.4c, Extended Data Fig. 3.10j-o). Taken together, the data suggest that increased colonization upon GR activation is mediated via ROR1.

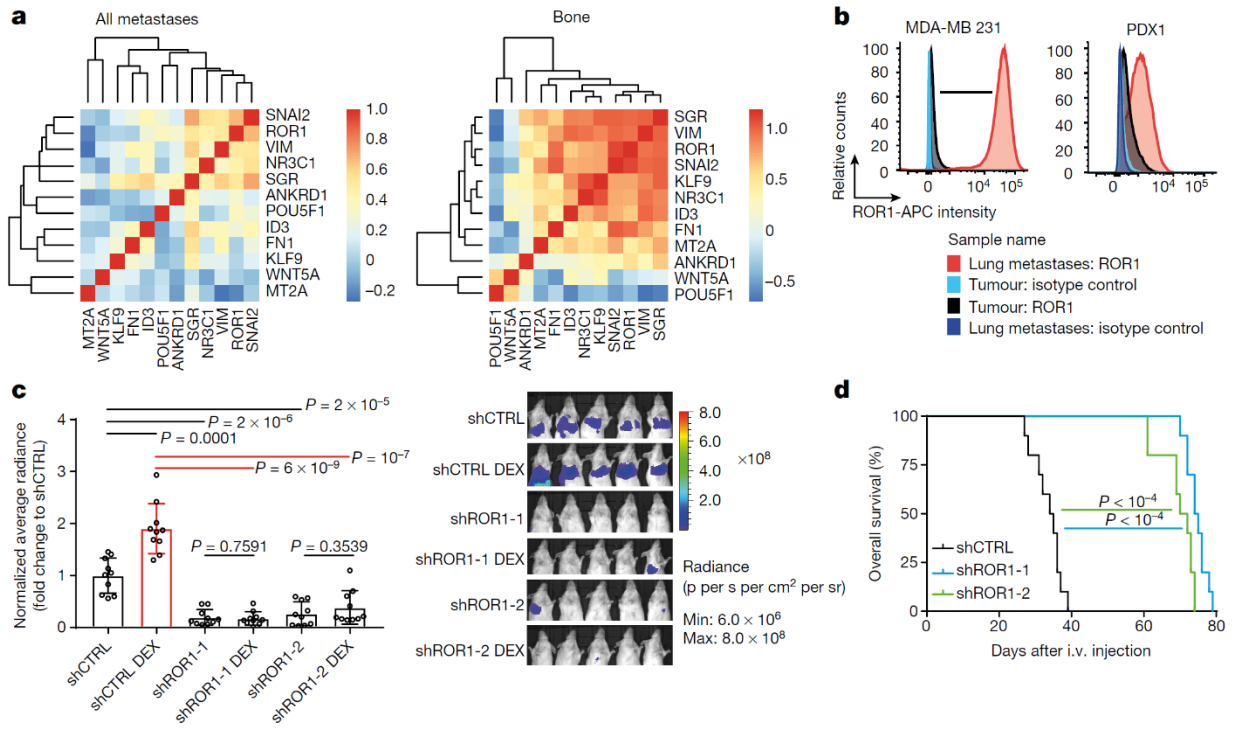
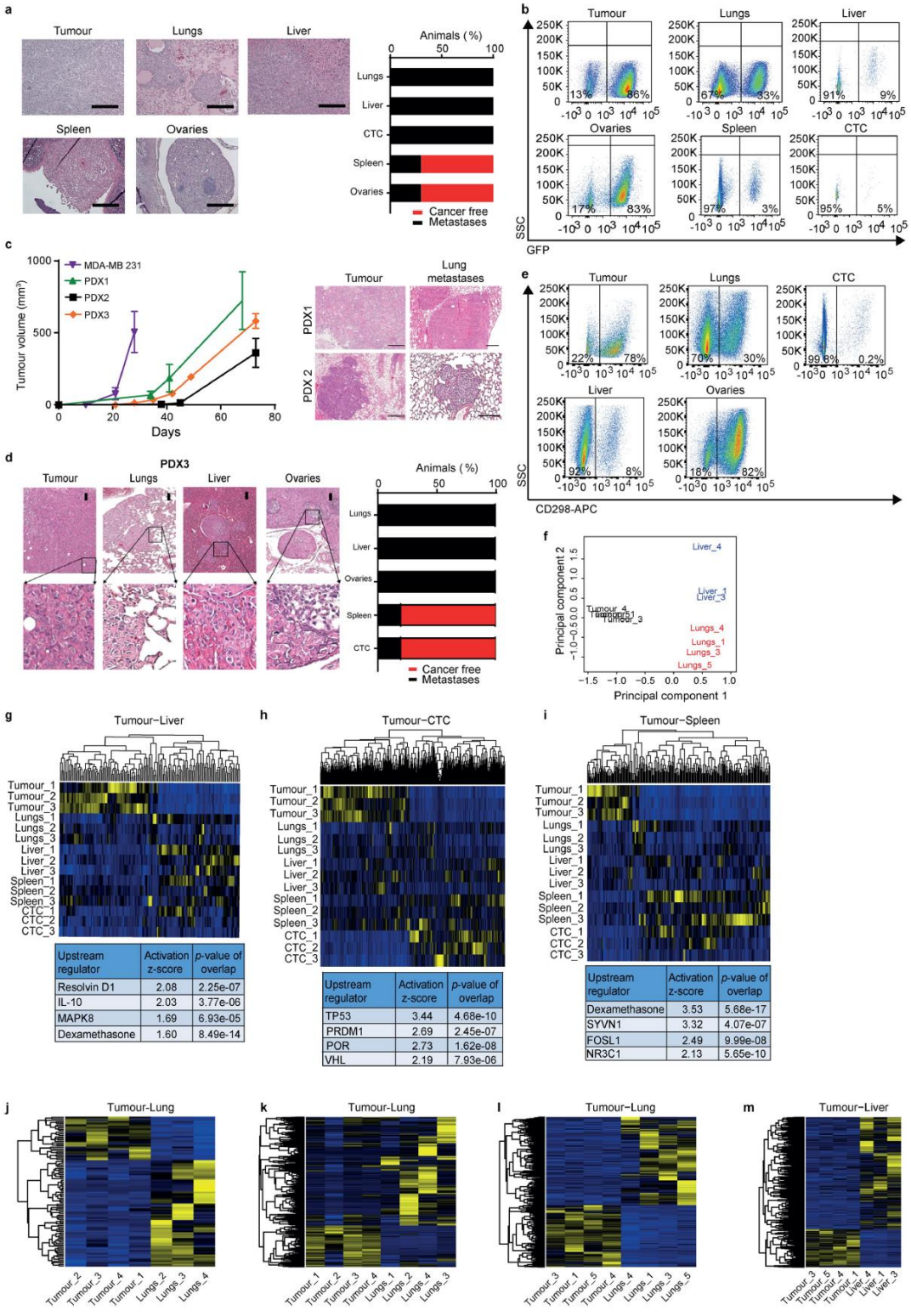


Figure 3.4. ROR1 mediates GR-induced lung metastatic colonization. a, GR activation signature (SGR) and ROR1 correlation in BC metastases, *n*=88 metastases among them *n*=7 bone metastases. Pearson correlation. **b**, Flow cytometry analysis of ROR1 levels in tumours and matched metastases of MDA-MB 231 and PDX1, *n*=3 biological replicates. **c**, Experimental metastases assay and *in vivo* bioluminescence imaging of animals inoculated *i.v.* with GR-activated or control MDA-MB 231 cells downregulated for ROR1 or shCTRL, means ± s.d., *n*=10 mice, pooled data from 2 independent experiments, *ns*= non-significant, two-tailed Student's *t*-test. **d**, Kaplan-Meier survival analysis of animals inoculated *i.v.* with control or shROR1 MDA-MB 231 cells, *n*=10 mice per group, two-tailed log-rank test.

3.6 Discussion

Tumour heterogeneity is one of the major obstacles to the treatment of metastatic BC. We show here that metastases display distinct phenotypes related to their growth sites. We also find that the stress hormone pathway is an effective inducer of colonization and the death of the animals, and that ROR1 ablation prevents the deleterious effect of GR activation. The data also reveal that GR activation decreases the efficacy of paclitaxel. Corticosteroids such as DEX are widely used in the treatment of BC to combat side-effects of chemotherapy and to treat symptoms related to advanced cancer. Given that cancer cell dissemination has already occurred by the time of primary tumour surgical resection in a substantial number of BC patients(Braun et al., 2005; Hosseini et al., 2016), and that GR activation fosters colonization at the distant sites, our results call for caution when administering corticosteroids to patients. Of note, GR has been shown to evoke adaptive resistance to anti-androgen receptor therapy in prostate tumours(Arora *et al.*, 2013). Thus, assessment of the effects of stress hormone pathways on metastasis and response to therapy is also warranted in other cancer types.

3.7 Extended figures

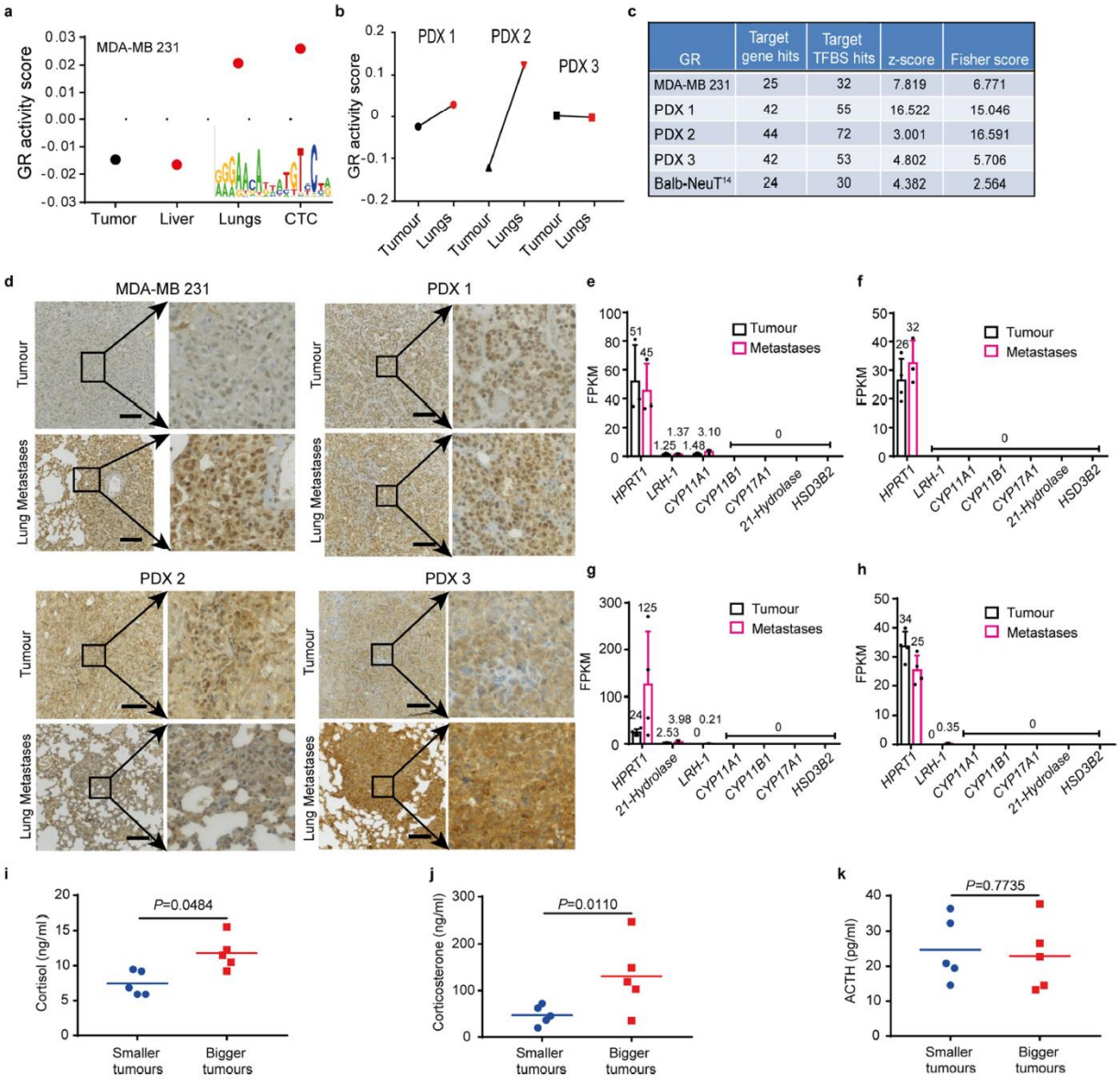


Upstream regulator	Activation z-score	*p*-value of overlap
Resolvin D1	2.08	2.25e-07
IL-10	2.03	3.77e-06
MAPK8	1.69	6.93e-05
Dexamethasone	1.60	8.49e-14

Upstream regulator	Activation z-score	*p*-value of overlap
TP53	3.44	4.68e-10
PRDM1	2.69	2.45e-07
POR	2.73	1.62e-08
VHL	2.19	7.93e-06

Upstream regulator	Activation z-score	*p*-value of overlap
Dexamethasone	3.53	5.68e-17
SYVN1	3.32	4.07e-07
FOSL1	2.49	9.99e-08
NR3C1	2.13	5.65e-10

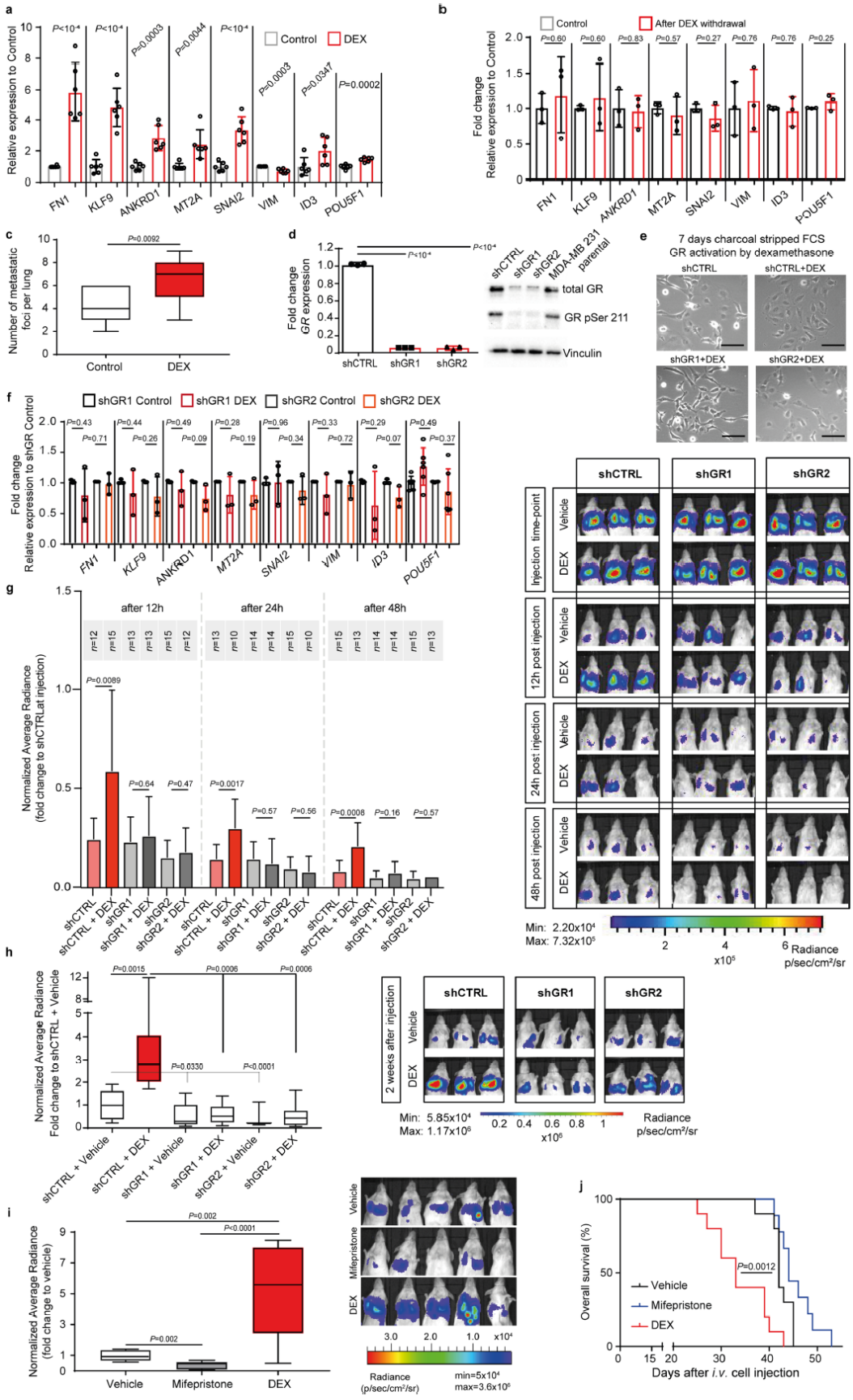
Extended Data Fig. 3.1 | Increase in GR activation in breast cancer metastases. a, Tumours and matched lung, liver, ovary and spleen metastases in the MDA-MB 231 model (haematoxylin and eosin staining). Scale bar, 200 μm. Right, frequency of metastases detected in distant organs upon tumour resection. n = 10 from 5 independent experiments. b, FACS analysis of organs affected with distant metastases in the MDA-MB 231 model. n = 10. c, Tumour growth kinetics after orthotopic transplantation in the MDA-MB 231 model (n = 9) and PDX1, PDX2 and PDX3 models (n = 5, 4 and 9 respectively). Mean ± s.e.m. Right, tumour and matched lung metastases in the PDX1 and PDX2 models (haematoxylin and eosin staining). Scale bar, 200 μm. d, Tumour and matched lung, liver or ovary metastases in the PDX3 model (haematoxylin and eosin staining). Scale bar, 200 μm. Right, frequency of metastases detected in distant organs upon tumour resection. n = 10. e, FACS analysis of PDX3 tumour and organs with matched metastases. n = 5. f, Principal component analysis of PDX3 tumours (n = 4) and matched liver (n = 3) and lung (n = 3) metastases. g–i, Heat maps of genes that are differentially expressed and upstream regulator analysis, for the MDA-MB 231 model. n = 3, Fisher's exact test. g, Tumours and liver metastases. h, Tumours and circulating tumour cells. i, Tumour and spleen. j–m, Heat maps of genes that are differentially expressed for PDX models. j, PDX1 tumour (n = 4) and lung metastases (n = 3). k, PDX2 tumour and lung metastases (n = 4). l, PDX3 tumour and lung metastases (n = 4). m, PDX3 tumour and liver metastases (n = 4). n represents biological replicates (mice) in all panels. Threshold criteria for all differential-expression heat map analyses are a fold-change ≥ 2 and FDR < 0.05. The statistical approach for differential expression analysis is provided in Methods.



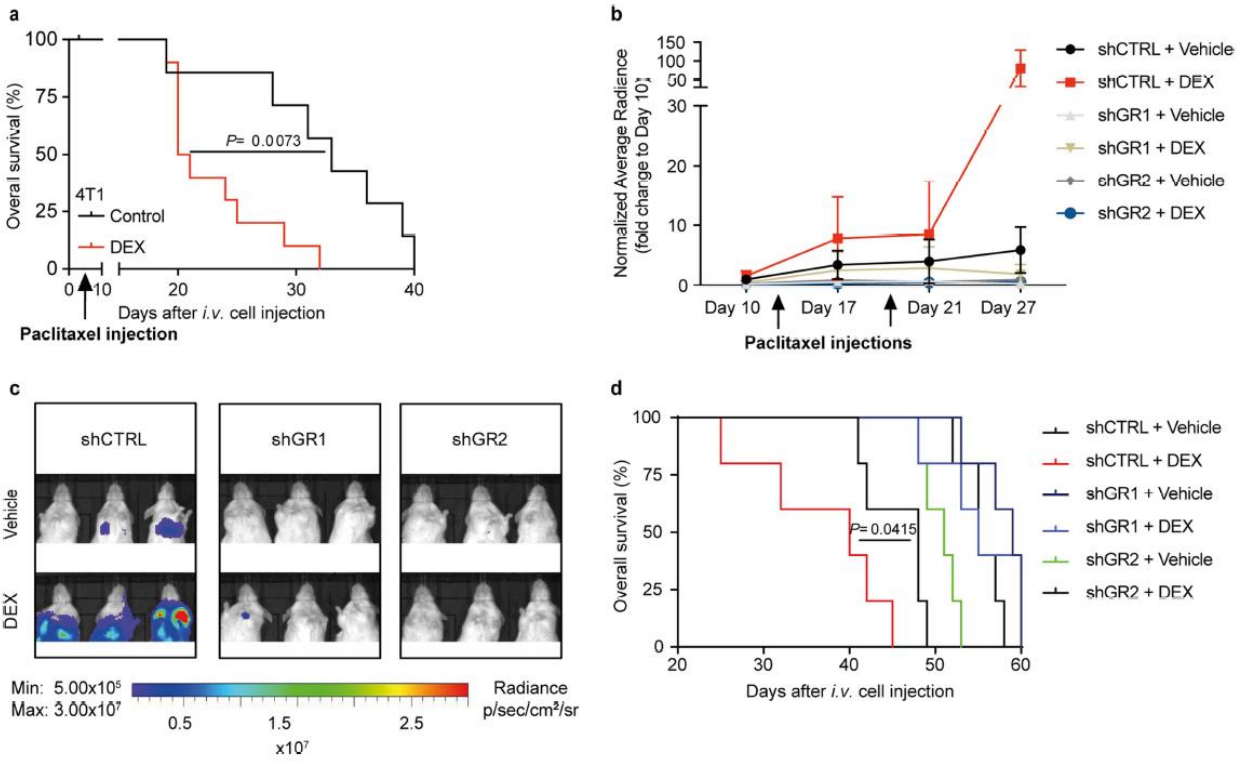
GR	Target gene hits	Target TFBS hits	z-score	Fisher score
MDA-MB 231	25	32	7.819	6.771
PDX 1	42	55	16.522	15.046
PDX 2	44	72	3.001	16.591
PDX 3	42	53	4.802	5.706
Balb-NeuT[14]	24	30	4.382	2.564

Extended Data Fig. 3.2 | GR activation in distant metastases and circulating tumour cells. a, ISMARA transcription-factor-activity plot of the tumour, lung, liver metastases and circulating tumour cells in the MDA-MB 231 model. *n* = 3 biological replicates. **b**, ISMARA transcription-factor activity of PDX models. *n* = 4 (apart from PDX1 lung metastases, *n* = 3) biological replicates (mice). **c**, GR transcription-factor binding sites in lung metastases of PDX1, PDX2, PDX3, MDA-MB 231 and BALB/c–NeuT models14. *n* = 4 (apart from PDX1 lung metastases, *n* = 3) biological replicates (mice). **d**, GR expression in tumours and matched lung metastases in MDA-MB 231, PDX1, PDX2 and PDX3 models. *n* = 3 biological replicates (mice). Scale bar,100 μm. **e**–**h**, Expression of genes involved in glucocorticoid synthesis, with *HPRT1* as an internal control. **e**, MDA-MB 231, *n* = 3. **f**, PDX1, *n* = 4 tumours and *n* =

