

## Title

High-dimensional mass cytometry reveals stemness state heterogeneity in pancreatic ductal adenocarcinoma

## Authors

Egle-Helene Ervin<sup>1</sup>, David Ahern<sup>2</sup>, Feng Liu<sup>1</sup>, Aniko Rendek<sup>3</sup>, Zahir Soonawalla<sup>4</sup>, Udo Oppermann<sup>1</sup>, Philippa Hulley<sup>1</sup> and Siim Pauklin<sup>1,\*</sup>

## Affiliations

<sup>1</sup> Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

<sup>2</sup> Kennedy Institute of Rheumatology, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

<sup>3</sup> Department of Histopathology, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

<sup>4</sup> Department of Hepatobiliary and Pancreatic Surgery, Oxford University Hospitals NHS, Oxford, UK

## Corresponding author

Siim Pauklin

[siim.pauklin@ndorms.ox.ac.uk](mailto:siim.pauklin@ndorms.ox.ac.uk)

Botnar Research Centre

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences

University of Oxford

Old Road, Headington, Oxford, OX3 7LD

United Kingdom

Phone: (+44) 01865226492

## Abstract

Stem-like cancer cells harbour high self-renewal capacity, exhibit enhanced tumourigenicity and have been associated with therapy resistance, metastasis and tumour relapse. Therefore, understanding the molecular features of stem-like cells is critical for targeting them effectively and improving treatment outcomes for cancer patients. Several markers have been used to isolate and study the putative stem-like cells of pancreatic ductal adenocarcinoma (PDAC), but the patterns of marker co-expression and overlap between identified individual subpopulations are yet to be comprehensively studied. Here we developed a mass cytometry antibody panel for simultaneous analysis of 33 stemness-associated markers at single-cell resolution. High-dimensional mass cytometry analysis of PDAC cell lines revealed molecularly heterogeneous stemness states and highlighted the role of genotype in determining the cell line-specific stemness signature. Stemness marker expression lie along a continuum in PDAC cell lines and patient samples indicative of stepwise phenotypic transitions. We also identified a subset of PDAC cells co-expressing high levels of Musashi-2, DCLK1 and CXCR4, and harbouring basal-like and EMT transcriptional programmes associated with highly plastic phenotype. This multiplexed analysis uncovers nuance and complexities of the stemness state in the PDAC.

## Introduction

Pancreatic cancer is the seventh leading cause of cancer death worldwide<sup>1</sup> and has the lowest survival of all common cancers with a 5-year survival rate of 7-12%<sup>2,3</sup>. Due to the increasing incidence and only an incremental improvement in survival, it is projected to become third leading cause of cancer death by 2025 in European countries<sup>1,4</sup>. The most common type of pancreatic cancer is pancreatic ductal adenocarcinoma (PDAC) which accounts for more than 90% of cases, and is associated with poor prognosis due to late diagnosis, aggressive growth and therapy resistance<sup>5,6</sup>. The molecular signature of PDAC involves near ubiquitous activating mutations in *KRAS* (approx. 90% of cases) and inactivating mutations in tumour suppressor genes *CDKN2A* (approx. 90%), *TP53* (50-75%) and *SMAD4* (approx. 55%) which are accompanied by substantial compendium of other, less frequent genomic, epigenomic and transcriptomic alterations<sup>7,8</sup>.

A number of reports have demonstrated that the key features of PDAC underlying the dismal response to treatment and enabling metastatic spread are the phenotypic heterogeneity and plasticity within the tumour cell compartment<sup>9,10</sup>. In particular, the PDAC cell ability to adopt a stem-like state, which is characterised by high self-renewal capacity and enhanced tumourigenicity, has been associated with the formation of metastasis and the tumour relapse<sup>11,12</sup>. Therefore, uncovering the molecular players facilitating the phenotype switching and governing the stem-like state is critical for improving outcomes for the patients with pancreatic cancer. Regarding the latter, numerous markers have been used to isolate and study the stem-like cells of PDAC<sup>13-27</sup>. For example, co-expression of cell surface proteins CD24, CD44 and EpCAM/ESA marks cells with enhanced tumourigenic potential<sup>13</sup>. In addition, cells positive for CD9<sup>14</sup>, CD133<sup>15</sup>, Nestin<sup>24,25</sup>, DCLK1<sup>21</sup>, ALDH<sup>28</sup> or Musashi-1/2<sup>26</sup> have also been shown to have stem cell properties. Nevertheless, due to the technical limitations, individual studies of the PDAC stem-like cells have focused only on a handful of factors and there is no overarching concept to explain the identification of many distinct stemness markers<sup>29,30</sup>. Thus, in-depth analysis of the stemness marker expression within the PDAC cell lines and tissue samples is warranted to further understanding of the stem-like state, which in turn enables specific targeting of the stem-like tumour cells and improves patient survival.

To this end, we developed and validated a 41-parameter stemness-centric suspension mass cytometry panel. We used this panel to describe the single-cell marker expression profiles of nine PDAC cell lines, four control cell lines and four PDAC patient tissue samples. The non-adherent (also known as 3D or floating sphere) culture system was used to enrich for stem-like cells and characterise the stemness phenotype. The study revealed considerable heterogeneity between PDAC cell lines and highlighted the role of genotype in determining the cell line-specific stemness signature. Importantly, analysis of the patient samples confirmed that the expression of stemness markers lies along a continuum. Collectively, our results demonstrate that there is no distinct population of universal stem-like cells and support the hypothesis of high plasticity enabling phenotypic transitions which manifest as diverse stemness spectra.

## Results

### Development of stemness-centric mass cytometry panel

Mass cytometry is an advanced molecular biology tool that has a capability to simultaneously quantify over 100 parameters at a single-cell resolution<sup>31</sup>. It has provided key insights into tumour

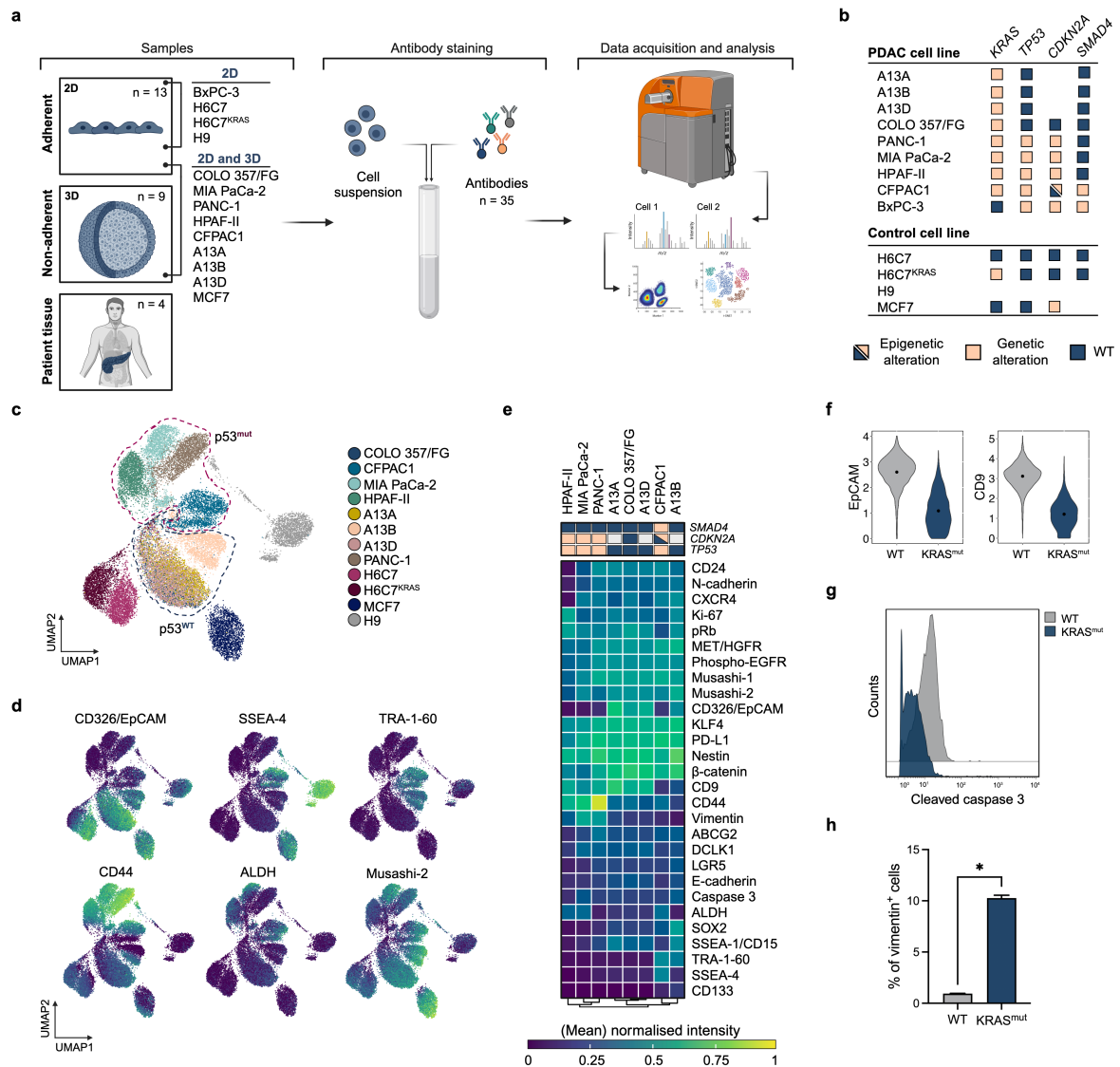


PANC-1 (ABCG<sup>pos</sup>)<sup>46</sup> cell lines (Fig. 1b). The cells harbouring high levels of phosphorylated Rb could also be distinguished from the ones where Rb is hypophosphorylated (Fig. 1b,c), and Wnt pathway intracellular signal transducer  $\beta$ -catenin could be detected in, for example, H9, MCF7 and A13A cells but not in unstimulated PBMCs (Fig. 1b and Supplementary Fig. 2a).

In contrast, antibodies detecting ROR $\gamma$  or phosphorylated SMAD2/3, respectively, were excluded from subsequent analysis due to lack of sufficient data to verify their specificity and selectivity (Fig. 1b and Supplementary Fig. 2a). The ROR $\gamma$  is a master transcriptional regulator of CD4 T cell differentiation and a marker of T helper 17 (Th17) cells<sup>47</sup>. However, distinct ROR $\gamma$ -positive Th17 cell population was not observed in the sample of CD4 T cells (Supplementary Fig. 2a). Lastly, the panel allows analysis of cells in different cell cycle phases (Fig. 1c), and discrimination between epithelial (EpCAM<sup>+</sup>, vimentin<sup>-</sup>), mesenchymal (EpCAM<sup>-</sup>, vimentin<sup>+</sup>) and intermediate EMT/MET phenotypes (Fig. 1d). Therefore, this novel, stemness-centric suspension mass cytometry panel has the potential to provide an unprecedented insight into the interplay between stemness, cell cycle, EMT and cell signalling.

### **Partly genotype-explained phenotypic differences between pancreatic cancer cell lines**

Following successful validation of the antibodies, the panel was employed to characterise the single-cell marker expression profiles of widely used PDAC cell lines (Fig. 2a). Of note, the genotypes of the selected cell lines are representative of the PDAC genomic landscape with the majority of cell lines harbouring an activating mutation in *KRAS* and five cell lines containing an additional alteration in at least one of the PDAC signature genes, namely *TP53*, *CDKN2A* and *SMAD4* (Fig. 2b and Supplementary Table 3)<sup>48</sup>. Considering the aim to describe the stemness phenotype, the culture was enriched for stem-like cancer cells using non-adherent (also known as 3D or floating sphere) culture system<sup>49–52</sup>.



**Fig. 2: Characterisation of PDAC cell lines cultured in the non-adherent conditions.** **a** Schematic of the experimental approach. **b** PDAC cell lines mutational status of *KRAS*, *TP53*, *CDKN2A* and *SMAD4*. Missing values represent genes whose mutational status has not been determined. For detailed information on the genotypes of PDAC cell lines, see Supplementary Table 3. **c** UMAP visualisation of PDAC cell line samples from non-adherent culture. Dimensionality reduction was calculated using subsampled data (2,200 cells/sample) and most variable stemness and EMT markers (see Methods). Each dot represents a cell and is coloured by the cell line. **d** UMAP visualisation as in **c** coloured by normalised expression of indicated markers. **e** Heatmap of mean normalised expression of indicated markers in PDAC cell lines with distinct genetic profiles. **f** Distribution of CD326/EpCAM and CD9 expression in samples of wild-type (WT) or KRAS mutant (KRAS<sup>mut</sup>) H6C7 cell line. Point represents the mean intensity. **g** Histograms of cleaved caspase 3 expression in the samples of wild-type (WT) or KRAS mutant (KRAS<sup>mut</sup>) H6C7 cell line. **h** Percentage of vimentin-positive cells in the samples of wild-type (WT) or KRAS mutant (KRAS<sup>mut</sup>) H6C7 cell line. Data are shown as mean ± SD (n = 2). Two-tailed Student's t-test, \* p-value < 0.05, t-score = 37.32, df = 1. **c-d, f-g** Live, single cells are shown. WT, wild-type.

In total, more than  $0.5 \times 10^6$  PDAC cells grown in the non-adherent conditions were analysed per biological replicate (Supplementary Table 5). In addition to the PDAC cells, four control cell lines cultured in adherent conditions were studied (Fig. 2a,b). The control cell lines included an

immortalised human pancreatic duct epithelial cell line H6C7 (also known as HPDE6c7), H6C7 cell line harbouring oncogenic mutation in *KRAS* (H6C7<sup>KRAS</sup>), human breast cancer cell line MCF7 and human embryonic stem cell line H9. To reveal structure in the data and investigate the role of genetic background in shaping cell phenotype, uniform manifold approximation and projection (UMAP) non-linear dimensionality reduction algorithm was used<sup>53</sup>. The analysis showed that based on the expression of the most variable stemness and EMT markers (Methods), the cells of a single cell line cluster together and separately from the cells of other cell lines (Fig. 2c and Supplementary Fig. 3a).

Therefore, there is significant phenotypic heterogeneity between PDAC cell lines. For instance, HPAF-II, MIA PaCa-2 and PANC-1 cells exhibit lower CD326/EpCAM but higher CD44 levels than the rest of the cells (Fig. 2d,e). While CFPAC1 cells are also CD326/EpCAM<sup>me/lo</sup>, their distinguishing feature is high expression of SSEA-4 and TRA-1-60. Another TRA-1-60<sup>hi</sup> cell line is A13B which is characterised by enriched expression of SOX2, and like COLO 357/FG, A13A and A13D displays high levels of CD326/EpCAM and medium or low expression of CD24 (Fig. 2d,e and Supplementary Fig. 3b). These phenotypic differences appear to be at least in part associated with the genetic features as the cells carrying a mutant p53 map to a UMAP space distinct from the areas occupied by the cells with a wild-type protein (Fig. 2c). The hierarchical clustering also confirmed that samples harbouring a mutation in *TP53* are more similar to each other than to wild-type (Fig. 2e). Nevertheless, an exception is CFPAC1 which in addition to *TP53* contains a methylated *CDKN2A*, and is the only cell line of the panel with both *KRAS* and *SMAD4* mutations<sup>54</sup>. CFPAC1 groups with A13B due to low CD9 and relatively high TRA-1-60 levels.

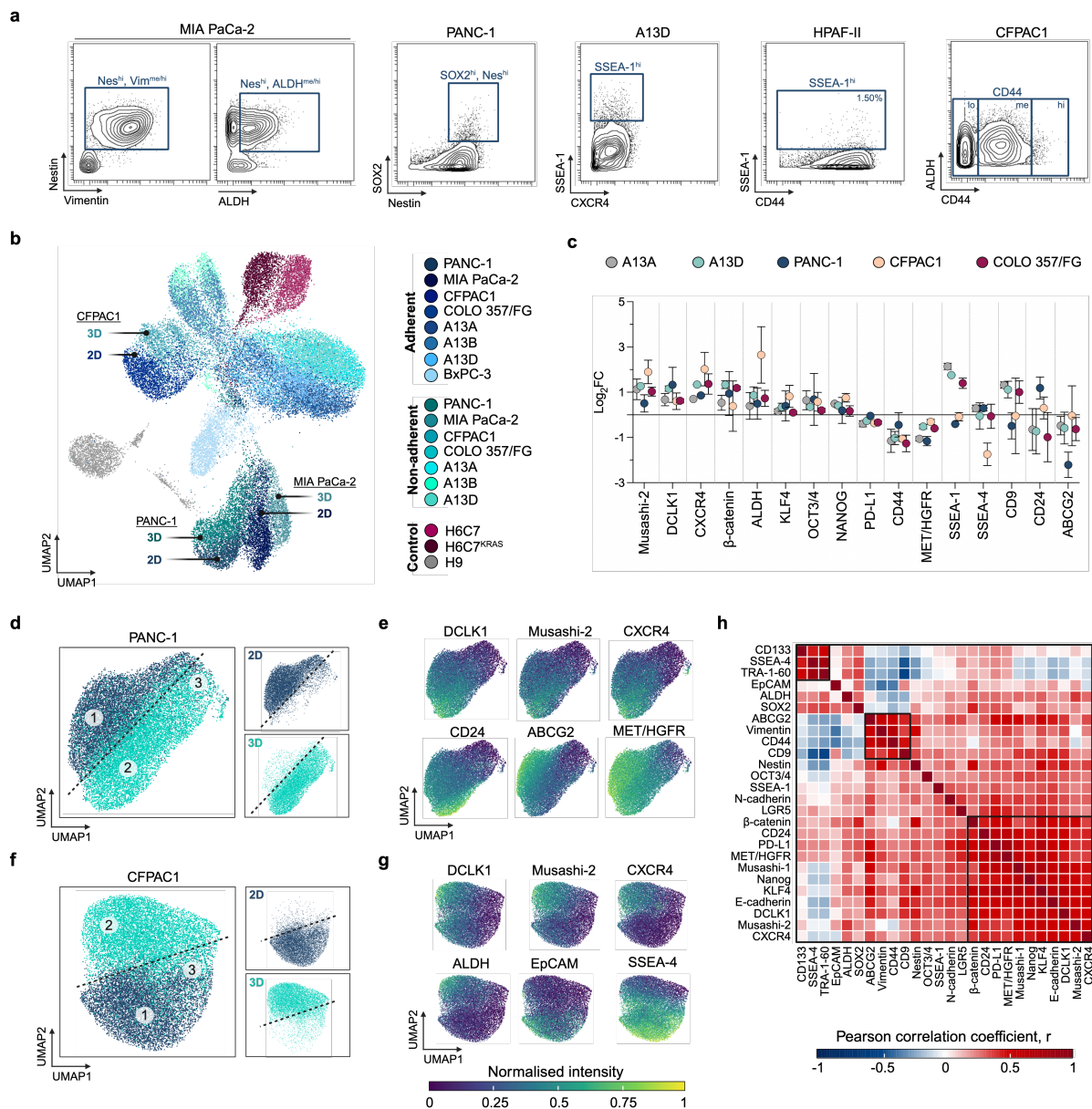
Given the clear pattern of mutational profile modulating the cell phenotype, and the previous evidence highlighting the central role of *KRAS* activation in the PDAC tumour initiation and progression, and thus, in acquisition of a highly tumourigenic state<sup>55</sup>, we investigated whether oncogenic *KRAS* has an impact on the expression of stemness markers. To this end, H6C7 and an isogenic cell line expressing *KRAS*<sup>G12V</sup> (H6C7<sup>KRAS</sup>) were used<sup>56</sup>. On the UMAP plot, H6C7 and H6C7<sup>KRAS</sup> cells form separate but neighbouring clusters (Fig. 2c). Analysis of the individual markers revealed that H6C7 cells harbouring a mutant *KRAS* express significantly lower levels of CD326/EpCAM and CD9 (Fig. 2f), and exhibit a reduced number of cells positive for cleaved caspase 3 (Fig. 2g and Supplementary Fig. 4). In contrast, the percentage of cells expressing elevated levels of vimentin is higher in H6C7<sup>KRAS</sup> sample than in control (Fig. 2h and Supplementary Fig. 4). These results are in line with previous studies reporting the involvement of *KRAS* in supporting cell survival and facilitating acquisition of a mesenchymal-like phenotype<sup>57,58</sup>. Notwithstanding, the link between the oncogenic activation of *KRAS* and regulation of the CD9 expression is yet to be established. Of note, six out of eight PDAC cell lines display CD9 levels comparable to the wild-type H6C7 despite activating mutation in *KRAS* (Supplementary Fig. 3b).

In summary, the marker expression profiles of the commonly used PDAC cell lines vary greatly and the mutational background has a profound impact on a cell phenotype.

## Heterogeneous stemness marker expression within pancreatic cancer cell lines

In addition to the differences between cell lines, there is also heterogeneity within individual cell lines. The heterogeneity exists in a form of distinct subpopulations or a low-to-high continuum.

As an example of the former, MIA PaCa-2 sample harbours subsets of Nestin<sup>hi</sup>, vimentin<sup>me/hi</sup> and Nestin<sup>hi</sup>, ALDH<sup>me/hi</sup> cells (Fig. 3a), respectively, while PANC-1 sample contains a subpopulation of SOX2<sup>hi</sup>, Nestin<sup>hi</sup> cells. Within A13D and HPAF-II samples, a small group of cells exhibiting elevated levels of SSEA-1 was detected. CFPAC1 cells can be divided into three subgroups based on the expression of CD44. Additionally, the data revealed that the cell line A13B displays greater intrasample diversity than A13A and A13D, and contains a small number of cells with phenotypes that constitute a dominant population in A13A and A13D samples (Supplementary Fig. 3a). Importantly, the A13B, A13A and A13D cell lines are derived from a primary tumour, local metastasis and distant metastasis, respectively, of the same PDAC patient<sup>59,60</sup>. This observation is concordant with the previous report noting greater heterogeneity within A13B and highlighting the selection events occurring during metastatic cascade<sup>60</sup>.



**Fig. 3: Heterogeneity within and stemness signatures of PDAC cell lines.** **a** Two-dimensional contour plots of indicated marker expression in the samples of MIA PaCa-2 (3,850 cells), PANC-1 (10,273 cells), A13D (10,183 cells), HPAF-II (10,215 cells) and CFPAC1 (9,307 cells) cell lines cultured in the non-adherent conditions. Rectangles represent examples of the cell subsets. **b** UMAP visualisation of PDAC



and control cell line samples from adherent and non-adherent culture. Dimensionality reduction was calculated using subsampled data (2,200 cells/sample). **c** Relative mean expression levels of indicated markers in non-adherent samples of PDAC cell lines compared to the respective adherent samples. Data are shown as mean  $\pm$  SD ( $n = 2$ ). **d** UMAP visualisation of samples of PANC-1 cell line from adherent (2D) and non-adherent (3D) culture. Dimensionality reduction was calculated using subsampled data (9,000 cells/sample). Areas 1-3 exhibit distinct proportions of cells from 2D and 3D culture. **e** UMAP visualisation as in **d** coloured by normalised expression of indicated markers. **f** UMAP visualisation of samples of CFPAC1 cell line from adherent (2D) and non-adherent (3D) culture. Dimensionality reduction was calculated using subsampled data (9,000 cells/sample). Areas 1-3 exhibit distinct proportions of cells from 2D and 3D culture. **g** UMAP visualisation as in **f** coloured by normalised expression of indicated markers. **h** Heatmap of the correlation coefficients for the indicated markers. Correlation analysis was performed on the non-adherent samples of PANC-1, CFPAC1 and COLO 357/FG cell lines (2,220 cells/sample). **a, b, d-g** Live, single cells are shown. **b, d-g** Dimensionality reduction was calculated using most variable stemness and EMT markers (see Methods). Each dot represents a cell and is coloured by the sample (**b, d, f**). FC, fold change.

Notwithstanding the above examples, the expression pattern of the majority of stemness markers forms a spectrum which is exemplified by CD44 expression in HPAF-II and Nestin expression in PANC-1 (Fig. 3a). Likewise, cells with low, intermediate and high levels of Musashi-2, CXCR4, CD24 or CD9 co-exist in individual samples of, for example, COLO-357/FG, PANC-1 and MIA PaCa-2 (Fig. 2d and Supplementary Fig. 3b). These continuous expression values agree with the recent notion that stemness is better described as a dynamic spectrum than as a binary feature<sup>11,61</sup>. Therefore, 1) despite predominantly mapping to a single UMAP area, cells at the opposite ends of a cell line cluster can be significantly dissimilar, and 2) the cells with low and high stemness can be distinguished based on the expression levels of pre-defined markers. The latter is different from a classical stem cell hypothesis stating that the stem cell can be identified solely by the expression (not expression levels) of specific markers that are absent in the non-stem cells<sup>61</sup>.

Regarding the markers characterising the stem-like cells, the fact that there is no single space on the UMAP plot where every cell line is represented suggests that universal, molecularly identical stemness<sup>hi</sup> phenotype shared by all cell lines does not exist (Supplementary Fig. 3a). Given the above findings, we next sought to describe and compare the stemness signatures of individual PDAC cell lines. Stemness signature is defined as a set of markers that are differentially expressed between stemness<sup>lo</sup> and stemness<sup>hi</sup> cells.

## Cell line-specific stemness phenotypes

To delineate the stemness signatures, cells cultured in the non-adherent (3D) system were compared to the respective cells grown in the adherent (2D) conditions (Fig. 2a). The non-adherent culture enriches for stem-like cells, whereas only a small fraction of cells in the adherent conditions are expected to be stem-like<sup>49-52</sup>.

First, UMAP visualisation of all samples demonstrated that neither adherent nor non-adherent samples form a distinct, spatially separated group (Fig. 3b), thereby further supporting the notion that the stemness<sup>hi</sup> phenotype is cell line-specific. Second, there are phenotypic differences between cells of the same cell line cultured in different (2D versus 3D) conditions as they occupy neighbouring and only minimally overlapping spaces (Fig. 3b). Next, to pinpoint which markers are differentially expressed in the non-adherent sample compared to the respective adherent sample, the mean

expression levels were studied. This analysis highlighted Musashi-2, DCLK1 and CXCR4 as markers which are upregulated in all tested non-adherent samples (Fig. 3c). In contrast, CD44 and MET/HGFR exhibit lower expression in non-adherent samples, while effect on the expression of  $\beta$ -catenin, ALDH, SSEA-1, SSEA-4, CD9, CD24 and ABCG2 is cell line-dependent, and levels of KLF4, OCT3/4, NANOG and PD-L1 remain unchanged or change negligibly. Therefore, not all markers previously associated with stemness become upregulated in the non-adherent culture.

We then focused on the individual cell lines at a single-cell resolution to uncover population dynamics and describe the phenotypes enriched in the non-adherent sample. In all three cell lines that were studied in detail, the adherent and non-adherent cells formed a single but clearly polarised cloud where three regions could be identified based on the abundance of cells from individual samples (Fig. 3d, f and Supplementary Fig. 5c). The spatial regions were the following: 1) dominated by the adherent cells, 2) dominated by the non-adherent cells, and 3) containing both adherent and non-adherent cells. This implies that the area 2 is more stem-like than the area 1, since phenotypes mapping to the area 2 are enriched in the non-adherent sample.

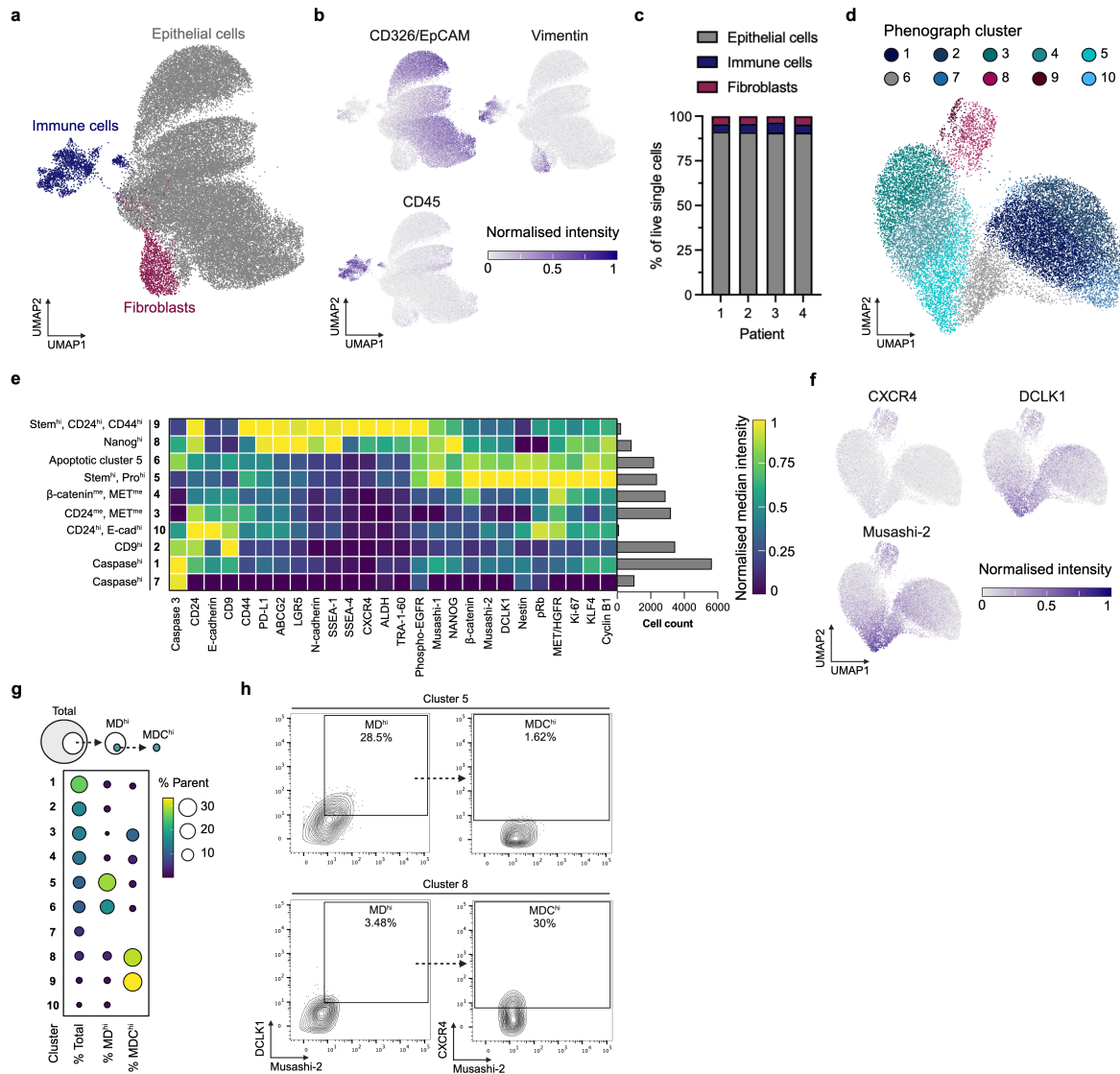
In PANC-1 cell line, areas occupied by non-adherent cells are characterised by the high expression of Musashi-2, DCLK1, CXCR4 and CD24, and low or lower levels of ABCG2 and MET/HGFR (Fig. 3e and Supplementary Fig. 6). In CFPAC1 cell line, non-adherent cells exhibit elevated levels of Musashi-1, Musashi-2, DCLK1, CXCR4 and ALDH, but express reduced levels of CD326/EpCAM, SSEA-4 and TRA-1-60 (Fig. 3g and Supplementary Fig. 6). Non-adherent cells of COLO 357/FG cell line demonstrate enhanced expression of Musashi-2 and CXCR4, relatively higher levels of DCLK1, and slightly lower levels of CD44 and ABCG2 than respective adherent cells (Supplementary Fig. 5d and 6). These patterns in the stemness marker expression suggest that while diverse stem-like phenotypes exist, there is a core signature of enhanced levels of Musashi-2, CXCR4 and DCLK1 that is shared between cell lines (Fig. 3c,e,g and Supplementary Fig. 5d). The accessory factors accompanying the core stemness proteins are cell line-specific.

Given the coordinated upregulation of Musashi-2, CXCR4 and DCLK1 in the non-adherent samples, we next asked whether they are co-expressed at a single-cell level. Correlation analysis showed that, among other relationships, there is a moderate positive correlation (Pearson correlation coefficient  $r$  of 0.5-0.59) between the expression levels of Musashi-2, CXCR4 and DCLK1 (Fig. 3h). Of note, a strong positive correlation was also observed between the expression of CD44 and vimentin ( $r = 0.68$ ), and TRA-1-60 and SSEA-4 ( $r = 0.81$ ). The latter form a co-expression cluster with CD133. In contrast, there is an inverse correlation between CD9 and TRA-1-60 ( $r = -0.52$ ), and CD9 and SSEA-4 ( $r = -0.44$ ). Analysis of the individual cell lines confirmed that the markers constituting stemness signature of the respective cell line are co-expressed (Supplementary Fig. 7). For example, Musashi-2, CXCR4, DCLK1 and CD24 group together in PANC-1 sample and there is moderate positive correlation between their expression ( $0.4 < r \leq 0.6$ ). In CFPAC1 sample, signal intensities corresponding to Musashi-1, Musashi-2, DCLK1, CXCR4 and ALDH exhibited positive correlation, while weak negative correlation was detected between the expression levels of ALDH and SSEA-4. This suggests that within an individual cell line is a single subset of cells expressing high levels of all stemness signature factors, as opposed to a model whereby individual subpopulations exist each showing enhanced levels of one or two factors.

## Characterisation of stemness marker expression in primary PDAC samples

To characterise the stemness marker expression in primary tumour tissue and investigate whether a subpopulation exhibiting elevated levels of core stemness factors (Musashi-2, CXCR4 and DCLK1) is also present in clinical samples, fresh or frozen surgical tissue from the primary tumours of four pancreatic cancer patients was analysed. In total, 495,815 live single cells were recorded with 38,803-200,685 cells acquired per patient.

First, UMAP dimensionality reduction was run to identify the main cell types and reveal a tumour cell cluster. Based on the canonical marker expression, the epithelial cells (CD326/EpCAM<sup>+</sup>), fibroblasts (vimentin<sup>+</sup>, CD326/EpCAM<sup>-</sup> and CD45<sup>-</sup>) and immune cells (CD45<sup>+</sup>) were identified within every patient sample with the epithelial cells being the most abundant cell type (Fig. 4a-c and Supplementary Fig. 8a). Second, we focused on the epithelial cells that were grouped into 10 clusters based on the expression of selected stemness, proliferation, EMT, signalling and other (cleaved caspase 3 and PD-L1) markers using unsupervised clustering via Phenograph (Fig. 4d,e and Supplementary Fig. 8b). Cluster 1 contained the highest number of cells and along with cluster 7 was characterised by elevated levels of cleaved caspase 3 but relatively low expression of other antigens. In contrast, clusters 9 and 10 were the smallest and had a distinguishing feature of elevated expression of CD24, CD44 and a set of stemness markers (cluster 9) or high CD24 and E-cadherin levels (cluster 10). Clusters 2 and 8 exhibited enhanced levels of CD9 and NANOG, respectively, while clusters 3 and 4 displayed intermediate expression of MET/HGFR coupled with intermediate levels of CD24 or  $\beta$ -catenin, respectively. Clusters 5 and 6 were among the most proliferative and showed highly similar marker expression profiles of (relatively) high/higher levels of Musashi-1, Musashi-2,  $\beta$ -catenin, DCLK1, Nestin, MET/HGFR and KLF4. The stemness marker expression was slightly lower in the cluster 6 which was also characterised by elevated levels of cleaved caspase 3. Therefore, the stemness marker expression within the epithelial compartment of patient tumours is heterogenous and several subpopulations of varying sizes exhibiting distinct signatures could be detected.



**Fig. 4: Mass cytometry analysis of primary PDAC samples.** **a** UMAP visualisation of all cells from four primary PDAC samples coloured by cell type. Dimensionality reduction was calculated using subsampled data (10,000 cells/patient). **b** UMAP visualisation as in **a** coloured by normalised expression of indicated cell type-specific markers. **c** Frequencies of epithelial cells (CD326/EpCAM<sup>+</sup>), immune cells (CD45<sup>+</sup>) and fibroblasts (vimentin<sup>+</sup>, CD326/EpCAM<sup>-</sup> and CD45<sup>-</sup>) in individual primary PDAC samples. **d** UMAP visualisation of epithelial cells from four primary PDAC samples coloured by Phenograph cluster. Dimensionality reduction was calculated using subsampled data (5,500 cells/patient). **e** Normalised median expression of indicated markers in the Phenograph clusters of epithelial cells. Cell count represents the number of cells assigned to respective Phenograph cluster. **f** UMAP visualisation as in **d** coloured by normalised expression of indicated markers. **g** Frequencies of MD<sup>hi</sup> (Musashi-2<sup>hi</sup> and DCLK1<sup>hi</sup>) and MDC<sup>hi</sup> (Musashi-2<sup>hi</sup>, DCLK1<sup>hi</sup> and CXCR4<sup>hi</sup>) cells in Phenograph clusters obtained through sequential gating. Column 1 – proportion of epithelial cells assigned to respective Phenograph cluster, Column 2 – proportion of MD<sup>hi</sup> cells in respective Phenograph cluster, Column 3 – proportion of MDC<sup>hi</sup> cells within MD<sup>hi</sup> population of respective Phenograph cluster. **h** Two-dimensional contour plots illustrating sequential gating of MD<sup>hi</sup> and MDC<sup>hi</sup> subpopulations in the Phenograph clusters 5 and 8 based on the indicated marker expression. **a-b, d, f** Dimensionality reduction was calculated using most variable markers (see Methods). Each dot represents a cell. **a-b, d, f, h** Live, single cells are shown.

Third, we aimed to delineate the expression patterns of core stemness factors (Musashi-2, CXCR4 and DCLK1). The unsupervised clustering revealed that the subpopulation exhibiting the highest mean expression of CXCR4 is cluster 9 (stem<sup>hi</sup>, CD24<sup>hi</sup>, CD44<sup>hi</sup>), while the highest Musashi-2 and DCLK1 levels map to cluster 5 (stem<sup>hi</sup>, proliferation<sup>hi</sup>) (Fig. 4e). This was confirmed by UMAP plot where signal intensities corresponding to Musashi-2 and DCLK1 showed significantly similar patterns, and the small number of CXCR4<sup>hi</sup> cells belonged predominantly to clusters 8 and 9 (Fig. 4f). Importantly, although the cells displaying the highest levels of Musashi-2 and DCLK1 are part of the clusters 5 and 6, a subpopulation of cells in clusters 8 and 9 are also positive for Musashi-2 and DCLK1. To establish whether a triple-high (Musashi-2<sup>hi</sup>, DCLK1<sup>hi</sup> and CXCR4<sup>hi</sup>) subset exists specifically within clusters 8 and 9, sequential gating was used.

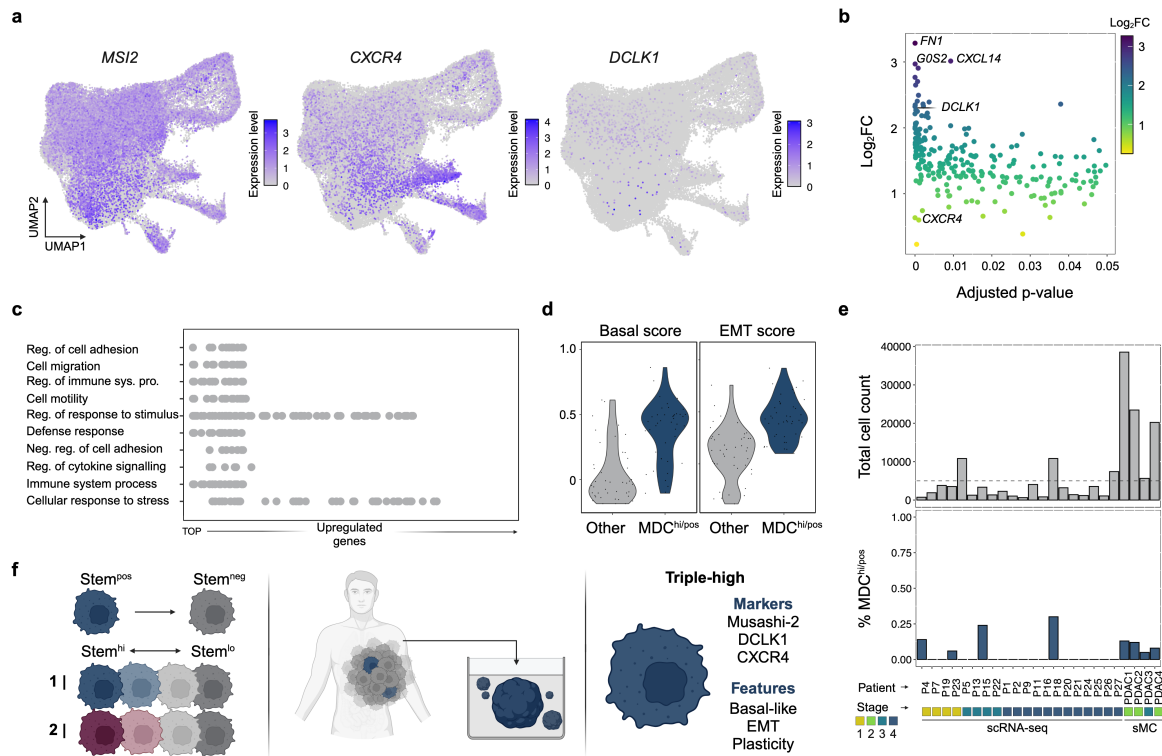
To this end, the gates were set based on cluster 3, a cluster exhibiting the lowest median expression of Musashi-2, DCLK1 and CXCR4 (Fig. 4e), so that less than 1.5% of cells were marker<sup>hi</sup> (Supplementary Fig. 9a). The analysis corroborated the findings that the highest proportion of double-high (Musashi-2<sup>hi</sup> and DCLK1<sup>hi</sup>) cells is present in clusters 5 and 6 (Fig. 4g), and clusters 8 and 9 harbour the greatest proportion of triple-high (Musashi-2<sup>hi</sup>, DCLK1<sup>hi</sup> and CXCR4<sup>hi</sup>) cells. However, the abundance of the latter is low (0.11% of total epithelial cells (Supplementary Fig. 9b)).

In sum, while the triple-high subset could be detected in both established cell lines and primary tumour tissue, the elevated levels of Musashi-2 and DCLK1 tend to be less commonly accompanied by high CXCR4 expression in the clinical samples than in the cell lines. This is indicative of in vitro culture system specifically selecting for triple-high cells or promoting activation of molecular networks driving co-expression of Musashi-2, DCLK1 and CXCR4. Nevertheless, only four clinical samples of early-stage, resectable disease were analysed, thus warranting further studies to capture the whole spectrum of PDAC, and determine which disease stages and subtypes the cell line models represent.

### **Transcriptional programmes and clinical relevance of triple-high cells**

The enrichment of Musashi-2<sup>hi</sup>, DCLK1<sup>hi</sup> and CXCR4<sup>hi</sup> cells in non-adherent culture system suggests these cells harbour molecular programmes critical for stemness properties. To delineate molecular networks underlying their stem-like phenotype, previously published single-cell RNA sequencing (scRNA-seq) data set was used<sup>62</sup>.

The scRNA-seq data set comprised 64,231 epithelial cells (identified based on the canonical marker signature (Supplementary Fig. 10a,b)) from 20 patients. The transcripts of *MSI2*, *DCLK1* and *CXCR4* could be detected (Fig. 5a), and like in the mass cytometry data, cells co-expressing detectable levels of *MSI2*, *DCLK1* and *CXCR4* are a rare subset constituting 0.1% of the epithelial cells. The transcriptomic profile of *MSI2*<sup>+</sup>, *DCLK1*<sup>+</sup> and *CXCR4*<sup>+</sup> cells was compared to the rest of cells to reveal differentially expressed genes. In total, 235 significantly upregulated transcripts were identified with the greatest increase in *FN1*, *CXCL14* and *GOS2* mRNA levels (Fig. 5b and Supplementary Fig. 10c).



**Fig. 5: Transcriptional profile of triple-high cells.** **a** UMAP visualisation of all epithelial cells from previously published scRNA-seq data set coloured by expression of indicated markers. Dimensionality reduction was calculated using most variable markers (see Methods). **b** Genes significantly (adjusted p-value < 0.05) upregulated in triple-high (*Musashi-2*<sup>hi</sup>, *DCLK1*<sup>hi</sup> and *CXCR4*<sup>hi</sup>) cells compared to the rest of epithelial cells. **c** Functional enrichment analysis of genes in **b**. Top 10 GO biological processes are shown and each dot represents a gene. **d** Basal-like and EMT expression signature scores of *MDC*<sup>hi</sup> (*Musashi-2*<sup>hi</sup>, *DCLK1*<sup>hi</sup> and *CXCR4*<sup>hi</sup>) and other epithelial cells. Subsampled data (43 cells/group) was used and each dot represents a cell. **e** Total cell count and percentage of *MDC*<sup>hi/pos</sup> (*Musashi-2*<sup>hi/pos</sup>, *DCLK1*<sup>hi/pos</sup> and *CXCR4*<sup>hi/pos</sup>) cells in the scRNA-seq and suspension mass cytometry (sMC) samples analysed in the current study. **f** Graphical summary of the following key findings: 1) instead of a universal, molecularly distinct stem-like phenotype, there is heterogeneous stemness continuum; 2) rare primary tumour cellular phenotypes/states may become enriched in in vitro culture, and 3) stemness<sup>hi</sup> phenotype co-expressing *Musashi-2*, *DCLK1* and *CXCR4*, and harbouring basal-like and EMT transcriptional programmes was identified. **b-e** Analysis of previously published scRNA-seq data set. FC, fold change; Neg., negative; Pro., processes; Reg., regulation; sMC, suspension mass cytometry; Sys., system

The *FN1* gene encodes fibronectin, a major extracellular matrix (ECM) component, which has recently been implicated in the PDAC tumour progression<sup>63</sup>. The mesenchymal cells enriched in the later disease stages exhibit higher expression of *FN1* and stemness-associated genes<sup>64</sup>. Additionally, an independent study found that the interaction between integrins and tumour-deposited fibronectin is involved in the niche-induced cancer cell stemness<sup>65</sup>. The *CXCL14* gene encodes a member of the CXC chemokine family whose role in the PDAC is relatively unexplored. However, Wente et al. have shown that *CXCL14* mRNA and protein levels are upregulated in the pancreatic cancer compared to the samples of normal pancreas and chronic pancreatitis. *CXCL14* is detected at the invasive front of the tumour and its elevated levels increase invasiveness of pancreatic cancer cells<sup>66</sup>. Similarly, the function

of *GOS2* gene product in the PDAC is relatively little studied, but it has been demonstrated to be upregulated together with *DCLK1* in the cells that exhibit greater metastatic capacity<sup>67</sup>.

The functional enrichment analysis confirmed that the genes showing elevated expression in *MSI2*<sup>+</sup>, *DCLK1*<sup>+</sup> and *CXCR4*<sup>+</sup> cells are associated with terms linked to the cell adhesion, migration and motility (Fig. 5c). Considering the previous reports highlighting the role of mesenchymal and basal-like expression programmes in the pancreatic cancer cell migratory capacity and metastatic propensity<sup>68,69</sup>, the EMT and basal signature scores were calculated for the individual cells. The *MSI2*<sup>+</sup>, *DCLK1*<sup>+</sup> and *CXCR4*<sup>+</sup> cells displayed significantly higher expression of EMT and basal signature genes than the rest of the tumour cells (Fig. 5d). Lastly, basal-like PDAC subtype has worse prognosis (Virtual microdissection), but no conclusion on the association between tumour grade and abundance of triple-positive cells could be drawn in the current study due to insufficient sampling (Fig. 5e). To illustrate, 50,000 cells need to be measured to capture 50 cells from a subset constituting 0.1% of the total population.

In conclusion, *MSI2*<sup>+</sup>, *DCLK1*<sup>+</sup> and *CXCR4*<sup>+</sup> cells exhibit an increase in the expression of genes implicated in the cell migration and harbour basal transcriptional programme. These molecular features suggest triple-high cells may represent a highly plastic and aggressive subset of PDAC stem-like cells (Fig. 5f)<sup>70</sup>.

## Discussion

Stemness state is an area of great interest in the field of tumour biology due to its involvement in the tumour development, metastatic spread, therapy resistance and relapse<sup>11,71</sup>. Several recent reports have demonstrated that, like the epithelial stem cell heterogeneity and plasticity in the non-cancerous tissues<sup>72</sup>, the population of cancer stem-like cells is heterogeneous and dynamic<sup>73–84</sup>. Here we show that the stemness state heterogeneity and plasticity feature also in the pancreatic cancer as several molecularly non-identical stemness<sup>hi</sup> states/phenotypes could be detected across PDAC cell lines. The single-cell expression levels of stemness-associated markers populate a continuum in both PDAC cell lines and primary PDAC samples, indicative of stepwise and potentially bidirectional cell state transitions.

With a goal to perform an in-depth analysis of the stemness marker expression within the PDAC cell lines and tissue samples, we opted for suspension mass cytometry which, to our knowledge, has not been applied to study the tumour cell compartment of pancreatic cancer. Suspension mass cytometry has several advantages over the standard single-cell RNA sequencing approaches that are commonly employed for high-content characterisation of cellular identities<sup>85</sup>. For example, it reads out the levels of protein antigens which better reflect the cell functional state than the transcriptomic profile. Additionally, suspension mass cytometry has higher throughput, thereby allowing analysis of rare or low abundant phenotypes<sup>86</sup>. Accordingly, simultaneous analysis of 33 stemness-associated markers at the single-cell resolution provided a glimpse into the functional state heterogeneity of pancreatic cancer stem-like cells and revealed rare co-expression phenotype.

First, the stemness-centric mass cytometry panel along with the adherent and non-adherent culture systems enabled deeper characterisation of widely used cell line models. Our results show that there is significant phenotypic heterogeneity between PDAC cell lines which is linked to cell genetic profile and transcriptomic subtype. To illustrate, HPAF-II, MIA PaCa-2 and PANC-1 cell lines all harbour a mutation in *KRAS*, *TP53* and *CDKN2A*, and exhibit lower CD326/EpCAM but higher CD44 levels than the rest of the cells. MIA PaCa-2 and PANC-1 have been previously shown to represent quasi-mesenchymal transcriptional subtype<sup>87,88</sup>, express EMT programme and display resistance to

chemotherapeutic agents<sup>89</sup>. In contrast, a more epithelial cell line COLO 357/FG is lacking a mutation in *TP53* and *CDKN2A*, shows high levels of CD326/EpCAM and relatively low expression of CD24, and is sensitive to chemotherapeutic agents<sup>89</sup>. These findings demonstrate that, expectedly, the molecular differences translate into divergent cell behaviour and highlight the importance of comprehensive characterisation of PDAC cell lines. The latter enables informed and consistent choice of *in vitro* models, thereby helping to unravel the PDAC biology. The single-cell proteomic data presented here complement previous reports describing primarily bulk monolayer cultures<sup>89–92</sup> and significantly advances our understanding of PDAC cell line characteristics.

Second, while the impact of genotype on the cell phenotype is widely accepted, the role of cancer cell mutational background in modulating its stemness state/phenotype is often overlooked. High-dimensional mass cytometry analysis of cultures enriched for stem-like cells showed that stemness<sup>hi</sup> state/phenotype is cell line-specific and no universal, molecularly identical stem-like state/phenotype exists. Similarly, the stemness signatures (set of markers that are differentially expressed between stemness<sup>lo</sup> and stemness<sup>hi</sup> cells) of the individual cell lines show incomplete overlap. Of note, Gil Vazquez et al. have found similar impact of tumour genotype and subtype on the stem phenotype in colorectal neoplasia<sup>93</sup>. This newly appreciated heterogeneity has a significant impact on our understanding of the stem-like cancer cell biology and their therapeutic targeting. The fact that stem-like features (i.e., self-renewal) are not necessarily associated with a fixed set of previously identified stemness markers implies intricate interplay between cell genetic profile, epigenetic landscape, signalling environment, molecular networks and self-renewal programme in determining the nuanced stem-like state. Thereby, strategies to disrupt the stemness networks have to reflect the complexity and heterogeneity of their target.

Third, although the stem-like cells of individual cell lines are molecularly heterogeneous at the global level, their shared feature is the upregulation of Musashi-2, CXCR4 and DCLK1, which we collectively term core stemness proteins. Musashi-2 is an RNA-binding protein that has been shown to promote pancreatic cancer development, progression and metastasis<sup>41,94–96</sup>. It contributes to drug resistance through a p53-dependent mechanism<sup>97</sup> and has been used as a reporter to identify the transcriptional programmes maintaining the stem-like state<sup>27</sup>. Musashi-2 expression is associated with poorer differentiation<sup>96</sup> and it mediates activation of EMT via ZEB1-ERK/MAPK signalling<sup>98</sup>. CXCR4 is a chemokine receptor whose best studied ligand is CXCL12 (also known as stromal derived factor-1, SDF-1)<sup>99</sup>. The CXCL12-CXCR4 axis rescues pancreatic cancer cells from gemcitabine-induced cytotoxicity by stimulating FAK, ERK and AKT signalling pathways, and promoting transcriptional activity of  $\beta$ -catenin and NF- $\kappa$ B<sup>100</sup>. CXCR4 expression is associated with PDAC progression<sup>101</sup> and malignant features (e.g., desmoplastic reaction and migration)<sup>102,103</sup>. Importantly, CXCR4 has been shown to mark a subpopulation of stem-like pancreatic cancer cells exhibiting enhanced invasive and metastatic properties<sup>15</sup>. DCLK1 is a bifunctional protein with kinase and microtubule-associated protein activity that identifies a rare population of long-lived, quiescent pancreatic progenitors facilitating regeneration and underlying tumour initiation<sup>104,105</sup>. DCLK1-positive cells detected in the preinvasive lesions and pancreatic cancer display features of gastrointestinal tuft cells, express genes linked to the cancer stem cells and harbour an active EMT programme<sup>21,43,106,107</sup>. Accordingly, DCLK1 is a putative marker of pancreatic cancer stem-like cells that is essential for their invasive and metastatic property<sup>108</sup>. It is also highly expressed by circulating tumour cells and metastatic tumours, and its inhibition reverses EMT programme and restores T cell activity<sup>43,109,110</sup>.

Given the previously described functions of the individual core stemness proteins, it is unsurprising that the cells co-expressing Musashi-2, CXCR4 and DCLK1 (triple-high/positive phenotype) are



transcriptionally characterised by elevated EMT and basal scores. EMT is a dynamic and reversible process that has been associated with metastasis, tumour-initiating potential, stemness and therapy resistance<sup>111</sup>. Basal cells have recently been demonstrated to represent a highly plastic state with the capacity to facilitate state transitions and promote intratumoural heterogeneity in PDAC<sup>112</sup>. These findings are in accordance with the notion that the stemness, EMT, plasticity, metastasis and therapy resistance are all closely interlinked<sup>111,113</sup>. The identified triple-high/positive stem-like cells may be a highly plastic tumour cell phenotype implicated in the metastatic spread and adaptive therapy resistance of the PDAC. Our data support the plasticity and state transitions surrounding the stem-like cells which are expected to result in the continuum of marker expression and spectrum of intermediate cell states<sup>112,114</sup>.

Future research will help to address the limitations of the current study and answer the outstanding questions. For example, (spatial) analysis of a larger number of patient samples is critical for confirming stemness state heterogeneity, understanding the clinical significance of the triple-high cells and shedding light on the stem-like cell niche/cellular neighbourhood. Mechanistic studies elucidating the origin, drivers, dynamics and functional differences of the heterogeneous stem-like cell subpopulations, and the role of triple-high cells in the tumour development, progression, metastasis and therapy resistance are also crucial next steps.

In conclusion, high-dimensional mass cytometry analysis reveals stemness state heterogeneity and plasticity in the PDAC. Our results add to the accumulating evidence describing stemness as a dynamic spectrum, not as a fixed binary feature. This paradigm shift highlights the need to develop new therapies that simultaneously target the individual cancer cell states and their plasticity machinery.

## Methods

### Cell lines

A13A, A13B and A13D cell lines were kindly provided by Prof. Christine Iacobuzio-Donahue (Memorial Sloan Kettering Cancer Center). COLO 357/FG cell line was obtained from MD Anderson Cancer Center Characterized Cell Line Core Facility. PANC-1, HPAF-II, CFPAC1 and MCF7 cell lines were kindly provided by Prof. Catherine Hogan (Cardiff University). MIA PaCa-2 and BxPC-3 cell lines were kindly provided by Prof. Eric O'Neill (University of Oxford). H6C7 cell line was obtained from Kerastat and H6C7<sup>KRAS</sup> cell line was kindly provided by Prof. Ming Tsao (University of Toronto). NIH3T3 cell line, PBMCs and CD4 T cells were kindly provided by Prof. Udo Oppermann, Prof. Claudia Monaco and Dr Liye Chen (all University of Oxford), respectively. H9 (WA09) cell line was obtained from WiCell. All cell lines were regularly tested for mycoplasma infection.

### Adherent cell culture

PDAC cell lines A13A, A13B, A13D, COLO 357/FG, PANC-1, HPAF-II, CFPAC1, breast cancer cell line MCF7 and mouse fibroblast cell line NIH3T3 were cultured in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% FBS, 1% Penicillin-Streptomycin, 1% MEM non-essential amino acids solution and 1% MEM vitamin solution (all Thermo Fisher). PDAC cell line MIA PaCa-2 was cultured in DMEM supplemented with 10% fetal bovine serum (FBS), 1% Penicillin-Streptomycin, 1% MEM non-essential amino acids solution, 1% MEM vitamin solution and 2.5% horse serum (all Thermo Fisher). PDAC cell line BxPC-3 was cultured in Roswell Park Memorial Institute (RPMI) 1640 Medium supplemented with 10% FBS (both Thermo Fisher). Human pancreatic duct epithelial cell lines H6C7 and H6C7<sup>KRAS</sup> cell lines were cultured in keratinocyte serum-free medium (SFM) supplemented with 5

ng/ml human recombinant epidermal growth factor (EGF) and 50 µg/ml bovine pituitary extract (BPE)(all Thermo Fisher). Above cells were split every 4-5 days using 5-10-minute incubation with 0.05% trypsin-EDTA (Thermo Fisher) at 37 °C. Human embryonic stem cell line H9 was cultured on the 6-well plate coated with the Vitronectin (5 µg/ml) and in the Essential 8 Flex medium (both Thermo Fisher). Cells were split every 7 days using 5-minute incubation with 0.5 mM EDTA-PBS (Thermo Fisher) at 37 °C. All cells were cultured in humidified incubator at 37 °C with 5% CO<sub>2</sub>.

### **Non-adherent cell culture**

PDAC cell lines A13A, A13B, A13D, COLO 357/FG, PANC-1, MIA PaCa-2, HPAF-II, CFPAC1 and breast cancer cell line MCF7 were cultured in non-adherent system of ultra-low attachment vessel and serum-free DMEM/Nutrient Mixture F-12 medium (Sigma-Aldrich) supplemented with 2% B-27, 1% L-glutamine, 1% Penicillin-Streptomycin (all Thermo Fisher) and bFGF (20 ng/ml, Peprotech). Spheroids were split every 5 days using 5-min incubation with TrypLE™ Express (Thermo Fisher) at 37 °C. Cells were passed through a cell strainer (40 µm, pluriSelect) followed by counting and reseeding single cells at a density of 2,000 cells/ml. Cells were cultured in humidified incubator at 37 °C with 5% CO<sub>2</sub>.

### **Patient tissue samples**

Primary PDAC tissue samples from a total of four patients were obtained from the Oxford Radcliffe Biobank. Written informed consent was obtained from all participants and study was conducted in compliance with the ethics permission (OCHRe ref: 21/A126, REC reference: 19/SC/0173; OCHRe ref: 19/A176, REC reference: 19/SC/0173). Tissue punch biopsies (size: 5 × 5 mm<sup>2</sup>) were cut immediately after resection and PDAC diagnosis in all samples was confirmed by certified pathologist. Fresh surgical tissue was stored in ice-cold DMEM (Thermo Fisher) before it was minced using scalpel, and cryopreserved in FBS supplemented with 10% dimethyl sulfoxide (DMSO) at -80 °C (samples for genotyping and mass cytometry) or processed directly using Tumor Dissociation Kit (human, Miltenyi Biotec) as described below (fresh sample for mass cytometry). To determine the genetic variants of key PDAC-associated genes (*KRAS*, *TP53*, *SMAD4* and *CDKN2A*), genomic DNA was extracted from cryopreserved patient tissue samples using GenElute Mammalian Genomic DNA Miniprep Kit (Sigma-Aldrich) according to manufacturer's instructions. Purified genomic DNA (10 ng, eluted in nuclease-free water) was prepared for next generation sequencing using Ion AmpliSeq Library Kit Plus, Ion AmpliSeq Cancer Hotspot Panel v2 and Ion Xpress Barcode Adapters according to manufacturer's instructions. Sequencing libraries were prepared with Ion 510 & Ion 520 & Ion 530 Kit – Chef using Ion Chef System, sequenced on Ion GeneStudio S5 System (all Thermo Fisher) and aligned to GRCh37 following manufacturer's recommended data processing pipeline. Customised workflow in Ion Reporter software 5.12 was used for variant calling and annotation.

### **Mass cytometry panel**

Mass cytometry panel allowed to measure 47 parameters/channels: 38 antigens (antibody targets), 6 barcodes, 1 S phase marker (IdU), 1 intercalator (Iridium) and 1 live/dead marker (Cisplatin)(Supplementary Table 1). Heavy metal-labelled antibodies were purchased from the Standard BioTools or prepared using Maxpar Antibody Labelling Kit (Standard BioTools) according to manufacturer's instructions. Validation of in-house-labelled antibodies was performed on known marker-negative and marker-positive cells (Supplementary Table 2). Optimal staining concentration was determined using 8-point (2-fold dilution series from 0 to 8 µg/ml) or 3-point (1:200, 1:100 and 1:67) titration of in-house-labelled and commercially available antibodies, respectively. If the 1:200 dilution resulted in a too high signal intensity, lower concentrations were tested. The list of antibodies,

clone information, metal isotype tags used and staining concentrations are provided in Supplementary Table 1.

### **Processing of cell line samples**

Spent culture medium was removed and cells were incubated in a serum-free medium containing Cell-ID™ 127 IdU (25 µM, Standard BioTools) for 30 min at 37 °C. Following a wash with PBS (Thermo Fisher), single-cell suspension was created by incubation in TrypLE™ Express (Thermo Fisher) supplemented with Benzonase nuclease (1:10,000, Santa Cruz Biotechnology) for 10 min at 37 °C. Single-cell suspension was washed with PBS (Thermo Fisher) and incubated in a serum-free medium containing viability reagent Cell-ID™ Cisplatin (1 µM, Standard BioTools) for 10 min at 37 °C. Cisplatin was quenched by washing the cells twice with serum-containing medium (DMEM supplemented with 10% FBS, adherent samples) or Maxpar Cell Staining Buffer (Standard BioTools, non-adherent samples). Cells were fixed in a 4% formaldehyde, methanol-free (Thermo Fisher) for 10 min at room temperature. Fixation solution was removed, cells were resuspended in cold Maxpar Cell Staining Buffer and counted. For mass cytometry staining, using  $0.5$  or  $0.75 \times 10^6$  cells per sample was preferred but lower number of cells was used if enough cells could not be obtained.

### **Processing of patient tissue samples**

Cryopreserved patient tissue samples were thawed rapidly into DMEM. Thawed and fresh samples were dissociated using the Tumor Dissociation Kit (human, Miltenyi Biotec) and tough dissociation protocol on gentleMACS Dissociator (Miltenyi Biotec) according to manufacturer's instructions. Following dissociation, samples were resuspended in DMEM (Thermo Fisher) and treated with Benzonase nuclease (1:10,000, Santa Cruz Biotechnology) for 5 min at room temperature. Benzonase-containing medium was removed by centrifugation and cells were incubated in a serum-free medium with Cell-ID™ 127 IdU (25 µM, Standard BioTools) for 20 min at 37 °C. Subsequently, viability reagent Cell-ID™ Cisplatin (1 µM, Standard BioTools) was added and cells were incubated for a further 10 min at 37 °C. Cisplatin was quenched by washing the cells once with serum-containing medium and once with Maxpar Cell Staining Buffer. Cells were fixed in a 4% formaldehyde, methanol-free (Thermo Fisher) for 10 min at room temperature. Fixation solution was removed, cells were resuspended in cold Maxpar Cell Staining Buffer and counted. For mass cytometry staining,  $0.1$ - $0.5 \times 10^6$  cells per sample were used.

### **Mass cytometry staining**

Cells were prepared as described above, washed with Maxpar Cell Staining Buffer, and stained with CXCR4 antibody for 30 min at 4 °C. Upon incubation, cells were washed with Maxpar Cell Staining Buffer and barcoded using Cell-ID™ 20-Plex Pd Barcoding Kit (Standard BioTools) according to manufacturer's instructions. Barcoded cells were combined into a composite sample before incubation in Human TruStain FcX Fc Receptor Blocking Solution (BioLegend) for 10 min at room temperature. Blocking solution was removed by centrifugation, volume of cell suspension was adjusted to achieve a concentration of  $3 \times 10^6$  cells in 100 µl and cells were stained with surface antibody cocktail for 30 min at 4 °C. After the incubation, cells were washed with Maxpar Cell Staining Buffer and chilled on ice for 10 min. For methanol fixation/permeabilisation, ice-cold 90% methanol (Thermo Fisher) was slowly added to cells and cells were incubated in methanol for 30 min on ice. Next, cells were washed twice with Maxpar Cell Staining Buffer supplemented with Benzonase nuclease (1:10,000, Santa Cruz Biotechnology). Volume of cell suspension was adjusted to achieve a concentration of  $3 \times 10^6$  cells in 100 µl and cells were stained with intracellular antibody cocktail for

30 min at room temperature. Following two washes with Maxpar Cell Staining Buffer, samples were incubated in intercalation solution (1.6% formaldehyde and 62.5 nM Cell-ID Intercalator-Ir (Standard BioTools) in Maxpar PBS (Standard BioTools)) overnight at 4 °C. On the day of acquisition, cells were washed once with Maxpar Cell Staining Buffer and once with Maxpar Water (Standard BioTools), filtered through cell strainer (35 µm, Falcon), resuspended in Maxpar Cell Acquisition Solution with 0.1X EQ beads and acquired on Helios mass cytometer (all Standard BioTools).

### Mass cytometry data analysis

Raw mass cytometry data were bead normalised and debarcoded on CyTOF Software v7.0. Population of live, single cells for further analysis was identified by manual gating on Cytobank as described in Supplementary Fig. 1. A summary of samples and corresponding number of events for all experiments are reported in Supplementary Table 5. Data were subsequently imported into R<sup>115</sup> environment and batch corrected by scaling to the 95<sup>th</sup> percentile of the anchor sample (adherent A13A) using BatchAdjust package<sup>116</sup>. Harmonised data were arcsinh transformed (cofactor 5) and downstream analysis was performed using tidytof package<sup>117</sup>. Uniform manifold approximation and projection (UMAP) dimensionality reduction was calculated on downsampled data with the following parameters: set.seed(0), neighbors = 15 and min\_dist = 0.5. The following channels, which showed the greatest variation between samples (visual inspection), were used for UMAP analysis: 1) Cell line samples: EpCAM, N-cadherin, SSEA-1, SSEA-4, ALDH, TRA-1-60, CD133, SOX2, Nestin, vimentin, Musashi-1, CXCR4, E-cadherin, DCLK1, LGR5, OCT3/4, CD24, Nanog, ABCG2, CD44, CD9, Musashi-2 and β-catenin; 2) Patient tissue samples (all cells): Ki-67, pRb, cyclin B1, caspase 3 (cleaved), ABCG2, β-catenin, CD9, CD24, CD45, CXCR4, DCLK1, Musashi-1, MET/HGFR, Nestin, EpCAM, E-cadherin, vimentin and PD-L1; 3) Patient tissue samples (epithelial cells): caspase 3 (cleaved), ABCG2, ALDH, β-catenin, CD9, CD24, CD44, CD133, CXCR4, DCLK1, KLF4, LGR5, Musashi-1, Musashi-2, MET/HGFR, Nestin, OCT3/4, SOX2, SSEA-1, SSEA-4, TRA-1-60, E-cadherin, N-cadherin, vimentin and PD-L1. Importantly, since distinct CD44<sup>+</sup> and CD44<sup>-</sup> subpopulations within adherent and non-adherent samples of CFPAC1 cell line hindered intersample comparisons, CD44 was excluded from UMAP analysis shown in Fig. 3f,g. Epithelial cells from patient tissue samples were clustered based on the expression of above markers (UMAP channels set number 3) using PhenoGraph algorithm with the following parameters: set.seed(0), num\_neighbors = 50 and distance\_function = "cosine". Pearson correlation coefficients were calculated using the R stats package<sup>115</sup>.

### Single-cell RNA-seq data analysis

Single-cell RNA-seq data on 20 treatment-naïve patients were obtained from a study reported by Werba et al. (GEO accession number GSE205013)<sup>118</sup>. Paired reads were aligned to the hg38 reference genome and gene expression matrices were generated using 10x Genomics Cell Ranger 7.0.0. Subsequent analysis on data matrices was performed using Seurat package (v4)<sup>119</sup>. Low quality cells were removed based on the following metrics: genes detected per cell (<400) and UMI counts per cell (<600). Cells with high mitochondrial counts were removed at the clustering step. Raw count data were normalised and variable genes for each sample were identified using NormalizeData and FindVariableFeatures functions, respectively. FastMNN algorithm was used for integration of individual samples. For cell type identification, dimensionality reduction was performed with RunUMAP on highly variable features followed by the clustering of cells based on their global transcriptional profiles using FindNeighbors and FindClusters functions. Cell types were annotated based on the marker genes in Supplementary Fig. 10b. For subsetting triple-positive cells, epithelial cells with feature count higher than 0 for *MSI2*, *DCLK1* and *CXCR4* were considered. Differential

expression analysis was performed with FindMarkers function and g:GOST tool on g:Profiler<sup>120</sup> web server was used for functional enrichment analysis of differentially expressed genes. Gene expression signature scores were obtained by dividing the sum of normalised and scaled counts of each feature/gene with the number of features/genes constituting the signature. Top 100 basal-like and EMT signature genes defined by Chang-Seng-Yue et al.<sup>121</sup> and Puram et al.<sup>122</sup>, respectively, were used.

### Visualisation

Plots were created using the GraphPad Prism version 10.0.2, ggplot2<sup>123</sup> and pheatmap<sup>124</sup>(both R packages). Schematic representations were created and figures were prepared with BioRender.

### Statistics

All statistical tests used in this study are described in detail in the corresponding figure legends. The experiment comparing single-cell profiles of PDAC cell lines grown in adherent versus non-adherent culture was repeated twice (n = 2).

### Data availability

The mass cytometry data are available on Zenodo. The existing, publicly available scRNA-seq data used in this paper are available on GEO with accession number GSE205013.

### Code availability

No custom code was used for processing the data presented in the manuscript.

### References

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209–249 (2021).
2. Siegel, R. L., Miller, K. D., Wagle, N. S. & Jemal, A. Cancer statistics, 2023. *CA Cancer J Clin* **73**, 17–48 (2023).
3. Baker, C. Cancer: summary of statistics. *Date accessed: 02/01/2023.* <https://commonslibrary.parliament.uk/research-briefings/sn06887>.
4. Ferlay, J., Partensky, C. & Bray, F. More deaths from pancreatic cancer than breast cancer in the EU by 2017. *Acta Oncol (Madr)* **55**, 1158–1160 (2016).
5. Sarantis, P., Koustas, E., Papadimitropoulou, A., Papavassiliou, A. G. & Karamouzis, M. V. Pancreatic ductal adenocarcinoma: Treatment hurdles, tumor microenvironment and immunotherapy. *World J Gastrointest Oncol* **12**, 173–181 (2020).
6. Yao, W., Maitra, A. & Ying, H. Recent insights into the biology of pancreatic cancer. *EBioMedicine* **53**, 102655 (2020).
7. Maitra, A. & Hruban, R. H. Pancreatic Cancer. *Annual Review of Pathology: Mechanisms of Disease* **3**, 157–188 (2008).
8. Lomberk, G., Dusetti, N., Iovanna, J. & Urrutia, R. Emerging epigenomic landscapes of pancreatic cancer in the era of precision medicine. *Nat Commun* **10**, 3875 (2019).
9. Evan, T., Wang, V. M.-Y. & Behrens, A. The roles of intratumour heterogeneity in the biology and treatment of pancreatic ductal adenocarcinoma. *Oncogene* **41**, 4686–4695 (2022).

10. Bärthel, S., Falcomatà, C., Rad, R., Theis, F. J. & Saur, D. Single-cell profiling to explore pancreatic cancer heterogeneity, plasticity and response to therapy. *Nat Cancer* **4**, 454–467 (2023).
11. Hermann, P. C. & Sainz, B. Pancreatic cancer stem cells: A state or an entity? *Semin Cancer Biol* **53**, 223–231 (2018).
12. Sancho, P., Alcalá, S., Usachov, V., Hermann, P. C. & Sainz, B. The ever-changing landscape of pancreatic cancer stem cells. *Pancreatology* **16**, 489–496 (2016).
13. Li, C. *et al.* Identification of pancreatic cancer stem cells. *Cancer Res* **67**, 1030–7 (2007).
14. Wang, V. M.-Y. *et al.* CD9 identifies pancreatic cancer stem cells and modulates glutamine metabolism to fuel tumour growth. *Nat Cell Biol* **21**, 1425–1435 (2019).
15. Hermann, P. C. *et al.* Distinct Populations of Cancer Stem Cells Determine Tumor Growth and Metastatic Activity in Human Pancreatic Cancer. *Cell Stem Cell* **1**, 313–323 (2007).
16. Amsterdam, A., Raanan, C., Schreiber, L., Polin, N. & Givol, D. LGR5 and Nanog identify stem cell signature of pancreas beta cells which initiate pancreatic cancer. *Biochem Biophys Res Commun* **433**, 157–162 (2013).
17. Li, C. *et al.* c-Met Is a Marker of Pancreatic Cancer Stem Cells and Therapeutic Target. *Gastroenterology* **141**, 2218-2227.e5 (2011).
18. Sasaki, N. *et al.* Stemness and anti-cancer drug resistance in ATP-binding cassette subfamily G member 2 highly expressed pancreatic cancer is induced in 3D culture conditions. *Cancer Sci* **109**, 1135–1146 (2018).
19. Rasheed, Z. A. *et al.* Prognostic Significance of Tumorigenic Cells With Mesenchymal Features in Pancreatic Adenocarcinoma. *JNCI: Journal of the National Cancer Institute* **102**, 340–351 (2010).
20. Kim, M. P. *et al.* ALDH Activity Selectively Defines an Enhanced Tumor-Initiating Cell Population Relative to CD133 Expression in Human Pancreatic Adenocarcinoma. *PLoS One* **6**, e20636 (2011).
21. Bailey, J. M. *et al.* DCLK1 Marks a Morphologically Distinct Subpopulation of Cells With Stem Cell Properties in Preinvasive Pancreatic Cancer. *Gastroenterology* **146**, 245–256 (2014).
22. Herreros-Villanueva, M. *et al.* SOX2 promotes dedifferentiation and imparts stem cell-like features to pancreatic cancer cells. *Oncogenesis* **2**, e61–e61 (2013).
23. Lu, Y. *et al.* Knockdown of Oct4 and Nanog expression inhibits the stemness of pancreatic cancer cells. *Cancer Lett* **340**, 113–123 (2013).
24. Matsuda, Y., Kure, S. & Ishiwata, T. Nestin and other putative cancer stem cell markers in pancreatic cancer. *Med Mol Morphol* **45**, 59–65 (2012).
25. Matsuda, Y. *et al.* Nestin Delineates Pancreatic Cancer Stem Cells in Metastatic Foci of NOD/Shi-scid IL2R $\gamma$ null (NOG) Mice. *Am J Pathol* **184**, 674–685 (2014).
26. Fox, R. G. *et al.* Image-based detection and targeting of therapy resistance in pancreatic adenocarcinoma. *Nature* **534**, 407–411 (2016).

27. Lytle, N. K. *et al.* A Multiscale Map of the Stem Cell State in Pancreatic Adenocarcinoma. *Cell* **177**, 572-586.e22 (2019).
28. Kim, M. P. *et al.* ALDH Activity Selectively Defines an Enhanced Tumor-Initiating Cell Population Relative to CD133 Expression in Human Pancreatic Adenocarcinoma. *PLoS One* **6**, e20636 (2011).
29. Jaiswal, K. R. *et al.* Comparative testing of various pancreatic cancer stem cells results in a novel class of pancreatic-cancer-initiating cells. *Stem Cell Res* **9**, 249–260 (2012).
30. Skoda, J. *et al.* Co-Expression of Cancer Stem Cell Markers Corresponds to a Pro-Tumorigenic Expression Profile in Pancreatic Adenocarcinoma. *PLoS One* **11**, e0159255 (2016).
31. Iyer, A., Hamers, A. A. J. & Pillai, A. B. CyTOF® for the Masses. *Front Immunol* **13**, (2022).
32. Brouwer, T. P. *et al.* Local and systemic immune profiles of human pancreatic ductal adenocarcinoma revealed by single-cell mass cytometry. *J Immunother Cancer* **10**, e004638 (2022).
33. Hutton, C. *et al.* Single-cell analysis defines a pancreatic fibroblast lineage that supports anti-tumor immunity. *Cancer Cell* **39**, 1227-1244.e20 (2021).
34. Yang, J. *et al.* Dynamic profiling of immune microenvironment during pancreatic cancer development suggests early intervention and combination strategy of immunotherapy. *EBioMedicine* **78**, 103958 (2022).
35. Gang, E. J., Bosnakovski, D., Figueiredo, C. A., Visser, J. W. & Perlingeiro, R. C. R. SSEA-4 identifies mesenchymal stem cells from bone marrow. *Blood* **109**, 1743–1751 (2007).
36. Wright, A. J. & Andrews, P. W. Surface marker antigens in the characterization of human embryonic stem cells. *Stem Cell Res* **3**, 3–11 (2009).
37. Zhang, P., Andrianakos, R., Yang, Y., Liu, C. & Lu, W. Kruppel-like factor 4 (Klf4) prevents embryonic stem (ES) cell differentiation by regulating Nanog gene expression. *Journal of Biological Chemistry* **285**, 9180–9189 (2010).
38. Qi, X. *et al.* KLF4 functions as an oncogene in promoting cancer stem cell-like characteristics in osteosarcoma cells. *Acta Pharmacol Sin* **40**, 546–555 (2019).
39. Domingo-Reinés, J. *et al.* Generation of a H9 Clonal Cell Line With Inducible Expression of NUP98-KDM5A Fusion Gene in the AAVS1 Safe Harbor Locus. *Front Cell Dev Biol* **10**, (2022).
40. Jaksch, M., Múnera, J., Bajpai, R., Terskikh, A. & Oshima, R. G. Cell Cycle–Dependent Variation of a CD133 Epitope in Human Embryonic Stem Cell, Colon Cancer, and Melanoma Cell Lines. *Cancer Res* **68**, 7882–7886 (2008).
41. Guo, K. *et al.* The Novel KLF4/MSI2 Signaling Pathway Regulates Growth and Metastasis of Pancreatic Cancer. *Clinical Cancer Research* **23**, 687–696 (2017).
42. Kang, M.-H. *et al.* Musashi RNA-binding protein 2 regulates estrogen receptor 1 function in breast cancer. *Oncogene* **36**, 1745–1752 (2017).
43. Ge, Y. *et al.* Inhibition of DCLK1 kinase reverses epithelial-mesenchymal transition and restores T-cell activity in pancreatic ductal adenocarcinoma. *Transl Oncol* **17**, 101317 (2022).

44. Wang, Y.-L., Li, Y., Ma, Y.-G. & Wu, W.-Y. DCLK1 promotes malignant progression of breast cancer by regulating Wnt/ $\beta$ -Catenin signaling pathway. *Eur Rev Med Pharmacol Sci* **23**, 9489–9498 (2019).
45. Zeng, H. *et al.* Lack of ABCG2 Expression and Side Population Properties in Human Pluripotent Stem Cells. *Stem Cells* **27**, 2435–2445 (2009).
46. Wang, F. *et al.* hsa-miR-520h downregulates ABCG2 in pancreatic cancer cells to inhibit migration, invasion, and side populations. *Br J Cancer* **103**, 567–574 (2010).
47. Crome, S. Q., Wang, A. Y., Kang, C. Y. & Levings, M. K. The role of retinoic acid-related orphan receptor variant 2 and IL-17 in the development and function of human CD4+ T cells. *Eur J Immunol* **39**, 1480–1493 (2009).
48. Waddell, N. *et al.* Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* **518**, 495–501 (2015).
49. Lonardo, E., Cioffi, M., Sancho, P., Crusz, S. & Heeschen, C. Studying Pancreatic Cancer Stem Cell Characteristics for Developing New Treatment Strategies. *Journal of Visualized Experiments* (2015) doi:10.3791/52801.
50. Ishiwata, T. *et al.* Electron microscopic analysis of different cell types in human pancreatic cancer spheres. *Oncol Lett* (2017) doi:10.3892/ol.2017.7554.
51. Domenichini, A. *et al.* Pancreatic cancer tumorspheres are cancer stem-like cells with increased chemoresistance and reduced metabolic potential. *Adv Biol Regul* **72**, 63–77 (2019).
52. Gaviraghi, M. *et al.* Pancreatic cancer spheres are more than just aggregates of stem marker-positive cells. *Biosci Rep* **31**, 45–55 (2011).
53. Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* **37**, 38–44 (2019).
54. Moore, P. S. *et al.* Genetic profile of 22 pancreatic carcinoma cell lines. *Virchows Archiv* **439**, 798–802 (2001).
55. Eser, S., Schnieke, A., Schneider, G. & Saur, D. Oncogenic KRAS signalling in pancreatic cancer. *Br J Cancer* **111**, 817–822 (2014).
56. Qian, J., Niu, J., Li, M., Chiao, P. J. & Tsao, M.-S. In vitro Modeling of Human Pancreatic Duct Epithelial Cell Transformation Defines Gene Expression Changes Induced by Kras Oncogenic Activation in Pancreatic Carcinogenesis. *Cancer Res* **65**, 5045–5053 (2005).
57. Singh, A. *et al.* A Gene Expression Signature Associated with “K-Ras Addiction” Reveals Regulators of EMT and Tumor Cell Survival. *Cancer Cell* **15**, 489–500 (2009).
58. Rhim, A. D. *et al.* EMT and Dissemination Precede Pancreatic Tumor Formation. *Cell* **148**, 349–361 (2012).
59. Embuscado, E. E. *et al.* Immortalizing the complexity of cancer metastasis: Genetic features of lethal metastatic pancreatic cancer obtained from rapid autopsy. *Cancer Biol Ther* **4**, 548–554 (2005).
60. Duesberg, P. *et al.* Origin of metastases: Subspecies of cancers generated by intrinsic karyotypic variations. *www.landesbioscience.com Cell Cycle 1151 Cell Cycle* **11**, 1151–1166 (2012).



61. Scott, J. G. *et al.* Recasting the Cancer Stem Cell Hypothesis: Unification Using a Continuum Model of Microenvironmental Forces. *Curr Stem Cell Rep* **5**, 22–30 (2019).
62. Werba, G. *et al.* Single-cell RNA sequencing reveals the effects of chemotherapy on human pancreatic adenocarcinoma and its tumor microenvironment. *Nat Commun* **14**, 797 (2023).
63. Ashok, G., Miryala, S. K., Saju, M. T., Anbarasu, A. & Ramaiah, S. FN1 encoding fibronectin as a pivotal signaling gene for therapeutic intervention against pancreatic cancer. *Molecular Genetics and Genomics* **297**, 1565–1580 (2022).
64. Chen, K. *et al.* Single-cell RNA-seq reveals dynamic change in tumor microenvironment during pancreatic ductal adenocarcinoma malignant progression. *EBioMedicine* **66**, 103315 (2021).
65. Wu, C. *et al.* Pancreatic cancer cells upregulate LPAR4 in response to isolation stress to promote an ECM-enriched niche and support tumour initiation. *Nat Cell Biol* (2023) doi:10.1038/s41556-022-01055-y.
66. Wente, M. N. *et al.* CXCL14 expression and potential function in pancreatic cancer. *Cancer Lett* **259**, 209–217 (2008).
67. Eguchi, D. *et al.* MAL2 expression predicts distant metastasis and short survival in pancreatic cancer. *Surgery* **154**, 573–582 (2013).
68. Gabitova-Cornell, L. *et al.* Cholesterol Pathway Inhibition Induces TGF- $\beta$  Signaling to Promote Basal Differentiation in Pancreatic Cancer. *Cancer Cell* **38**, 567-583.e11 (2020).
69. Moffitt, R. A. *et al.* Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet* **47**, 1168–1178 (2015).
70. Pitter, K. L. *et al.* Systematic comparison of pancreatic ductal adenocarcinoma models identifies a conserved highly plastic basal cell state. *Cancer Res* (2022) doi:10.1158/0008-5472.CAN-22-1742.
71. Sancho, P., Alcalá, S., Usachov, V., Hermann, P. C. & Sainz, B. The ever-changing landscape of pancreatic cancer stem cells. *Pancreatology* **16**, 489–496 (2016).
72. Donati, G. & Watt, F. M. Stem Cell Heterogeneity and Plasticity in Epithelia. *Cell Stem Cell* **16**, 465–476 (2015).
73. Gil Vazquez, E. *et al.* Dynamic and adaptive cancer stem cell population admixture in colorectal neoplasia. *Cell Stem Cell* **29**, 1213-1228.e8 (2022).
74. Biddle, A., