Generation and characterisation of two D2A1 mammary cancer sublines to model spontaneous and experimental metastasis in a syngeneic BALB/c host

[1,5]Ute Jungwirth, [1]Antoinette van Weverwijk, [1,2]Miriam J. Melake, [3]Ann F. Chambers, [1]Qiong Gao, [1,4]Marc Fivaz and [1,5]Clare M. Isacke

[1]The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, UK

[2]German Cancer Research Center, DKFZ, Heidelberg, Germany

[3] University of Western Ontario, London, Ontario, Canada

[4]Current address. Faculty of Engineering & Science, University of Greenwich at Medway, Kent, ME4 4TB

[5]Correspoondence to Ute Jungwirth (ute.jungwirth@icr.ac.uk) or Clare M. Isacke (clare.isacke@icr.ac.uk)

Running Title: Metastatic mammary cancer model

Keywords: mammary cancer, metastatic sublines, syngeneic, spontaneous metastasis, D2A1, BALB/c

Summary statement: We describe two D2A1 mouse mammary cancer sublines with enhanced spontaneous metastasis in a syngeneic host, and highlight the limitations of *in vitro* assays to predict *in vivo* metastatic behaviour.

Abstract

Studying the complex mechanisms underlying breast cancer metastasis and therapy response necessitates relevant *in vivo* models, particularly syngeneic models with an intact immune system. Two syngeneic spontaneously metastatic sublines, D2A1-m1 and D2A1-m2, were generated from the poorly metastasising BALB/c-derived D2A1 cell line by serial *in vivo* passaging. *In vivo* and *in vitro* analyses revealed distinct and shared characteristics of the metastatic D2A1-m1 and D2A1-m2 sublines. In particular, D2A1-m1 cells are more aggressive in experimental metastasis assays, while D2A1-m2 cells are more efficient at disseminating from the primary tumour in spontaneous metastasis assays. Surprisingly, classical metastasis-associated *in vitro* phenotypes such as enhanced proliferation, migration and invasion are reduced in the sublines compared to the parental cell line. Further, evasion of immune control cannot fully explain their enhanced metastatic properties. By contrast, both sublines show increased resistance to apoptosis when cultured in non-adherent conditions and, for the D2A1-m2 subline, increased 3D tumour spheroid growth. Moreover, the enhanced spontaneous metastatic phenotype of the D2A1-m2 subline is associated with an increased ability to recruit an activated tumour stroma. The metastatic D2A1-m1 and D2A1-m2 cell lines provide additional syngeneic models for investigating the different steps of the metastatic cascade and thereby represent valuable tools for breast cancer researchers. Finally, this study highlights that morphology and cell behaviour in 2D cell-based assays cannot be used as a reliable predictor of metastatic behaviour *in vivo*.

Introduction

To study the complex mechanisms of breast cancer metastasis and therapy response in advanced disease, it is important to have relevant *in vivo* models. Ideally, the model recapitulates the full metastatic cascade including growth of a primary tumour, dissemination of tumour cells into the circulation, colonisation of secondary sites and the development of macrometastatic disease. In addition, to assess the impact of the immune system, an immunocompetent syngeneic model is required. A recent study has molecularly characterised 12 mouse mammary cancer cell lines and performed phenotypic analysis of the primary tumours grown in syngeneic hosts (Yang et al., 2017). To date, the best characterised spontaneous breast cancer metastasis model is the BALB/c-derived 4T1 cell line (Aslakson and Miller, 1992) and the 4T1 sublines selected for increased metastasis to the bone and lung (Lelekakis et

al., 1999; Tester et al., 2000) or brain (Lockman et al., 2010). More recently Johnstone and colleagues have derived and characterised a spontaneously metastasising variant of the C57BL/6-derived murine medullary mammary adenocarcinoma cell line E0771 (Johnstone et al., 2015), allowing for metastasis studies to be performed in an alternative mouse strain. However, there is still an increasing demand for independent models both for study validation and to address the inter- and intra-tumour heterogeneity of human disease.

In this study we describe the generation of two breast cancer cell sublines, D2A1-m1 and D2A1-m2, derived from parental D2A1 cells. The parental D2A1 cell line was derived from a mouse mammary tumour in a BALB/c mouse implanted with the transplantable D2 hyperplastic alveolar nodule cell line (Mahoney et al., 1985; Miller et al., 1989; Morris et al., 1993). In a recent comprehensive analysis of 12 mouse mammary cancer cell lines (Yang et al., 2017), D2A1 cells are classified as oestrogen receptor (ER) and ErbB2/HER2-negative, *Pik3ca* and *Trp53* wildtype, having a 'claudin-low' transcriptional profile and assignment to the luminal B subtype. *In vivo*, D2A1 tumours have a spindle-cell histology with ~18% Ki67-positive cells and ~20% caspase 3-positive nuclei. It was originally reported that 4 weeks after orthotopic inoculation into immunocompromised nude BALB/c mice, 2 out of the 5 mice showed visible macrometastatic disease in the lungs with the remaining mice showing lung micrometastatic disease upon histological examination (Morris et al., 1993). Subsequently it was reported that parental D2A1 cells can colonise the lungs of immunocompetent BALB/c mice if injected via the tail vein in an experimental metastasis assays (Shibue and Weinberg, 2009) and that rates of spontaneous metastasis in BALB/c can be increased by pre-irradiation of the mammary gland (Bouchard et al., 2013), albeit still with a relatively low metastatic burden.

The D2A1-m1 and D2A1-m2 sublines were derived by serial inoculation of tumour cells in BALB/c mice followed by recovery from the lung tissue *ex vivo*. Characterisation of these two metastatic sublines demonstrates them to have both overlapping and distinct phenotypes in *in vivo* assays, *in vitro* assays and by gene expression profiling. In particular, the D2A1-m1 subline displays an enhanced ability to colonise the lungs and other tissues in experimental metastasis assays whereas the D2A1-m2 subline shows a robust and reproducible ability to colonise the lungs in a spontaneously metastasis assay (inoculation into the mammary fat pad), associated with an increased ability to recruit an activated tumour stroma. Consequently, these two D2A1 sublines provide useful and complementary models to interrogate the different stages of the metastatic cascade.

Results

Generation of spontaneously metastatic D2A1 sublines

The scheme for the generation of the D2A1 sublines is shown in Fig. 1A. The two sublines were derived independently. In brief, for each subline, parental D2A1 cells were inoculated orthotopically into the 4th mammary fat pad of an immunocompetent BALB/c mouse. When the primary tumour reached 10 - 12 mm in diameter, the lungs were harvested individually from each mouse at necropsy, dissociated, and placed into culture. Tumour cells that grew out were expanded and inoculated into the tail vein of a recipient mouse and 11 - 13 days later, lungs removed at necropsy. In total, three rounds of intravenous inoculation were performed resulting in the selection of the independent metastatic sublines, D2A1-m1 and D2A1-m2.

In an initial experiment to assess the metastatic competency of the selected sublines, parental D2A1, D2A1-m1 and D2A1-m2 cells were inoculated orthotopically into the mammary fat pad of recipient BALB/c mice. Animals were culled individually between days 31 and 39 when the primary tumours reached 12 - 14 mm in diameter (Fig. 1B). Consistent with the published literature (Bouchard et al., 2013; Morris et al., 1993; Yang et al., 2017) parental D2A1 cells readily formed primary tumours but, with the exception of a single large metastatic nodule in one mouse, gave rise to only small metastatic nodules in the lungs and with the majority of mice having undetectable metastatic disease. All mice inoculated with D2A1-m1 and D2A1-m2 cells had metastatic disease as monitored both by tumour area and number of metastatic nodules, with the D2A1-m2 subline giving rise to the highest metastatic burden in this spontaneous metastasis assay. Though there is a difference in the number of metastatic nodules in mice inoculated with the D2A1-m1 and D2A1-m2 sublines and an increase in the % tumour area in the lungs, the average size of the lung nodules is similar. Of note, although the D2A1-m1 and D2A1-m2 primary tumours grew slightly faster than the parental D2A1 tumours, in subsequent experiments no differences in primary tumour growth was observed (see below).

To better mimic the clinical setting in breast cancer, spontaneous metastasis assays in BALB/c mice were then performed in which the primary tumour was surgically resected at a small size (~ 6 mm in diameter; day 11 - 13) and all mice were culled on day 43 (Fig. 1C). Despite the increased length of time of the experiment, when lungs were examined at autopsy, there was only limited metastatic disease in the parental D2A1 inoculated mice whereas the D2A1-m2 subline gave rise to extensive tumour burden in the lungs. The D2A1-m1 subline gave mixed results with 2 out of 5 mice showing no metastatic disease and the remaining 3 mice

having readily detectable macrometastatic nodules. The enhanced metastatic burden observed in mice where the primary tumour is resected (Fig. 1C) likely reflects two factors. First, in experiments where the primary tumour was not removed, mice were culled when the primary tumour reached its maximum allowable size (31 - 39 days (Fig. 1B). By contrast, following tumour resection the study was conducted over a longer time period (43 days) allowing increased growth of metastatic deposits (Francia et al., 2011). Second, in experimental mouse models, there have been numerous reports that removal of a primary tumour can enhance growth of metastatic lesions (O'Reilly et al., 1994). In addition, these observations suggest that dissemination of cells from the primary tumour is an early event.

To assess whether the immune system plays a role in limiting D2A1 tumour growth and metastasis, parental D2A1 cells and the metastatic D2A1-m2 subline were inoculated orthotopically into immunocompetent BALB/c (Supplementary Figure S1) or immunocompromised NOD scid gamma (NSG) (Fig. 2) mice. In both mouse strains, there was no significant difference in primary tumour growth but, again, the D2A1-m2 cells gave rise to a significantly increased metastatic burden in the lung. Moreover, although the primary tumours did not grow at a noticeably faster rate in the NSG compared to the BALB/c mice (Fig. 2A, Supplementary Figure S1A), it is important to note that the NSG mice experiment had to be terminated at an earlier time (and when the primary tumours were of a much smaller size) due to the extensive burden of metastatic disease in the D2A1-m2 inoculated mice (Fig. 2B). In addition, although the D2A1 parental cells showed a much reduced metastatic burden compared to the D2A1-m2 subline, the amount of D2A1 metastatic burden in the lung was greater in the immunocompromised NSG mice compared to the BALB/c mice. Although, differences in tumour growth can vary between different mouse strains (Hunter, 2006) these data suggest that both the D2A1 parental cells and their metastatic derivatives are under immune control, particularly in the metastatic setting, and that evasion of immune control cannot fully explain the enhanced metastatic properties of the metastatic subline.

Experimental metastasis assays

Next we assessed the ability of the two metastatic sublines to colonise secondary sites in different experimental metastasis assays. First, untagged (Fig. 3A) or luciferase tagged (Supplementary Figure S2) tumour cells were inoculated intravenously into recipient BALB/c mice. Consistent with the spontaneous metastasis assays, injection of both sublines resulted in a greater lung tumour burden than the parental cells. Interestingly, although the D2A1-m2 subline gave rise

to a higher burden of spontaneous lung metastases compared to the D2A1-m1 subline (Fig. 1B), in this experimental metastasis assay it was the D2A1-m1 subline that gave rise to the greatest lung tumour burden (Fig. 3A).

Intravenous inoculation results in the majority of tumour cells lodging in the lungs. To investigate other sites of metastasis, we next performed intracardiac inoculation of tumour cells (Fig. 3B), which favours dissemination via the arterial system to the bones, brain and other organs (Bos et al., 2009; Khanna and Hunter, 2005). *Ex vivo* IVIS imaging revealed that, again, the D2A1-m1 subline gave rise to the greatest tumour burden in the bones. Neither the D2A1 parental cells nor the metastatic sublines showed evidence of brain colonisation.

Finally, we performed intrasplenic inoculations to assess colonisation of the liver (Khanna and Hunter, 2005) (Fig. 3C). 50% of mice inoculated with parental D2A1 cells had undetectable tumour burden in the liver but both the D2A1-m1 and D2A1-m2 sublines gave rise to extensive disease with only one mouse in each group remaining tumour free.

Together these *in vivo* experiments, indicate that the two D2A1 sublines have different properties with the D2A1-m1 subline having a superior ability to grow in the lungs and the bones but, compared to the D2A1-m2 subline, a reduced ability to disseminate from the primary tumour. Consequently, we next addressed whether these two sublines displayed altered properties using a panel of *in vitro* proliferation, migration and invasion assays.

D2A1 metastatic sublines in adherent culture.

When growing as adherent cultures on tissue culture plastic the D2A1 metastatic sublines have distinct morphologies, with D2A1-m1 cells having a more elongated shape and D2A1-m2 cells being more rounded, compared to the parental cells (Fig. 4A; Supplementary Fig. S3). These different morphologies did not impact on the plating efficiency as exemplified by all three cell lines giving rise to the same number of colonies in a colony formation assay (Fig. 4B). By contrast, in this assay, it was notable that the two metastatic sublines gave rise to smaller colonies, indicating a lower proliferation rate *in vitro*. Consistent with this, both sublines displayed a lower proliferation rate as continuously monitored in the IncuCyte live cell analysis (Fig. 4C) and reduced cell number over time as monitored in the CellTiter-Glo assay (Fig. 4D). Together with the observation that there was no consistent difference in primary tumour growth between the parental line and the sublines, these data indicate that the increased metastatic ability of the sublines is not due to the acquisition of a more proliferative phenotype.

Next we tested whether the sublines show a difference in their migratory abilities, a characteristic linked to an increased metastatic potential. To our surprise, both sublines displayed a reduced migratory and chemotactic capacity *in vitro* as monitored in the IncuCyte scratch wound assay (Fig. 4E) and a Transwell chemotactic assay (Fig. 4F). These data highlight how morphology and cell behaviour in 2D adherent assays cannot be used as a predictor of metastatic behaviour *in vivo*. Consequently we moved to assays under non-adherent conditions.

D2A1 metastatic sublines in non-adherent culture.

As an example of how non-adherent assays can reveal properties of cells that are not evident in adherent assays, D2A1 parental cells and the D2A1 metastatic sublines D2A1 were seeded into tissue culture or low adherence plates and monitored for apoptosis after 24 hours (Fig. 5A). In adherent culture, essentially all cells were viable as monitored by the lack of annexin V (AV) staining and propidium iodide (PI) uptake. After 24 hours in non-adherent culture more than 50% of the D2A1 parental cells were classed as early apoptotic (AV+/PI-), late apoptotic (AV+/PI+) or necrotic (AV-/PI+), while both the D2A1-m1 and D2A1-m2 sublines showed a significantly lower level of anoikis under these conditions. To explore this further, we next plated cells into low-adherence U bottom plates (Fig. 5B). The parental cells readily assembled into dense tumour spheroids with well delineated margins, and expanded in size over time. Of the two metastatic sublines, the D2A1-m1 subline again assembled into tumour spheroids but, consistently, they were less regular in shape. However, there was no difference in D2A1-m1 spheroid growth compared to the parental cells. By contrast, the D2A1-m2 subline formed less well aggregated spheroids with loosely attached cells associated with the main core. Moreover, as monitored by CellTiter-Glo, these disorganised spheroids showed a significant increase in cell number over time compared to the parental D2A1 and D2A1-m1 spheroids.

Given the less regular spheroid shape formed by the D2A1 sublines, we anticipated that the more loosely attached tumour cells might be more invasive. To address this, tumour spheroids were transferred after 3 days into a collagen matrix. At the indicated time points, collagen plugs were fixed, stained with DAPI and cell invasion into the 3D collagen matrix quantified. Surprisingly, the parental D2A1 cells were significantly more invasive than either of the metastatic sublines. Despite the limitation of using *in vitro* assays to model *in vivo* metastasis, these data point to the metastatic sublines having an enhanced metastatic capacity due to their ability to resist stress-induced anoikis.

D2A1-m2 cells promote stromal cell activation

To investigate further how these differing behaviours in non-adherent culture might impact on tumour growth *in vivo*, primary tumours in BALB/c mice (Fig. 1B, Supplementary Figure S1A) and NSG mice (Fig. 1C) were sectioned and stained for αSMA, a marker of activated fibroblasts and endomucin, a marker of endothelial cells. The elevated levels of αSMA-positive fibroblasts and blood vessels in the D2A1-m2 primary tumours in both the BALB/c immunocompetent (Fig. 6A, Supplementary Figure S1D,E) and NSG immunocompromised (Fig. 6B) mice suggest that this subline has an enhanced ability to promote stromal cell recruitment activation and this may, in part, provide a mechanistic explanation for the enhanced spontaneous dissemination of tumour cells to secondary sites.

Although stromal cell activation and recruitment observed in *in vivo* tumours cannot be fully recapitulated *in vitro*, we examined whether the metastatic sublines had an increased ability to promote directional migration of fibroblasts using the 'ibidi' µ-slide chemotaxis system. 3T3 mouse fibroblasts were plated into the central viewing chamber that is connected to two larger reservoirs on either side (Supplementary Figure S4A). By plating different cell populations into the two reservoirs, conditioned medium gradients are set up and cell migration can be monitored by time-lapse microscopy. As expected, if D2A1 cells are plated into both reservoirs, there is no bias in the directional migration of the fibroblasts (Fig. 6C; left panel). However, if D2A1 cells are plated into one reservoir and either D2A1-m1 or D2A1-m2 cells are plated into the other, the fibroblasts show preferential displacement towards the metastatic sublines over the parental cells (see Supplementary Figure S4B for individual fibroblast migration tracks). However, albeit with relatively small effects, only the D2A1-m2 subline induced a significant difference in fibroblast migration speed (velocity) and Euclidean distance migrated (Fig. 6C; middle and right panels) and a significant fibroblast chemotaxis (Rayleigh test; Supplementary Figure S4C). Finally, to assess the capacity of the metastatic sublines to promote fibroblast activation, fibroblasts were embedded into a Matrigel/collagen gel and treated with conditioned media from D2A1, D2A1-m1 or D2A1-m2 cells or with TGFβ as a positive control (Fig. 6D). Only D2A1-m2 conditioned medium and TGFβ treatment resulted in a significant induction of fibroblast contractility. Consistent with the *in vivo* observations (Fig. 6A) these data indicate that the D2A1-m2 subline has an enhanced ability to both recruit and activate stromal fibroblasts.

Gene expression profiling and analysis of human datasets

Finally, we subjected the parental D2A1 cells and metastatic sublines to gene expression profiling. Principle component analysis revealed that each cell line was qualitatively unique, with the metastatic sublines being distinct but more closely related to each other than to the parental cells (Fig. 7A). Overall, there were 890 genes differentially expressed between the D2A1-m1 cells and parental D2A1 cells, from which 323 genes were with a fold change of ≥1.5-fold. In the case of the D2A1-m2 cells, 1339 genes were differentially expressed, from which 318 had a fold change ≥1.5 (Fig 7B). A comparison of the differentially expressed genes showed that approximately 50% were unique to each subline and 50% overlapped as shown in the Venn diagram (Fig 7B bottom panel). Unsupervised two-dimensional hierarchical clustering of the D2A1, D2A1-m1 and D2A1-m2 cells (based on genes with a fold change ≥1.5, using their log2 gene expression) confirmed the close relationship of the two sublines but also distinct gene expression patterns (Fig 7C). Top differentially expressed genes for D2A1-m1 *vs.* D2A1 and D2A1-m2 *vs.* D2A1 are shown in Supplementary Table S1 and Supplementary Table S2, respectively. Top genes that were differentially changed in both sublines (Fig. 7B) and their average fold change are shown in Supplementary Table S3. Similarly, Ingenuity Pathway Analysis demonstrated distinct top canonical pathways, while the upstream regulators and molecular and cellular functions were highly overlapping (Fig. 7D, Supplementary Tables S4 and S5). Even though in both, D2A1-m1 and D2A1-m2 cells, TGFβ1 is identified as upstream regulator, only in D2A1-m2 subline is the predictive z-score >2 fold, indicating a stronger activation of this pathway in D2A1-m2 cells than in the parental and D2A1-m1 cells, and potentially accounting for the observed ability of the D2A1-m2 subline to promote stromal activation.

Discussion

Studying the complex mechanisms driving breast cancer metastasis and therapeutic response of advanced disease has been hampered by the lack of robust *in vivo* models, and in particular spontaneously metastatic syngeneic models where the interaction with an intact immune system can be studied. To date, there has been heavy reliance on the BALB/c-derived 4T1 cell line (Aslakson and Miller, 1992; Fantozzi and Christofori, 2006) as 4T1 cells, and selected 4T1 sublines, readily form primary orthotopic tumours that can spontaneously metastasise to the lungs and other secondary sites (Eckhardt et al., 2005; Fantozzi and Christofori, 2006; Yang et al., 2004). More recently Johnstone and colleagues have described the generation of a spontaneously metastatic E0771.LMB subline from the C57BL/6-derived E0771

mouse medullary mammary adenocarcinoma E0771 cell line (Johnstone et al., 2015) that provides a valuable model in a different mouse strain (Chen et al., 2017). The molecular and phenotypic characterisation of 12 mouse mammary carcinoma cell lines (Yang et al., 2017), provides valuable information for the research community but, as the authors note, the efficiency of metastatic spread with these lines lacks reproducibility, for example only 45% of the mice inoculated with parental D2A1 cells had detectable metastatic lesions in the lungs. Alternatives to syngeneic mouse cell lines are tumour prone genetically modified mice (Fantozzi and Christofori, 2006) such as those expressing a polyoma middle T (PyMT) (Guy et al., 1992a) or ErbB2/neu (Guy et al., 1992b) transgene in the mammary gland. Genetically modified models have provided powerful insights into tumour initiation and progression, but they have limitations when studying the development of distant metastases due to the variable latency and often late presentation of secondary disease, and the inability to rapidly perform additional genetic manipulation. In addition to the mouse syngeneic models, there are many human breast cancer cell lines that can be grown as xenografts in immunocompromised mice, and more recently, an increasing number of patient-derived xenografts where primary tumour material is transplanted directly into immunocompromised mice and then serially passaged. Such models have been extremely valuable in studying the heterogeneity of human disease and for monitoring response to tumour targeting agents (Bruna et al., 2016; Byrne et al., 2017; Holen et al., 2017; Neve et al., 2006). However, few of these models reproducibly spontaneous metastasise. Further, the lack of a species matched stroma and the lack of an intact immune system limit the interrogation of breast cancer biology and/or the assessment of stromal targeting agents. Consequently, there remains an urgent need to develop a wider variety of syngeneic models where the different stages of the metastatic cascade can be investigated.

In this study we describe the generation and characterisation of two sublines of the poorly metastatic mouse mammary tumour cell line, D2A1. These sublines, D2A1-m1 and D2A1-m2, reproducibly give rise to spontaneous metastases from the primary tumour to the lungs, and in experimental metastasis assays readily colonise the lung, liver, and in case of the D2A1-m1 subline, bone. However, there are notable differences in the behaviour of these two sublines *in vivo*. In particular, the D2A1-m1 subline, compared to the D2A1-m2 subline, is less efficient at disseminating from the primary tumour in a spontaneous metastasis assay but more efficient at colonising the lungs, bone and liver when inoculated via the tail vein, the left ventricle of the heart or the spleen, respectively. This would suggest that the D2A1-m1 subline has acquired properties that better enable it to extravasate from the circulation and/or

efficiently proliferate at these secondary sites and that, conversely, the D2A1-m2 has acquired properties that better enable it to disseminate from the primary tumours.

To address the potential mechanism underlying these distinct phenotypes, first we extensively characterised these cells lines in *in vitro* assays. In adherent cultures, the two sublines have distinct morphologies with the D2A1-m1 cells having a more elongated and the D2A1-m2 cells have a more rounded shape. However, these morphological differences did not relate to their behaviour with both sublines, surprisingly, showing a lower rate of proliferation and migration compared to the parental D2A1 cells. More revealing was their behaviour in non-adherent conditions. When normal epithelial cells are detached from a matrix, a programme of caspase-mediated apoptosis, known as anoikis, is activated (Frisch and Screaton, 2001; Paoli et al., 2013). A hallmark of cancer is 'resisting cell death' (Hanahan and Weinberg, 2011) and D2A1 parental cells show evidence of such resistance as evidenced by their ability to form and proliferate as free floating tumour spheroids when plated into U bottom low attachment plates. However, when plated under conditions that do not promote cell:cell contact between the detached cells (low adherent flat bottom plates) within 24 hours over 50% of the D2A1 cells show markers of apoptotic cell death, compared to <10% of the cells plated onto adherent flat bottomed plates. In the same assays, both metastatic sublines show increased resistance to anoikis, and in spheroid culture the D2A1-m2 cells show enhanced proliferative ability. Consequently, we conclude that the increased metastatic potential of the two sublines is due, at least in part, to their ability to survive in a more hostile environment. However, these *in vitro* studies were relatively uninformative as to why the D2A1-m1 subline shows a more aggressive behaviour in experimental metastasis assays. Some clues are provided by gene expression profiling where examination of genes differentially expressed between the D2A1-m1 subline and the parental cells revealed that upregulation of canonical pathways involved in glutathione regulation and the actin cytoskeleton, features that may provide an advantage for the cells in extravasating from the circulation and/or surviving in the foreign metastatic environments. Future studies will be required to fully address this hypothesis and it will certainly be of interest to compare the transcriptional profile of the D2A1 sublines freshly isolated from tumours, as this may well be more informative that profiling the sublines grown in culture.

More informative was the analysis of the D2A1-m2 subline. In addition to showing an increased resistance to anoikis, the D2A1-m2 primary tumours, in both immunocompetent BALB/c mice and immunocompromised NSG mice, showed striking infiltration of αSMA-positive stromal fibroblasts and pericytes. At least in part,

this phenotype could be recapitulated *in vitro* where we show that when presented with an option for directional migration, cultured fibroblasts preferentially migrate, at an increased velocity, towards the D2A1-m2 subline and D2A1-m2 conditioned medium, but not conditioned medium from D2A1 or D2A1-m1 cells, promoted fibroblast activation as monitored in a gel contraction assay. A number of studies have demonstrated that an activated stroma, in particular the presence of activated pericytes on the blood vessels, is required for efficient intravasation of tumour cells from the primary tumour into the circulation (Harney et al., 2015; Viski et al., 2016; Xian et al., 2006; Yang et al., 2016) and an activated cancer-associated fibroblasts can prime tumour cells for metastatic growth (Zhang et al., 2013). The data presented here suggest that the increased efficiency of the D2A1-m2 subline to spontaneously metastasise may, at least in part, result from the ability to recruit and activate stromal cells, particularly stromal fibroblasts and pericytes.

In summary, the data presented here describes the derivation and characterisation of two new syngeneic metastatic mouse mammary carcinoma cell lines that have both overlapping and distinct behaviours when introduced into mice. Surprisingly, the ability of these cells to give rise to both spontaneous and experimental metastases *in vivo*, is not reflected by a more aggressive phenotype in *in vitro* proliferation, migration and invasion assays. However, some clues as to the mechanisms driving the metastatic phenotype have been revealed by assessing the cell lines in non-adherent culture and by examination of the tumour stroma.

Methods

In vivo studies

All *in vivo* studies were performed under UK Home Office Project Licenses 70/7413 and P6AB1448A granted under the Animals (Scientific Procedures) Act 1986. All studies were performed at The Institute of Cancer Research (Establishment Licence, X702B0E74 70/2902). Ethical permission was granted by the Institute of Cancer Research "Animal Welfare and Ethical Review Body" (AWERB). Female BALB/c and NSG mice (Charles River) between 6 and 12 weeks of age were housed in individually ventilated cages, monitored on a daily basis for signs of ill health and had food and water ad libitum. In all cases, experiments were terminated if the primary tumour reached a maximum allowable diameter of 15 mm or if a mouse showed signs of ill health.

The two metastatic sublines were generated independently (see Fig. 1A). In each case, 5 x 10^4 D2A1 cells were injected into the 4th mammary fat pad of a single BALB/c mouse under general anaesthesia. When the primary tumour reached 12 - 14 mm in diameter, lungs were dissected post-mortem, mechanically dissociated and placed into culture in DMEM (Gibco by ThermoFisher) plus 10% FBS (Gibco by ThermoFisher) and 1% penicillin/streptomycin. Medium was changed after 24 hours and then twice a week. After two weeks, when tumour cell colonies were visible, cells were replated and expanded. Once expanded, 5 x 10^4 cells were injected via the tail vein into a single BALB/c mouse. After 11-13 days, lungs were processed as before. The intravenous inoculation was repeated a further 2 times, each time into a single mouse. This procedure resulted in the two independently selected D2A1 sublines, D2A1-m1 and D2A1-m2.

For spontaneous metastasis assays, 5 x 10^4 cells were injected into the 4th mammary fat pad of BALB/c or NSG mice under general anaesthesia. Tumour growth was monitored using callipers and tumour volume calculated as 0.5236 x [(width + length)/2]^3 (Janik et al., 1975). Primary tumours and lung tissue were formalin-fixed and paraffin-embedded (FFPE) prior to sectioning. FFPE sections of primary tumours were stained with αSMA (Sigma, clone 1A4, 1:1000 dilution) or endomucin (Santa Cruz SC-65495; 1:2000 dilution). Detection was achieved with the VectaStain ABS system and sections were scanned on the NanoZoomer Digital Pathology (Hamamatsu). HRP staining was analysed in ImageJ from ≥6 fields of view per tumour, avoiding areas of necrosis. To quantify lung tumour burden, sections taken midway through the lung and stained with haematoxylin and eosin (H&E). Sections were scanned and analysed using the NanoZoomer Digital Pathology, file names were blinded and the number of macroscopic tumour nodules (defined as having a minimum area of 1000 μm^2/nodule) counted manually or quantified as % tumour burden. The % lung tumour burden was defined as (total tumour area)/(lung area) x 100 from a coronal H&E section of the lung. Where indicated, primary tumours were surgically resected under general anaesthesia.

For intravenous inoculation, 4 x 10^5 cells were injected into the tail vein of BALB/c mice. After 11 days mice were sacrificed. Lungs were processed as described in the spontaneous metastasis assay. For intracardiac inoculation, 2 x 10^5 luciferase expressing cells were injected into the left ventricle of BALB/c mice under general anaesthesia. Tumour burden was analysed by *ex vivo* IVIS imaging. Mice were injected intraperitoneally with 150 mg/kg D-luciferin (Caliper Life Sciences) in 100 μL. After 5 minutes, dissected hind limb long bones and brains were imaged using an IVIS imaging chamber (IVIS Illumina II). Luminescence measurements

(photons/second/cm^2) were acquired over 1 minute and analysed using the Living Image software (PerkinElmer) by placing a constant size region of interest over the tissues. For intrasplenic inoculation, 2 x 10^5 luciferase expressing cells were inoculated into the spleen parenchyma of BALB/c mice under general anaesthesia. After 10 minutes, a splenectomy was performed to avoid growth of splenic tumours. Tumour burden in the liver was assessed by *ex vivo* IVIS imaging as described above.

Cell lines

D2A1 cells were from Chambers lab stocks, from stocks originally obtained from Dr. Fred Miller (Mahoney et al., 1985; Miller et al., 1989). NIH-3T3 cell were from Isacke laboratory stocks. Cells were maintained in DMEM plus 10% FBS and 1% penicillin/streptomycin. Cells were luciferase transduced with lentiviral expression particles containing a firefly luciferase gene and a blasticidin-resistance gene (Amsbio, LVP326). All cells were routinely subject to mycoplasma testing.

In vitro studies

For cell shape analysis, 1x10^3 cells/well were seeded into a ViewPlate-96 Black with optically clear bottom (PerkinElmer). After 24 hours, cells were fixed in 4% paraformaldehyde (PFA), permeabilised with 0.5% Triton X-100 prior to staining with DAPI (molecular probes, D1306, 1:10,000) and Alexa488-labeled phalloidin (molecular probes, A12379, 1:500). Automated image acquisition was performed on an Operetta high content imaging system (Perkin Elmer). Cell shapes were analysed using basic algorithms in the Harmony high content analysis software package (Perkin-Elmer). Cells were initially defined using the DAPI channel to identify the nucleus and the cytoplasm was segmented using the Alexa488 channel. After this, the Harmony software allows the extraction of cell shape parameters, including “cell roundness”. On average 1110 cells were analysed per well (n=6 wells/cell line).

For cell proliferation/viability assays, 1x10^3 cells/well were seeded into 96-well plates. Cell viability was quantified either by CellTiter-Glo (Promega) at the indicated time points or by time-lapse imaging and quantification of cell confluence using the Live-Cell Analysis System IncuCyte (EssenBioscience).

For colony formation assay, 50 cells/well were seeded into a 6-well plate. Tumour cell colonies were stained 7 days later with crystal violet. Plates were scanned using the GelCount (Oxford Optronix) and image analysis performed using GelCount software and ImageJ.

For scratch wound migration assays, cells were seeded at high density to form a confluent layer in a 96 Image Lockplate (EssenBioscience). Scratches were created using a WoundMaker tool (EssenBioscience). Plates were imaged for 72 hours and analysed using the Live-Cell Analysis System IncuCyte (EssenBiosience).

For the Transwell migration assay, $1x10^3$ cells/well were seeded into IncuCyte ClearView 96-well chemotaxis plates in DMEM supplemented with 3% FBS. The bottom well contained DMEM supplemented with 10% FBS. Chemotactic migration was monitored and analysed using the IncuCyte Chemotaxis System (EssenBiosience).

For competitive fibroblast attraction assays the µ-Slide Chemotaxis system (ibidi) was used. In brief, 1.2×10^4 3T3 fibroblasts were seeded into the central observation chamber. 3 hours later, 3×10^4 D2A1 cells were seeded into the left reservoir and either D2A1, D2A1-m1 or D2A1-m2 cells into the right reservoir. Migration of 3T3 fibroblasts was imaged over 8 hours (20 minute intervals) using Slidebook 6 (3i) and a Nikon Eclipse TE2000-5 widefield microscope equipped with a plan flour 10x/0.3 ph1 wd 16 objective (Nikon) and a temperature- and CO_2-controlled chamber. Cell migration was analysed using the ImageJ Manual tracking and Chemotaxis and Migration Tool (ibidi) plugin.

To collect conditioned media, cells were seeded in DMEM supplemented with 10% FBS and cultured until 70-80% confluency. The media was then changed to DMEM supplemented with 2% FBS. Conditioned media was collected after 24 hours and filtered prior to use. For fibroblast contraction assays, 7×10^4 3T3 fibroblasts were embedded in 100 µl of a Matrigel (final concentration 2 mg/ml, Corning) and rat tail Collagen I (final concentration 4 mg/ml, Corning) mixture and seeded onto a glass bottom dish (P24G-1.0-13-F, MatTek Corporation). After the gel was set at 37°C, conditioned medium or DMEM supplemented with 2% FBS with or without recombinant TGFβ1 (R&D Systems)(5 ng/ml) was added. After 14 days, plates were scanned and the contracted gel area was measured using ImageJ.

For apoptosis assay, 5×10^4 cells/well were plated into either a tissue culture treated or low adherence 6-well plates. 24 hours after seeding, cells were stained with the Annexin V-APC/ PI Apoptosis Detection Kit (eBioscience) and analysed using a BD Biosciences LSRII flow cytometer with FACSDIVA and FlowJo software.

For spheroid growth assays, 1×10^2 cells were seeded into ultra-low adherence 96-well round bottom plates (Corning). Spheroid growth was monitored using the Celigo Image Cytometer (Nexcelom Bioscience). Cell viability was monitored with CellTiter-Glo at the indicated time points. To ensure proper lysis of

the spheroids, the incubation time with the CellTiter-Glo reagent was extended from 10 to 30 minutes before recording luminescence.

For invasion assays, 2 x 10^4 cells were seeded into ultra-low adherence 96-well round bottom plates. 3 days later, spheroids were transferred into fresh round bottom plates containing 2mg/ml collagen (rat tail collagen, Corning). After 24, 48 and 72 hours collagen plugs were fixed in 4% PFA, permeabilised with 0.5% Triton X-100 and stained with DAPI. Confocal z-stacks were acquired using the ImageXpress Micro Confocal High-Content Analysis System (Molecular Devices) equipped with a 60µm pinhole spinning-disk, and a plan apo λ 10x/0.45 objective (Nikon). z-stacks were maximally projected and DAPI images were quantified using a custom-written MATLAB script. In brief, spheroids were binarised using an intensity threshold while individual cells invading into the collagen, away from the spheroid, were marked using the watershed transform, which allows separation of contacting cells. The centroid coordinates of the spheroid and individual cells were extracted and the distance *d* between each cell and the spheroid was calculated using Euclidean trigonometry. The mean square displacement (MSD), a global measure of invasion, was obtained using the following equation:

$$\mathrm{MSD} = \frac{1}{N}\sum_{k=1}^{N} d_k^2 \quad (1)$$

where d_k is the distance (in µm) between the *k* cell and the spheroid and *N* is the total number of cells.

Gene expression profiling and analysis of clinical datasets

RNA from D2A1 and D2A1-m2 cultured cells (n = 3 independent biological replicates) and D2A1-m1 cultured cells (n = 2 independent biological replicates) was extracted using the RNeasy Mini kit (Qiagen). Microarray experiments were performed in two independent batches at Cambridge Genomic Services, University of Cambridge. In brief, RNA was assessed for concentration and quality using a SpectroStar (BMG Labtech) and a Bioanalyser (Agilent Technologies). The 8 RNA samples were amplified, labelled and hybridised on a MouseWG-6 v2.0 Expression BeadChip array (Illumina) following the manufacturer's instructions. Raw expression data were extracted in R using lumi package (http://www.bioconductor.org). Data were filtered to remove any non-expressed probes (detection $p > 0.01$) across all samples, transformed using variance-stabilising transformation, normalised using the robust spline normalisation method, and then batch-corrected using the function (ComBat) in the R package (sva). Sample relations were estimated using principle component analysis based on 7970 genes with coefficient of variance (standard deviation/mean > 0.1). Two-sample *t*-tests were used to identify differentially expressed genes

between (a) D2A1 and D2A1-m1, (b) D2A1 and D2A1-m2 and (c) D2A1-m1 and D2A1-m2, using the BRB-Array Tools (https://brb.nci.gov/BRB-ArrayTools/) with a threshold of parametric *p*-value < 0.001. Differentially expressed genes with a fold change ≥ 1.5 were subject to Ingenuity Pathway Analysis (IPA) to identify altered pathways, where $p < 0.05$ were considered as significant. When multiple probes were mapped to the same gene, the most variable probe measured by interquartile range (IQR) across the samples was selected to represent the gene. Gene expression data from this study are deposited at deposited in NCBI Gene Expression Omnibus (GSE101579).

Statistics

Statistics were performed using GraphPad Prism 6. Unless otherwise stated, all numerical data are expressed as mean ±s.e.m. Significant outliers were identified using the Grubb's test (α = 0.05, GraphPad Prism) and are shown as white symbols in the graphs. If not indicated otherwise, all comparisons between two groups were made using two-tailed, unpaired Student's *t*-test. If there was a significant difference in the variance of samples a Welch's correction was applied. Where multiple groups over time were compared, a two-way ANOVA followed by Bonferroni post-hoc testing was performed. *, $p < 0.05$; ** $p < 0.01$; ***, $p < 0.001$.

List of abbreviations

a.u., arbitrary units
AV, annexin V
cps, counts per second
DMEM, Dulbecco's Modified Eagle's Medium
FBS, foetal bovine serum
IPA, Ingenuity Pathway Analysis
Luc, luciferase
MSD, mean square displacement
NSG, NOD scid gamma
PFA, paraformaldehyde
PI, propidium iodide

Declarations

Acknowledgements. We are grateful to the expertise in the Breast Cancer Now Histopathology Facility and the ICR FACS and Light Microscopy Facility. Cambridge Genomic Services performed the gene expression analysis. We thank Slavomir Wantuch for lab support, David Vicente for help with the *in vivo* assays, Fredrik Wallberg for help with the imaging assays.

Competing interests. The authors declare that they have no competing interests

Funding. This work was funded by a grant from Worldwide Cancer Research to CMI and UJ, funding from Breast Cancer Now, working in partnership with Walk the Walk to CMI. We acknowledge NHS funding to the NIHR Biomedical Research Centre at The Royal Marsden and the ICR. UJ was a recipient of a Schrödinger fellowship of the Austrian Science Fund (FWF): J3434-B13. MJM was a recipient of a Helmholtz International Graduate School for Cancer Research (HIGS) fellowship and part of the EU-funded ERASMUS Placement Programme.

Availability of data and materials. Microarray data have been deposited in NCBI Gene Expression Omnibus (GSE101579).

References

Aslakson, C. J. and Miller, F. R. (1992). Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. *Cancer Res* **52**, 1399-405.

Bos, P. D., Zhang, X. H., Nadal, C., Shu, W., Gomis, R. R., Nguyen, D. X., Minn, A. J., van de Vijver, M. J., Gerald, W. L., Foekens, J. A. et al. (2009). Genes that mediate breast cancer metastasis to the brain. *Nature* **459**, 1005-9.

Bouchard, G., Bouvette, G., Therriault, H., Bujold, R., Saucier, C. and Paquette, B. (2013). Pre-irradiation of mouse mammary gland stimulates cancer cell migration and development of lung metastases. *Br J Cancer* **109**, 1829-38.

Bruna, A., Rueda, O. M., Greenwood, W., Batra, A. S., Callari, M., Batra, R. N., Pogrebniak, K., Sandoval, J., Cassidy, J. W., Tufegdzic-Vidakovic, A. et al. (2016). A Biobank of Breast Cancer Explants with Preserved Intra-tumor Heterogeneity to Screen Anticancer Compounds. *Cell* **167**, 260-274 e22.

Byrne, A. T., Alferez, D. G., Amant, F., Annibali, D., Arribas, J., Biankin, A. V., Bruna, A., Budinska, E., Caldas, C., Chang, D. K. et al. (2017). Interrogating open issues in cancer precision medicine with patient-derived xenografts. *Nat Rev Cancer* **17**, 254-268.

Chen, P., Zuo, H., Xiong, H., Kolar, M. J., Chu, Q., Saghatelian, A., Siegwart, D. J. and Wan, Y. (2017). Gpr132 sensing of lactate mediates tumor-macrophage interplay to promote breast cancer metastasis. *Proc Natl Acad Sci U S A* **114**, 580-585.

Eckhardt, B. L., Parker, B. S., van Laar, R. K., Restall, C. M., Natoli, A. L., Tavaria, M. D., Stanley, K. L., Sloan, E. K., Moseley, J. M. and Anderson, R. L. (2005). Genomic analysis of a spontaneous model of breast cancer metastasis to bone reveals a role for the extracellular matrix. *Mol Cancer Res* **3**, 1-13.

Fantozzi, A. and Christofori, G. (2006). Mouse models of breast cancer metastasis. *Breast Cancer Res* **8**, 212.

Francia, G., Cruz-Munoz, W., Man, S., Xu, P. and Kerbel, R. S. (2011). Mouse models of advanced spontaneous metastasis for experimental therapeutics. *Nat Rev Cancer* **11**, 135-41.

Frisch, S. M. and Screaton, R. A. (2001). Anoikis mechanisms. *Curr Opin Cell Biol* **13**, 555-62.

Guy, C. T., Cardiff, R. D. and Muller, W. J. (1992a). Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol Cell Biol* **12**, 954-61.

Guy, C. T., Webster, M. A., Schaller, M., Parsons, T. J., Cardiff, R. D. and Muller, W. J. (1992b). Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. *Proc Natl Acad Sci U S A* **89**, 10578-82.

Hanahan, D. and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* **144**, 646-74.

Harney, A. S., Arwert, E. N., Entenberg, D., Wang, Y., Guo, P., Qian, B. Z., Oktay, M. H., Pollard, J. W., Jones, J. G. and Condeelis, J. S. (2015). Real-Time Imaging Reveals Local, Transient Vascular Permeability, and Tumor Cell Intravasation Stimulated by TIE2hi Macrophage-Derived VEGFA. *Cancer Discov* **5**, 932-43.

Holen, I., Speirs, V., Morrissey, B. and Blyth, K. (2017). In vivo models in breast cancer research: progress, challenges and future directions. *Dis Model Mech* **10**, 359-371.

Hunter, K. (2006). Host genetics influence tumour metastasis. *Nat Rev Cancer* **6**, 141-6.

Janik, P., Briand, P. and Hartmann, N. R. (1975). The effect of estrone-progesterone treatment on cell proliferation kinetics of hormone-dependent GR mouse mammary tumors. *Cancer Res* **35**, 3698-704.

Johnstone, C. N., Smith, Y. E., Cao, Y., Burrows, A. D., Cross, R. S., Ling, X., Redvers, R. P., Doherty, J. P., Eckhardt, B. L., Natoli, A. L. et al. (2015). Functional and molecular characterisation of EO771.LMB tumours, a new C57BL/6-mouse-derived model of spontaneously metastatic mammary cancer. *Dis Model Mech* **8**, 237-51.

Khanna, C. and Hunter, K. (2005). Modeling metastasis in vivo. *Carcinogenesis* **26**, 513-23.

Lelekakis, M., Moseley, J. M., Martin, T. J., Hards, D., Williams, E., Ho, P., Lowen, D., Javni, J., Miller, F. R., Slavin, J. et al. (1999). A novel orthotopic model of breast cancer metastasis to bone. *Clin Exp Metastasis* **17**, 163-70.

Lockman, P. R., Mittapalli, R. K., Taskar, K. S., Rudraraju, V., Gril, B., Bohn, K. A., Adkins, C. E., Roberts, A., Thorsheim, H. R., Gaasch, J. A. et al. (2010). Heterogeneous blood-tumor barrier permeability determines drug efficacy in experimental brain metastases of breast cancer. *Clin Cancer Res* **16**, 5664-78.

Mahoney, K. H., Miller, B. E. and Heppner, G. H. (1985). FACS quantitation of leucine aminopeptidase and acid phosphatase on tumor-associated macrophages from metastatic and nonmetastatic mouse mammary tumors. *J Leukoc Biol* **38**, 573-85.

Miller, F. R., McEachern, D. and Miller, B. E. (1989). Growth regulation of mouse mammary tumor cells in collagen gel cultures by diffusible factors produced by normal mammary gland epithelium and stromal fibroblasts. *Cancer Res* **49**, 6091-7.

Morris, V. L., Tuck, A. B., Wilson, S. M., Percy, D. and Chambers, A. F. (1993). Tumor progression and metastasis in murine D2 hyperplastic alveolar nodule mammary tumor cell lines. *Clin Exp Metastasis* **11**, 103-12.

Neve, R. M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F. L., Fevr, T., Clark, L., Bayani, N., Coppe, J. P., Tong, F. et al. (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer cell* **10**, 515-27.

O'Reilly, M. S., Holmgren, L., Shing, Y., Chen, C., Rosenthal, R. A., Moses, M., Lane, W. S., Cao, Y., Sage, E. H. and Folkman, J. (1994). Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell* **79**, 315-28.

Paoli, P., Giannoni, E. and Chiarugi, P. (2013). Anoikis molecular pathways and its role in cancer progression. *Biochim Biophys Acta* **1833**, 3481-98.

Shibue, T. and Weinberg, R. A. (2009). Integrin beta1-focal adhesion kinase signaling directs the proliferation of metastatic cancer cells disseminated in the lungs. *Proc Natl Acad Sci U S A* **106**, 10290-5.

Tester, A. M., Ruangpanit, N., Anderson, R. L. and Thompson, E. W. (2000). MMP-9 secretion and MMP-2 activation distinguish invasive and metastatic sublines of a mouse mammary carcinoma system showing epithelial-mesenchymal transition traits. *Clin Exp Metastasis* **18**, 553-60.

Viski, C., Konig, C., Kijewska, M., Mogler, C., Isacke, C. M. and Augustin, H. G. (2016). Endosialin-Expressing Pericytes Promote Metastatic Dissemination. *Cancer Res* **76**, 5313-25.

Xian, X., Hakansson, J., Stahlberg, A., Lindblom, P., Betsholtz, C., Gerhardt, H. and Semb, H. (2006). Pericytes limit tumor cell metastasis. *J Clin Invest* **116**, 642-51.

Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., Savagner, P., Gitelman, I., Richardson, A. and Weinberg, R. A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117**, 927-39.

Yang, Y., Andersson, P., Hosaka, K., Zhang, Y., Cao, R., Iwamoto, H., Yang, X., Nakamura, M., Wang, J., Zhuang, R. et al. (2016). The PDGF-BB-SOX7 axis-modulated IL-33 in pericytes and stromal cells promotes metastasis through tumour-associated macrophages. *Nat Commun* **7**, 11385.

Yang, Y., Yang, H. H., Hu, Y., Watson, P. H., Liu, H., Geiger, T. R., Anver, M. R., Haines, D. C., Martin, P., Green, J. E. et al. (2017). Immunocompetent mouse allograft models for development of therapies to target breast cancer metastasis. *Oncotarget* **8**, 30621-30643.

Zhang, X. H., Jin, X., Malladi, S., Zou, Y., Wen, Y. H., Brogi, E., Smid, M., Foekens, J. A. and Massague, J. (2013). Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. *Cell* **154**, 1060-73.

Figure Legends

Fig. 1 Generation of syngeneic spontaneously metastatic D2A1 sublines. **A** Diagram outlining the strategy for selection of the metastatic sublines (see Methods and text for details). **B,C** 5 x 10^4 D2A1, D2A1-m1 or D2A1-m2 cells were inoculated into the 4th mammary fat pad of BALB/c mice (n = 5 - 6 mice per group). Growth of tumours in individual mice is shown. **B** Mice were culled between days 31 - 39 when the primary tumours reached a diameter of 12 - 14 mm (left panel). Metastatic burden in the lungs was monitored by % tumour area and number of lung nodules per lung section. Significant outliers are shown as white symbols. Data are mean values per mouse ±s.e.m. Shown are representative images of lung metastases (arrowheads). Scale bar, 200 μm. **C** Primary tumours were surgically excised when they reached ~6 mm in diameter (day 13 - 15). All mice were culled on day 43 when the first mouse showed signs of ill health. Data shown are mean values per mouse ±s.e.m. for primary tumour growth and metastatic burden in the lungs as assessed by % tumour area per lung section. Significant outliers are shown as white symbols. Representative images of lung sections are shown. Scale bar, 2mm.

Fig. 2 Spontaneous metastasis in an immunocompromised setting. 5 x 10^4 D2A1 or D2A1-m2 cells were inoculated into the 4th mammary fat pad of NOD scid gamma (NSG) mice (n = 5 mice per group). **A** Mean tumour volume ±s.e.m., non-significant at all time points. All mice were culled on day 28 when the first mouse showed signs of ill health. **B** Metastatic burden in the lung assessed as % tumour area and number of metastatic nodules per lung section. Shown are representative images. Data

shown are mean values per mouse ±s.e.m. Significant outlier is shown as white symbol. Scale bar, 2 mm.

Fig. 3 Experimental metastasis assays**. A** 4 x 10^5 D2A1, D2A1-m1 or D2A1-m2 cells were inoculated via the tail vein into BALB/c mice (n = 8 mice per group). 11 days later, lungs were removed at necropsy. Data show quantification of tumour burden as monitored by % tumour area and number of lung nodules per lung section. Right panel, representative lung images. Scale bar, 2 mm. **B** 2 x 10^5 D2A1-Luc, D2A1-m1-Luc or D2A1-m2-Luc cells were inoculated into the left ventricle of BALB/c mice (n = 6 mice per group). 10 days later, bones and brains were removed at necropsy and IVIS imaged *ex vivo*. Representative *ex vivo* bone IVIS images are shown. Scale bar, 1 cm. **C** 2 x 10^5 D2A1-Luc, D2A1-m1-Luc or D2A1-m2-Luc cells were inoculated into the spleen of BALB/c mice (n = 8 mice per group). 13 days later livers were removed at necropsy and IVIS imaged *ex vivo*. Shown are representative IVIS images Scale bar, 1 cm. Significant outliers are shown as white symbols. All data are mean values per mouse ±s.e.m.

Fig. 4 Characterisation of the D2A1 metastatic sublines *in vitro*. **A** Upper panel, phase contrast images of D2A1, D2A1-m1 and D2A1-m2 cells. Scale bar, 200 μm. Lower panel, cells cultured on coverslips were stained with Alexa488-phalloidin and DAPI. Scale bar, 100 μm. Data shows mean values (n = 6 samples per cell line) ±s.e.m. of cell roundness (see Methods). Equivalent results were obtained on 2 separate occasions. **B** Colony formation assay. Data shows mean number of colonies and average colony area ±s.e.m., n = 3 wells per cell line. Representative images of crystal violet stained wells are shown below. Equivalent results were obtained on 3 separate occasions. **C** Cell proliferation monitored using the IncuCyte Live-Cell Analysis System. Data shown are mean values ±s.d., n = 6 wells per cell line. Equivalent results were obtained on 2 separate occasions. **D** Cell viability measured by CellTiter-Glo at the indicated time points. Data shown are mean values ±s.d., n = 6 wells per cell line per time point. Equivalent results were obtained on 2 separate occasions. **E** Wound healing assay monitored using the IncuCyte Live-Cell Analysis System. Data represents % wound closure ±s.d., n = 8 wells per cell line. Shown are representative false coloured images of cells time at 0 and at 24 hours. Scale bar, 300 μm**. F** Transwell chemotaxis assay. Data shown are mean values ±s.e.m. (n = 3) for cells transmigrated to the lower side of the filter relative to the initial plated cells. **C-F** Statistical differences were determined using 2-way ANOVA and Bonferroni post-hoc testing, with time and cell lines as independent variables. In

all cases both D2A1-m1 and D2A1-m2 are significantly different to the parental D2A1 cells ($P < 0.0001$).

Fig. 5 Characterisation of the D2A1 metastatic sublines in non-adherent culture. **A** D2A1, D2A1-m1 and D2A1-m2 cells were plated onto adherent tissue culture or non-adherent 6-well plates and apoptosis monitored 24 hours later by annexin V/PI staining. Data shows the proportions of live (AV-/PI-), early apoptotic (AV+/PI-), late apoptotic (AV+/PI+) and necrotic (AV-/PI+) cells, mean of 3 experiments ±s.e.m., n = 3 per cell line per experiment. D2A1 cells had significantly less viable cells compared to D2A1-m1 ($P = 0.0114$) and D2A1-m2 ($P = 0.0119$) cells. **B** Spheroid assay. Cells were plated in U bottom low adherence plates (n = 6 per cell line) and cultured for 14 days. At stated time points, cell viability in the tumour spheroids was monitored by CellTiter-Glo, mean values ±s.e.m. Examples of tumour spheroid at day 14 are shown. Scale bar, 400 μm. Equivalent results were obtained on 2 separate occasions. **c** Invasion into a 3D collagen matrix. Spheroids were embedded into collagen and single cell invasion was monitored 24, 48 and 72 hours later by measuring the mean square displacement (MSD). Data shown are mean values ±s.e.m., n = 3 - 6 spheroids per time point. Representative images at 48 hours are shown. Scale bar, 400 μm. All panels, statistical differences were determined using 2-way ANOVA and Bonferroni post-hoc testing. Equivalent results were obtained on 2 separate occasions.

Fig. 6 Stromal cell recruitment in primary tumours**. A, B** Primary tumours from the spontaneous metastasis experiments shown in Fig. 1B and Fig. 2 were sectioned and stained for the αSMA or endomucin. **A** BALB/c mice from Fig. 1B, **B** NSG mice from Fig. 2. Left panels, representative images. Scale bar, 100 μm. Right panels, quantification of staining from at least 4 tumours per group ±s.e.m. (n > 6 fields of view per tumour). Significant outlier is shown as white symbol. **C** Competitive fibroblast attraction assay (see Supplementary Figure S4A for experimental set up). Migration of individual 3T3 cells towards D2A1 *vs.* D2A1, D2A1 *vs.* D2A1-m1 or D2A1 *vs.* D2A1-m2 cells was monitored over 8 hours (n > 60 cells tracked/sample). Left panel, percentage net displacement of fibroblasts (see Supplementary Figure S4B for individual cell tracks and quantification). Middle and right panels, velocity (μm/min) and Euclidean distance (μm) of fibroblast migration. Box and whisker graphs, box extends from the 25th to 75th percentiles, middle line plotted at median, whiskers are minimum and maximum, + indicates mean of data. **D** Fibroblast contraction assay. Data shown are % contraction of the fibroblast containing

Matrigel/collagen gel after 14 days treatment with DMEM, TGFβ or conditioned medium from D2A1, D2A1-m1 and D2A1-m2 cells, n = 2 wells/condition. Representative images of the gels are shown.

Fig. 7 Gene expression profiling. RNA was isolated from independent biological replicates (D2A1.1, .2 and .3; D2A1-m1.1 and .2; D2A1-m2.1, .2 and .3) and subject to gene expression profiling. **A** Principle component analysis estimating the relations of D2A1, D2A1-m1 and D2A1-m2 cells based on the genes with coefficient of variance > 0.1 across the 8 samples. **B** Upper panel, number of differentially expressed genes ($p < 0.001$, ≥ 1.5-fold change). Lower panel, number of unique and shared differentially expressed genes between D2A1-m1 and D2A1-m2 sublines. Venn diagram depicting percentages. **C** Dendrogram shows correlation-centred hierarchical clustering based on average linkage. Shown are tumour cell expression data of 481 genes that were significantly differentially expressed between D2A1 and D2A1-m1 or D2A1 and D2A1-m2 cells with ≥ 1.5-fold change, $p < 0.001$. **D** Top canonical pathways changed between D2A1-m1 and D2A1 as well as D2A1-m2 and D2A1 identified using IPA ($p<0.05$) based on the differentially expressed genes ($p <$ 0.001, ≥ 1.5-fold change).

Figure 1

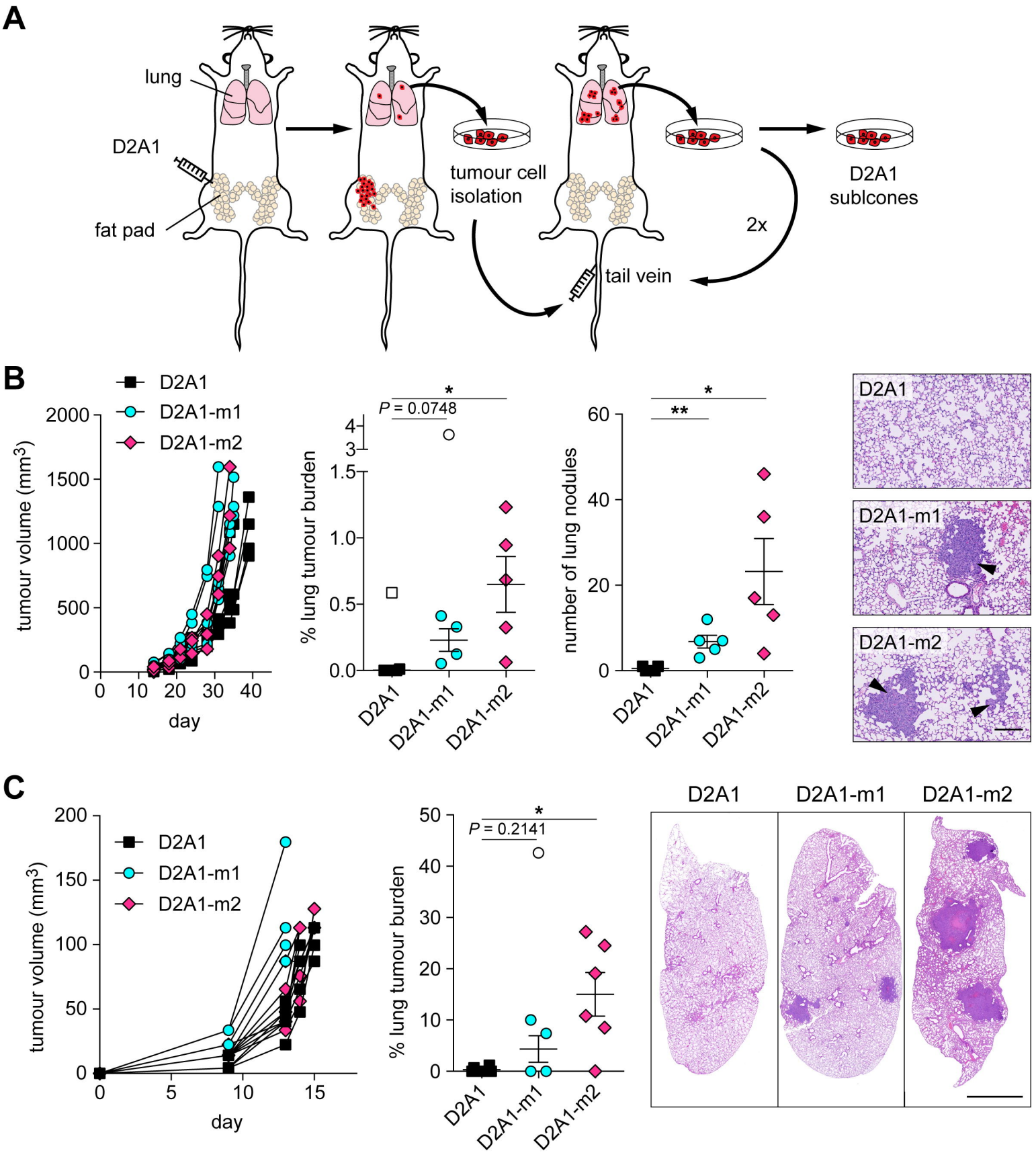

A
lung
D2A1
fat pad
tumour cell isolation
2x
tail vein
D2A1 sublcones
B
D2A1
D2A1-m1
D2A1-m2
tumour volume (mm3)
day
% lung tumour burden
P = 0.0748
number of lung nodules
C
D2A1
D2A1-m1
D2A1-m2
tumour volume (mm3)
day
% lung tumour burden
P = 0.2141

Figure 2

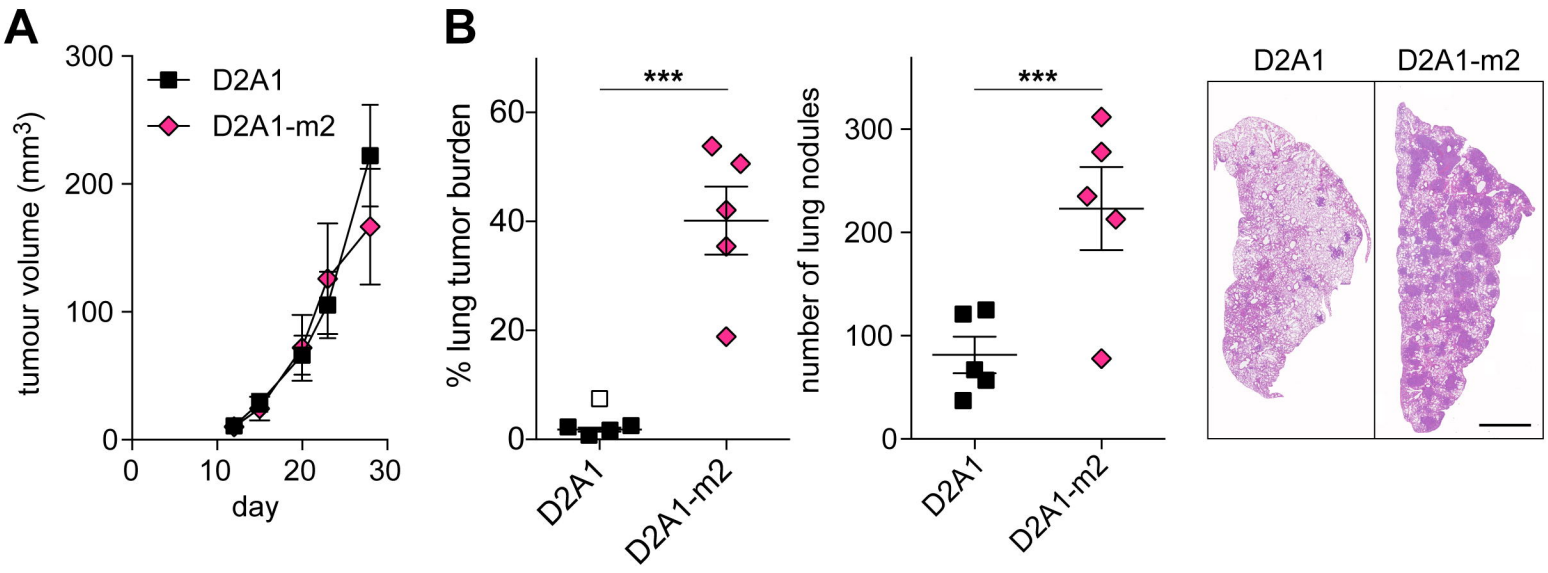

A
B
tumour volume (mm3)
300
200
100
0
0
10
20
30
day
D2A1
D2A1-m2
% lung tumor burden
60
40
20
0
***
D2A1
D2A1-m2
number of lung nodules
300
200
100
0
***
D2A1
D2A1-m2
D2A1
D2A1-m2

Figure 3

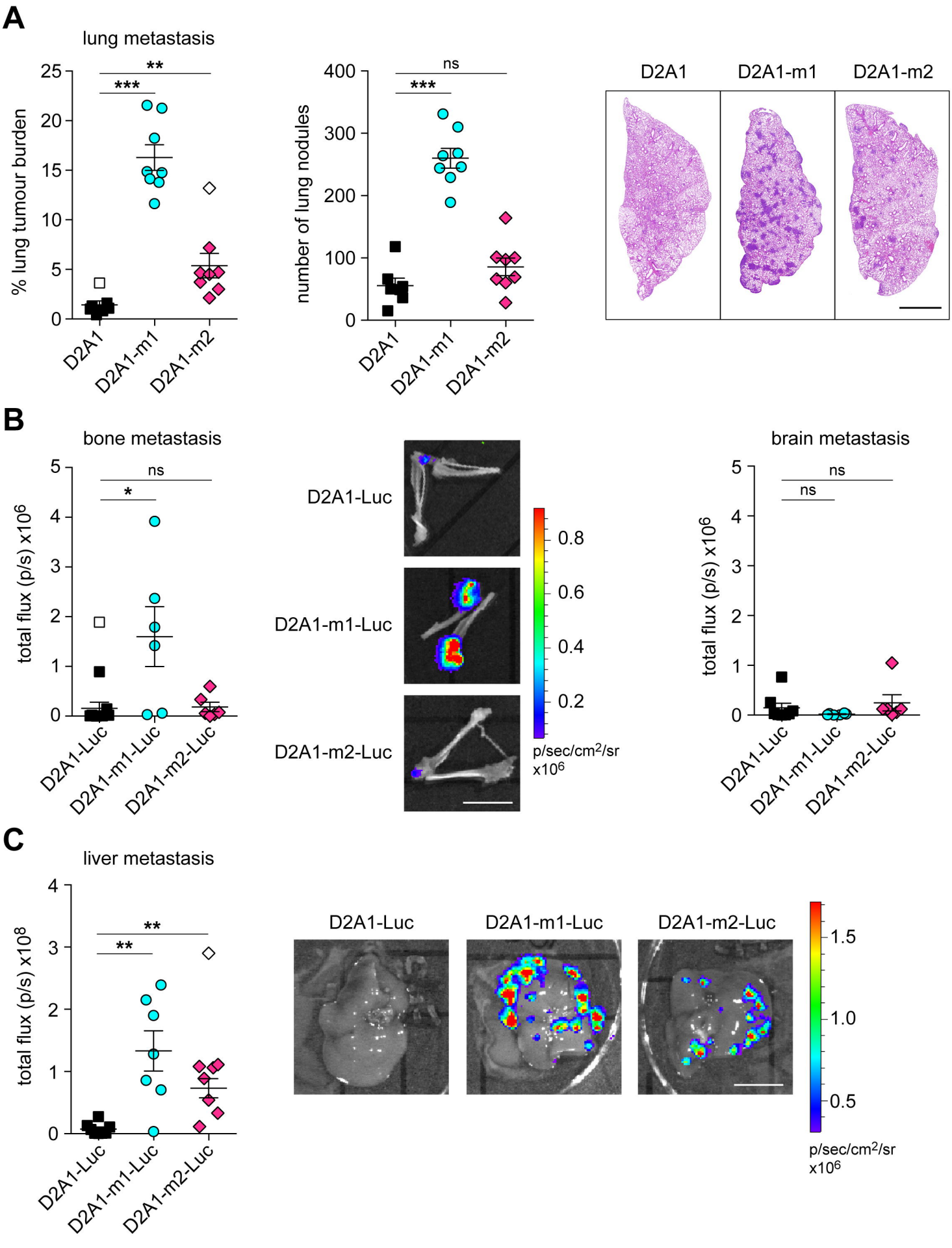

A
lung metastasis
% lung tumour burden
25
20
15
10
5
0
**
***
D2A1
D2A1-m1
D2A1-m2
number of lung nodules
400
300
200
100
0
ns
***
D2A1
D2A1-m1
D2A1-m2
D2A1
D2A1-m1
D2A1-m2
B
bone metastasis
total flux (p/s) x10^6
5
4
3
2
1
0
ns
*
D2A1-Luc
D2A1-m1-Luc
D2A1-m2-Luc
D2A1-Luc
D2A1-m1-Luc
D2A1-m2-Luc
0.8
0.6
0.4
0.2
p/sec/cm2/sr
x10^6
brain metastasis
total flux (p/s) x10^6
5
4
3
2
1
0
ns
ns
D2A1-Luc
D2A1-m1-Luc
D2A1-m2-Luc
C
liver metastasis
total flux (p/s) x10^8
4
3
2
1
0
**
**
D2A1-Luc
D2A1-m1-Luc
D2A1-m2-Luc
D2A1-Luc
D2A1-m1-Luc
D2A1-m2-Luc
1.5
1.0
0.5
p/sec/cm2/sr
x10^6

Figure 4

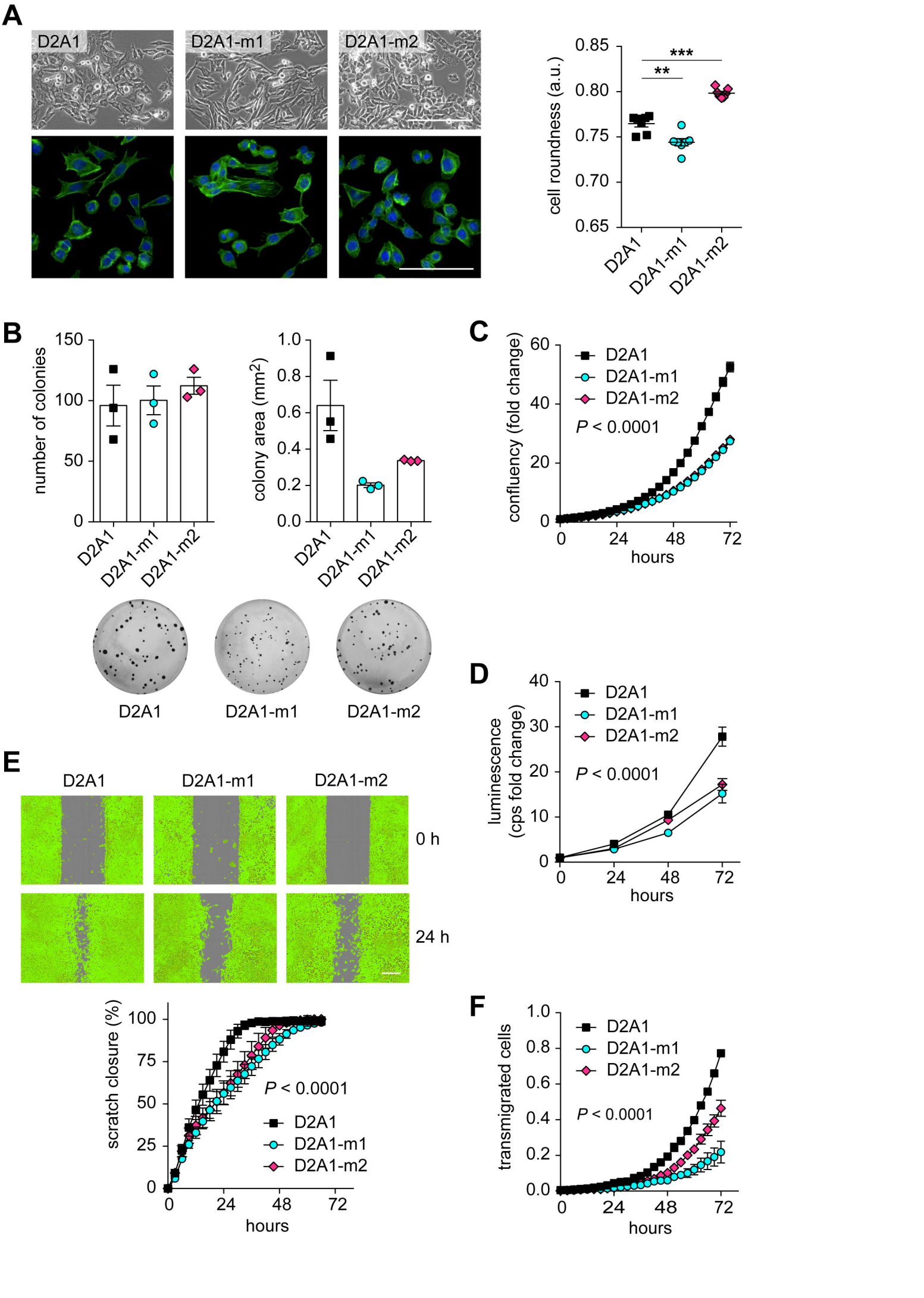

A
D2A1
D2A1-m1
D2A1-m2
cell roundness (a.u.)
0.85
0.80
0.75
0.70
0.65
***
**
B
number of colonies
150
100
50
0
colony area (mm2)
1.0
0.8
0.6
0.4
0.2
0.0
C
confluency (fold change)
60
40
20
0
P < 0.0001
0
24
48
72
hours
D
luminescence (cps fold change)
40
30
20
10
0
E
0 h
24 h
scratch closure (%)
100
75
50
25
F
transmigrated cells
1.0
0.8
0.6
0.4
0.2
0.0

Figure 5

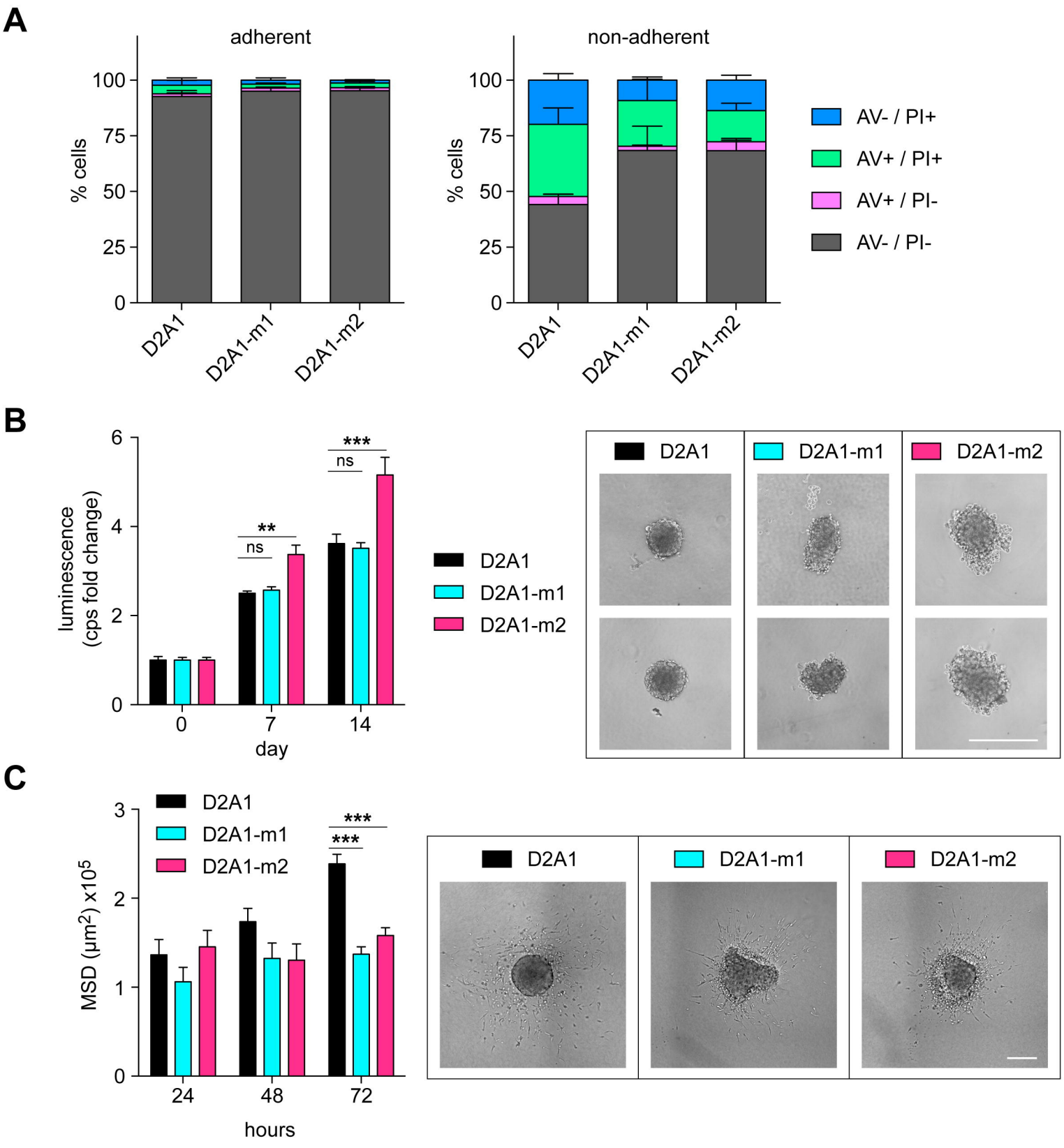

A
adherent
non-adherent
% cells
100
75
50
25
0
D2A1
D2A1-m1
D2A1-m2
AV- / PI+
AV+ / PI+
AV+ / PI-
AV- / PI-
B
luminescence (cps fold change)
6
4
2
0
**
ns
***
ns
0
7
14
day
D2A1
D2A1-m1
D2A1-m2
C
MSD (µm²) x10⁵
3
2
1
0
***
***
24
48
72
hours

Figure 6

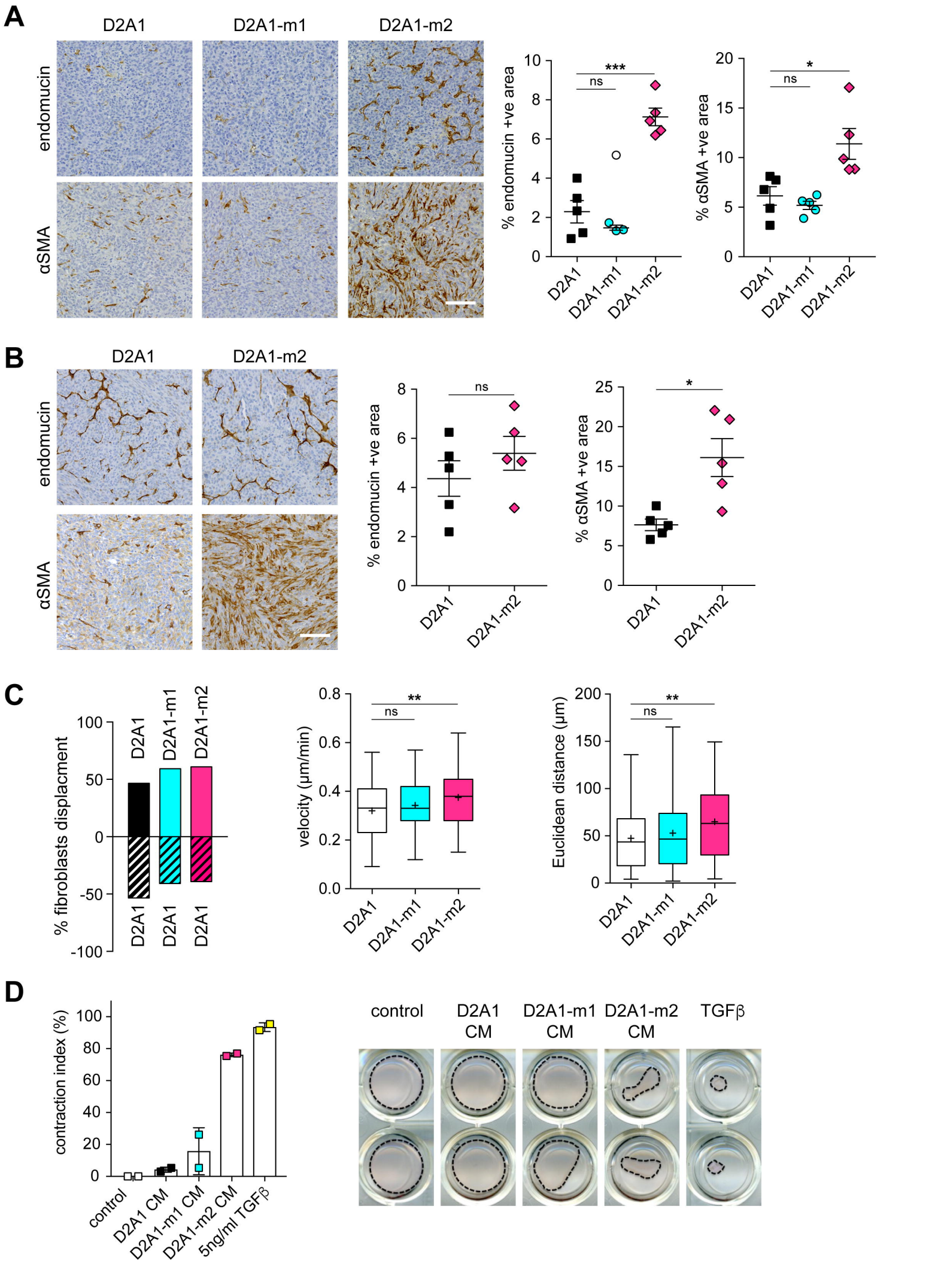
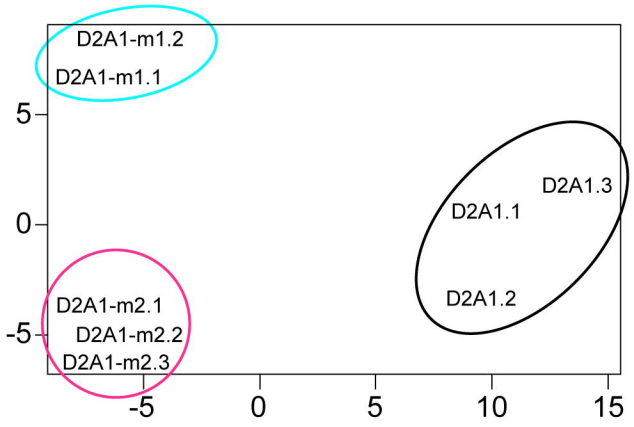

A
D2A1
D2A1-m1
D2A1-m2
endomucin
αSMA
% endomucin +ve area
% αSMA +ve area
***
*
ns
B
D2A1
D2A1-m2
endomucin
αSMA
% endomucin +ve area
% αSMA +ve area
ns
*
C
% fibroblasts displacment
velocity (μm/min)
Euclidean distance (μm)
**
ns
D
contraction index (%)
control
D2A1 CM
D2A1-m1 CM
D2A1-m2 CM
5ng/ml TGFβ
TGFβ

Figure 7

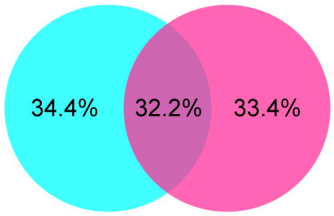
A



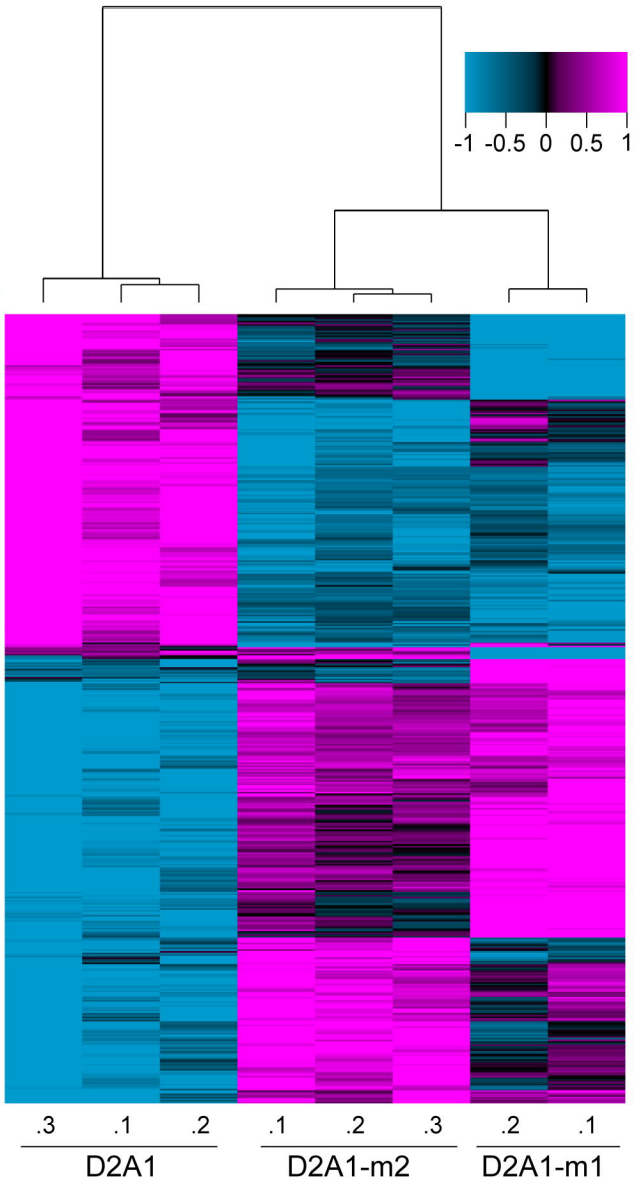
B

	D2A1-m1 vs D2A1	D2A1-m2 vs D2A1
sign. changes > 1.5 fold	323	318
up-regulated genes	176	178
down-regualed genes	147	140

	number of genes
genes changed in both m1 & m2	156
genes only changed in m1	167
genes only changed in m2	162

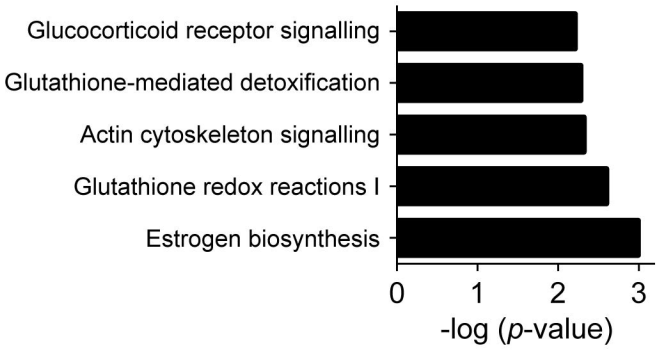


C

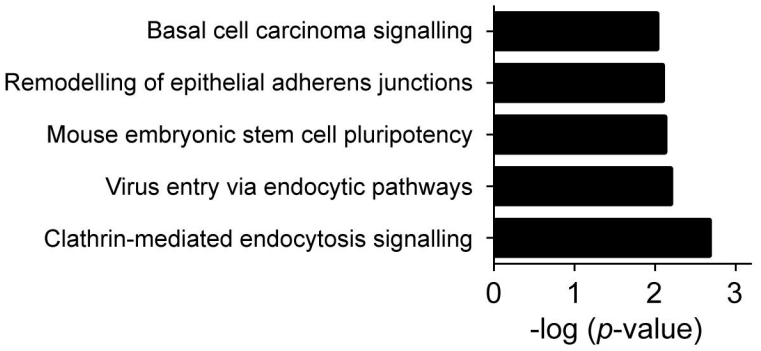


D

Top canonical pathways D2A1-m1 *vs.* D2A1



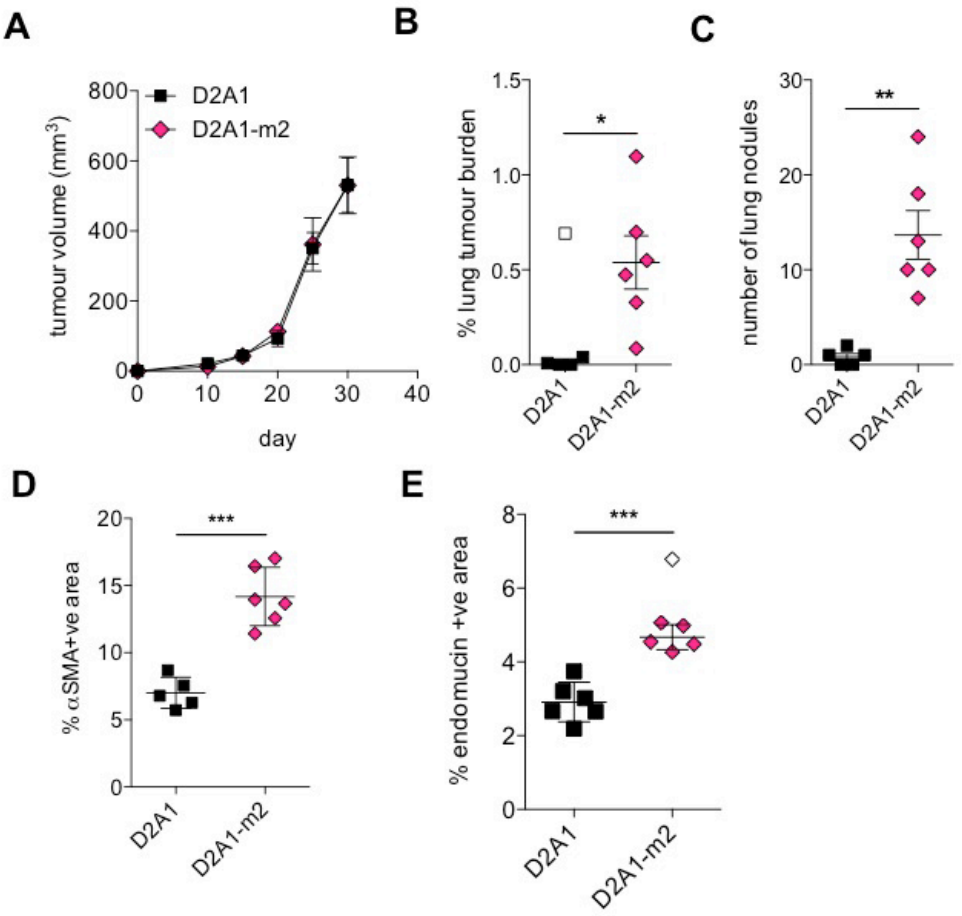
Top canonical pathways D2A1-m2 *vs.* D2A1



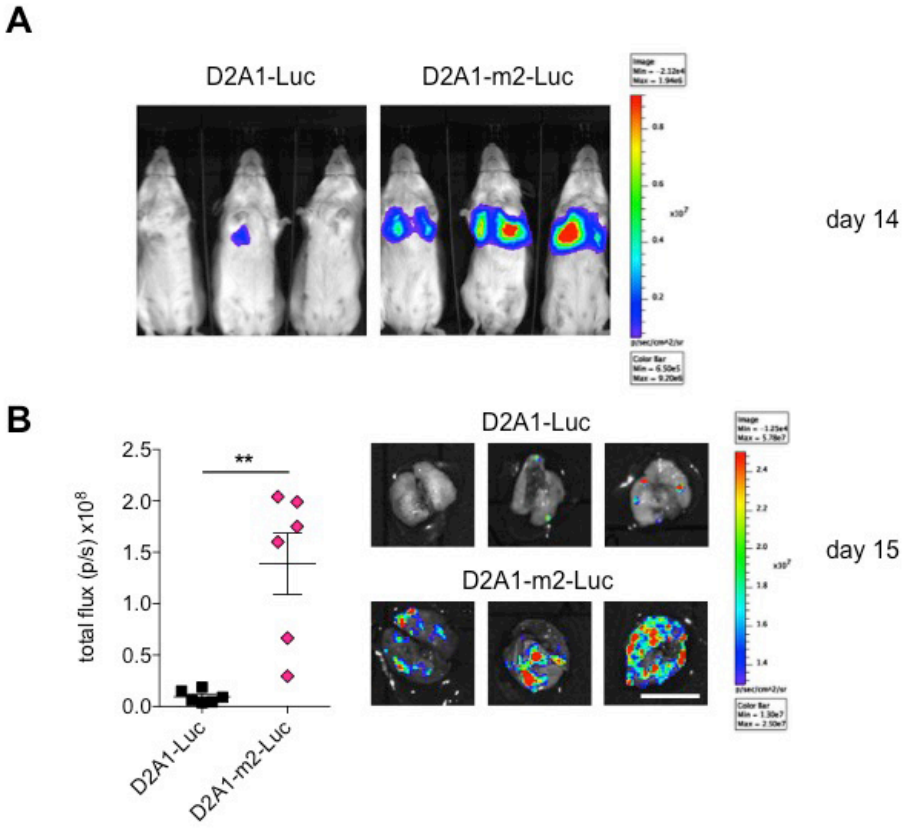
Generation and characterisation of two D2A1 mammary cancer sublines to model spontaneous and experimental metastasis in a syngeneic BALB/c host

Ute Jungwirth, Antoinette van Weverwijk, Miriam J. Melake, Ann F. Chambers, Qiong Gao, Marc Fivaz and Clare M. Isacke

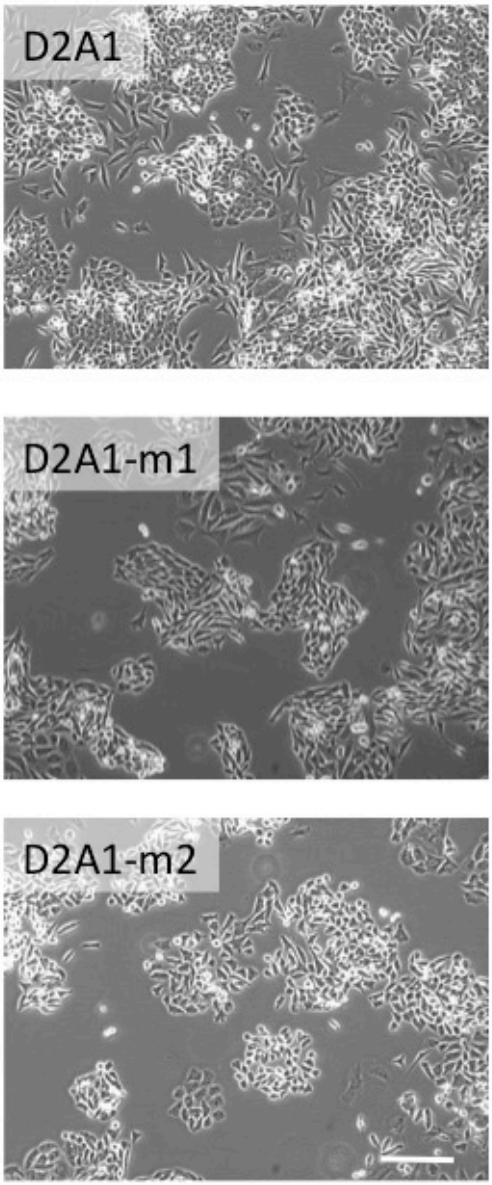
Supplementary Material



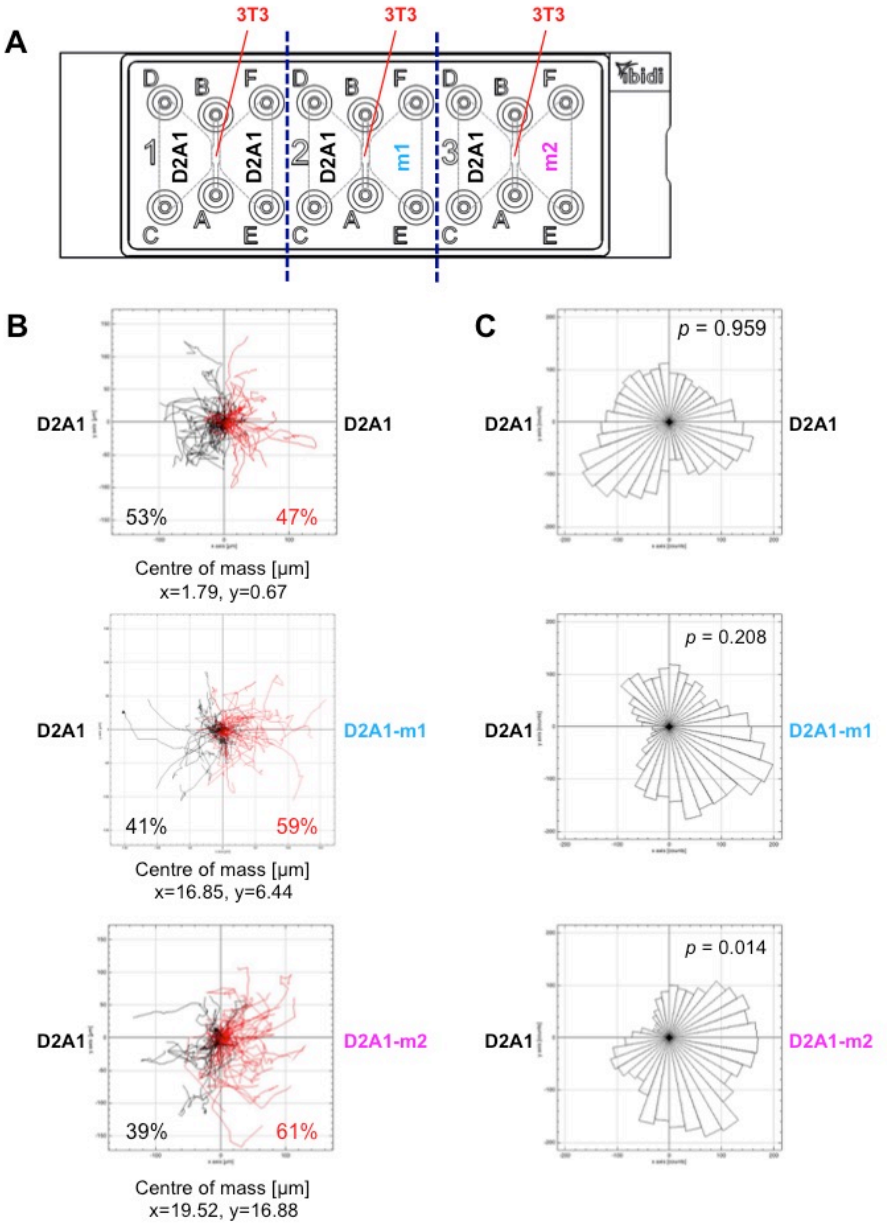
Supplementary Figure S1 Spontaneous metastasis assay in BALB/c mice. 5 x 10^4 D2A1 or D2A1-m2 cells were inoculated into the 4th mammary fat pad of BALB/c mice (n = 5 or 6 mice per group) and culled on day 30. **A** Primary tumour growth. **B,C** Spontaneous metastasis to the lung assessed by % tumour burden and number of metastatic nodules per lung section. Data shown are mean values per mouse +/- SEM. **D**,**E** Primary tumours were sectioned and stained for **D** the activated fibroblast marker αSMA, or **E** the endothelial marker endomucin. Data shows quantification of staining of 6 tumours per group ±s.e.m. (n > 6 fields of view per tumour).



Supplementary Figure S2 Experimental lung metastasis assay with luciferase tagged cells. 4 x 10^5 D2A1-Luc or D2A1-m2-Luc cells were inoculated intravenously into BALB/c mice (n = 6 mice per group). **A** Representative *in vivo* IVIS imaging on day 14. **B** Mice were culled on day 15 when the first mouse showed signs of ill health. Quantification of tumour burden in the lungs via *ex vivo* IVIS imaging. Mean values per mouse ±s.e.m.. Representative *ex vivo* lung IVIS images are shown. Scale bar, 1 cm.



Supplementary Figure S3 Phase contrast images of D2A1, D2A1-m1 and D2A1-m2 cells grown on tissue culture plastic. Scale bar, 400 μm.



Supplementary Figure S4 Competitive fibroblasts migration assay. **A** Diagram of the experimental setup for the 'ibidi' µ-slide chemotaxis system. Following the manufacturer's instructions, fibroblasts were injected into the central viewing chamber via port A. 3 hours later, D2A1 cells were injected into the left reservoir via port C and either D2A1, D2A1-m1 or D2A1-m2 cells were injected into the right reservoir via port E. Fibroblast migration in the central chamber was imaged over 8 hours (see Methods for further details). **B** Paths taken by individual 3T3 fibroblasts, where all initial starting positions (t=0) are plotted at the origin (x=0, y=0). Red indicates paths with a positive net displacement towards tumour cells seeded on the right hand side. Black paths indicate net displacement towards the left. The percentage of cells exhibiting movement to the left (black) or right (red) is indicated and shown graphically in Fig. 6C, left panel. Coordinates of centre of mass are given below the panels. **C** Corresponding circular histograms, with *p*-values calculated using the Rayleigh test for vector data. Only D2A1-m2 cells induce a significant inhomogeneous distribution of fibroblast trajectories.

Supplementary Table S1. Top 40 genes differentially expressed in D2A1-m1 cells *vs.* D2A1 cells

Symbol	Gene name	Accession Number	Fold change	Parametric *p*-value
Mgp	Matrix Gla protein	NM_008597	16.95	< 1e-07
Arhgdib	Rho, GDP dissociation inhibitor beta	NM_007486	13.33	5.10E-05
Fgf7	Fibroblast Growth Factor 7	NM_008008	12.35	1.71E-05
Itgbl1	Integrin Subunit Beta Like 1	NM_145467	6.67	2.21E-05
Thbs2	Thrombospondin 2	NM_011581	6.25	1.30E-06
Pdgfrl	Platelet-derived growth factor receptor-like	NM_026840	5.88	2.70E-06
Id4	Inhibitor of DNA binding 4	NM_031166	5.88	7.40E-06
Tnfrsf11b	TNF receptor superfamily member 11b	NM_008764	5.56	2.93E-05
Lgals7	Galectin 7	NM_008496	5.56	3.59E-05
Akr1c18	Aldo-keto reductase family 1, member C18	NM_134066	5.00	2.40E-06
Ly6a	Lymphocyte antigen 6 complex, locus A	NM_010738	5.00	4.80E-06
Tpm2	Tropomyosin 2, beta	NM_009416	5.00	1.10E-05
Chst1	Carbohydrate (keratan sulfate Gal-6) sulfotransferase 1	NM_023850	4.17	5.00E-07
Ifi27	Interferon, alpha-inducible protein 27	NM_029803	4.17	8.64E-04
AA467197		NM_001004174	4.00	3.00E-07
Casp1	Caspase 1	NM_009807	4.00	6.88E-04
Aqp1	Aquaporin 1	NM_007472	3.45	5.17E-05
Krt8	Keratin 8	NM_031170	3.45	1.19E-04
Ptrf	Caveolae associated 1	NM_008986	3.23	2.70E-05
Vcam1	Vascular cell adhesion molecule 1	NM_011693	3.23	4.65E-04
Aqp5	Aquaporin 5	NM_009701	3.13	4.00E-07
Ank	Progressive ankylosis	NM_020332	3.03	1.00E-07
Idb4	Inhibitor of DNA binding 4	AK041164	3.03	7.50E-06
Aldh3a1	Aldehyde dehydrogenase 3 family member A1	NM_007436	3.03	4.91E-05
Lxn	Latexin	NM_016753	2.94	3.20E-06
D14Ertd668e	PHD finger protein 11D	NM_199015	2.70	5.70E-06
Tmem86a	Transmembrane protein 86A	NM_026436	2.70	1.77E-05
LOC100046 616		XM_001476512	2.63	2.00E-07
AI467606		NM_178901	2.63	1.28E-05
Ecm1	Extracellular matrix protein 1	NM_007899	2.63	1.61E-04
Cgn	Cingulin	XM_001001375	-2.59	1.10E-06
Tuft1	Tuftelin 1	NM_011656	-2.59	4.90E-05
Ak3	Adenylate kinase 3	NM_021299	-2.87	1.61E-04
Bgn	Biglycan	NM_007542	-2.98	3.05E-05
1810015A11 Rik	YdjC homolog	NM_026940	-3.01	5.45E-05
Ogn	Osteoglycin	NM_008760	-3.10	5.62E-05
Tpd52	Tumor protein D52	NM_009412	-3.15	5.40E-06
Chchd10	Coiled-coil-helix-coiled-coil-helix domain containing 10	NM_175329	-3.52	4.25E-04
4732462B05 Rik		AK028848	-4.13	1.85E-05
Ptx3	Pentraxin related gene	NM_008987	-4.85	3.93E-05

Supplementary Table S2. Top 40 genes differentially expressed in D2A1-m2 cells *vs.* D2A1 cells

Symbol	Gene name	Accession number	Fold change	Parametric *p*-value
Casp1	Caspase 1	NM_009807	10.99	4.30E-06
Arhgdib	Rho, GDP dissociation inhibitor beta	NM_007486	10.99	1.22E-05
Lgals7	Galectin 7	NM_008496	9.09	7.00E-07
Fgf7	Fibroblast Growth Factor 7	NM_008008	7.69	6.40E-06
Itgbl1	Integrin Subunit Beta Like 1	NM_145467	5.00	6.20E-06
Tpm2	Tropomyosin 2, beta	NM_009416	4.76	6.50E-06
Mgp	Matrix Gla protein	NM_008597	4.55	5.50E-06
Tnfrsf11b	TNF receptor superfamily member 11b	NM_008764	4.17	1.45E-05
Krt8	Keratin 8	NM_031170	3.70	5.90E-06
Vcam1	Vascular cell adhesion molecule 1	NM_011693	3.57	4.64E-05
Id4	Inhibitor of DNA binding 4	NM_031166	3.45	3.70E-06
Thbs2	Thrombospondin 2	NM_011581	3.45	1.31E-05
Pdgfrl	Platelet-derived growth factor receptor-like	NM_026840	3.03	6.40E-06
Ptrf	Caveolae associated 1	NM_008986	2.86	2.80E-06
Gadd45g	Growth arrest and DNA-damage-inducible 45 gamma	NM_011817	2.78	7.21E-05
Rab32	RAB32, member RAS oncogene family	NM_026405	2.70	8.90E-06
Lxn	Latexin	NM_016753	2.63	9.00E-07
Csrp2	Cysteine and glycine-rich protein 2	NM_007792	2.56	1.10E-06
Eno3	Enolase 3	NM_007933	2.56	3.00E-06
Sparc	Secreted acidic cysteine rich glycoprotein	NM_009242	2.56	6.02E-05
Id2	Inhibitor of DNA binding 2	NM_010496	2.56	1.15E-04
Nuak1	NUAK family, SNF1-like kinase, 1	NM_001004363	2.50	1.40E-06
Tmem86a	transmembrane protein 86A	NM_026436	2.44	3.50E-06
Dap	death-associated protein	NM_146057	2.33	4.40E-05
AA467197		NM_001004174	2.33	6.03E-05
Gpnmb	Glycoprotein (transmembrane) nmb	NM_053110	2.33	5.42E-04
AI467606		NM_178901	2.27	1.40E-06
Akr1c18	Aldo-keto reductase family 1, member C18	NM_134066	2.27	1.59E-04
Csnk	Casein kappa	NM_007786	-2.28	1.63E-04
Sox12	SRY (sex determining region Y)-box 12	NM_011438	-2.32	6.10E-06
2810003C17 Rik	Allograft inflammatory factor 1-like	NM_145144	-2.32	1.38E-05
Slpi	Secretory leukocyte peptidase inhibitor	NM_011414	-2.36	2.79E-04
Wfdc2	WAP four-disulfide core domain 2	NM_026323	-2.63	3.76E-05
1810015A11 Rik	YdjC homolog	NM_026940	-2.69	1.23E-05
Chchd10	Coiled-coil-helix-coiled-coil-helix domain containing 10	NM_175329	-2.74	3.09E-04
Bgn	Biglycan	NM_007542	-2.83	2.00E-07
Olfml2b	Olfactomedin-like 2B	NM_177068	-2.87	7.26E-05
LOC100048 733		XM_001481081	-3.07	1.50E-06
Ak3	Adenylate kinase 3	NM_021299	-3.27	1.93E-05
4732462B05 Rik		AK028848	-4.13	2.70E-06

Supplementary Table S3. Top 40 genes differentially expressed in both D2A1-m1 and D2A1-m2 *vs.* D2A1 cells

Symbol	Gene name	Accession number	Average fold change
Arhgdib	Rho, GDP dissociation inhibitor beta	NM_007486	12.13
Mgp	Matrix Gla protein	NM_008597	10.63
Fgf7	Fibroblast Growth Factor 7	NM_008008	10.10
Casp1	Caspase 1	NM_009807	7.46
Lgals7	Galectin 7	NM_008496	7.25
Itgbl1	Integrin Subunit Beta Like 1	NM_145467	5.88
Tpm2	Tropomyosin 2, beta	NM_009416	4.97
Thbs2	Thrombospondin 2	NM_011581	4.93
Tnfrsf11b	TNF receptor superfamily member 11b	NM_008764	4.89
Id4	Inhibitor of DNA binding 4	NM_031166	4.62
Pdgfrl	Platelet-derived growth factor receptor-like	NM_026840	4.54
Akr1c18	Aldo-keto reductase family 1, member C18	NM_134066	3.58
Krt8	Keratin 8	NM_031170	3.58
Vcam1	Vascular cell adhesion molecule 1	NM_011693	3.40
AA467197		NM_001004174	3.20
Ptrf	Caveolae associated 1	NM_008986	3.03
Chst1	Carbohydrate (keratan sulfate Gal-6) sulfotransferase 1	NM_023850	3.00
Lxn	Latexin	NM_016753	2.78
Aqp1	Aquaporin 1	NM_007472	2.76
Idb4	Inhibitor of DNA binding 4	AK041164	2.59
Tmem86a	Transmembrane protein 86A	NM_026436	2.59
Gadd45g	Growth arrest and DNA-damage-inducible 45 gamma	NM_011817	2.49
Ank	Progressive ankylosis	NM_020332	2.45
AI467606		NM_178901	2.44
Dap	death-associated protein	NM_146057	2.44
Aqp5	Aquaporin 5	NM_009701	2.38
Axl	AXL receptor tyrosine kinase	NM_009465	2.27
Ecm1	Extracellular matrix protein 1	NM_007899	2.25
Gli2	GLI-Kruppel family member GLI2	NM_001081125	-2.31
Tuft1	Tuftelin 1	NM_011656	-2.37
Csnk	Casein kappa	NM_007786	-2.39
Cgn	Cingulin	XM_001001375	-2.41
Ogn	Osteoglycin	NM_008760	-2.67
LOC100048733		XM_001481081	-2.81
1810015A11Rik	YdjC homolog	NM_026940	-2.85
Bgn	Biglycan	NM_007542	-2.91
Ak3	Adenylate kinase 3	NM_021299	-3.07
Chchd10	Coiled-coil-helix-coiled-coil-helix domain containing 10	NM_175329	-3.13
Ptx3	Pentraxin related gene	NM_008987	-3.46
4732462B05Rik		AK028848	-4.13

Supplementary Table S4. Top cellular and molecular functions identified by IPA

	p-value	Activation z-score	# genes
D2A1-m1 *vs.* D2A1			
Cellular Movement	3.63E-03 - 5.69E-09		79
migration of cells*	5.69E-09	2.105	67
migration of breast cancer cell lines*	4.28E-05	2.179	14
Cell Death and Survival	3.49E-03 - 7.79E-08		101
cell survival*	2.83E-03	2.012	38
Cellular Assembly and Organisation	3.19E-03 - 6.74E-07		65
organization of cytoplasm*	6.74E-07	2.355	52
organization of cytoskeleton*	8.45E-06	2.567	46
Cellular Function and Maintenance	3.77E-03 - 6.74E-07		103
Cellular Development	3.50E-03 - 1.03E-06		92
D2A1-m2 *vs.* D2A1			
Cellular Movement	2.19E-03 - 1.54E-11		82
cell movement*	1.54E-11	2.414	79
migration of cells*	4.53E-11	2.602	72
Cell Death and Survival	1.73E-03 - 1.70E-08		105
cell death of connective tissue cells*	5.10E-04	2.113	21
cell survival*	8.36E-04	2.535	40
cell viability*	1.73E-03	2.604	37
Cellular Assembly and Organisation	2.04E-03 - 4.56E-08		76
Cellular Function and Maintenance	1.93E-03 - 4.56E-08		94
Cellular Development	2.22E-03 - 7.64E-08		102
Common in D2A1-m1 and D2A1-m2 *vs.* D2A1			
Cellular Movement	5.76E-03 - 7.79E-07		41
migration of cells*	7.49E-07	2.125	21
Cell Morphology	5.62E-03 - 7.49E-05		21
Cellular Assembly and Organisation	5.62E-03 - 7.49E-05		32
Small Molecule Biochemistry	5.62E-03 - 9.36E-05		16
Cell-To-Cell Signaling and Interaction	5.62E-03 - 1.05E-04		24
binding of connective tissue cells*	1.16E-03	2.394	6
adhesion of connective tissue cells*	4.13E-03	2.198	5

*Predicted activation, z-score >2 or inactivation, z-score <-2

Supplementary Table S5. Top upstream regulators identified by IPA

	p-value of overlap	Activation z-score
D2A1-m1 *vs.* D2A1		
TP53	7.17E-13	1.962
TGFB1	2.44E-12	0.532
HRAS	8.46E-11	1.551
CTNNB1	1.20E-10	1.281
ERBB2	3.15E-10	0.665
KRAS	9.12E-10	0.561
IKBKB	2.01E-08	1.227
HTT*	**2.25E-08**	**2.128**
MYC	2.28E-08	-0.576
LH	5.19E-08	1.567
D2A1-m2 *vs.* D2A1		
TGFB1*	**8.58E-19**	**2.651**
TP53	2.95E-12	1.681
MYC	6.61E-11	0.869
TNF	3.91E-09	1.962
TGFBR2	4.28E-09	1.083
SMARCA4*	**9.13E-09**	**2.159**
CTNNB1	7.10E-08	1.610
IKBKB	8.30E-08	-0.077
HRAS	8.39E-08	0.597
KRAS	2.52E-07	0.077
Common in D2A1-m1 and D2A1-m2 *vs.* D2A1		
TGFB1	1.50E-11	1.129
KRAS	1.87E-08	0.218
IKBKB	3.43E-08	0.705
CTNNB1	5.50E-07	1.370
NFKBIA	1.02E-06	-1.021
MYC	2.58E-06	0.036
SMO	4.62E-06	
HRAS	5.24E-06	0.186
SMARCA4	7.66E-06	1.718
TP53	1.06E-05	1.176

*Predicted activation, z-score >2 or inactivation, z-score <-2