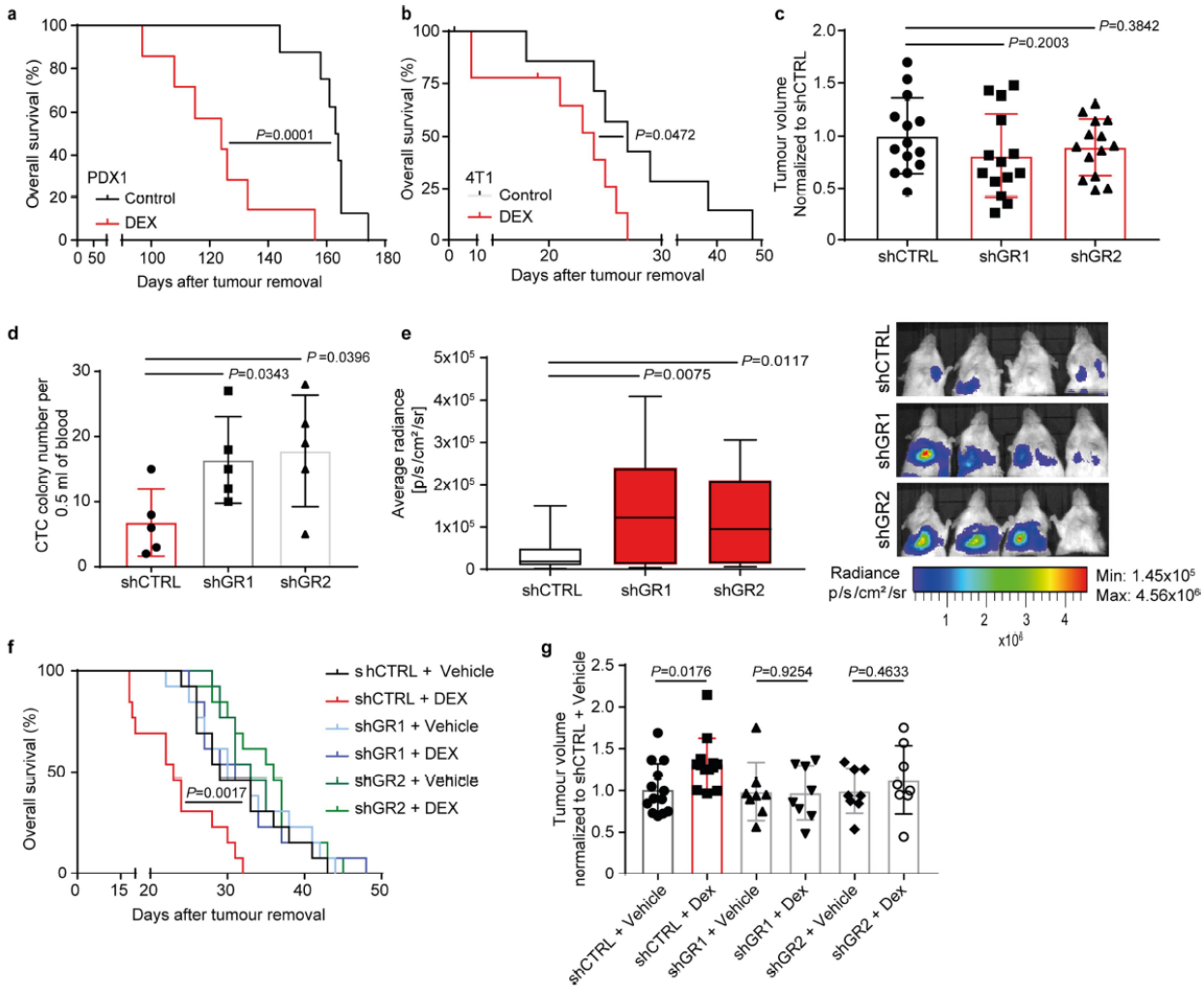
3 lung metastases. **g**, PDX2, $n = 4$. **h**, PDX3 models, $n = 4$. Mean ± s.d., *n* indicates biological replicates (mice). **i**–**k**, Subgroup analysis of plasma hormone levels in mice of the MDA-MB 231 model before tumour resection (M0 mice). The M0 group has been split at the median into two groups, one with smaller tumours (mean volume, 446 mm3) and the other with larger tumours (mean volume, 692 mm3). **i**, Cortisol levels. **j**, Corticosterone levels. **k**, Adrenocorticotropic hormone levels. Means and single data points are represented. $n = 5$ biological replicates (mice), two-tailed Student's *t*-test.


a
Relative expression to Control
Control
DEX
P<10⁻⁴
P<10⁻⁴
P=0.0003
P=0.0044
P<10⁻⁴
P=0.0003
P=0.0347
P=0.0002
FN1
KLF9
ANKRD1
MT2A
SNAI2
VIM
ID3
POU5F1
b
Fold change
Relative expression to Control
Control
After DEX withdrawal
P=0.60
P=0.60
P=0.83
P=0.57
P=0.27
P=0.76
P=0.76
P=0.25
FN1
KLF9
ANKRD1
MT2A
SNAI2
VIM
ID3
POU5F1
c
Number of metastatic foci per lung
P=0.0092
Control
DEX
d
Fold change GR expression
P<10⁻⁴
P<10⁻⁴
shCTRL
shGR1
shGR2
shCTRL
shGR1
shGR2
MDA-MB 231 parental
total GR
GR pSer 211
Vinculin
e
7 days charcoal stripped FCS
GR activation by dexamethasone
shCTRL
shCTRL+DEX
shGR1+DEX
shGR2+DEX
f
shGR1 Control
shGR1 DEX
shGR2 Control
shGR2 DEX
Fold change
Relative expression to shGR Control
P=0.43
P=0.71
P=0.44
P=0.26
P=0.49
P=0.09
P=0.28
P=0.19
P=0.96
P=0.34
P=0.33
P=0.72
P=0.29
P=0.07
P=0.49
P=0.37
FN1
KLF9
ANKRD1
MT2A
SNAI2
VIM
ID3
POU5F1
g
Normalized Average Radiance
(fold change to shCTRLat injection)
after 12h
after 24h
after 48h
n=12
n=15
n=13
n=13
n=15
n=12
n=13
n=10
n=14
n=14
n=15
n=10
n=15
n=13
n=14
n=14
n=15
n=13
P=0.0089
P=0.64
P=0.47
P=0.0017
P=0.57
P=0.56
P=0.0008
P=0.16
P=0.57
shCTRL
shCTRL + DEX
shGR1
shGR1 + DEX
shGR2
shGR2 + DEX
shCTRL
shGR1
shGR2
Injection time-point
12h post injection
24h post injection
48h post injection
Vehicle
DEX
Min: 2.20x10⁴
Max: 7.32x10⁵
x10⁵
Radiance
p/sec/cm²/sr
h
Normalized Average Radiance
Fold change to shCTRL + Vehicle
P=0.0015
P=0.0006
P=0.0006
P=0.0330
P<0.0001
shCTRL + Vehicle
shCTRL + DEX
shGR1 + Vehicle
shGR1 + DEX
shGR2 + Vehicle
shGR2 + DEX
2 weeks after injection
shCTRL
shGR1
shGR2
Vehicle
DEX
Min: 5.85x10⁴
Max: 1.17x10⁶
x10⁶
Radiance
p/sec/cm²/sr
i
Normalized Average Radiance
(fold change to vehicle)
P=0.002
P<0.0001
P=0.002
Vehicle
Mifepristone
DEX
Vehicle
Mifepristone
DEX
x10⁴
Radiance
(p/sec/cm²/sr)
min=5x10⁴
max=3.6x10⁵
j
Overall survival (%)
P=0.0012
Vehicle
Mifepristone
DEX
Days after i.v. cell injection

Extended Data Fig. 3.3 | Glucocorticoids promote colonization via GR. a, Expression of GR targets (qPCR). Mean ± s.d., n = 6 biological replicates, in technical duplicates; two-tailed Student's t-test. **b**, Expression of GR targets three weeks after discontinuation of GR activation by dexamethasone. Mean ± s.d., n = 3 biological replicates in technical duplicates, two-tailed Student's t-test. **c**, Number of metastatic foci in lungs of mice injected with 4T1 dexamethasone or vehicle-treated cells. n = 9 mice, two-tailed Student's t-test. **d**, GR downregulation in MDA-MB 231 cells (left, qPCR; right, immunoblotting). Mean ± s.d., n = 3 biological replicates, two-tailed Student's t-test. **e**, MDA-MB 231 cells were propagated in the presence of dexamethasone or vehicle for seven days. n = 7 biological replicates. **f**, GR-downregulated MDA-MB 231 cells did not express the GR-activation marker gene set upon dexamethasone treatment. Mean ± s.d., n = 3 biological replicates in technical duplicates, two-tailed Student's t-test. **g**, Bioluminescence imaging 12 h, 24 h and 48 h after intravenous injection of control and dexamethasone-treated MDA-MB 231 cells transduced with control or *GR* shRNA. n represents biological replicates (mice), mean ± s.d., two-tailed Student's t-test. **h**, Bioluminescence imaging of mice two weeks after intravenous injection. n = 15 for vehicle + control shRNA, dexamethasone + control shRNA, and dexamethasone + *GR* shRNA 2 (shGR2); n = 14 for dexamethasone + *GR* shRNA 1 (shGR1); n = 13 for vehicle + *GR* shRNA 2; n = 12 for vehicle + *GR* shRNA 1. Three independent experiments; n represents biological replicates (mice), two-tailed Student's t-test. **i**, Bioluminescence imaging two weeks after intravenous injection of GR-activated, mifepristone- or vehicle-treated MDA-MB 231 cells. n = 10 mice, 2 independent experiments, two-tailed Student's t-test. **j**, Kaplan–Meier survival analysis of mice upon intravenous injection of GR-activated, mifepristone- or vehicle-treated MDA-MB 231 cells. n = 10 mice for vehicle and dexamethasone groups; n = 9 for mifepristone group. Two independent experiments, two-tailed log-rank test, vehicle-treated versus mifepristone-treated groups, P = 0.054. In box plots, the centre line indicates the median, the box extends from the 25th to 75th percentiles and whiskers extend to the most extreme data points.



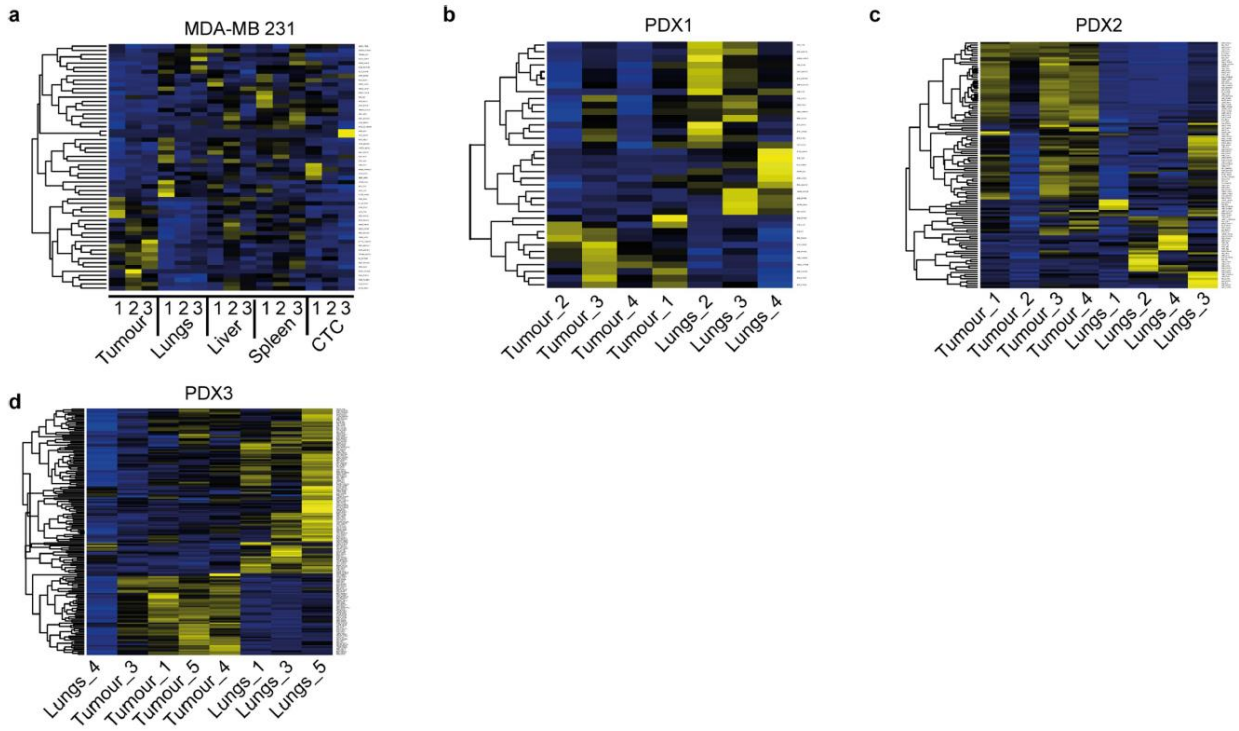
Extended Data Fig. 3.4 | Dexamethasone offsets the response to paclitaxel. a, Kaplan–Meier survival analysis of mice intravenously injected with 4T1 GR-activated or control cells. Dexamethasone decreases 4T1 response to paclitaxel *in vivo*. Paclitaxel was administrated five days after 4T1 cell inoculation. n = 9 control; n = 10 dexamethasone-treated biological replicates (mice) per group; 2 mice were censored; two-tailed log-rank test. **b**, Dexamethasone offsets paclitaxel effect in the MDA-MB 231 model. Analysis of colonization potential under paclitaxel treatment of MDA-MB 231 cells transduced with one of the two *GR* shRNAs or control shRNA, and treated with dexamethasone or vehicle, intravenously injected into NSG mice. Two paclitaxel injections (15 and 22 days after cell injection). Mean ± s.d., n = 5 mice per group, two-tailed Student's t-test. **c**, Bioluminescence imaging corresponding to day 21 of **b**. **d**, Kaplan–Meier survival analysis of mice shown in **b**. n = 5 biological replicates (mice) per group, two-tailed log-rank test.



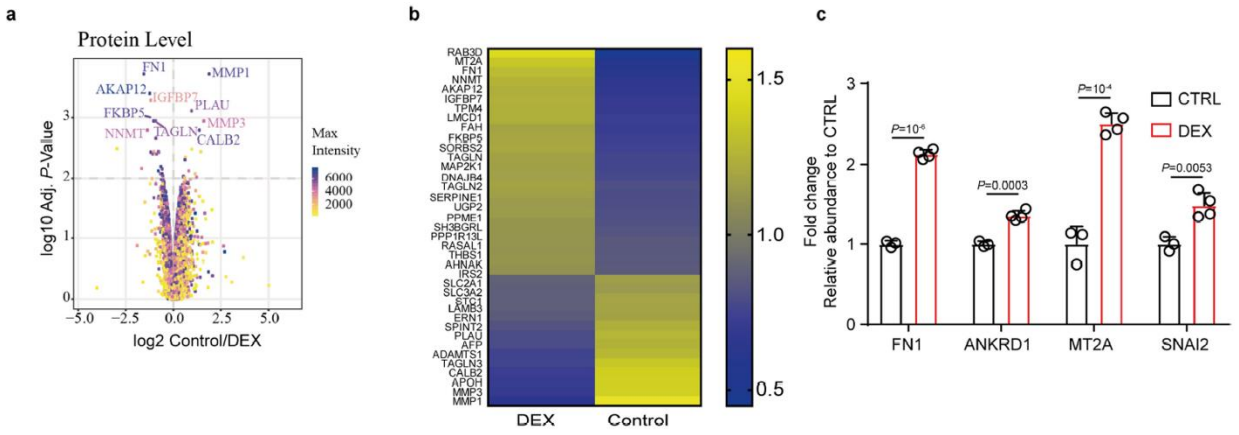
Extended Data Fig. 3.5 | Dexamethasone reduces overall survival and GR downregulation increases cancer cell dissemination from the primary site. a, b, Upon tumour removal from the 4th mammary gland, randomized mice were treated with dexamethasone or vehicle on 5 consecutive days (intraperitoneal injection of 0.1 mg kg−1 dexamethasone once daily). Kaplan–Meier survival analysis in PDX1 model (n = 8 control; n = 7 dexamethasone-treated biological replicates (mice)) (a) or 4T1 model (n = 8 control; n = 9 dexamethasone-treated biological replicates (mice)) (b). Two 4T1 mice were censored. Two-tailed log-rank test (a, b). c, GR downregulation in MDA-MB 231 cells does not affect tumour volume, relative to tumour cells transduced with control shRNA, at resection. Mean ± s.d., n = 14 biological replicates (mice), pooled data from 3 independent experiments, two-tailed Student's t-test. d, Circulating tumour cell count measured by the number of in vitro propagated colonies upon circulating tumour cell isolation from peripheral blood of tumour-bearing mice at tumour resection time point. Mean ± s.d., n = 5 biological replicates (mice), two-tailed Student's t-test. e, *In vivo* bioluminescence imaging upon tumour removal. In the boxplot, the centre line is the median, the box extends from the 25th to 75th percentiles, and whiskers extend to the most extreme

data points. *n* = 17 control shRNA; *n* = 10 *GR* shRNA 1 and 2 biological replicates (mice), pooled data from 3 independent experiments, two-tailed Student's *t*-test. f, Kaplan–Meier survival analysis of mice upon of removal of MDA-MB 231 tumours transduced with control shRNA or one of the two *GR* shRNAs,

and treatment with dexamethasone or vehicle. *n* = 13 biological replicates (mice) per group, pooled data from 3 independent experiments, two-tailed log-rank test. g, Tumour volumes at resection. In vitro dexamethasone- or vehicle-treated MDA-MB 231 cells transduced with control shRNA, *GR* shRNA 1 or *GR* shRNA 2 inoculated into the mammary fat pad of NSG mice. *n* = 13 control shRNA; *n* = 8 *GR* shRNA 1 and 2 biological replicates (mice), pooled data from 2 independent experiments, two-tailed Student's *t*-test. All tumours in all experiments were resected at the same time point.



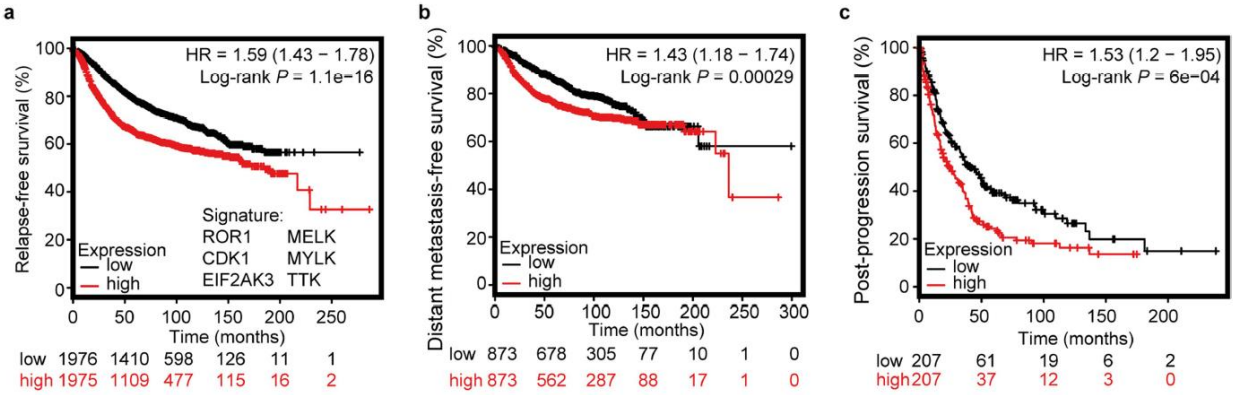
Extended Data Fig. 3.6 | Differential expression of protein kinases in tumours and matched metastases. a–d, Expression of protein kinases in MDA-MB 231 model, *n* = 3 biological replicates (mice) (a); PDX1 model, *n* = 4 tumour and *n* = 3 matched lung metastases; biological replicates (mice) (b); PDX2 model, *n* = 4 biological replicates (mice) (c); and PDX3 model, *n* = 4 biological replicates (mice) (d). The threshold criteria used for the analysis are fold-change ≥ 2 and $P < 0.05$. Further details of the statistical analysis are provided in Methods.

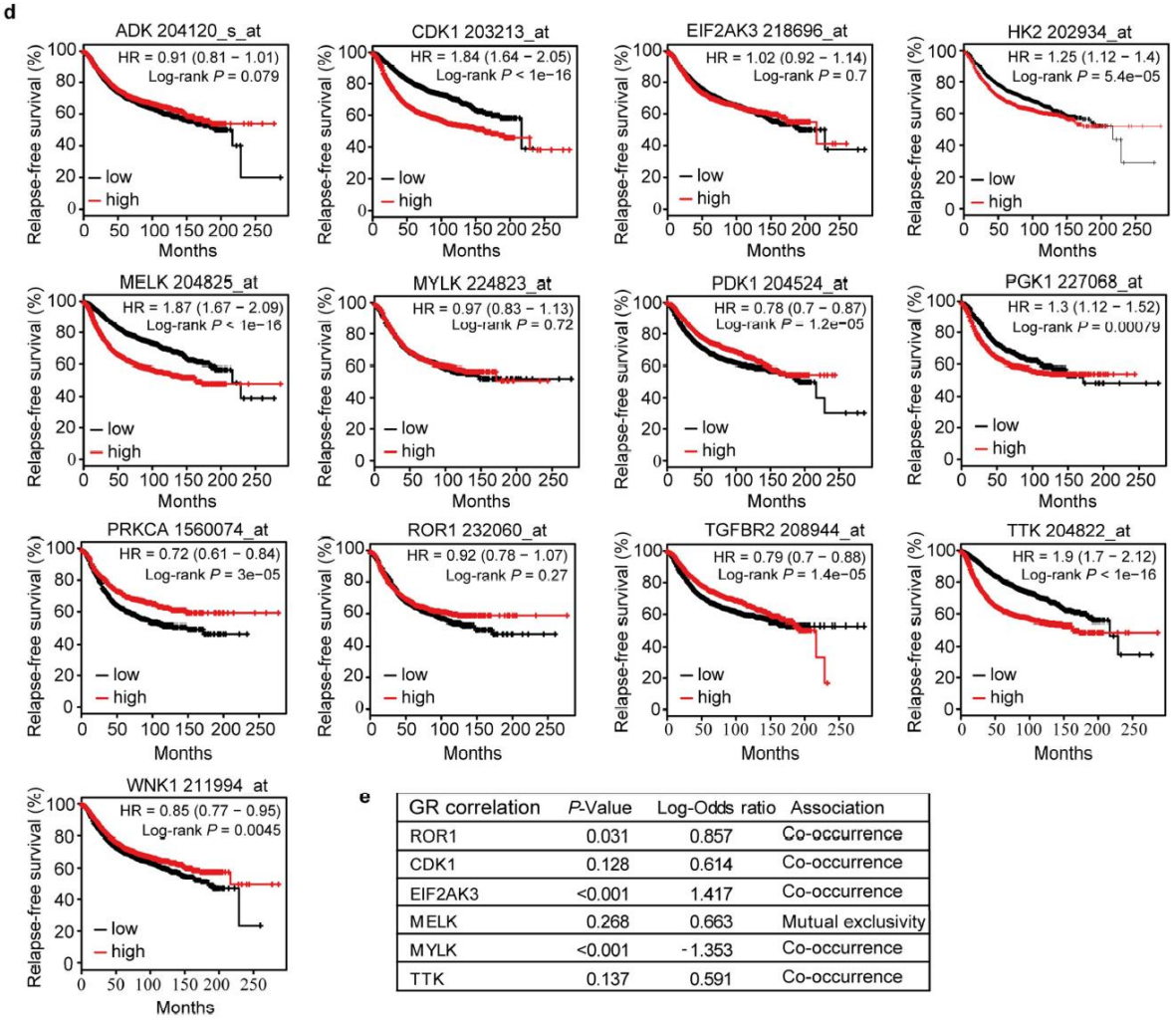


d

Process	*P*-value	FDR	Ratio
Development_EGFR signalling pathway	1.293e-7	3.361e-5	8/71
Signal transduction_ERK1/2 signalling pathway	6.474e-6	5.271e-4	5/32
Immune response_IL-33 signalling pathway	8.726e-6	5.672e-4	6/58
Immune response_IL-9 signalling pathway	1.180e-5	6.605e-4	5/36
Development_VEGF signalling via VEGFR2 - generic cascades	1.270e-5	6.605e-4	7/93
Development_HGF signalling pathway	4.464e-5	1.548e-3	5/47
Pathways of EMT in cancer cells	6.658e-5	1.902e-3	5/51
Cytoskeleton remodeling_FAK signalling	1.140e-4	2.578e-3	5/57
Development_Prolactin receptor signalling	1.240e-4	2.623e-3	5/58
Immune response_IL-6 signalling pathway via MEK/ERK and PI3K/AKT cascades	3.921e-4	4.634e-3	5/74
Development_G-CSF signalling	7.565e-4	7.868e-3	4/49

Extended Data Fig. 3.7 | Differential protein abundance upon GR activation. a, Volcano plot of protein abundance after GR activation in MDA-MB 231 cells. *n* = 3 control; *n* = 4 dexamethasone-treated biological replicates, Bayes-moderated *t*-statistics, *P* values corrected for multiple testing using the Benjamini–Hochberg method, calculations performed in R using the LIMMA package, Bioconductor. **b**, Heat map of differentially abundant proteins in dexamethasone-treated and vehicle-treated (control) cells. *n* = 3 control; *n* = 4 dexamethasone-treated biological replicates; FDR < 0.05, Bayes-moderated *t*-statistics; *P* values were corrected for multiple testing using the Benjamini–Hochberg method, calculations performed in R using the LIMMA package, Bioconductor. **c**, Abundance of proteins used for generation of the GR activation signature. *n* = 3 control; *n* = 4 dexamethasone-treated biological replicates, mean ± s.d., two-tailed Student's *t*-test. **d**, Pathway enrichment analysis of all phosphoproteins with significant abundance changes against all phospho-proteins quantified as a background using MetaCore (Clarivate Analytics). Enrichment *P* values and FDRs were determined by the software-specific algorithms using default parameters.

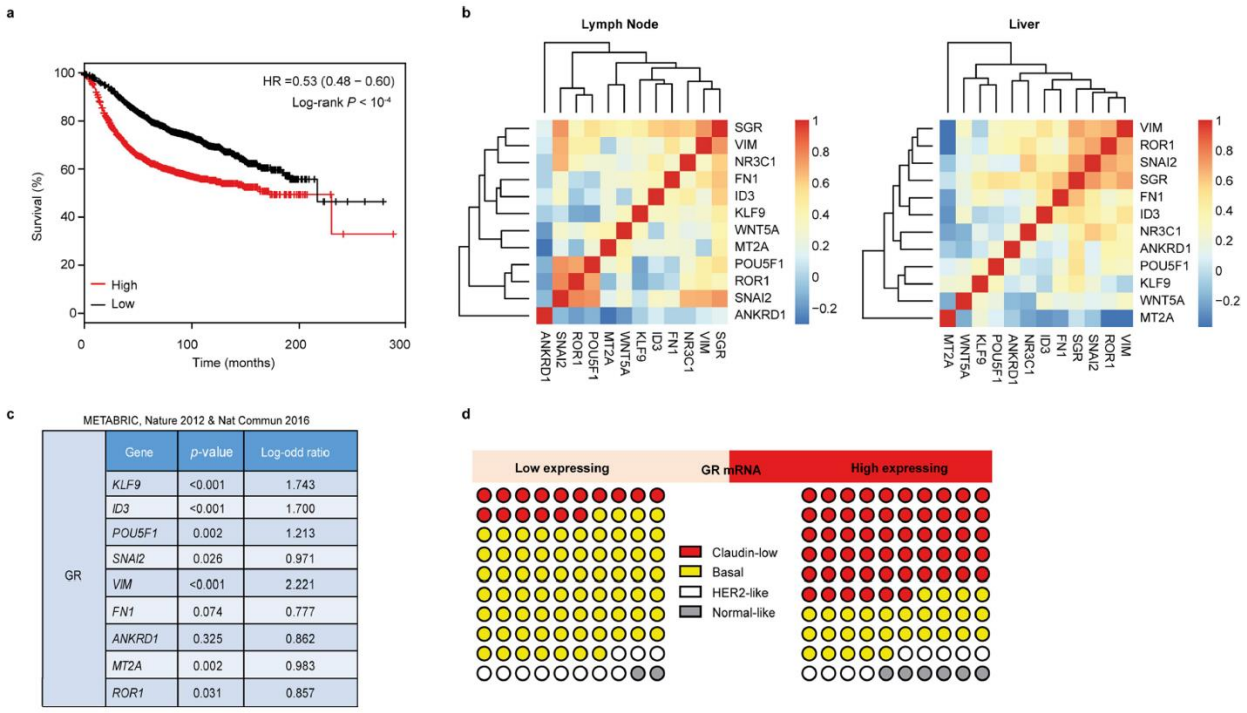




e

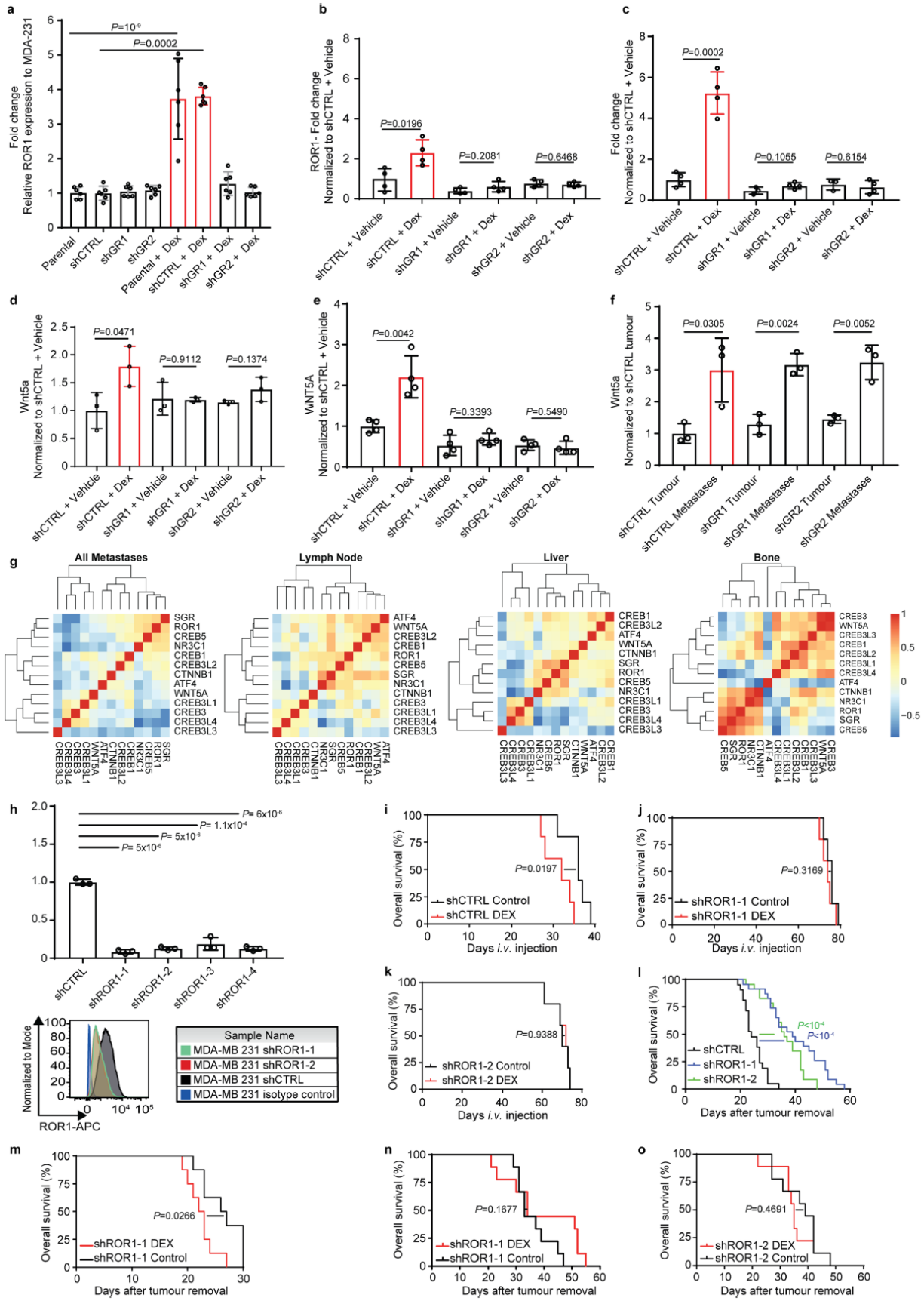
GR correlation	*P*-Value	Log-Odds ratio	Association
ROR1	0.031	0.857	Co-occurrence
CDK1	0.128	0.614	Co-occurrence
EIF2AK3	<0.001	1.417	Co-occurrence
MELK	0.268	0.663	Mutual exclusivity
MYLK	<0.001	-1.353	Co-occurrence
TTK	0.137	0.591	Co-occurrence

Extended Data Fig. 3.8 | GR activation increases the expression of kinases that are predictive of survival in breast cancer. Survival based on the expression of the protein kinase signature that is upregulated in the metastases. **a**, Relapse-free survival, two-tailed log-rank test. **b**, Distant metastasis-free survival, two-tailed log-rank test. **c**, Postprogression survival, two-tailed log-rank test. Number of patients (*n*) and *P* values are presented in the panels. **d**, Individual protein kinases, relapse-free survival, *n* = 1,764, two-tailed log-rank test. **e**, Co-occurrence of GR and protein kinases in publically available breast cancer datasets, Fisher's exact test, *n* = 2,509 (refs. 27,28).



	Gene	*p*-value	Log-odd ratio
GR	*KLF9*	<0.001	1.743
	ID3	<0.001	1.700
	POU5F1	0.002	1.213
	SNAI2	0.026	0.971
	VIM	<0.001	2.221
	FN1	0.074	0.777
	ANKRD1	0.325	0.862
	MT2A	0.002	0.983
	ROR1	0.031	0.857

Extended Data Fig. 3.9 | ROR1 expression in breast cancer and metastases. a, Relapse-free survival analysis of patients with the ROR1 signature (G-2-0, Kaplan–Meier), *n* = 4,029, two-tailed log-rank test. **b**, GR-activation signature correlates with increased levels of ROR1 in breast cancer metastases. *n* = 21 lymph node; *n* = 34 liver metastases. Pearson correlation. **c**, Co-occurrence of GR-activation gene signature with GR and ROR1, *n* = 2,509 (refs. 27,28), Fisher's exact test. **d**, Breast cancers that express high levels of *GR* mRNA were enriched in the claudinlow profile, *n* = 299 (refs. 27,28).


a
Fold change
Relative ROR1 expression to MDA-231
P=10-9
P=0.0002
Parental
shCTRL
shGR1
shGR2
Parental + Dex
shCTRL + Dex
shGR1 + Dex
shGR2 + Dex
b
ROR1- Fold change
Normalized to shCTRL + Vehicle
P=0.0196
P=0.2081
P=0.6468
shCTRL + Vehicle
shCTRL + Dex
shGR1 + Vehicle
shGR1 + Dex
shGR2 + Vehicle
shGR2 + Dex
c
Fold change
Normalized to shCTRL + Vehicle
P=0.0002
P=0.1055
P=0.6154
d
Wnt5a
Normalized to shCTRL + Vehicle
P=0.0471
P=0.9112
P=0.1374
e
WNT5A
Normalized to shCTRL + Vehicle
P=0.0042
P=0.3393
P=0.5490
f
Wnt5a
Normalized to shCTRL tumour
P=0.0305
P=0.0024
P=0.0052
shCTRL Tumour
shCTRL Metastases
shGR1 Tumour
shGR1 Metastases
shGR2 Tumour
shGR2 Metastases
g
All Metastases
Lymph Node
Liver
Bone
SGR
ROR1
CREB5
NR3C1
CREB1
CREB3L2
CTNNB1
ATF4
WNT5A
CREB3L1
CREB3
CREB3L4
CREB3L3
h
P= 6x10-6
P= 1.1x10-4
P= 5x10-6
P= 5x10-6
shCTRL
shROR1-1
shROR1-2
shROR1-3
shROR1-4
Normalized to Mode
ROR1-APC
Sample Name
MDA-MB 231 shROR1-1
MDA-MB 231 shROR1-2
MDA-MB 231 shCTRL
MDA-MB 231 isotype control
i
Overall survival (%)
P=0.0197
shCTRL Control
shCTRL DEX
Days i.v. injection
j
P=0.3169
shROR1-1 Control
shROR1-1 DEX
k
P=0.9388
shROR1-2 Control
shROR1-2 DEX
l
P<10-4
P<10-4
shCTRL
shROR1-1
shROR1-2
Days after tumour removal
m
P=0.0266
shROR1-1 DEX
shROR1-1 Control
n
P=0.1677
shROR1-1 DEX
shROR1-1 Control
o
P=0.4691
shROR1-2 DEX
shROR1-2 Control

Extended Data Fig. 3.10 | Dexamethasone increases metastases and precipitates death, via ROR1. a, ROR1 expression in in vitro propagated control and GR-downregulated cells. Mean ± s.d., *n* = 6 biological replicates, two-tailed Student's *t*-test. **b**, **c**, *ROR1* qPCR. RNA from tissues of mice injected with control or GR-downregulated cells in tumours (*n* = 4 biological replicates (mice)) (**b**) and lung metastases (*n* = 4, vehicle + control shRNA, dexamethasone + control shRNA and dexamethasone +*GR* shRNA 1 or 2; *n* = 3 vehicle + *GR* shRNA 1 or 2, biological replicates (mice) in technical duplicates or triplicates) (**c**). Mean ± s.d., two-tailed Student's *t*-test. **d**, Levels of WNT5A protein in supernatant of dexamethasone-treated or vehicle-treated MDA-MB 231 cells transduced with control or one of the two *GR* shRNAs. Mean ± s.d., *n* = 3 biological replicates, two-tailed Student's *t*-test. **e**, *WNT5A* qPCR in dexamethasone-treated or vehicle-treated MDA-MB 231 cells transduced with control or one of the two *GR* shRNAs. Mean ± s.d., *n* = 4 biological replicates in technical triplicates, two-tailed Student's *t*-test. **f**, Levels of WNT5A protein in tumours transduced with control or one of the two *GR* shRNAs, and their matched metastases. Mean ± s.d., *n* = 3 biological replicates (mice) in technical duplicates, two-tailed Student's *t*-test. **g**, Pearson correlation of GR activation, ROR1, WNT5A and Wnt signalling pathway members in breast cancer metastases. *n* = 88 breast cancer metastases; *n* = 21 lymph node; *n* = 34 liver; *n* = 7 bone metastases. **h**, ROR1 downregulation in MDA-MB 231 cells. Mean ± s.d., *n* = 3 biological replicates in technical duplicates, two-tailed Student's *t*-test, qPCR (top) and flow cytometry (bottom). **i**–**k**, Kaplan–Meier survival analysis of mice intravenously inoculated with vehicle-treated or dexamethasone-treated MDA-MB 231cells transduced with control shRNA (**i**), *ROR1* shRNA 1 (shROR1-1) (**j**) or *ROR1* shRNA 2 (shROR1-2) (**k**). *n* = 5, two-tailed log-rank test. **l**, Kaplan–Meier survival analysis of mice injected in the mammary fat pad with MDA-MB 231 cells transduced with control or one of the two *ROR1* shRNAs. *n* = 21 control shRNA; *n* = 23 *ROR1* shRNA 1 or 2, two-tailed log-rank test. **m**–**o**, Kaplan–Meier survival analysis of mice inoculated in the mammary fat pad with MDA-MB 231 cells transduced with control shRNA (**m**), *ROR1* shRNA 1 (**n**) or *ROR1* shRNA 2 (**o**), propagated in the presence of dexamethasone or vehicle. *n* = 8 control shRNA; *n* = 9 *ROR1* shRNA 1 or 2, two-tailed log-rank test.

Materials and correspondence

Correspondence and requests for materials should be addressed to Mohamed Bentires-Alj, m.bentires-alj@unibas.ch.

This manuscript contains 4 Figures, 10 Extended Data Figures, 5 Supplementary Tables, and 3 Supplementary Videos.

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Data availability. All mass spectrometry raw data files have been deposited to the ProteomeXchange Consortium, accession code PXD009102, http://proteomecentral.proteomexchange.org. The mRNA sequencing data are deposited at the Gene Expression Omnibus (GEO) database and have been currently under evaluation. The anticipated date of final validation: January 10, 2019. Datasets generated during this study are available in the source files and from the corresponding author on reasonable request.

Author contributions. M.M.S.O. conceived the study, designed and performed the experiments, analysed the data, interpreted the results and wrote the manuscript. B.H. performed experiments on gene expression and helped with ROR1 shRNA and mouse experiments, analysed the data and interpreted the results. N.M. established methods and measured stress hormones levels in plasma, analysed the data and interpreted the results. J.P.C. designed experiments, analysed the data and interpreted the results. S.M. performed histopathological analysis of the PDX models, analysed the data and interpreted the results. R.O. characterized the metastatic potential of PDX models, analysed the data and interpreted the results. A. Sethi performed computational analysis of metastatic breast cancer samples, analysed the data and interpreted the results. H.K. performed fluorescence-activated cell sorting experiments, analysed the data and interpreted the results. M.-M.C. performed intravital imaging, analysed the data and interpreted the results. A. Schmidt performed proteomics and phosphoproteomics experiments, analysed the data and interpreted the results. M.B.-A. conceived the study, designed the experiments and interpreted the results. All authors read and approved the final manuscript.

3.8 Material and Methods

***In vivo* experiments.** All *in vivo* experiments were performed in accordance with the Swiss animal welfare ordinance and approved by the cantonal veterinary office Basel Stadt. Female severe combined NOD-*scid IL2rγnull* (NSG) and Balb/c animals were maintained in the Friedrich Miescher Institute for Biomedical Research and the Department of Biomedicine animal facilities in accordance with Swiss guidelines on animal experimentation. Animals were maintained in a

sterile environment with light, humidity and temperature-control (light / dark cycle with light from 7:00 to 17:00, with a gradual change from light to dark), temperature 21-25°C, humidity 45-65%.). Before each experiment, animals were allowed to acclimatize for a minimum of 7 days. MDA-MB 231 cells (10,000 cells) were resuspended in 40 µl Matrigel:PBS (1:1) and injected into the pre-cleared mammary fat-pads of 4- to 8-week-old female NSG mice. PDX models were transplanted into the pre-cleared 4th mammary fat-pads of NSG mice, while 4T1 cells were injected into the mammary fat-pads of 4- to 8-week-old female Balb/c mice. Tumours were resected when the largest diameter reached 10 mm and mice were monitored regularly for signs of metastatic outgrowth and distress. In none of the experiments tumour volumes exceeded approved limits. All orthotropic experimental procedures (tumour resection and tumour cell implantation) were undertaken on anesthetized animals by a single investigator according to protocols approved by the cantonal veterinary office Basel Stadt. Experimental metastases assays were performed by injecting 100,000 cells into tail veins. After *i.v.* injection of MDA-MB 231, we performed *in vivo* bioluminescence imaging to confirm injection and to monitor metastatic outgrowth. Bioluminescence imagining was performed using an IVIS Lumina *XR* (Caliper LifeSciences) upon injection of luciferin (Biosynth; L8220).

In vivo DEX treatment (water soluble DEX; D2915- Sigma) was performed with clinically relevant doses of 0.1 mg/kg for 5 consecutive days as previously described(Pang *et al.*, 2006). Mifepristone (Selleckchem, S2606) treatment was performed for 5 consecutive days as previously described(Skor *et al.*, 2013). Paclitaxel (T7191- Sigma) was administrated intraperitoneal once a week at 20 mg/kg for 2 weeks in the MDA-MB 231 model and 5 days after *i.v.* cell injection in the 4T1 model. DEX was administrated for 5 consecutive days after each paclitaxel injection. Paclitaxel, dexamethasone and mifepristone were administrated by a single investigator. For overall survival experiments, independent assessments of mouse fitness were performed by multiple investigators.

Cell lines and PDX models.

Cell lines MDA-MB 231, 4T1 and HEK293T were purchased from ATCC and cultured according to ATCC protocols. Cell line identity was confirmed and routinely tested using short tandem repeat

(STR) sequencing; all cell lines were routinely tested for mycoplasma contamination. MDA-MB 231 and 4T1 cells were propagated in monolayer cultures in DMEM supplemented with 10% FCS.

GR activation experiments were performed in monolayer cultures in DMEM supplemented with 2.5% charcoal-stripped FCS (Thermo Fisher Scientific; cat. no. 12676029) in the presence of water-soluble dexamethasone (700 nM, Sigma, D2915) or vehicle for 7 consecutive days. All experiments were performed with 70-90% confluent cells.

The PDX used in this study were described earlier(Gao, Joshua M Korn, *et al.*, 2015),(DeRose *et al.*, 2011). PDX1, PDX2 and PDX4-11 originated from primary breast tumours(Gao, Joshua M Korn, *et al.*, 2015). PDX3 was obtained from a pleural effusion of a BC patient(DeRose *et al.*, 2011). The metastatic potentials of the examined PDX models were analysed by H&E staining and expression of a human-specific CD298 marker(Lawson *et al.*, 2015).

Lentiviral vectors, lentivirus and infection. For GR downregulation, we tested 6 shRNA constructs: V3LHS_404051, V3LHS_404052, V2LHS_239186, V2LHS_82796, V2LHS_82797, V3LHS_326099 (pGIPZ vector from Dharmacon). ROR1 downregulation was performed using 4 shRNA constructs: V3THS_349217, V3THS_306714, V3THS_306715, V3THS_240995 (Dharmacon, pTRIPZ). Non-targeting shRNAs (pGIPZ or pTRIPZ) were used as controls. Lentiviral batches were produced using PEI transfection on 293T cells as previously described(Britschgi et al., 2017). The titre of each lentiviral batch was determined in MDA-MB 231 cells. Cells were infected for 8 h in the presence of polybrene (8 µg/ml). Selection with 2 µg/ml puromycin (Sigma) was applied 48 h after infection.

Immunoblotting. The cells were lysed in RIPA buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with 1× protease inhibitor cocktail (Complete Mini, Roche), 0.2 mM sodium orthovanadate, 20 mM sodium fluoride, and 1 mM phenylmethylsulfonyl fluoride. The BCA protein assay kit (Thermo Scientific; cat no. 23227) was used to measure and equalize the concentrations of extracted proteins. Whole-cell lysates, immunoprecipitates or nuclear cell lysates (40 µg) were subjected to 6% SDS–PAGE, transferred to PVDF membranes (Immobilon-P, Millipore), and blocked for 1 h at room temperature with 5%

milk in PBS–0.1% Tween 20. Membranes were then incubated overnight with antibodies as indicated and exposed to secondary HRP-coupled anti-mouse or -rabbit antibody at 1:5,000–10,000 for 1 h at room temperature. The results shown for each of the blots presented are representative of at least three independent experiments. The following antibodies were used: glucocorticoid receptor antibody (GeneTex; cat. no. GTX101120), pSer GR antibody (Thermo Fisher Scientific; cat. no. PA5-17668).

Fluorescence-activated cell sorting. Tumours and matched metastases were mechanically and enzymatically digested using a collagenase/hyaluronidase solution (Stemcell Technologies; cat. no. 07912) at 37°C. Tumour cells were isolated using FACS based on the expression of GFP in MDA-MB 231 and the expression of human-specific marker CD298 in PDX models (Biolegend; cat. no. 341706). Prior to CTC sorting, erythrocytes were eliminated using "Red blood cell lysis buffer" (Sigma; cat. no. R7757). Cells were filtered twice through 40-μm cell strainers (Falcon) to obtain single cells. FACS was carried out with a BD FACSAria III (Becton Dickinson) using a 70-μm nozzle. Single cells were gated on the basis of their forward and side-scatter profiles and pulse-width was used to exclude doublets. Dead cells (DAPI bright) were gated out. Antibodies: APC anti-human CD298 antibody (Biolegend; cat. no. 341706), APC anti-human ROR1 antibody (Biolegend; cat. no. 357805).

RNA preparation, qPCR and sequencing. Isolated cells were sorted in the extraction buffer of the Arctutus PicoPure RNA Isolation Kit (cat. no. 12204-01) and mRNA isolated using the manufacturer's protocol. RNA was depleted of rRNA using the Ribo-Zero Magnetic Kit (MRZ11124C) from Epicentre and the column purified with the RNA Cleanup & Concentrator from Zymo Research. RNA integrity was measured on an Agilent 2100 Bioanalyzer using RNA Pico reagents (Agilent Technologies). The library was prepared using the ScriptSeq v2 RNA-Seq Library Preparation Kit (Epicentre). Library quality was measured on an Agilent 2100 Bioanalyzer for product size and concentration. Single-end libraries were sequenced by an Illumina HiSeq 2500 (50-nt read length). Quantitative PCR analysis was performed after mRNA isolation (Qiagene, RNeasy Plus Mini kit; cat. no. 74136). We used 1 μg of mRNA for cDNA generation (BioRad, iScript cDNA synthesis kit; cat. no. 170-8891) and the IDT master mix. *HPRT1* was used as a housekeeping gene (IDT Hs.PT.58v.45621572). IDT predesigned qPCR

assay Ids: *NR3C1*: Hs.PT.58.27480377, *ROR1*: Hs.PT.58.39481678, *FN1*: Hs.PT.58.40005963, *KLF9*: Hs.PT.56A.15636661, *SNAI2*: Hs.PT.58.177250059, *VIM*: Hs.PT.58.38906895, *POU5F1*: Hs.PT.58.14494169g, *MT2A*: Hs.PT.5046709g, human *WNT5a*: Hs.PT.58.22221435, mouse *WNT5a*: Mm.PT.58.16402801 and taqman probes (Thermo Fisher Scientific) *ANKRD1*: Hs00923599_m1, *ID3*: Hs00954037_g1. All measurements were performed in technical duplicate and triplicate as previously described(Britschgi et al., 2017). Technical duplicates were preferred and duplicates were used only in case of limited cDNA material, i.e. lung metastatic cells. The arithmetic mean of the Ct values was used for calculations: target gene mean Ct values were normalized to the respective housekeeping genes (*HPRT1*), mean Ct values (internal reference gene, Ct), and then to the experimental control. The values obtained were $2^{-\Delta\Delta Ct}$ expressed as fold changes in regulation compared to the experimental control using the $2^{-\Delta\Delta Ct}$ method of relative quantification.

Two-photon intravital microscopy. Tumour-bearing animals were anesthetized with isofluran and mounted on a custom-made stage 24 days after orthotopic injection of MDA-MB 231 shCTRL, shGR1 or shGR2 cells. Tumours were exposed by skin flap surgery(Wyckoff *et al.*, 2011) using a custom-made multiphoton microscope(Bonapace *et al.*, 2012). Imaging was performed as previously described(Bonapace *et al.*, 2014). Briefly, tumours were imaged at 880 nm with a 25X/1.05NA water immersion objective (Olympus). Cell motility was observed by time-lapse imaging over 30 min at 2-min intervals, where a 100 μm z-series with 5-μm intervals was recorded for each frame. Videos were generated using Image J (National Institutes of Health). For visualizing vasculature, 100 μl of 20 mg/ml 70-kDa Texas Red-dextran (Invitrogen, Molecular Probes, Switzerland) was injected in PBS into the tail vein of the mice before surgery.

Quantification of cortisol and corticosterone in mouse plasma. Cortisol (≥99%), corticosterone (≥98.5%), and formic acid (≥98%) were purchased from Sigma-Aldrich (Buchs, Switzerland). Diazepam (≥98%) was acquired from Toronto Research Chemicals (Toronto, Canada). Oasis HLB cartridges (1 cc, 30 mg) were from Waters (Milford, MA). Solvents of LC-MS or higher purity grade were used. Stock solutions of compounds were prepared in DMSO or methanol and stored at –20°C until use.

Mouse blood samples were collected by a single investigator in EDTA-coated tubes and mice plasma prepared by centrifugation for 15 min at 2,000 *g*. All subsequent measurements were performed blind. Plasma samples were stored at –80 °C until analysis. To precipitate proteins and extract cortisol and corticosterone from mice plasma, 50 µl of plasma was mixed with 950 µl of acetonitrile. Samples were shaken for 30 min at 10°C, placed at –20°C for 30 min, and then centrifuged at 5°C for 20 min at 24,000 *g*. Supernatants were separated and the pellets reconstituted in 100 µl of water and extracted for a second time with 900 µl of acetonitrile. After a second centrifugation (20 min at 24,000 *g* and 5°C), supernatants from two plasma extraction steps were combined and vacuum concentrated to an approximate volume of 100 µl. Oasis HLB solid-phase extraction cartridges (SPE; 1 cc, 30 mg) were activated with acetonitrile (1 ml) and conditioned with 5% acetonitrile in 0.1% formic acid in water (1 ml). Plasma extracts were applied to the SPE cartridges and washed with 0.1% formic acid in water (3 × 1 ml). Analytes were eluted from SPE cartridges with acetonitrile (3×0.5 ml). Samples were vacuum concentrated, spiked with a solution of internal standard (500 nM diazepam) and made up to 500 µl with HPLC mobile phase.

For quantification of cortisol and corticosterone in mouse plasma, samples were analysed with a Quattro Ultima triple-quadruple mass spectrometer equipped with an electro-spray source (Waters, Milford, MA) coupled to an Agilent 1200 HPLC system (Agilent, Santa Clara, CA). The analytical column was a HALO C18 (100 × 2.1 mm, 2.7 µm; Advanced Materials Technology, Wilmington, Delaware). The column flow rate and temperature were 400 $\mu l \cdot min^{-1}$ and 50°C. Eluents A and B were 0.1% formic acid in water and acetonitrile, respectively. Gradient elution was as follows: 0–1 min, 5% B; 1–10 min, 5→100% B; 10–11 min, 100% B, 11–12 min, 100→5% B; 12–15 min, 5% B. Source and desolvation temperatures were 140 and 240°C, respectively; cone and desolvation gas flows were 50 and 500 l/h. Capillary voltage was 2.5 kV, cone voltage 70 V, and collision energy 16–28 eV. Analyte quantification was performed in the positive ionization mode, relative to the internal standard (diazepam) using a multiple-reaction monitoring mode. The following transitions were used (*m/z*): cortisol (363.1→121.1, 327.2, 309.2, 345.2), corticosterone (347.1→329.2, 311.2, 293.2, 121.1), and diazepam (285→222, 228, 257; internal standard). Calibration curves were prepared by spiking the mobile phase with authentic metabolite standards in the concentration range of 0.8–1000 nM. Chromatograms were analysed with MassLynx 4.1

software (Waters, Milford, MA). Lower limits of detection (LLOD) and quantification (LLOQ) in mouse plasma samples were assessed based on signal-to-noise ratios of 3 and 10, respectively. Observed LLOD (lower limit of detection) for cortisol and corticosterone were 1.46 and 0.85 ng/ml. Observed LLOQ (lower limit of quantification) for cortisol and corticosterone were 4.56, 0.84, and 2.55 ng/ml.

Phosphoproteomic analysis. MDA-MB 231 cells propagated for 7 days in charcoal-stripped FCS in the presence of DEX or vehicle were mechanically detached, washed and snap frozen. Cells were lysed in 8 M Urea (Sigma) and 0.1 M ammonium bicarbonate in the presence of phosphatase inhibitors (Sigma P5726&P0044) using strong ultra-sonication (Bioruptor, 10 cycles, 30 s on/off, Diagenode, Belgium). Protein concentration was determined by BCA assay (Thermo Fisher Scientific) using a small sample aliquot. Aliquots of 250 µg of protein were digested as described previously(Ahrné *et al.*, 2016), reduced with 5 mM TCEP for 60 min at 37°C and alkylated with 10 mM chloroacetamide for 30 min at 37°C. After dilution of samples with 100 mM ammonium bicarbonate buffer to a final urea concentration of 1.6 M, proteins were digested by incubation with sequencing-grade modified trypsin (1/50 w/w; Promega, Madison, WI) overnight at 37°C. After acidification by 5% TFA, peptides were desalted on C18 reversed-phase spin columns according to the manufacturer's instructions (Macrospin, Harvard Apparatus) and dried under vacuum.

Peptide samples were enriched for phosphorylated peptides using Fe(III)-IMAC cartridges on an AssayMAP Bravo platform as described recently(Post *et al.*, 2017). The setup of the µRPLC-MS system was as described previously(Ahrné *et al.*, 2016). Chromatographic separation of peptides was carried out using an EASY nano-LC 1000 system (Thermo Fisher Scientific) equipped with a heated RP-HPLC column (75 µm x 37 cm) packed in-house with 1.9 µm C18 resin (Reprosil-AQ Pur, Dr. Maisch). Aliquots of 1 µg total peptides were analysed per LC-MS/MS run using a linear gradient ranging from 95% solvent A (0.15% formic acid, 2% acetonitrile) and 5% solvent B (98% acetonitrile, 2% water, 0.15% formic acid) to 30% solvent B over 90 min at a flow rate of 200 nl/min. Mass spectrometry analysis was performed on a Q-Exactive HF mass spectrometer equipped with a nanoelectrospray ion source (both Thermo Fisher Scientific). Each MS1 scan was followed by high-collision-dissociation (HCD) of the 10 most abundant precursor ions with

dynamic exclusion for 20 s. Total cycle time was approximately 1 s. For MS1, 3e6 ions were accumulated in the Orbitrap cell over a maximum time of 100 ms and scanned at a resolution of 120,000 FWHM (at 200 m/z). MS2 scans were acquired at a target setting of 1e5 ions, accumulation time of 100 ms and a resolution of 30,000 FWHM (at 200 m/z). Singly charged ions and ions with unassigned charge state were excluded from triggering MS2 events. The normalized collision energy was set to 27%, the mass isolation window was set to 1.4 m/z, and one microscan was acquired for each spectrum.

The acquired raw-files were imported into the Progenesis QI software (v2.0, Nonlinear Dynamics Limited), which was used to extract peptide precursor ion intensities across all samples applying the default parameters. The generated mgf-files were searched using MASCOT as above using the following search criteria: full tryptic specificity was required (cleavage after lysine or arginine residues, unless followed by proline); 3 missed cleavages were allowed; carbamidomethylation (C) was set as fixed modification; oxidation (M) and phosphorylation (STY) were applied as variable modifications; mass tolerance of 10 ppm (precursor) and 0.02 Da (fragments). The database search results were filtered using the ion score to set the false discovery rate (*FDR*) to 1% on the peptide and protein level, respectively, based on the number of reverse protein sequence hits in the datasets. The relative quantitative data were normalized and statistically analysed using our in-house script as above(Ahrné *et al.*, 2016).

Proteomics analysis using tandem mass tags. The peptide sample flow-through obtained after IMAC enrichment was dried and 25 µg of peptides labelled with tandem mass isobaric tags (TMT 10-plex, Thermo Fisher Scientific) according to the manufacturer's instructions. To control for ratio distortion during quantification, a peptide calibration mixture consisting of 6 digested standard proteins mixed in different amounts was added to each sample before TMT labelling as described recently(Ahrné *et al.*, 2016). After pooling the TMT-labelled peptide samples, peptides were again desalted on C18 reversed-phase spin columns according to the manufacturer's instructions (Macrospin, Harvard Apparatus) and dried under vacuum. TMT-labeled peptides were fractionated by high-pH reversed phase separation using a XBridge Peptide BEH C18 column (3,5 µm, 130 Å, 1 mm x 150 mm, Waters) on an Agilent 1260 Infinity HPLC system. Peptides were loaded on columns in buffer A (20 mM ammonium formate in water, pH 10) and eluted using a

two-step linear gradient starting from 2% to 10% in 5 min and then to 50% (v/v) buffer B (90% acetonitrile / 10% ammonium formate: 20 mM, pH 10) over 55 min at a flow rate of 42 µl/min. Elution of peptides was monitored with a UV detector (215 nm, 254 nm). A total of 36 fractions were collected, pooled into 12 fractions using a post-concatenation strategy as described previously(Wang *et al.*, 2011) and dried under vacuum. Aliquots (1 µg) of peptides were LC-MS analysed as described above with the following changes: the normalized collision energy was set to 35% and the mass isolation window was set to 1.1 m/z.

The acquired raw-files were converted to the mascot generic file (mgf) format using the msconvert tool (part of ProteoWizard, version 3.0.4624 (2013-6-3)). Using the MASCOT algorithm (Matrix Science, Version 2.4.1), the mgf files were searched against a decoy database containing normal and reverse sequences of the predicted SwissProt entries of *Homo sapiens* (www.ebi.ac.uk, release date 2014/11/24), the 6 calibration mix proteins(Ahrné *et al.*, 2016), and commonly observed contaminants (in total 84,610 sequences for *Homo sapiens*) generated using the SequenceReverser tool from the MaxQuant software (Version 1.0.13.13). The precursor ion tolerance was set to 10 ppm and fragment ion tolerance was set to 0.02 Da. The search criteria were set as follows: full tryptic specificity was required (cleavage after lysine or arginine residues unless followed by proline), three missed cleavages were allowed, carbamidomethylation (C), TMT6plex (K and peptide n-terminus) were set as fixed modification and oxidation (M) as a variable modification. Next, the database search results were imported to the Scaffold Q+ software (version 4.3.2, Proteome Software Inc., Portland, OR) and the protein false identification rate was set to 1% based on the number of decoy hits. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. Acquired reporter ion intensities in the experiments were employed for automated quantification and statistically analysed using a modified version of our in-house developed SafeQuant R script(Ahrné *et al.*, 2016). This analysis included adjustment of reporter ion intensities, global data normalization by equalizing the total reporter ion intensity across all channels, summation of reporter ion intensities per protein and channel, calculation of protein abundance ratios, and testing for differential

abundance using empirical Bayes moderated t-statistics. Finally, the calculated *P*-values were corrected for multiple testing using the Benjamini−Hochberg method.

Immunohistochemistry. Tissue was fixed in FormalFix for 24 h at 4°C, washed with 70% ethanol, and embedded in paraffin. Sections (3 μm) were prepared and processed for haematoxylin and eosin staining and for immunohistochemistry.

Wnt5a ELISA. Human Wnt5a BioAssay Elisa kits were purchased from US Biological. Elisa was performed according to the manufacturer's protocol.

Computational analysis. Sequenced reads were aligned against the human genome (hg19) using the R/Bioconductor package QuasR(Gaidatzis *et al.*, 2015) with the unspliced option and allowing only uniquely mapping reads unless differently specified. Raw gene counts were obtained by using the TxDb.Hsapiens.UCSC.hg19.known Gene package in QuasR, counting only alignments on the same strand as the query region. Differential gene expression was determined using edgeR(Robinson, McCarthy and Smyth, 2010), a cut-off of a linear fold change ≥2 and an adjusted *FDR*≤0.05 (corrected with the Benjamini–Hochberg algorithm method) were used. An integrated system for motif activity response analysis (ISMARA) was performed as described(Britschgi et al., 2017). For transcription-factor binding-site enrichment, we used oPOSSUM (v1) (http://opossum.cisreg.ca/oPOSSUM3/). We used Ingenuity Pathway Analysis (IPA) in the search for the Upstream Regulators. MetaCore pathway analysis was used for the analysis of phosphoproteomic data and we used String(Szklarczyk *et al.*, 2017) (https://string-db.org/cgi/input.pl?UserId=input_page_show_search=on) for analysis of proteomic data. Statistical analysis for String-related data was derived from the software integrated Fisher's exact test followed by a correction for multiple testing. GSEA was performed using the JAVA application from the Broad Institute v2.0 (http://www.broadinstitute. org/gsea). We used cBioPortal(Gao *et al.*, 2013),(Cerami et al., 2012) for the GR expression correlation study with publicly available data(Curtis *et al.*, 2012),(Pereira et al., 2016).

RFS, DMFS and PPS were generated using the 2017 version of KMplotter(Győrffy *et al.*, 2010) (http://kmplot.com/analysis/index.php?p=service&cancer=breast) and ROR1 and kinase-based signature relapse-free survival was generated using g-2-o platform (http://www.g-2-

o.com/?q=G2OBreast). The ROR1 signature represents a set of genes found to be overexpressed in breast tumours with ROR1 amplification(Lőrinc Pongor *et al.*, 2015).

Gene expression data of 88 metastases of breast cancer were obtained from Robinson et al., 2017(Robinson *et al.*, 2017) (website: https://met500.path.med.umich.edu/). After removing genes encoding for mitochondrial and ribosomal proteins, we performed library size normalization and log2 transformation of the remaining 18,115 genes. We created a signature of GR activation (SGR) by adding the expression of 8 genes (FN1, KLF9, ANKRD1, MT2A, VIM, SNAI2, POU5F1 and ID3) and explored the dataset utilizing Pearson correlation.

Statistical data analysis. The standard laboratory practice randomization procedure was used for cell line groups and animals of the same age and sex. The investigators were not blinded to allocation during experiments and outcome assessment. The required number of mice was calculated by performing power analysis using data from small pilot experiments. Values represent means ± s.d, unless differently stated. *P* values were determined using unpaired two-tailed *t*-tests and statistical significance was set at $P=0.05$. The variance was similar between the groups compared. Experimental replicates are independent experiments. Technical replicates are tests or assays run on the same sample multiple times. Data were tested for normal distribution and Student's *t*-tests (if normally distributed) or nonparametric Mann–Whitney U/Wilcoxon-tests were applied unless stated otherwise. Kaplan–Meier plots were generated using the survival calculation tool from Graphpad Prism and significance was calculated using the two-tailed log-rank test at $P<0.05$

4. Results Part III: Single-cell analysis reveals inter- and intra-tumour heterogeneity in metastatic breast cancer

Research Article – Submission to Journal of Mammary Gland Biology and Neoplasia

Single-cell analysis reveals inter- and intratumour heterogeneity in metastatic breast cancer

Baptiste Hamelin[1], Milan Obradović[1,3,6], Atul Sethi[1,3,5,6], Michal Kloc[1,5], Simone Muenst[2], Christian Beisel[4], Katja Eschbach[4], Hubertus Kohler[3], Savas Soysal[1], Marcus Vetter[1], Walter P. Weber[7], Michael B. Stadler[3,5], Mohamed Bentires-Alj[1,3]

[1]Department of Biomedicine, Department of Surgery, University Hospital Basel, University of Basel, Switzerland

[2]Institute of Pathology, University Hospital Basel, University of Basel, Switzerland

[3]Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland

[4]Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland

[5]Swiss Institute of Bioinformatics, Basel, Switzerland

[6]Roche, Basel, Switzerland

[7]Breast Center, Department of Surgery, University Hospital Basel, Switzerland

*Correspondence to: Mohamed Bentires-Alj, Department of Biomedicine, Department of Surgery, University Hospital Basel, University of Basel, Switzerland, m.bentires-alj@unibas.ch

4.1 Abstract

Metastasis is the leading cause of cancer-related deaths of breast cancer patients. Some cancer cells in a tumour go through successive steps, referred to as the metastatic cascade, and give rise to metastases at a distant site. We know that the plasticity and heterogeneity of cancer cells play critical roles in metastasis but the precise underlying molecular mechanisms remain elusive. Here we aimed to identify molecular mechanisms of metastasis during colonization, one of the most important yet poorly understood steps of the cascade. We performed single-cell RNA-Seq (scRNA-Seq) on tumours and matched lung macrometastases of patient-derived xenografts of breast cancer. After correcting for confounding factors such as the cell cycle and the percentage of detected genes (PDG), we identified cells in three states in both tumours and metastases. Gene-set Enrichment Analysis revealed biological processes specific to proliferation and invasion in two states. Our findings suggest that these states are a balance between epithelial-to-mesenchymal (EMT) and mesenchymal-to-epithelial transitions (MET) traits that results in so-called partial EMT phenotypes. Analysis of the top Differentially Expressed Genes (DEGs) between these cell states revealed a common set of partial EMT Transcriptions Factors (TFs) controlling gene expression, including ZNF750, OVOL2, TP63, TFAP2C and HEY2. Our data suggest that the TFs related to EMT delineate different cell states in tumours and metastases. The results highlight the marked interpatient heterogeneity of breast cancer but identify common features of single cells from five models of metastatic breast cancer.

4.2 Interpatient heterogeneity is dominant over intrapatient heterogeneity

Breast cancer is the most frequent cancer type in women worldwide, causing about 700,000 deaths per year(Sung *et al.*, 2021). Most of these fatalities result from metastasis(Hanahan and Robert A. Weinberg, 2011), a multi-step process in which cells from the tumour disseminate and colonize distant organs. Previous work has shed light on the different stages undergone by these cancer cells: invasion of the tissue surrounding the tumour, intravasation and dissemination as circulating tumour cells, extravasation, and colonization of the distant site. This process(Massagué and Obenauf, 2016) involves phenotypic changes that increase the resistance of specific cells to the conditions of the "foreign" environment and result in metastasis(Fares *et al.*, 2020). Important in this regard is the plasticity and stemness of cancer cells, which reversibly result in epithelial, mesenchymal or stem cell-like states(Jehanno *et al.*, 2022). Secondly, the inherent heterogeneity of cancer cell populations (Almendro, Marusyk and Polyak, 2013),(Koren and Bentires-Alj, 2015) at the genetic, epigenetic and microenvironmental levels predisposes some cells to the foreign environment. Thus, identifying the molecular mechanisms underlying metastatic progression is paramount to understanding this currently incurable disease and improving patient care. New technologies have been leveraged to better characterize the drivers of metastasis at various stages of the cascade, but some remain elusive due to the lack of granularity of bulk-sequencing approaches.

To better understand heterogeneity at the single-cell level, we orthotopically implanted four patient-derived xenografts (PDX)(Gao, Joshua M. Korn, *et al.*, 2015) and a cell line (see supplementary table) known for their lung metastatic potential into NOD-SCID-Il2rg^{null} (NSG) mice (Fig. 4.1a). Tumours were resected from the mammary fat pad and the animals were monitored for metastasis. Once the animals showed signs of distress (i.e., weight loss, difficulty to breath), lungs presenting metastatic lesions were collected and processed for single-cell transcriptional profiling. To exclude murine cells from the downstream analysis, human cancer cells from the tumours and matched lung metastases were purified via FACS gating GFP-positive MDA-MB-231 cells (Fig. 4.1b) or CD298-positive cells for PDX models (Fig. 4.1c). Single cells were isolated using a microfluidic device (Fluidigm C1) ahead of library preparation and

sequencing. This workflow yielded a total of 1,523 single cells (Supplementary Fig. 4.1a) after RNA-Seq and quality control.

Initial clustering of the quality-controlled data revealed that the cells formed groups (clusters) according to the donor models (Fig. 4.1d). Within each cluster, cells did not clearly separate based on their origin, tumour, or metastasis (Fig. 4.1e). These observations highlight the importance of interpatient over intrapatient heterogeneity.

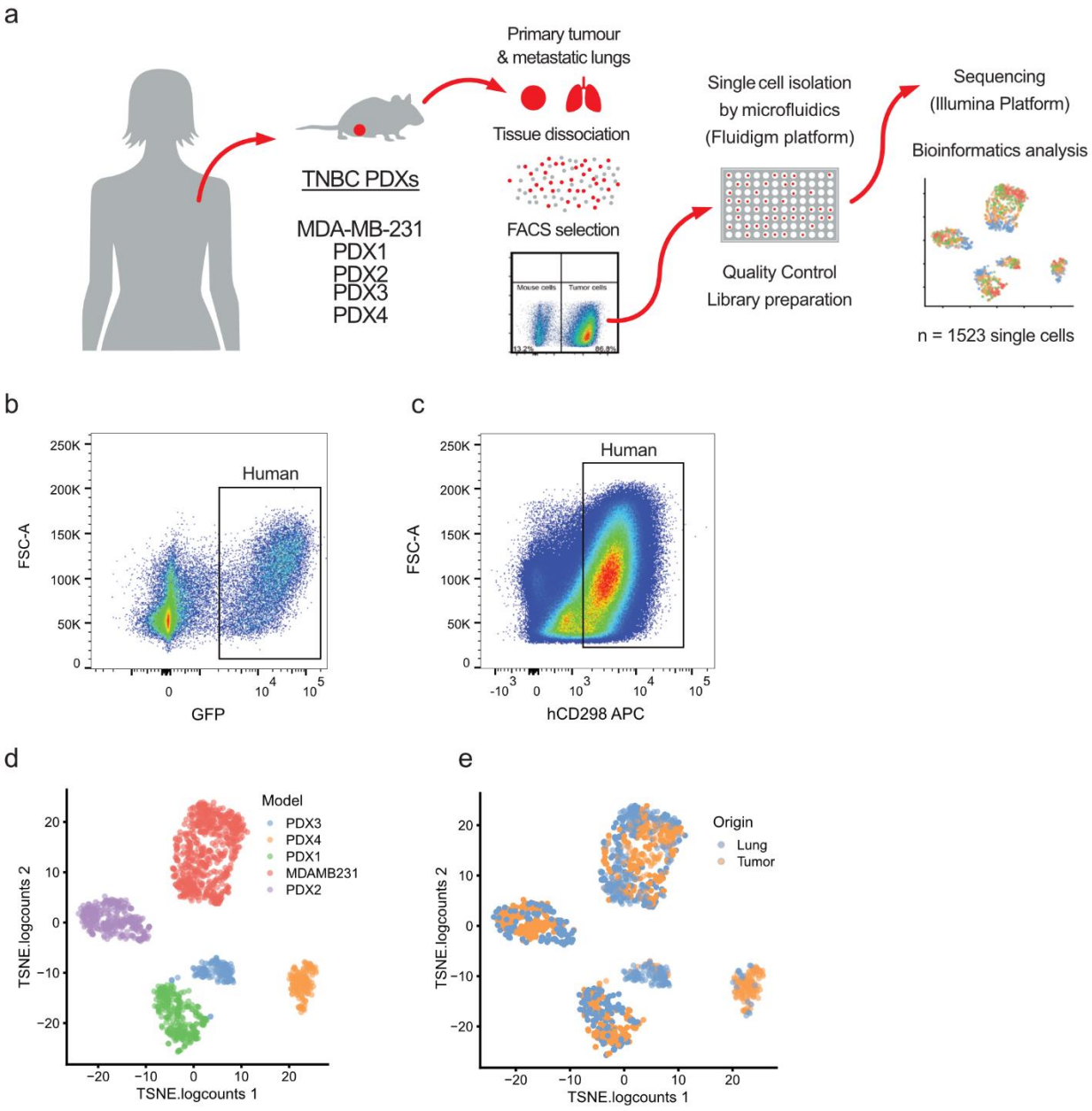


Figure 4.1. Interpatient heterogeneity is dominant over intrapatient heterogeneity. **a**, Overview of the experimental setting and models used. Human breast cancer models were implanted in the mammary fat pad of NSG mice. Tumours and lung macrometastases were harvested and mechanically and enzymatically dissociated. Human cells were purified by FACS using GFP or CD298 staining. Single cells were isolated with the Fluidigm C1 microfluidic platform and then sequenced. **b**, Representative FACS strategy for the

isolation of MDA-MB-231 GFP positive. **d**, tSNE plot showing initial clustering of all the sequenced cells, according to models of origin. **e,** tSNE plot showing initial clustering of all the sequenced cells, according to the site of origin.

4.3 Cell cycle and percentage of detected genes delineate cell clustering

We then asked whether cells gather within each donor-specific cluster by known biological or technical features. We projected the percentage of detected genes (PDG) in each single-cell library (Fig. 4.2a) onto tSNE and found that the PDG influences the clustering of the cells within each given model. This variable potentially represents both biological and technical effects. Next, we assessed whether the cell cycle stage influenced the analysis of the data, as this biological variable has a broad impact on gene expression(Bristow, Leman and Haase, 2014). We developed a method to infer cell cycle stages in single cells (in r package gripgh) and each cell was labelled with one of the four labels (G1, G1/S, S/G2, G2/M). We found that the cell cycle stage does influence the clustering of the cells within each model (Fig. 4.2b).

We then applied graph-based clustering to the different models and obtained 16 subclusters (Fig. 4.2c). These subclusters were mostly composed of cells in a similar cell cycle stage (Fig. 4.2d, left bar graph) and with similar PDGs (Fig. 4.2d, left bar graph). The influences of both the cell cycle and the PDG were also observed when the models were analysed individually (Supplementary Fig. 4.2a, top bar graphs for each model). Marked interpatient heterogeneity also led to the clustering of the cells according to models (Supplementary Fig. 4.2b). Altogether the data suggest that, in contrast to the site of origin (i.e., primary tumour or metastasis), the PDG and the cell cycle both influence the clustering of cells (Fig. 4.1e).

For the cell cycle stage prediction we considered that the cell cycle is not composed of discrete stages but is more a continuum of states. Gene expression is gradually modulated as a cell progresses in the cycle. Using this riche single-cell RNA sequencing data and genesets whose expression varies in different cell cycle stages, we created a circular trajectory and placed the cells in a cycle (Fig. 4.2e). This precise allocation along the cell cycle continuum allowed precise cell cycle staging.

To observe underlying biological processes involved in different tumour cells and to (beyond donor effect, PDG and cell cycle stage) and group biologically similar cells, we needed to remove the confounding factors of donor effect, PDG and cell cycle stage. To remove biases attributed to these factors, we used a generalized linear model (GLM)(Risso *et al*, 2014), correcting gene expression according to the position of the cell on the cell cycle spectrum and the complexity of the RNA-Seq library that it yielded (Fig. 4.3a).

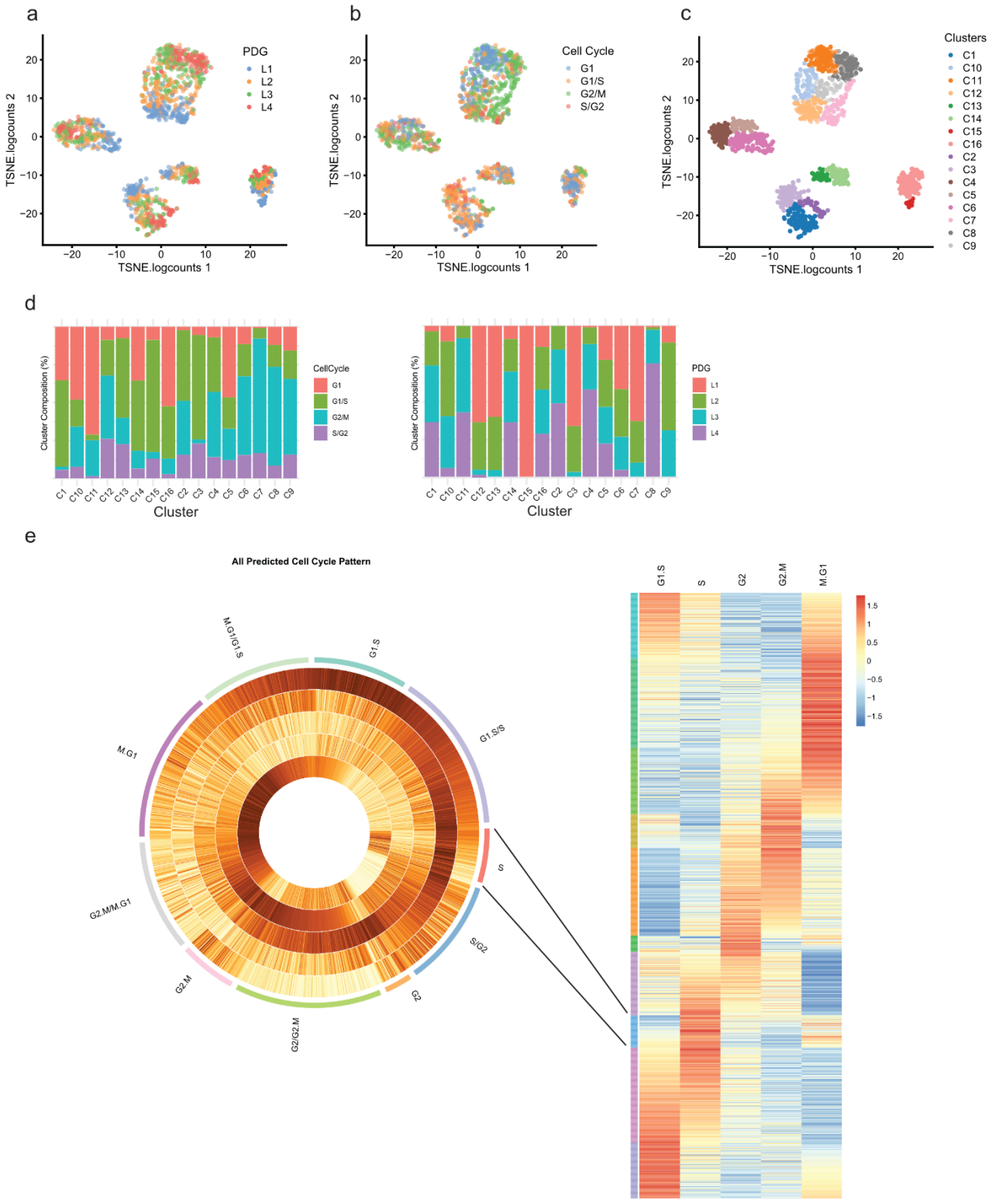


Figure 4.2. Cell cycle and percentage of detected genes delineate cell clustering. Single cells within PDX models cluster by **a**, library complexity (PDG) and **b**, cell cycle similarity. **c**, Clustering performed via t-SNE produced 16

cell clusters. **d**, Cluster composition according to cell cycle (left) and library complexity (right) **e**, An in-silico cell cycle scoring prediction to arrange single cells on a continuous cell cycle spectrum in addition to distinct cell cycle stages.

4.4 Removal of cell cycle variations and percentage of detected genes reveals 3 major biological clusters

Initially we performed gene set enrichment analysis (GSEA) with the Hallmarks gene set (Liberzon *et al.*, 2015) (Supplementary Fig. 4.3a) on clusters defined for each model on the corrected data (Supplementary Fig. 4.2a, bottom bar graphs for each model, named Ax, Bx, Fx, Dx, Ex). We observed that cell clusters from different models show enrichment in a common set of biological processes. This suggests that each model contains cells in closely related biological states that could now be visualized after bias correction.

We then analysed the cells irrespective of their models or sites of origin. Using the corrected gene expression data, the cells formed 14 new clusters (Cx). The compositions of these groups were then analysed according to cell cycle stage (Fig. 4.3b, left), PDG (Fig. 4.3b, centre), and model (Fig. 4.3b, right). Correcting for these variables led to unbiased clustering of the cells, with clusters composed of cells from various cell cycle stages, library complexities, and models. We also plotted the composition of each cluster in terms of site of origin before and after correction (Supplementary Fig. 4.3b). Corrected clusters exhibited a more balanced composition, with roughly equal proportions of cells originating from the tumour and lung metastases. The data suggest that populations of cells clustering together due to their biological similarities can be found in both the tumour and the metastatic sites, without specificity to one or the other.

To further investigate the different biological states suggested by the data in Supplementary Fig. 4.3a, we performed GSEA on the 14 clusters formed by all the cells, regardless of their model of origin. The 14 clusters formed three “super” biological clusters (Fig. 4.3c). Supercluster A (C8 to C6, Fig. 4.3c left) is characterized by low enrichment for most of the Hallmark geneset pathways. Supercluster B (C13 to C2, Fig. 4.3c centre) displays the most heterogeneous regulatory landscape, with highly, moderately, and minimally enriched processes. This supercluster is defined by highly enriched Hypoxia and TNFα signalling via NFκB. The moderately enriched pathways

include relevant processes such as EMT, TGFβ signalling, Interferon Gamma response, or the P53 Pathway. Finally, the least enriched genesets include MYC signalling, G2M checkpoint, and E2F targets. Interestingly these processes are most enriched in supercluster C (C7 to C4, Fig. 4.3c right), which also displays marked enrichment of genesets pertaining to oxidative phosphorylation, mTOR signalling, fatty acid metabolism, and DNA repair. Genesets enriched in supercluster B suggest a phenotype related to EMT, while cells in supercluster C appear to be proliferating while still partially enriched for EMT-related pathways.

We then plotted the repartition of the cells according to supercluster allocation (Supplementary Fig. 4.4a), noticing that cells from superclusters B and C were the most distant, with cells from supercluster A in between. We also assessed repartition according to the site of origin (Supplementary Fig. 4.4b) and once again found a relative balance between tumour origin and lung metastases origin of the cells forming the superclusters.

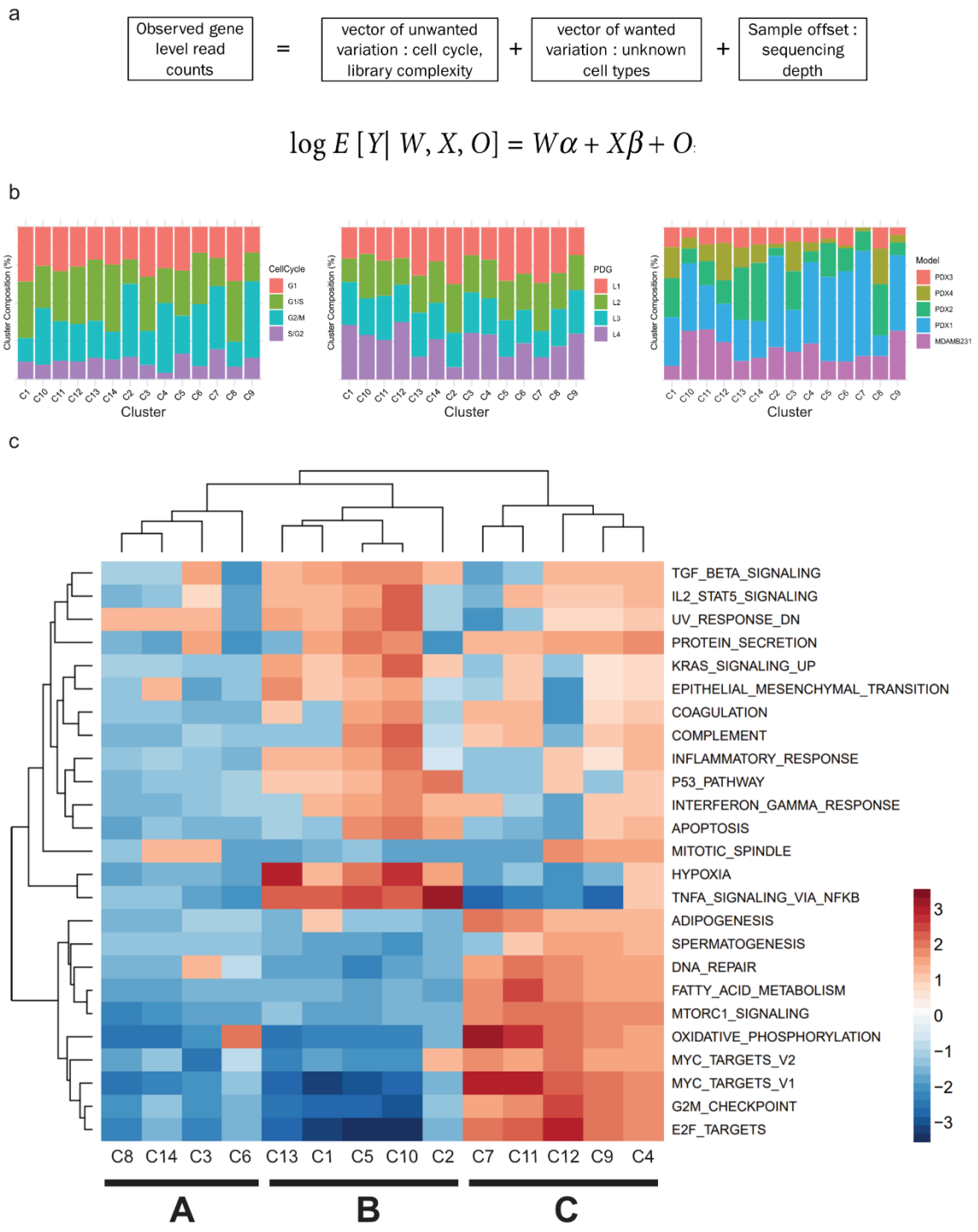


Figure 4.3. Removal of cell cycle variation and percentage of detected genes reveals three major biological clusters. a, Generalized linear model approach used to remove biases (cell cycle, and library complexity). **b**, Post-correction the new cell clusters have a more balanced distribution of cells from different cell cycles, library complexity (PDG), and cell source (PDX model). **c**, Gene Set Enrichment Analyses (GSEA) of individual clusters reveal common

and different Hallmark genesets enriched among clusters. Top 25 of these genesets according to the absolute values of the normalized enrichment score (NES) are shown and were used to identify the superclusters.

4.5 Comparison of major biological superclusters reveal partial EMT states regulators

Next, we selected EMT and Proliferation markers significantly altered (FDR<0.05) in pairwise comparisons between the superclusters (Fig. 4.4a). Proliferation markers Ki67, MCM3, and PCNA(Juríková *et al.*, 2016) confirmed that supercluster C is the most proliferative, followed by cluster A, while supercluster B expresses these markers the least. EMT markers indicated that this process was taking place at varying levels across the different superclusters. Supercluster A displayed the least engagement in the transition according to its low expression of several EMT markers (ZEB1, SOX9, SNAI1, FN1, TGFBR1). Superclusters B and C showed increased expression of these markers but at varying levels. Such heterogeneity suggests that these superclusters may undergo EMT but could be at different stages of the process. Such partial EMT has previously been described(Saitoh, 2018; Pastushenko and Blanpain, 2019) and may reflect the balance between proliferative potential and migratory capability, both properties being typical of different stages of the metastatic cascade.

To investigate partial EMT states of the superclusters, we selected the top 250 up- and downregulated genes of each cluster and performed GSEA as well as Transcription Factor Enrichment Analysis (TFEA) with the EnrichR(Xie *et al.*, 2021) platform (Fig. 4.4a, b, c bar plots).

The “TF Perturbations followed by expression” geneset, which was generated by the curation of experiments altering TFs before measuring gene expression, revealed that different TFs govern the top differentially expressed genes (DEGs) across the superclusters. These uncommon TFs paint a complex picture of the EMT states existing in superclusters A, B, and C. HEY2, OVOL2, TFAP2C, and TP63 are TFs described as regulators of partial EMT states in mouse models(Pastushenko and Blanpain, 2019). ZNF750 was recently described as an EMT repressor in breast cancer(Cassandri *et al.*, 2020), an activity shared by OVOL2(Watanabe *et al.*, 2014), TP63(Pastushenko and Blanpain, 2019), and FOXO1(Jiramongkol and Lam, 2020). HEY2, TFAP2C (Cyr *et al.*, 2015) (Kim *et al.*, 2016), SOX4, and SOX9(Grimm *et al.*, 2020) are known

to promote EMT. U2AF1 is a splicing factor fine-tuning translation with reported effects in development and EMT.

These TFs control the EMT and proliferation state of the superclusters shown in Fig. 4.4a. Supercluster A, shown to be mildly proliferative, is controlled by TFs evocating differentiated slowly proliferating cells. E2F4 is known to be abundant in differentiated cells and to repress proliferative genes(Sun *et al.*, 2019). TP63 and TFAP2C have been described as controlling early hybrid EMT states, with cells close to an epithelial state and more prone to proliferate than mesenchymal cells. Additionally, KDM5B has been reported to characterize a slow-cycling cell subpopulation in melanoma(Roesch *et al.*, 2010), which fits the traits of supercluster A. Supercluster B, the least proliferative, is also the one with the strongest expression of canonical EMT markers/TFs such as FN1, SNA1, MYC, and SOX9. When compared to superclusters A and C, DEGs from supercluster B appear to be under the control of HEY2, a member of the basic Helix-Loop-Helix (bHLH) transcription factor family. bHLH have been described as late hybrid EMT regulators(Pastushenko and Blanpain, 2019) responsible for mesenchymal phenotypes that favour migration and have little proliferation potential. Supercluster C, the most proliferative, is controlled by ZNF750, TFAP2C, and OVOL2. These TFs have been described as either EMT repressors or regulators of early hybrid EMT stages, corresponding to epithelial-like phenotypes permitting proliferation. It should however be noted that the EMT markers are high in this supercluster, indicating that the process is ongoing yet likely oriented towards the proliferation of cells rather than their migration.

Next, we analysed individual hits from the pairwise supercluster comparisons (Fig. 4.4a,b,c, volcano plots). Several members of the Activator Protein 1 (AP1) family of transcription factors (JUN, FOS, ATF3) were consistently altered in the different superclusters. This is of importance as AP1 has been reported as one of the “core” TFs binding enhancers regulating epithelial and mesenchymal states(Sethuraman *et al.*, 2018; Pastushenko and Blanpain, 2019). This core is subsequently modulated by other TFs such as those described in the previous paragraph. Another element strongly differing between superclusters was BHLHE40, a member of the bHLH TF family, which was found strongly downregulated in supercluster A. BHLHE40 has been reported to induce EMT as well as tumour growth and lung metastases via HBEGF exosomal

release(Sethuraman *et al.*, 2018). Different types of RNA-coding genes are modulated during EMT, some of which (NEAT1, MALAT1) figured in the top DEGs across the superclusters. NEAT1 was found upregulated in superclusters A and B compared to cluster C. NEAT1 has been described as promoting chemoresistance and cancer stemness(Jiang *et al.*, 2018; Shin *et al.*, 2019; Park *et al.*, 2021). More importantly, it was found to enhance glycolysis as a scaffold of key glycolytic enzymes. Its downregulation in supercluster C, which is enriched for fatty acid metabolism and the oxidative phosphorylation process, is notable. This observation may reflect a metabolic shift away from glycolysis and towards oxidative phosphorylation that was shown to exacerbate breast cancer lung micrometastases. Like NEAT1, MALAT1 is upregulated in superclusters A and B compared to C. Its effects are manifold and some controversy exists about its activities in different cancer types and settings(Chen, Zhu and Jin, 2020; Shaath *et al.*, 2021). It is also interesting to note that MALAT1 has been described to interact with TEAD, which is part of the core TFs controlling the EMT process.

Finally, to relate each of the biological states to patient outcome, we selected the upregulated transcripts characteristic of each supercluster by overlapping the different contrasts (see Supplementary Table 2). The transcripts defining supercluster A correlate with better relapse-free survival (RFS) in patients suffering from basal-like breast cancer (Supplementary Fig. 4.4c). On the contrary, upregulated transcripts found in superclusters B and C (Supplementary Fig. 4.4d and 4.4e respectively) correlated with worse RFS. These results suggest that these cell states may exhibit different levels of aggressiveness driven by different sets of transcripts.

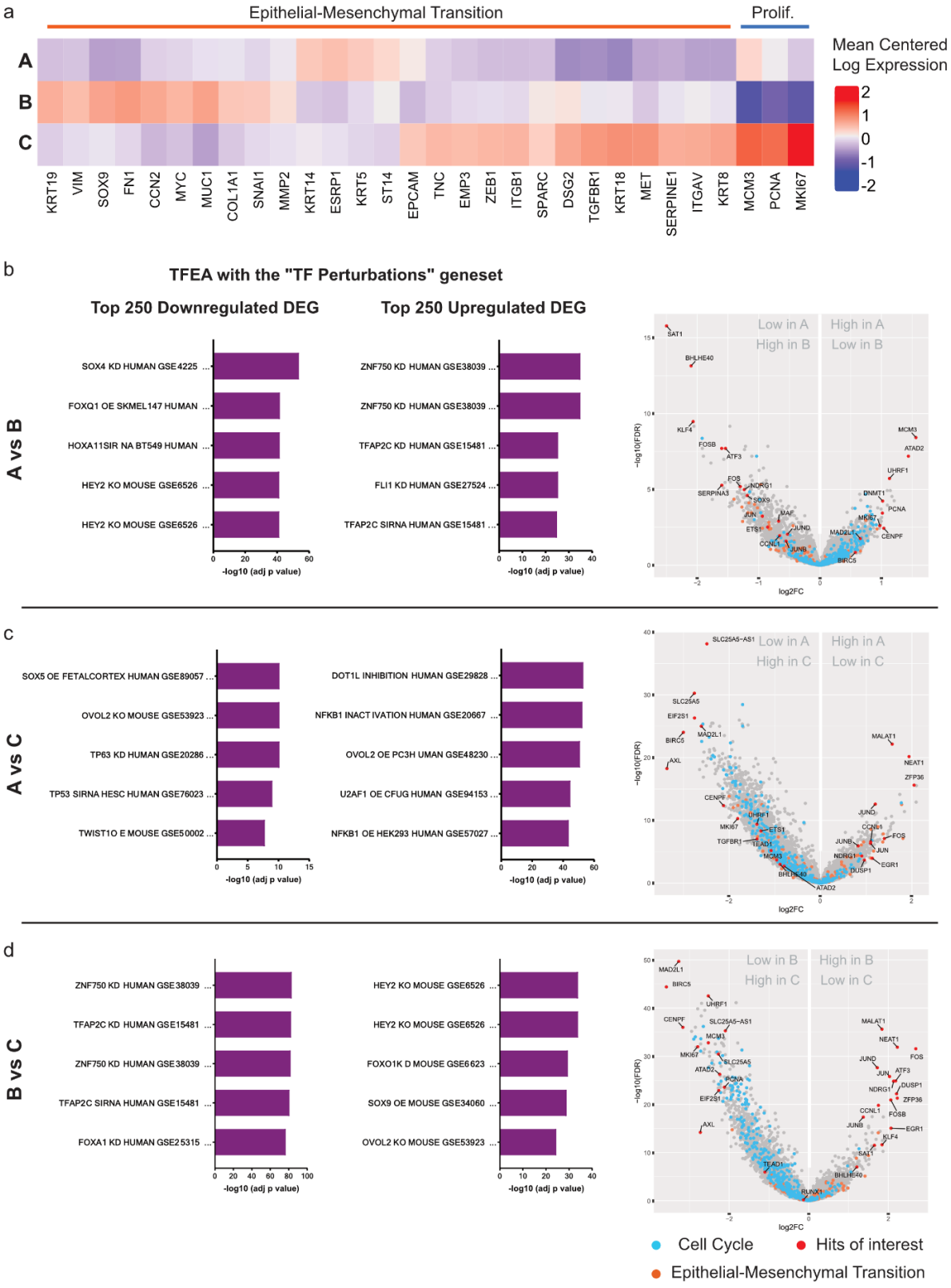

a
Epithelial-Mesenchymal Transition
Prolif.
A
B
C
KRT19 VIM SOX9 FN1 CCN2 MYC MUC1 COL1A1 SNAI1 MMP2 KRT14 ESRP1 KRT5 ST14 EPCAM TNC EMP3 ZEB1 ITGB1 SPARC DSG2 TGFBR1 KRT18 MET SERPINE1 ITGAV KRT8 MCM3 PCNA MKI67
Mean Centered Log Expression
2
1
0
-1
-2
b
TFEA with the "TF Perturbations" geneset
Top 250 Downregulated DEG
Top 250 Upregulated DEG
A vs B
SOX4 KD HUMAN GSE4225 ...
FOXQ1 OE SKMEL147 HUMAN ...
HOXA11SIR NA BT549 HUMAN ...
HEY2 KO MOUSE GSE6526 ...
HEY2 KO MOUSE GSE6526 ...
ZNF750 KD HUMAN GSE38039 ...
ZNF750 KD HUMAN GSE38039 ...
TFAP2C KD HUMAN GSE15481 ...
FLI1 KD HUMAN GSE27524 ...
TFAP2C SIRNA HUMAN GSE15481 ...
-log10 (adj p value)
Low in A
High in B
High in A
Low in B
SAT1
BHLHE40
KLF4
FOSB
ATF3
FOS
NDRG1
SERPINA3
SOX9
JUN
MAF
ETS1
JUND
CCNL1
JUNB
MAD2L1
BIRC5
MCM3
ATAD2
UHRF1
DNMT1
PCNA
MKI67
CENPF
-log10(FDR)
log2FC
c
A vs C
SOX5 OE FETALCORTEX HUMAN GSE89057 ...
OVOL2 KO MOUSE GSE53923 ...
TP63 KD HUMAN GSE20286 ...
TP53 SIRNA HESC HUMAN GSE76023 ...
TWIST1O E MOUSE GSE50002 ...
DOT1L INHIBITION HUMAN GSE29828 ...
NFKB1 INACT IVATION HUMAN GSE20667 ...
OVOL2 OE PC3H UMAN GSE48230 ...
U2AF1 OE CFUG HUMAN GSE94153 ...
NFKB1 OE HEK293 HUMAN GSE57027 ...
-log10 (adj p value)
Low in A
High in C
High in A
Low in C
SLC25A5-AS1
SLC25A5
EIF2S1
MAD2L1
BIRC5
AXL
CENPF
UHRF1
MKI67
ETS1
TGFBR1
TEAD1
MCM3
BHLHE40
ATAD2
MALAT1
NEAT1
ZFP36
JUND
CCNL1
JUNB
FOS
JUN
NDRG1
EGR1
DUSP1
-log10(FDR)
log2FC
d
B vs C
ZNF750 KD HUMAN GSE38039 ...
TFAP2C KD HUMAN GSE15481 ...
ZNF750 KD HUMAN GSE38039 ...
TFAP2C SIRNA HUMAN GSE15481 ...
FOXA1 KD HUMAN GSE25315 ...
HEY2 KO MOUSE GSE6526 ...
HEY2 KO MOUSE GSE6526 ...
FOXO1K D MOUSE GSE6623 ...
SOX9 OE MOUSE GSE34060 ...
OVOL2 KO MOUSE GSE53923 ...
-log10 (adj p value)
Low in B
High in C
High in B
Low in C
MAD2L1
BIRC5
UHRF1
CENPF
SLC25A5-AS1
MCM3
MKI67
SLC25A5
ATAD2
PCNA
EIF2S1
AXL
TEAD1
RUNX1
MALAT1
NEAT1
JUND
FOS
JUN
ATF3
DUSP1
NDRG1
ZFP36
CCNL1
FOSB
JUNB
EGR1
KLF4
SAT1
BHLHE40
-log10(FDR)
log2FC
Cell Cycle
Hits of interest
Epithelial-Mesenchymal Transition

Figure 4.4. Comparison of major biological superclusters reveals partial EMT state regulators. a, EMT and Proliferation markers significantly altered (FDR<0.05) in supercluster pairwise comparisons. **b**, **c** and **d**, Right: Transcription Factors Enrichment Analysis of the top 250 up- and downregulated differentially expressed genes of each supercluster. Left: Volcano plots highlighting the cell cycle gene set (blue, Bioplanet 2019), the EMT gene set (orange, mSigDB) and hits of interest (red) in corresponding supercluster pairwise comparisons.

4.6 Discussion

The cellular plasticity arising from the transitions between epithelial and mesenchymal states has been recognized as instrumental to metastatic progression(Jehanno *et al.*, 2022). The characterization of cancer cells at different stages of the metastatic cascade in terms of canonical EMT markers and TFs has recently given rise to the concept of partial EMT(Zhang and Weinberg, 2018; Pastushenko and Blanpain, 2019). The balanced epithelial and mesenchymal traits in partial EMT states results in positive reactions to the conditions imposed by the metastatic process(Lüönd, Tiede and Christofori, 2021) such as dissemination as CTCs or colonization of a foreign microenvironment. Characterization of these partial EMT states is however delicate, model- and context-dependent, and a task for which the use of canonical EMT markers and TFs is not sufficient.

Using scRNA-Seq on tumours and lung macrometastases of PDX models of TNBC, we observed marked heterogeneity at the intra- and interpatient levels. After correcting for the two confounding factors cell cycle and PDG, we identified three cell states (superclusters) present at both sites. These superclusters differ from each other in EMT and proliferation markers. Analysis of the top DEGs of each supercluster revealed that the variation between them is controlled by a set of TFs (i.e., TFAP2C, ZNF750, OVOL2, TP63, HEY2, bHLHs), which have been recently shown to finely modulate partial EMT states in other models(Pastushenko *et al.*, 2018). These TFs act by modulating core TFs such as AP-1 (composed of its JUN and FOS subunits), which we also found deeply altered across the superclusters. Our findings highlight the presence of partial EMT states controlled by the aforementioned TFs in breast cancer PDX models, both in the tumour and lung macrometastases. While our data have not yet allowed identification of factors driving lung macrometastases specifically, they still shed light on breast cancer tumour and metastases biology at this specific stage. Additionally, our results highlight the importance for breast cancer of several elements such as NEAT1 or MALAT1. These lncRNAs have been recently implicated in breast cancer initiation, growth, metastasis, and chemoresistance(Jiang *et al.*, 2018; Shin *et al.*, 2019; Chen, Zhu and Jin, 2020; Shaath *et al.*, 2021; Zhang *et al.*, 2021). Their downregulation in the most proliferative supercluster questions their roles in these cells and whether choosing them as

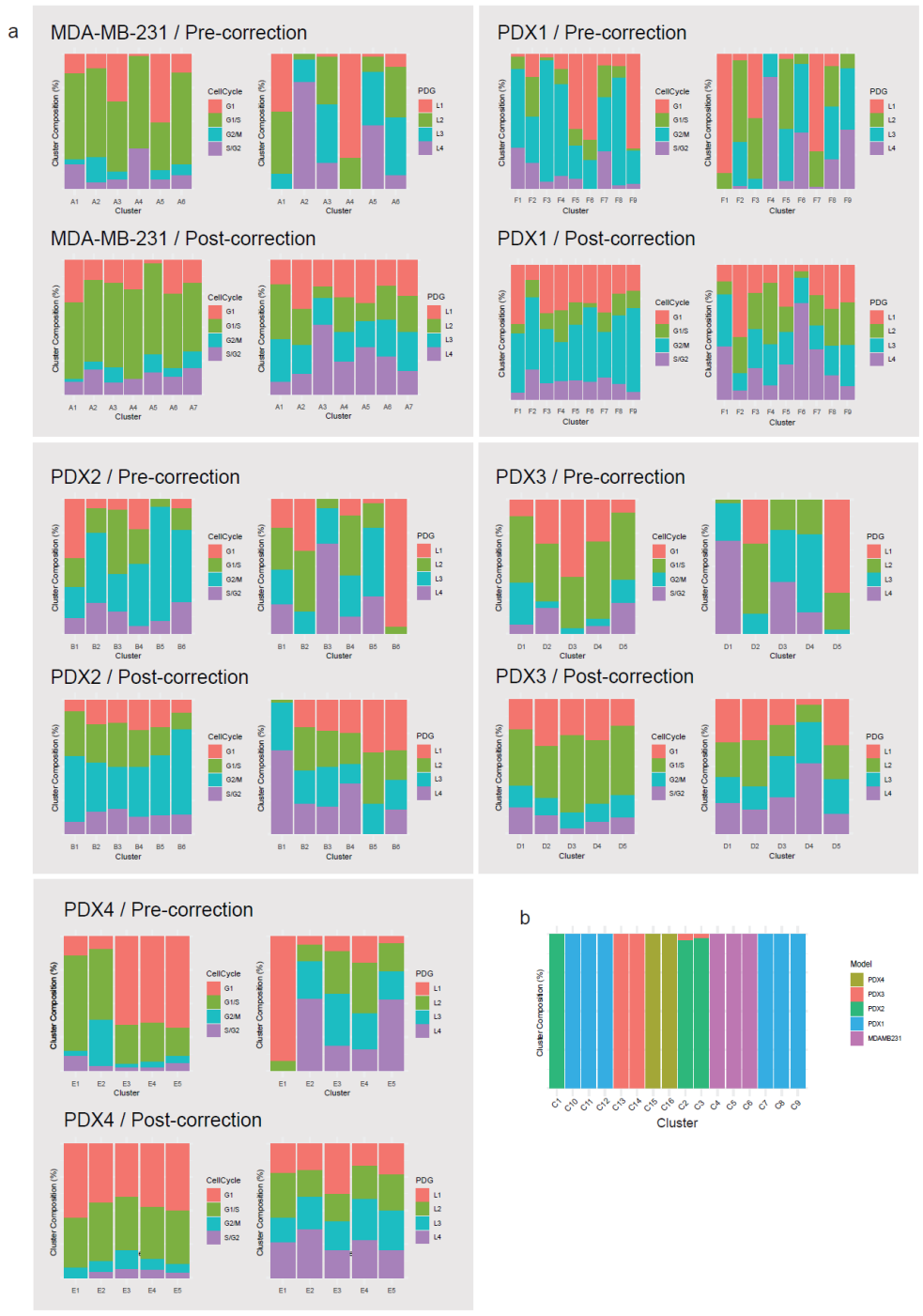
therapeutic targets could affect the different cell states identified here. Such dynamics may ultimately affect how tumours and metastases respond to treatment. Our results highlight the marked heterogeneity of breast cancer cells. They call for further studies at the single-cell level to better characterize the different partial EMT states at different stages of the metastatic cascade. Future studies, especially those including the stromal and immune compartments, will further our knowledge of drivers of metastasis and how to tackle them.

4.7 Extended figures

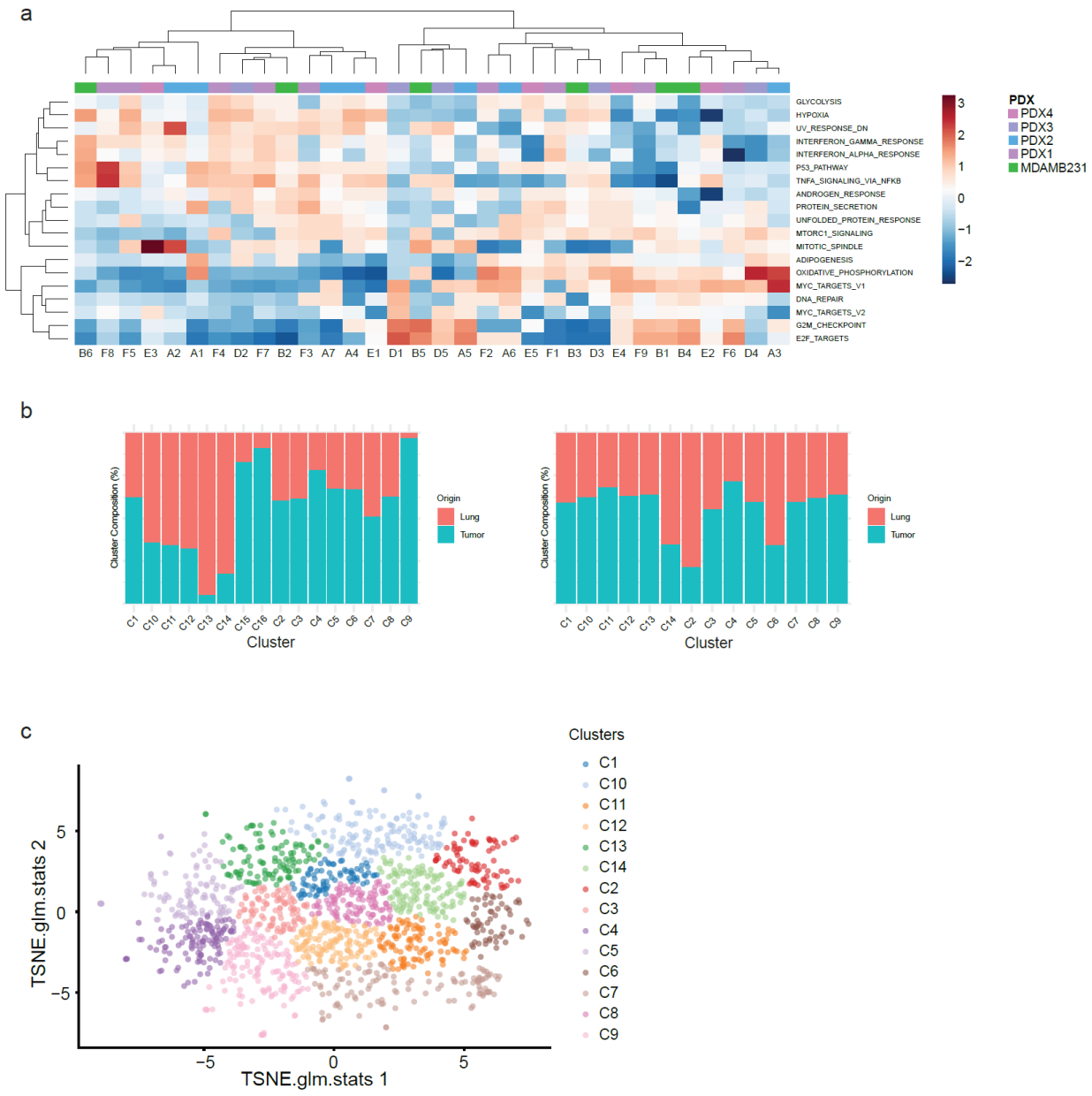
a

Model	Tumor	Lung Metastases
PDX4	151	17
PDX3	23	115
PDX2	177	114
PDX1	298	314
MDA-MB-231	221	93
Total	870	653

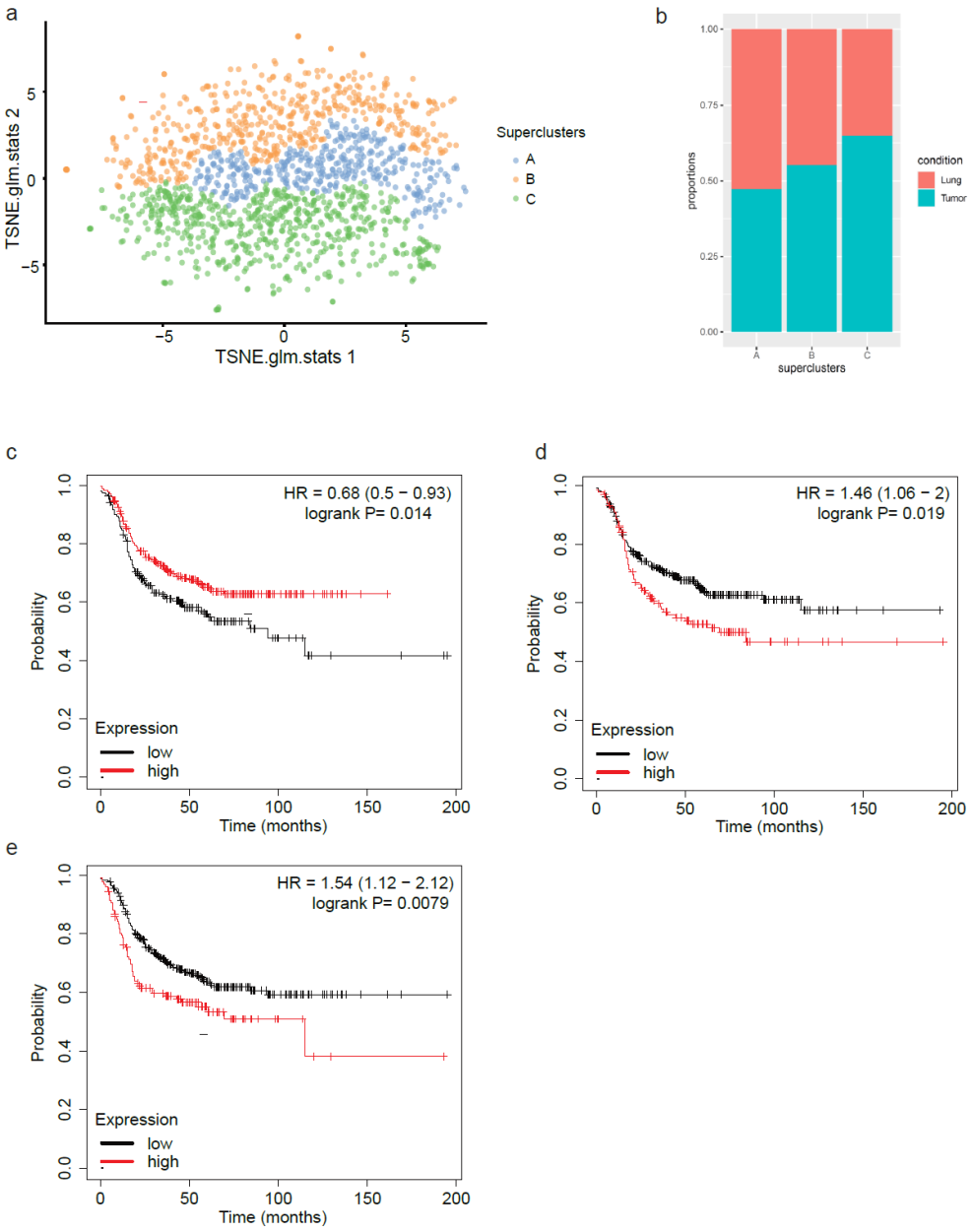
Extended Data Figure 4.1. a, Table summarizing the numbers of cells sequenced and processed for analysis for each model / site of origin after quality control



Extended Data Figure 4.2. a, Clustering and distribution of cells from different cell cycle stages and library complexity (PDG) before correction (top row) and after correction (lower row) for each model. b, Clustering performed over tSNE produces 16 cell clusters before bias correction



Extended Data Figure 4.3. a, Cell clusters identified separately from each individual model, show overlapping enrichment in Hallmark genets. b, Corrected clusters generated from all the cells irrespectively of the model of origin, showing the cluster composition according to site of origin. c, tSNE plot showing clustering of the corrected clusters



Extended Data Figure 4.4. a, tSNE plot showing the cell repartition across the 3 superclusters defined via GSEA. b, Barplot showing the supercluster composition according to site of origin. c, Kaplan-Meyer plot showing the RFS in patients with basal-like BC (n=442) according to expression of the upregulated transcripts defining supercluster A. d, Kaplan-Meyer plot showing the RFS in patients with basal-like BC (n=442) according to expression of the upregulated transcripts defining supercluster B. e, Kaplan-Meyer plot showing the RFS in patients with basal-like BC (n=442) according to expression of the upregulated transcripts defining supercluster C

4.8 Material and Methods

In vivo experiments.

All *in vivo* experiments were performed in accordance with the Swiss animal welfare ordinance and were approved by the cantonal veterinary office, Basel Stadt. Female NSG and BALB/c mice were maintained in the Friedrich Miescher Institute for Biomedical Research and the Department of Biomedicine animal facilities, in accordance with Swiss guidelines on animal experimentation. Mice were maintained in a sterile environment with light, humidity, and temperature control (light-dark cycle with light from 7:00 to 17:00, with a gradual change from light to dark), temperature 21–25 °C, and humidity 45–65%). Before each experiment, mice were allowed to acclimatize for a minimum of seven days. MDA-MB 231 cells (10,000 cells) were resuspended in 40 µl Matrigel:PBS (1:1) and injected into the pre-cleared mammary fat pads of 4- to 8-week-old female NSG mice. PDX models were transplanted into the pre-cleared 4th mammary fat pads of NSG mice. Tumours were resected when the largest diameter reached 10 mm, and mice were monitored regularly for signs of metastatic outgrowth and distress. In none of the experiments did tumour volumes exceed approved limits. All orthotropic experimental procedures (tumour resection and tumour cell implantation) were undertaken on anesthetized mice by a single investigator, according to protocols approved by the cantonal veterinary office Basel Stadt. Experimental metastasis assays were performed by injecting 100,000 cells into tail veins. After intravenous injection of MDA-MB 231 cells, we performed *in vivo* bioluminescence imaging to confirm injection and to monitor metastatic outgrowth. Bioluminescence imagining was performed using an IVIS Lumina XR (Caliper LifeSciences) upon injection of luciferin (Biosynth; L8220).

Cell lines and PDX models.

The cell lines MDA-MB 231 and HEK293T were purchased from ATCC and cultured according to ATCC protocols. Cell line identity was confirmed and routinely tested using short tandem repeat sequencing; all cell lines were routinely tested for mycoplasma contamination. MDA-MB 231 were propagated in monolayer cultures in DMEM supplemented with 10% FCS. All experiments

were performed with 70–90% confluent cells. The PDXs used in this study have previously been described(Gao, Joshua M. Korn, *et al.*, 2015). PDX models were passaged in NSG animals, via tumour piece implantation in the 4th mammary fat pad. Immunohistochemistry staining was performed to ensure the ER/PR/HER2 status.

Lentiviral vectors, lentivirus, and infection

Lentiviral batches were produced using PEI transfection on HEK293T cells as previously described(Britschgi, Duss, Kim, Couto, Brinkhaus, Koren, de Silva, *et al.*, 2017).The titre of each lentiviral batch was determined in MDA-MB 231 cells. Cells were infected for 8 h in the presence of polybrene (8 μg/ml). Selection with 2 μg/ml puromycin (Sigma) was applied 48 h after infection.

Sample preparation

Tumours and matched lung metastases were dissociated into single cells using mechanical disruption followed by enzymatic digestion using a collagenase-hyaluronidase solution (StemCell Technologies; 07912) for 1 h at 37°C without mechanical agitation. The resulting material was filtered twice on 40μm cell strainers and depleted of erythrocytes using a red blood cell lysis buffer (Sigma, R7757).

Fluorescence-activated cell sorting.

MDA-MB-231 expressing a GFP-Luciferase construct were selected based on GFP expression. PDX models cells selection relied on the CD298 human-specific marker. Sorts were performed on a BD FACS BD Aria III (70μm nozzle). DAPI staining was used to gate out dead cells. Single cells and doublets were respectively selected and excluded based on their forward and side scatter profiles and pulse width. The APC anti-human CD298 antibody (Biolegend, 341706) was used.

scRNA-Seq

Human cells were processed for single cell isolation and library preparation using the Fluidigm C1 platform. Single cell capture was performed by microfluidics on medium and large Fluidigm C1 integrated fluidics chip (IFC) (Fluidigm; PN100-5760 and PN100-5761). Visual quality control by

microscopy allowed the assessment of capture efficiency. cDNA was generated from the captured cells as per the manufacturer's protocol using SMART-Seq Ultra Low RNA (Takara Bio; 634833) before being processed for Illumina sequencing via the Nextera XT DNA Library Preparation kit (Illumina; FC-131-1096). Sequencing was performed on an Illumina platform.

Computational analysis.

RNA sequencing data analyses

Reads were aligned to the human genome (UCSC version hg38AnalysisSet) with STAR. The output was sorted and indexed with samtools. Strand-specific coverage tracks per sample were generated by tiling the genome in 20 bp windows and counting 5'end of reads per window using the function bamCount from the bioconductor package bamsignals. These window counts were exported in bigWig format using the bioconductor package rtracklayer. The rsubread::featureCounts function was used to count the number of reads (5'ends) overlapping with the exons of each gene assuming an exon union model (gene annotation: ensembldb_Homo_sapiens_GRCh38_ensembl_96.sqlite).

Removal of potential doublets

Following observation under the microscope, wells containing more than 1 cell or debris were marked and removed from further analyses. To further remove potential human and mouse multiplets, we used fastq_screen to count reads mapping uniquely to human and mouse genomes (human-to-mouse ratio). Libraries with a human-to-mouse ratio >=5 were considered human cells and retained for further analyses. Libraries with <= 100,000 reads mapping to human transcriptome were also filtered out. Libraries expressing less than 2,346 genes and more than 9,884 genes were also filtered out, retaining the cells that expressed 10% - 40% of the total number of unique genes observed in all models (23,459 genes). This step removed libraries with low complexity and a few outlier libraries where very high number of genes were observed.

Inference and correction of cell cycle signal

The cell cycles of individual cells were inferred using sets of cell cycle-regulated genes known to peak in transcription at given cell cycle stages obtained from (PMID: 12058064, PMID:11416145). For each cell, five normalized cell cycle stage scores (G1.S, G2, G2.M, M.G1, S) were calculated, and the cell assigned to that most closely resembling the expected profile based on correlation. Optionally, labels were further refined by iteratively estimating new cell cycle score profiles based on estimated cell labels, and re-assigning cells to the profile with the highest correlation. The function call predictCellCycle(org = "human.Whitfield", cor_thr = 0.2, refine_iter = 200) is available as part of R package griph (https://github.com/ppapasaikas/griph). The cell cycle and library complexity (percentage of detected genes) were modeled as covariates and regressed out of the log-normalized counts using glm.fit (R package glm).

Dimensionality reduction, clustering, differential gene expression

Each library was normalized to 100,000 reads and log transformed adding pseudocount of 1. PCA, tSNE, and UMAP projections (on log-normalized data or residuals) were computed with R package scater using default parameters. Nearest neighbour graphs were computed using function buildSNNGraph (R package scran) with tSNE and UMAP as inputs separately. Function cluster_louvain (R package igraph) was used for graph-based clustering. Differential gene expression between single cell clusters was performed using pairwise t-test implemented in FindMarkers function from R package scran. The values pval.type="some", min.prop=0.2 were used as arguments in FindMarkers indicating that the genes considered as marker genes are those differentially expressed in at least 20% of the pairwise comparisons, i.e. combined p-value from the pairwise comparisons was calculated by taking the minimal value of the top 20% Holm-corrected p-values for each gene. Genes were ranked by the respective combined logFC (summary.logFC) and gene set enrichment was performed with bioconductor packages fgsea and msigdb (collections H, C2, C5).

RFS Analysis

Kaplan-Meier plots were generated from the KMplotter database using the mRNA Gene Chip dataset. Upregulated transcripts specific to a supercluster were isolated (with the Venny online

tool) from the different contrasts generated from the analysis of the Top 250 DEGs used for the rest of the analysis. The generated lists then served as input to the KMplotter tool, via the Mean Expression for Multiple Genes function. RFS analysis was performed on the patients with basal-like breast cancer (PAM50 classifier, n=442) and with the Autoselect Best Cutoff parameter.

Supplementary Information

Acknowledgements

We thank the members of the Bentires-Alj team who participated in discussions about this project. We are also grateful to the core facilities of the FMI and DBM core facilities for supporting this work. We also thank Tim Roloff for his insights about single-cell RNA-seq. The Bentires-Alj laboratory research is supported by the Swiss Initiative for Systems Biology (SystemsX), the Swiss National Science Foundation, the European Research Council (ERC advanced grant 694033 STEM-BCPC), the Krebsliga Beider Basel, the Swiss Cancer League, the Swiss Personalized Health Network, and the Department of Surgery of the University Hospital Basel.

Author contributions

B.H analysed the data, interpreted the results, and wrote the manuscript. M.M.S.O conceived the study, designed, and performed the experiments, and analysed the data. S.M performed histopathological analysis of the PDX models. A.S performed computational analysis, designed the cell cycle/PDG correction method as well as analysis of the data and interpretated the results. M.K performed data analysis and result interpretation. C.B and K.E, helped with the experimental design and use of the single-cell platform. H.K performed fluorescence-activated cell sorting experiments, analysis, and interpretation of results. M.S advised on experiment design and data analysis. S.S, M.V, W.P.W provided patient material. M.B-A conceived the study, designed the experiments, and interpreted the results. All authors read and provided feedback on the manuscript.

Author information

B.H, M.M.S.O, A.S, M.K declare no competing interests. M.M.S.O and A.S are employees of Roche. M.B.-A. owns equities in Novartis.

Materials and correspondence

Correspondence and requests for materials should be addressed to Mohamed Bentires-Alj, m.bentires-alj@unibas.ch.

This manuscript contains 4 Figures, 4 Extended Data Figures, 2 Supplementary Tables.

Data availability.

The sequencing data have been deposited in the Gene Expression Omnibus (GEO) and are accessible under the accession code GSE202695. Processed transcriptomic data is available upon request to the corresponding author.

5. Conclusion & Perspectives

The phenotypic heterogeneity and plasticity of cancer cells have been described as an important driver of the metastatic cascade. In the two projects described we explored the molecular basis of heterogeneity and plasticity in breast cancer metastasis using bulk and single-cell RNA-Seq.

5.1 Glucocorticoids promote breast cancer metastasis

In the first project we leveraged bulk RNA-Seq to compare tumour and matched metastases gene expression. We identified an increase in GR signalling in lung metastases. We show that GR activation led to the induction of signalling networks implicated in disease progression. Among these ROR1 was demonstrated to mediate the effect of GR activation on metastatic colonization.

These findings are relevant at several levels. First, they demonstrate the heterogeneity of breast cancer. Our results show that while GR has little influence on the primary tumour, TNBC metastases display a different phenotype and benefit from GR activation in a distant site, leading to increased lung colonization.

Secondly, they raise pertinent questions regarding the proper characterization and care of breast cancers. Our work highlights the effect of the stress hormone pathway in BC progression. This signalling is mediated via the glucocorticoid receptor (NR3C1). Upon binding of its ligands, GR, a nuclear receptor, is internalized and translocated to the nucleus where it binds GR response elements (GRE) in the genome(Scheschowitsch, Leite and Assreuy, 2017). GR signalling is implicated in many physiological and pathological settings(Kadmiel and Cidlowski, 2013), with sometimes unclear roles(Mayayo-Peralta, Zwart and Prekovic, 2021).

Stress hormones (such as cortisol) are found to be elevated in BC patients(Van Der Pompe, Antoni and Heijnen, 1996). They are a first mean of GR stimulation and represents an environmental variable influencing metastatic progression. Stress, which can take many forms, is therefore a parameter to consider in the prevention of BC and the care of its patients.

More importantly, synthetic glucocorticoids are routinely administered in various clinical settings. In BC dexamethasone is associated with chemotherapy in the aim of alleviating secondary

effects such as nausea. Dex acts by modulating the inflammation caused by the chemotherapy. Our findings are particularly relevant in this setting, as Dex is administered to patients in which cancer cells most likely already disseminated to distant sites(Hosseini *et al*., 2016;Braun *et al*., 2005). The administration of Dex in this context may precipitate the relapse of the patient. This is of particular concern when treating TNBC patients, who already suffer from a particularly aggressive subtype mostly treated with chemotherapy.

In recent years development of new treatments provided new alternatives but also raised new questions. Nab-Paclitaxel, a variant of the classic Paclitaxel heavily used in TNBC treatment, is slowly becoming available. It combines the drug with albumin, improving the solubility of the taxol molecule and dispensing the need for the inflammation-causing vehicle previously used. This led to better tolerance of the chemotherapeutic treatment by the patient hereby dispensing the use of Dex. Clinical trials have demonstrated marginal outcome improvements(Lu *et al.*, 2021) compared to classic chemotherapies. These improvements may be in part because Dexamethasone is not administered anymore to patients receiving Nab-Paclitaxel. Our findings call for caution when using Dex to decrease side effects of chemotherapies.

Another aspect to consider is the influence of GR signalling on the microenvironment, as our work mostly focused on cell-autonomous mechanisms. This is especially important when considering that immunotherapy is explored to treat TNBCs. Glucocorticoids have anti-inflammatory and immunosuppressive properties. Notably Dex is used to treat auto-immune disorders and some haematological malignancies. Like for chemotherapy, glucocorticoids may be used to manage the side effects of immunotherapies. It is therefore of great importance to consider the effect of Dex on the immune compartment when exploring the possibility of combining standard chemotherapies with immunotherapies in the treatment of TNBCs(Adorisio *et al*., 2021;Kalfeist *et al*., 2022). Studying such dynamics is however arduous (especially considering that Dex is used to make treatments tolerable by the patients), and complicated by the heterogeneity of BCs immune compartment(Yuan *et al.*, 2021).

Finally, as our work focused on TNBC, it would be of interest to elucidate the effect of GR signalling in luminal BCs. GR and ER are both nuclear receptors and their respective response

elements have been shown to overlap (West *et al.*, 2016; Noureddine *et al.*, 2021). While some studies assessing the interactions between the two receptors and their respective pathways in BC have been done(Noureddine *et al.*, 2021), studies in the metastatic setting remain to be performed.

5.2 Single-cell analysis reveals inter- and intra-tumour heterogeneity in metastatic breast cancer

Similar in its approach, the second project leveraged single-cell RNA-Seq to gain more detailed insights about breast cancer tumours and lung metastatic cells. After performing cell cycle correction and compensating for technical biases, we highlight three major biological clusters, present in both primary site and metastases. We showed that each of these clusters top differentially expressed genes were under the control of transcription factors related to partial EMT. These results highlight the importance of this process in breast cancer metastases and shed light on its molecular basis in PDX models.

One disappointment with this work is that we were not able to identify metastases-specific populations or gene expression pattern. This may be due to several factors. First, we chose to study macrometastases. These fully developed lesions may be the stage at which metastases resemble the tumour the most. Another study using the same setting and technologies as our own focused on micrometastases and identified oxidative phosphorylation as a driver and vulnerability of TNBC micrometastases(Davis *et al.*, 2020). Secondly our data were generated in the early days of single cell RNA-seq, with a technology now obsolete. While allowing relatively high sequencing depth and high number of detected transcripts, the Fluidigm platform was limited in throughput, limiting the number of cells that could be analyzed. This drawback, coupled to the fact that we included 5 different models in the study led to a limited number of cells per model and stage (tumour or metastasis) potentially decreasing the resolution of our data.

This point leads to the current and future state of single cell studies. The Fluidigm platform became obsolete compared to droplet-based technologies such as the one commercialized by 10X Genomics. Droplet-based solutions allow to capture and analyse a much higher number of cells,

at the expense of a marginally lower depth of sequencing and number of detected transcripts. This approach nonetheless became the dominating one in the field over recent years. Studies took advantages of the high number of cells to paint a broader landscape of breast cancers, including the microenvironment and more importantly the immune system(Azizi *et al.*, 2018; Alshetaiwi *et al.*, 2020; Ren *et al.*, 2021). The information yielded by such studies may help better understand the immune compartment of BCs and enable the use of immunotherapies for at least some patients(Yao, Li and Wang, 2022).

The field is now taking advantage of increasing powerful spatial single-cell transcriptomics to analyse gene expression in cells in the context of their neighbouring interactions(Longo *et al.*, 2021; Marx, 2021). Such approach could yield great insights about critical steps of the metastatic cascade such as dormancy. Immune cells have been shown to maintain DTCs dormant and contribute to their awakening. Being able to investigate the fine interactions between the cells at such stage may also yield tremendous insights into the process and how to harness its therapeutic potential.

It however remains too soon to know if single-cell sequencing will one day be applicable in the clinic. Such type of analysis remains relatively expensive and work-intensive, and the interpretation is hardly straightforward. The inclusion of patient samples in single-cell studies is however increasingly common, allowing to compare single-cell landscapes of clinically relevant samples with cell lines and PDX models(Gambardella *et al.*, 2022). One can only contemplate the potential of analysing fine-needle biopsy via single-cell sequencing, which could potentially yield a lot of information on both cancer and its microenvironment from a very limited amount of material(Ahmadi *et al.*, 2022).

With regards to EMT, our study brought into light transcription factors whose role in partial EMT is starting to emerge. Our work, using human models, confirms previous findings in mouse models of breast cancer(Pastushenko *et al.*, 2018). It highlights the importance of this process in the plasticity of cancer cell populations and its importance in metastases. More mechanistic and functional studies are warranted to unravel the effects of each of these less-known TFs in metastasis.

In addition, we found several lncRNAs in the top differentially expressed genes across the different superclusters. LncRNAs effects have been relatively unknown until several studies found them to exert control over EMT and be of relevance in TNBC metastasis(Zhang *et al.*, 2021). As lncRNAs module signalling pathways involved in metastasis, one could envision their use as biomarkers guiding targeted therapies. Novel delivery systems are also considered to deliver siRNA targeting the lncRNA themselves(Vaidya *et al.*, 2019). While these options remain in early development, they may one day expand the arsenal against metastatic TNBCs.

6. References

Aceto, N. *et al.* (2014) "Circulating Tumor Cell Clusters Are Oligoclonal Precursors of Breast Cancer Metastasis," *Cell*, 158(5), pp. 1110–1122. doi:10.1016/j.cell.2014.07.013.

Adorisio, S. *et al.* (2021) "Glucocorticoid and PD-1 Cross-Talk: Does the Immune System Become Confused?," *Cells*, 10(9), p. 2333. doi:10.3390/cells10092333.

Ahmadi, S. *et al.* (2022) "The landscape of receptor-mediated precision cancer combination therapy via a single-cell perspective," *Nature Communications*, 13(1), p. 1613. doi:10.1038/s41467-022-29154-2.

Ahrné, E. *et al.* (2016) "Evaluation and Improvement of Quantification Accuracy in Isobaric Mass Tag-Based Protein Quantification Experiments.," *Journal of proteome research*, 15(8), pp. 2537–47. doi:10.1021/acs.jproteome.6b00066.

Albrengues, J. *et al.* (2018) "Neutrophil extracellular traps produced during inflammation awaken dormant cancer cells in mice," *Science*, 361(6409). doi:10.1126/science.aao4227.

Alizadeh, A.A. *et al.* (2015) "Toward understanding and exploiting tumor heterogeneity," *Nature Medicine*, 21(8), pp. 846–853. doi:10.1038/nm.3915.

Almendro, V., Marusyk, A. and Polyak, K. (2013) "Cellular heterogeneity and molecular evolution in cancer," *Annual Review of Pathology: Mechanisms of Disease*. Annual Reviews Inc., pp. 277–302. doi:10.1146/annurev-pathol-020712-163923.

Alshetaiwi, H. *et al.* (2020) "Defining the emergence of myeloid-derived suppressor cells in breast cancer using single-cell transcriptomics," *Science Immunology*, 5(44). doi:10.1126/sciimmunol.aay6017.

Arora, V.K. *et al.* (2013) "Glucocorticoid Receptor Confers Resistance to Antiandrogens by Bypassing Androgen Receptor Blockade," *Cell*, 155(6), pp. 1309–1322. doi:10.1016/j.cell.2013.11.012.

Azizi, E. *et al.* (2018) "Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment," *Cell*, 174(5), pp. 1293-1308.e36. doi:10.1016/j.cell.2018.05.060.

Balwierz, P.J. *et al.* (2014) "ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs," *Genome Research*, 24(5), pp. 869–884. doi:10.1101/gr.169508.113.

Barchiesi, G. *et al.* (2021) "Emerging Role of PARP Inhibitors in Metastatic Triple Negative Breast Cancer. Current Scenario and Future Perspectives," *Frontiers in Oncology*, 11. doi:10.3389/fonc.2021.769280.

Bardia, A. *et al.* (2019) “Sacituzumab Govitecan-hziy in Refractory Metastatic Triple-Negative Breast Cancer,” *New England Journal of Medicine*, 380(8), pp. 741–751. doi:10.1056/NEJMoa1814213.

Baumann, Z., Auf der Maur, P. and Bentires-Alj, M. (2022) “Feed-forward loops between metastatic cancer cells and their environment - the stage of escalation,” *In press* [Preprint].

Bianchini, G. *et al.* (2022) “Treatment landscape of triple-negative breast cancer — expanded options, evolving needs,” *Nature Reviews Clinical Oncology*, 19(2), pp. 91–113. doi:10.1038/s41571-021-00565-2.

Black, J.R.M. and McGranahan, N. (2021) “Genetic and non-genetic clonal diversity in cancer evolution,” *Nature Reviews Cancer*. Nature Research, pp. 379–392. doi:10.1038/s41568-021-00336-2.

Bonapace, L. *et al.* (2012) “If You Don’t Look, You Won’t See: Intravital Multiphoton Imaging of Primary and Metastatic Breast Cancer,” *Journal of Mammary Gland Biology and Neoplasia*, 17(2), pp. 125–129. doi:10.1007/s10911-012-9250-8.

Bonapace, L. *et al.* (2014) “Cessation of CCL2 inhibition accelerates breast cancer metastasis by promoting angiogenesis,” *Nature*, 515(7525), pp. 130–133. doi:10.1038/nature13862.

Braun, S. *et al.* (2005) “A Pooled Analysis of Bone Marrow Micrometastasis in Breast Cancer,” *New England Journal of Medicine*, 353(8), pp. 793–802. doi:10.1056/NEJMoa050434.

Bristow, S. L., Leman, A. R., & Haase, S. B. (2014). *Cell Cycle-Regulated Transcription: Effectively Using a Genomics Toolbox* (pp. 3–27). https://doi.org/10.1007/978-1-4939-0888-2_1

Britschgi, A., Duss, S., Kim, S., Couto, J.P., Brinkhaus, H., Koren, S., De Silva, D., *et al.* (2017) “The Hippo kinases LATS1 and 2 control human breast cell fate via crosstalk with ERα,” *Nature Publishing Group*, 541, pp. 4–1. doi:10.1038/nature20829.

Cardoso, F. *et al.* (2019) “Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up,” *Annals of Oncology*, 30(8), pp. 1194–1220. doi:10.1093/annonc/mdz173.

Cassandri, M. *et al.* (2020) “ZNF750 represses breast cancer invasion via epigenetic control of prometastatic genes,” *Oncogene*, 39(22), pp. 4331–4343. doi:10.1038/s41388-020-1277-5.

Cerami, E. *et al.* (2012) “The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.,” *Cancer discovery*, 2(5), pp. 401–4. doi:10.1158/2159-8290.CD-12-0095.

Chaffer, C.L. *et al.* (2016) “EMT, cell plasticity and metastasis,” *Cancer and Metastasis Reviews*, 35(4), pp. 645–654. doi:10.1007/s10555-016-9648-7.

Chen, Q., Zhu, C. and Jin, Y. (2020) "The Oncogenic and Tumor Suppressive Functions of the Long Noncoding RNA MALAT1: An Emerging Controversy," *Frontiers in Genetics*. Frontiers Media S.A. doi:10.3389/fgene.2020.00093.

Chien, H.-P. *et al.* (2016) "Expression of ROR1 has prognostic significance in triple negative breast cancer," *Virchows Archiv*, 468(5), pp. 589–595. doi:10.1007/s00428-016-1911-3.

Clarke, M.F. *et al.* (2006) "Cancer Stem Cells—Perspectives on Current Status and Future Directions: AACR Workshop on Cancer Stem Cells," *Cancer Research*, 66(19), pp. 9339–9344. doi:10.1158/0008-5472.CAN-06-3126.

Correia, A.L. *et al.* (2021) "Hepatic stellate cells suppress NK cell-sustained breast cancer dormancy," *Nature*, 594(7864), pp. 566–571. doi:10.1038/s41586-021-03614-z.

Cui, B. *et al.* (2013) "Targeting ROR1 inhibits epithelial-mesenchymal transition and metastasis.," *Cancer research*, 73(12), pp. 3649–60. doi:10.1158/0008-5472.CAN-12-3832.

Curtis, C. *et al.* (2012) "The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups.," *Nature*, 486(7403), pp. 346–52. doi:10.1038/nature10983.

Cyr, A.R. *et al.* (2015) "TFAP2C governs the luminal epithelial phenotype in mammary development and carcinogenesis," *Oncogene*, 34(4), pp. 436–444. doi:10.1038/onc.2013.569.

Davis, R.T. *et al.* (2020) "Transcriptional diversity and bioenergetic shift in human breast cancer metastasis revealed by single-cell RNA sequencing," *Nature Cell Biology*, 22(3), pp. 310–320. doi:10.1038/s41556-020-0477-0.

Denais, C.M. *et al.* (2016) "Nuclear envelope rupture and repair during cancer cell migration.," *Science (New York, N.Y.)*, 352(6283), pp. 353–8. doi:10.1126/science.aad7297.

DeRose, Y.S. *et al.* (2011) "Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes," *Nature Medicine*, 17(11), pp. 1514–1520. doi:10.1038/nm.2454.

Dillekås, H., Rogers, M.S. and Straume, O. (2019) "Are 90% of deaths from cancer caused by metastases?," *Cancer Medicine*, 8(12), pp. 5574–5576. doi:10.1002/cam4.2474.

Eikesdal, H.P. *et al.* (2021) "Olaparib monotherapy as primary treatment in unselected triple negative breast cancer," *Annals of Oncology*, 32(2), pp. 240–249. doi:10.1016/j.annonc.2020.11.009.

Fares, J. *et al.* (2020) "Molecular principles of metastasis: a hallmark of cancer revisited," *Signal Transduction and Targeted Therapy*, 5(1), p. 28. doi:10.1038/s41392-020-0134-x.

Fouad, Y. A., & Aanei, C. (2017). Revisiting the hallmarks of cancer. *American journal of cancer research*, *7*(5), 1016.

Gaidatzis, D. *et al.* (2015) “QuasR: quantification and annotation of short reads in R,” *Bioinformatics*, 31(7), pp. 1130–1132. doi:10.1093/bioinformatics/btu781.

Gambardella, G. *et al.* (2022) “A single-cell analysis of breast cancer cell lines to study tumour heterogeneity and drug response,” *Nature Communications*, 13(1), p. 1714. doi:10.1038/s41467-022-29358-6.

Gao, H., Korn, Joshua M, *et al.* (2015) “High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response,” *Nature Medicine*, 21(11), pp. 1318–1325. doi:10.1038/nm.3954.

Gao, J. *et al.* (2013) “Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal,” *Science Signaling*, 6(269), pp. pl1–pl1. doi:10.1126/scisignal.2004088.

Gay, L.J. and Felding-Habermann, B. (2011) “Contribution of platelets to tumour metastasis,” *Nature Reviews Cancer*, pp. 123–134. doi:10.1038/nrc3004.

Grimm, D. *et al.* (2020) “The role of SOX family members in solid tumours and metastasis,” *Seminars in Cancer Biology*. Academic Press, pp. 122–153. doi:10.1016/j.semcancer.2019.03.004.

Grosse-Wilde, A. *et al.* (2015) “Stemness of the hybrid Epithelial/Mesenchymal State in Breast Cancer and Its Association with Poor Survival,” *PLOS ONE*, 10(5), p. e0126522. doi:10.1371/journal.pone.0126522.

Gupta, M.K. and Qin, R.-Y. (2003) “Mechanism and its regulation of tumor-induced angiogenesis MECHANISM OF ANGIOGENESIS,” *World Journal of Gastroenterology World J Gastroenterol*, 9(6), pp. 1144–1155. Available at: http://www.wjgnet.com/1007-9327/9/1144.asp.

Györffy, B. *et al.* (2010) “An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients,” *Breast Cancer Research and Treatment*, 123(3), pp. 725–731. doi:10.1007/s10549-009-0674-9.

van de Haar, J. *et al.* (2021) “Limited evolution of the actionable metastatic cancer genome under therapeutic pressure,” *Nature Medicine*, 27(9), pp. 1553–1563. doi:10.1038/s41591-021-01448-w.

Hanahan, D. (2022) “Hallmarks of Cancer: New Dimensions,” *Cancer Discovery*, 12(1), pp. 31–46. doi:10.1158/2159-8290.CD-21-1059.

Hanahan, D. and Weinberg, R.A. (2000) “The Hallmarks of Cancer,” *Cell*, 100(1), pp. 57–70. doi:10.1016/S0092-8674(00)81683-9.

Hanahan, D. and Weinberg, Robert A. (2011) “Hallmarks of cancer: The next generation,” *Cell*, pp. 646–674. doi:10.1016/j.cell.2011.02.013.

Harbeck, N. *et al.* (2019) “Breast cancer,” *Nature Reviews Disease Primers*, 5(1). doi:10.1038/s41572-019-0111-2.

Hosseini, H. *et al.* (2016) “Early dissemination seeds metastasis in breast cancer,” *Nature*, 540(7634), pp. 552–558. doi:10.1038/nature20785.

Hu, Z. *et al.* (2020) “Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases,” *Nature Genetics*, 52(7), pp. 701–708. doi:10.1038/s41588-020-0628-z.

Jehanno, C. *et al.* (2022) “Phenotypic Plasticity during metastatic colonization,” *Trends in Cell Biology* [Preprint].

Jiang, X. *et al.* (2018) “NEAT1 contributes to breast cancer progression through modulating miR-448 and ZEB1,” *Journal of Cellular Physiology*, 233(11), pp. 8558–8566. doi:10.1002/jcp.26470.

Jin, X. and Mu, P. (2015) “Targeting Breast Cancer Metastasis,” *Breast Cancer: Basic and Clinical Research*, 9s1, p. BCBCR.S25460. doi:10.4137/BCBCR.S25460.

Jiramongkol, Y. and Lam, E.W.F. (2020) “FOXO transcription factor family in cancer and metastasis,” *Cancer and Metastasis Reviews*. Springer, pp. 681–709. doi:10.1007/s10555-020-09883-w.

Juríková, M. *et al.* (2016) “Ki67, PCNA, and MCM proteins: Markers of proliferation in the diagnosis of breast cancer,” *Acta Histochemica*. Elsevier GmbH, pp. 544–552. doi:10.1016/j.acthis.2016.05.002.

Kadmiel, M. and Cidlowski, J.A. (2013) “Glucocorticoid receptor signaling in health and disease,” *Trends in Pharmacological Sciences*, 34(9), pp. 518–530. doi:10.1016/j.tips.2013.07.003.

Kalfeist, L. *et al.* (2022) “Impact of Glucocorticoid Use in Oncology in the Immunotherapy Era,” *Cells*, 11(5), p. 770. doi:10.3390/cells11050770.

Kim, W. *et al.* (2016) “TFAP2C-mediated upregulation of TGFBR1 promotes lung tumorigenesis and epithelial-mesenchymal transition,” *Experimental & molecular medicine*, 48(11), p. e273. doi:10.1038/emm.2016.125.

Koren, S. and Bentires-Alj, M. (2015) “Breast Tumor Heterogeneity: Source of Fitness, Hurdle for Therapy,” *Molecular Cell*, 60(4), pp. 537–546. doi:10.1016/j.molcel.2015.10.031.

Lamouille, S., Xu, J. and Derynck, R. (2014) “Molecular mechanisms of epithelial-mesenchymal transition,” *Nature Reviews Molecular Cell Biology*, pp. 178–196. doi:10.1038/nrm3758.

Langley, R.R. and Fidler, I.J. (2011) “The seed and soil hypothesis revisited-The role of tumor-stroma interactions in metastasis to different organs,” *International Journal of Cancer*, 128(11), pp. 2527–2535. doi:10.1002/ijc.26031.

Lawson, D.A. *et al.* (2015) “Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells.,” *Nature*, 526(7571), pp. 131–5.

Li, C. *et al.* (2017) “A ROR1–HER3–lncRNA signalling axis modulates the Hippo–YAP pathway to regulate bone metastasis,” *Nature Cell Biology*, 19(2), pp. 106–119. doi:10.1038/ncb3464.

Lindström, L.S. *et al.* (2012) “Clinically used breast cancer markers such as estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 are unstable throughout tumor progression.,” *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 30(21), pp. 2601–8. doi:10.1200/JCO.2011.37.2482.

Liu, J. *et al.* (2018) “Synaptopodin-2 suppresses metastasis of triple-negative breast cancer via inhibition of YAP/TAZ activity,” *The Journal of Pathology*, 244(1), pp. 71–83. doi:10.1002/path.4995.

Longo, S.K. *et al.* (2021) “Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics,” *Nature Reviews Genetics*. Nature Research, pp. 627–644. doi:10.1038/s41576-021-00370-8.

Lu, H. *et al.* (2021) “A systematic review and meta-analysis of nab-paclitaxel mono-chemotherapy for metastatic breast cancer,” *BMC Cancer*, 21(1), p. 830. doi:10.1186/s12885-021-08441-z.

Lüönd, F., Tiede, S. and Christofori, G. (2021) “Breast cancer as an example of tumour heterogeneity and tumour cell plasticity during malignant progression,” *British Journal of Cancer*. Springer Nature, pp. 164–175. doi:10.1038/s41416-021-01328-7.

Marotta, L.L.C. *et al.* (2011) “The JAK2/STAT3 signaling pathway is required for growth of $CD44^{+}CD24^{-}$ stem cell-like breast cancer cells in human tumors.,” *The Journal of clinical investigation*, 121(7), pp. 2723–35. doi:10.1172/JCI44745.

Marusyk, A., Almendro, V. and Polyak, K. (2012) “Intra-tumour heterogeneity: a looking glass for cancer?,” *Nature Reviews Cancer*, 12(5), pp. 323–334. doi:10.1038/nrc3261.

Marx, V. (2021) “Method of the Year: spatially resolved transcriptomics,” *Nature Methods*, 18(1), pp. 9–14. doi:10.1038/s41592-020-01033-y.

Massagué, J. and Obenauf, A.C. (2016) “Metastatic colonization by circulating tumour cells,” *Nature*, 529(7586), pp. 298–306. doi:10.1038/nature17038.

Mayayo-Peralta, I., Zwart, W. and Prekovic, S. (2021) “Duality of glucocorticoid action in cancer: tumor-suppressor or oncogene?,” *Endocrine-Related Cancer*, 28(6), pp. R157–R171. doi:10.1530/ERC-20-0489.

Meacham, C.E. and Morrison, S.J. (2013) “Tumour heterogeneity and cancer cell plasticity,” *Nature*, 501(7467), pp. 328–37. doi:10.1038/nature12624.

Nguyen, L. v. *et al.* (2012) “Cancer stem cells: an evolving concept,” *Nature Reviews Cancer*, 12(2), pp. 133–143. doi:10.1038/nrc3184.

Noureddine, L.M. *et al.* (2021) “Glucocorticoid Receptor: A Multifaceted Actor in Breast Cancer,” *International Journal of Molecular Sciences*, 22(9), p. 4446. doi:10.3390/ijms22094446.

Obenauf, A.C. and Massagué, J. (2015) “Surviving at a distance: organ specific metastasis.,” *Trends in cancer*, 1(1), pp. 76–91. doi:10.1016/j.trecan.2015.07.009.

Ocaña, O.H. *et al.* (2012) “Metastatic Colonization Requires the Repression of the Epithelial-Mesenchymal Transition Inducer Prrx1,” *Cancer Cell*, 22(6), pp. 709–724. doi:10.1016/j.ccr.2012.10.012.

Pang, D. *et al.* (2006) “Dexamethasone decreases xenograft response to paclitaxel through inhibition of tumor cell apoptosis,” *Cancer Biology & Therapy*, 5(8), pp. 933–940. doi:10.4161/cbt.5.8.2875.

Park, H.-A., Brown, S.R. and Kim, Y. (2020) “Cellular Mechanisms of Circulating Tumor Cells During Breast Cancer Metastasis.,” *International journal of molecular sciences*, 21(14). doi:10.3390/ijms21145040.

Park, M.K. *et al.* (2021) “NEAT1 is essential for metabolic changes that promote breast cancer growth and metastasis,” *Cell Metabolism*, 33(12), pp. 2380-2397.e9. doi:10.1016/j.cmet.2021.11.011.

Parker, J.S. *et al.* (2009) “Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes,” *Journal of Clinical Oncology*, 27(8), pp. 1160–1167. doi:10.1200/JCO.2008.18.1370.

Pastushenko, I. *et al.* (2018) “Identification of the tumour transition states occurring during EMT,” *Nature*, 556(7702). doi:10.1038/s41586-018-0040-3.

Pastushenko, I. and Blanpain, C. (2019) “EMT Transition States during Tumor Progression and Metastasis,” *Trends in Cell Biology*. Elsevier Ltd, pp. 212–226. doi:10.1016/j.tcb.2018.12.001.

Pereira, B. *et al.* (2016) “The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes,” *Nature Communications*, 7, p. 11479. doi:10.1038/ncomms11479.

Perou, C.M. *et al.* (2000) “Molecular portraits of human breast tumours,” *Nature*, 406(6797), pp. 747–752. doi:10.1038/35021093.

Polyak, K. and Weinberg, R.A. (2009) “Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits,” *Nature Reviews Cancer*, 9(4), pp. 265–273. doi:10.1038/nrc2620.

Polzer, B. and Klein, C.A. (2013) “Metastasis Awakening: The challenges of targeting minimal residual cancer,” *Nature Medicine*, 19(3), pp. 274–275. doi:10.1038/nm.3121.

Van Der Pompe, G., Antoni, M.H. and Heijnen, C.J. (1996) “Elevated basal cortisol levels and attenuated ACTH and cortisol responses to a behavioral challenge in women with metastatic breast cancer,” *Psychoneuroendocrinology*, 21(4), pp. 361–374. doi:10.1016/0306-4530(96)00009-1.

Pongor, Lorinc *et al.* (2015) “A genome-wide approach to link genotype to clinical outcome by utilizing next generation sequencing and gene chip data of 6,697 breast cancer patients,” *Genome Medicine* [Preprint]. doi:10.1186/s13073-015-0228-1.

Post, H. *et al.* (2017) “Robust, Sensitive, and Automated Phosphopeptide Enrichment Optimized for Low Sample Amounts Applied to Primary Hippocampal Neurons.,” *Journal of proteome research*, 16(2), pp. 728–737. doi:10.1021/acs.jproteome.6b00753.

Ren, L. *et al.* (2021) “Single cell RNA sequencing for breast cancer: present and future,” *Cell Death Discovery*, 7(1), p. 104. doi:10.1038/s41420-021-00485-1.

Risso, D., Ngai, J., Speed, T. P., & Dudoit, S. (2014). Normalization of RNA-seq data using factor analysis of control genes or samples. *Nature Biotechnology*, *32*(9), 896–902. https://doi.org/10.1038/nbt.2931

Risson, E. *et al.* (2020) “The current paradigm and challenges ahead for the dormancy of disseminated tumor cells,” *Nature Cancer*, 1(7), pp. 672–680. doi:10.1038/s43018-020-0088-5.

Robinson, D.R. *et al.* (2017) “Integrative clinical genomics of metastatic cancer,” *Nature*, 548(7667), pp. 297–303. doi:10.1038/nature23306.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) “edgeR: a Bioconductor package for differential expression analysis of digital gene expression data,” *Bioinformatics*, 26(1), pp. 139–140. doi:10.1093/bioinformatics/btp616.

Roesch, A. *et al.* (2010) “A Temporarily Distinct Subpopulation of Slow-Cycling Melanoma Cells Is Required for Continuous Tumor Growth,” *Cell*, 141(4), pp. 583–594. doi:10.1016/j.cell.2010.04.020.

Saitoh, M. (2018) “JB special review-cellular plasticity in epithelial homeostasis and diseases: Involvement of partial EMT in cancer progression,” *Journal of Biochemistry*, 164(4), pp. 257–264. doi:10.1093/jb/mvy047.

Saxena, M. and Christofori, G. (2013) “Rebuilding cancer metastasis in the mouse,” *Molecular Oncology*, 7(2), pp. 283–296. doi:10.1016/J.MOLONC.2013.02.009.

Scheschowitsch, K., Leite, J.A. and Assreuy, J. (2017) “New Insights in Glucocorticoid Receptor Signaling—More Than Just a Ligand-Binding Receptor,” *Frontiers in Endocrinology*, 8. doi:10.3389/fendo.2017.00016.

Sephton, S.E. (2000) “Diurnal Cortisol Rhythm as a Predictor of Breast Cancer Survival,” *Journal of the National Cancer Institute*, 92(12), pp. 994–1000. doi:10.1093/jnci/92.12.994.

Sethuraman, A. *et al.* (2018) “BHLHE40 confers a pro-survival and pro-metastatic phenotype to breast cancer cells by modulating HBEGF secretion,” *Breast Cancer Research*, 20(1). doi:10.1186/s13058-018-1046-3.

Shaath, H. *et al.* (2021) “Single-cell long noncoding RNA (lncRNA) transcriptome implicates MALAT1 in triple-negative breast cancer (TNBC) resistance to neoadjuvant chemotherapy,” *Cell Death Discovery*, 7(1). doi:10.1038/s41420-020-00383-y.

Shin, V.Y. *et al.* (2019) “Long non-coding RNA NEAT1 confers oncogenic role in triple-negative breast cancer through modulating chemoresistance and cancer stemness,” *Cell Death and Disease*, 10(4). doi:10.1038/s41419-019-1513-5.

Siddhartha Mukherjee (2010) *The Emperor of All Maladies*. Scribner.

Skor, M.N. *et al.* (2013) “Glucocorticoid receptor antagonism as a novel therapy for triple-negative breast cancer.,” *Clinical cancer research : an official journal of the American Association for Cancer Research*, 19(22), pp. 6163–72. doi:10.1158/1078-0432.CCR-12-3826.

Smith, B. and Bhowmick, N. (2016) “Role of EMT in Metastasis and Therapy Resistance,” *Journal of Clinical Medicine*, 5(2), p. 17. doi:10.3390/jcm5020017.

Sosa, M.S., Bragado, P. and Aguirre-Ghiso, J.A. (2014) “Mechanisms of disseminated cancer cell dormancy: an awakening field,” *Nature Reviews Cancer*, 14(9), pp. 611–622. doi:10.1038/nrc3793.

Strauss, R. *et al.* (2009) “Epithelial Phenotype Confers Resistance of Ovarian Cancer Cells to Oncolytic Adenoviruses,” *Cancer Research*, 69(12), pp. 5115–5125. doi:10.1158/0008-5472.CAN-09-0645.

Sun, C.C. *et al.* (2019) “Comprehensive Analysis of the Expression and Prognosis for E2Fs in Human Breast Cancer,” *Molecular Therapy*, 27(6), pp. 1153–1165. doi:10.1016/j.ymthe.2019.03.019.

Sung, H. *et al.* (2021) “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA: A Cancer Journal for Clinicians*, 71(3), pp. 209–249. doi:10.3322/caac.21660.

Szczerba, B.M. *et al.* (2019) “Neutrophils escort circulating tumour cells to enable cell cycle progression,” *Nature*, 566(7745), pp. 553–557. doi:10.1038/s41586-019-0915-y.

Szklarczyk, D. *et al.* (2017) “The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible,” *Nucleic Acids Research*, 45(D1), pp. D362–D368. doi:10.1093/nar/gkw937.

Tang, K.H. *et al.* (2012) “CD133+ liver tumor-initiating cells promote tumor angiogenesis, growth, and self-renewal through neurotensin/interleukin-8/CXCL1 signaling,” *Hepatology*, 55(3), pp. 807–820. doi:10.1002/hep.24739.

Tarantino, P. *et al.* (2022) “Immunotherapy for early triple negative breast cancer: research agenda for the next decade,” *npj Breast Cancer*, 8(1). doi:10.1038/s41523-022-00386-1.

Vaidya, A.M. *et al.* (2019) “Systemic Delivery of Tumor-Targeting siRNA Nanoparticles against an Oncogenic LncRNA Facilitates Effective Triple-Negative Breast Cancer Therapy,” *Bioconjugate Chemistry*, 30(3), pp. 907–919. doi:10.1021/acs.bioconjchem.9b00028.

Vanharanta, S. and Massagué, J. (2013) “Origins of Metastatic Traits,” *Cancer Cell*, 24(4), pp. 410–421. doi:10.1016/J.CCR.2013.09.007.

Wang, Yuexi *et al.* (2011) “Reversed-phase chromatography with multiple fraction concatenation strategy for proteome profiling of human MCF10A cells.,” *Proteomics*, 11(10), pp. 2019–26. doi:10.1002/pmic.201000722.

Watanabe, K. *et al.* (2014) “Mammary morphogenesis and regeneration require the inhibition of EMT at terminal end buds by ovol2 transcriptional repressor,” *Developmental Cell*, 29(1), pp. 59–74. doi:10.1016/j.devcel.2014.03.006.

Wei, S.C. *et al.* (2015) “Matrix stiffness drives epithelial–mesenchymal transition and tumour metastasis through a TWIST1–G3BP2 mechanotransduction pathway,” *Nature Cell Biology*, 17(5), pp. 678–688. doi:10.1038/ncb3157.

Weinberg, R.A. (Robert A. (2014) *The biology of cancer*. Gerald Science.

West, D.C. *et al.* (2016) “GR and ER Coactivation Alters the Expression of Differentiation Genes and Associates with Improved ER $^{+}$ Breast Cancer Outcome,” *Molecular Cancer Research*, 14(8), pp. 707–719. doi:10.1158/1541-7786.MCR-15-0433.

West, D.C. *et al.* (2018) “Discovery of a Glucocorticoid Receptor (GR) Activity Signature Using Selective GR Antagonism in ER-Negative Breast Cancer.,” *Clinical cancer research : an official*

journal of the American Association for Cancer Research, 24(14), pp. 3433–3446. doi:10.1158/1078-0432.CCR-17-2793.

WHO Classification of Tumours Editorial Board (2019) *WHO Classification of Tumours*. 5th edn.

Wyckoff, J. *et al.* (2011) "High-Resolution Multiphoton Imaging of Tumors In Vivo," *Cold Spring Harbor Protocols*, 2011(10), p. pdb.top065904-pdb.top065904. doi:10.1101/pdb.top065904.

Xie, Z. *et al.* (2021) "Gene Set Knowledge Discovery with Enrichr," *Current Protocols*, 1(3). doi:10.1002/cpz1.90.

Xu, Yixiang *et al.* (2017) "Breast tumor cell-specific knockout of *Twist1* inhibits cancer cell plasticity, dissemination, and lung metastasis in mice," *Proceedings of the National Academy of Sciences*, 114(43), pp. 11494–11499. doi:10.1073/pnas.1618091114.

Yamashita, N. *et al.* (2018) "Epithelial Paradox: Clinical Significance of Coexpression of E-cadherin and Vimentin With Regard to Invasion and Metastasis of Breast Cancer," *Clinical Breast Cancer*, 18(5), pp. e1003–e1009. doi:10.1016/j.clbc.2018.02.002.

Yao, J., Li, S. and Wang, X. (2022) "Identification of Breast Cancer Immune Subtypes by Analyzing Bulk Tumor and Single Cell Transcriptomes," *Frontiers in Cell and Developmental Biology*, 9. doi:10.3389/fcell.2021.781848.

Ye, X. *et al.* (2015) "Distinct EMT programs control normal mammary stem cells and tumour-initiating cells," *Nature*, 525(7568), pp. 256–260. doi:10.1038/nature14897.

Yu, M. *et al.* (2013) "Circulating Breast Tumor Cells Exhibit Dynamic Changes in Epithelial and Mesenchymal Composition," *Science*, 339(6119), pp. 580–584. doi:10.1126/science.1228522.

Yuan, X. *et al.* (2021) "Single-Cell Profiling to Explore Immunological Heterogeneity of Tumor Microenvironment in Breast Cancer," *Frontiers in Immunology*, 12. doi:10.3389/fimmu.2021.643692.

Zhang, H. *et al.* (2021) "The role of EMT-related lncRNA in the process of triple-negative breast cancer metastasis," *Bioscience Reports*, 41(2). doi:10.1042/BSR20203121.

Zhang, Y. and Weinberg, R.A. (2018) "Epithelial-to-mesenchymal transition in cancer: complexity and opportunities.," *Frontiers of medicine*, 12(4), pp. 361–373. doi:10.1007/s11684-018-0656-6.

7. Acknowledgements

I would like to thank Mohamed Bentires-Alj for giving me the opportunity to work in his laboratory, for the projects he entrusted me and the insights he delivered to me during his supervision. The Bentires-Alj lab was an exciting and interesting environment to work in over the past 5 years. I also thank my PhD Advisory Committee, including Primo Schär, Berend Snijder and Thomas Radimerski for their support, time, and inputs.

A PhD has plenty of hard times, but I was lucky that none of these were due to colleagues (as dirty hoods and scales do not really count as hard times). I thank all of them, past and present, for their constant feedback and support, even when presented with an excessive number of confusing western blots. Thank you Markus Ackerknecht, Romain Amante, Priska Auf der Maur, Zora Bauman, Marie-May Coissieux, Ana Correia, Duvini De Silva, Maren Diependruck, Sherlyn Foo Sok Lin, Aurélie Godel, Carolina Hager, Juliane Heilig, Charly Jehanno, Evrim Ceren Kabak, Michal Kloc, Shany Koren, Nicolas Kramer, Madhuri Manivannan, Valentina Mele, Jorge Gomez Miragaya, Alexandra Musch, Ryoko Okamoto, Marta Palafox, Catherine Piggin, Tiberius Preca, Veronica Richina, Marion Salvador, Loïc Sauteur, Atul Sethi, Joana Silva, Katrinn Volkmann and Federica Zilli. Last but not least I would like to particularly thank Milan Obradović for involving me in his work, teaching me most of what I learned in the lab and investing of his time to mentor me over the years. I thank you all for your pertinent advice and critical input at every step of my academic journey.

I am also grateful to the facilities of the Department of Biomedicine, Biozentrum (especially Alexander Schmidt of the proteomics core facility) and the University Hospital of Basel for their support and contributions to my work.

Finally, I would like to thank my family for supporting me over the years, in my choices and studies. Among them Zora Baumann, colleague turned dearest one, who especially supported in the most difficult moments of my PhD.

8. Appendix

8.1 Abbreviations

ACTH	Adrenocorticotropic hormone
AP1	Activator Protein 1
ATF3	Activating Transcription Factor 3
BC	Breast Cancer
bHLH	basic helix–loop–helix
BHLHE40	Basic Helix-Loop-Helix Family Member E40
CTC	Circulating Tumor Cell
DEG	Differencially expressed genes
DEX	Dexamethasone
EMT	Epithelial-Mesenchymal Transition
ER	Estrogen receptor
FACS	Fluorescence-Activated Cell Sorting
FN1	Fibronectin 1
FOS	Fos Proto-Oncogene
GFP	Green fluorescent protein
GLM	Generalized linear model
GR	Glucocorticoid Receptor
HBEGF	Heparin Binding EGF Like Growth Factor
HER2	Human epidermal growth factor receptor 2
HEY2	Hes Related Family BHLH Transcription Factor With YRPW Motif 2
i.v	intra-veinously
JUN	Jun Proto-Oncogene
KM	Kaplan-Meier
lncRNA	Long non coding RNA
MALAT1	Metastasis-Related Lung Adenocarcinoma Transcript 1
MCM3	Minichromosome Maintenance Complex Component 3
MET	Mesenchymal-Epithelial Transition
MYC	My Proto-Oncogene
NEAT1	Nuclear Enriched Abundant Transcript 1
NSG	NOD-scid IL2rγnull
OVOL2	Ovo Like Zinc Finger 2
PCA	Principal Component Analysis
PCNA	Proliferating Cell Nuclear Antigen

PDG	Percentage of detected genes
PDX	Patient-Derived Xenograft
PR	Progesterone receptor
qPCR	quantitative Polymerase Chain Reaction
RFS	Relapse free survival
RNA	Ribonucleic acid
ROR1	Receptor Tyrosine Kinase Like Orphan Receptor 1
s.d	standard deviation
sc	single cell
Seq	sequencing
shRNA	short hairpin RNA
SNAI1	Snail Family Transcriptional Repressor 1
SOX4	SRY-Box Transcription Factor 4
SOX9	SRY-Box Transcription Factor 9
TF	Transcription factor
TGFBR1	Transforming Growth Factor Beta Receptor 1
TGFβ	Transforming Growth Factor β
TNBC	Triple-Negative Breast Cancer
TNFα	Tumour Necrosis Factor α
TP53	Tumor Protein P53
TP63	Tumor Protein P63
t-SNE	t-distributed neighbour embedding method
U2AF1	U2 Small Nuclear RNA Auxiliary Factor 1
ZEB1	Zinc Finger E-Box Binding Homeobox 1
ZNF750	Zinc Finger Protein 750

8.2 List of figures

8.2.1 Introduction

Note: Figures 1.1 to 1.7 were created with BioRender.com.

8.2.2 Glucocorticoids promote breast cancer metastasis

Extended Data Figure 3.5 – Dexamethasone reduces overall survival and GR downregulation increases cancer cell dissemination from the primary site

Extended Data Figure 3.6 – Differential expression of protein kinases in tumours and matched metastases

Extended Data Figure 3.7 – Differential protein abundance upon GR activation

Extended Data Figure 3.8 – GR activation increases the expression of kinases that are predictive of survival in breast cancer

Extended Data Figure 3.9 – ROR1 expression in breast cancer and metastases

Extended Data Figure 3.10 – Dexamethasone increases metastases and precipitates death, via ROR1

8.2.3 Single-cell analysis reveals inter- and intra-tumour heterogeneity in metastatic breast cancer

Figure 4.1 – Interpatient heterogeneity is dominant over intrapatient heterogeneity.

Figure 4.2 – Cell cycle and percentage of detected genes delineate cell clustering.

Figure 4.3 – Removal of cell cycle variations and percentage of detected genes reveals 3 major biological clusters.

Figure 4.4 – Figure 4. Comparison of major biological superclusters reveals partial EMT states regulators.

Extended Data Figure 4.1

Extended Data Figure 4.2

Extended Data Figure 4.4

Extended Data Figure 4.4

About me

I am a technology-driven, hard working person looking to overcome challenges and provide solutions to my collaborators. Curious and fast-learner, I thrive in a dynamic and interactive work environment.

Languages

- **French** | Native
- **English** | Fluent
- **German** | Intermediate
  Goethe B1
  Actively learning

Additional Training

- Advanced R Course: Interpretation of genomic data with R/ Bioconductor
  University of Basel / SIB
- LTK1 - Module 1
  Animal Experimentation
  FELASA, UZH
- Project Management
  Business Management
  Negotiation
  University of Basel
  Transferable Skills Program

Baptiste Hamelin

Ph.D. - Group leader in molecular pathology

Basel Area | E-Mail: hamelin.baptiste@gmail.com | Phone: +33 6 59 32 09 35

Research and Professional Experience

Present — **Group Head - Molecular Pathology**
Kantonsspital Baselland | Liestal, Switzerland | With Pr. Kirsten Mertz
Routine and research molecular biology for diagnostic in oncology and infectiology

2017 — **Ph.D. Candidate - Molecular Biology | Breast Cancer**
Department of Biomedicine | Basel, Switzerland | Bentires-Alj Group
Investigation of drivers of metastasis in triple negative breast cancer

2016 — **Engineer / MSc. Internship | Research Associate**
New England Biolabs (NEB) | Boston Area, USA | Shengxi Guan Group
Development of a CRISPR-Cas9-based ribosomal RNA depletion kit integrated to NEB RNA-Seq library preparation kits

2016 and prior / See further experiences on my LinkedIn profile

Education

Present — **Ph.D. in Molecular Biology**
University of Basel | Department of Biomedicine | Bentires-Alj Group
Basel, Switzerland

2016 — **Engineering Degree in Biotechnology**
ESBS | Strasbourg, France | Basel, Switzerland | Freiburg, Germany
+ Master Degree in High-throughput Biotechnology
University of Strasbourg | Strasbourg, France

2013 — **Bachelor in Medical Laboratory Techniques**
IFTLM | Lyon, France

Skills

- **Molecular Biology**
  PCR | Cloning | CRISPR-Cas9 & related
- **Next Generation Sequencing (NGS)**
  DNA & RNA-Seq | Bulk and Single-cell Library preparation | Illumina sequencing
- **High-Dimensional Data Interpretation**
  NGS data | (Phospho)Proteomics
- **Use of repositories of clinical data**
  Genomic and transcriptomic profiles
- **Software & Programming**
  Proficient in R | Notions of Python | Graphpad Prism
  Adobe | Affinity | Microsoft suites
- **Tissue Culture**
  2D and 3D | Cell lines and patient samples
- **In vivo Experimentation**
- **Flow Cytometry and High-Throughput Microscopy**
- **Good communication skills and team spirit**
- **Project Management | Equipment Supervision**

Publications

- **Single cell analysis reveals inter and intratumour heterogeneity in metastatic breast cancer.**
  PhD project publication (main author) | https://doi.org/10.1101/2022.10.16.512415
- **Glucocorticoids promote breast cancer metastasis.**
  Obradović, M.M.S., Hamelin B., Manevski N., et al. Nature 567, 540-544 (2019)
- Additional internal and external collaborations articles to be published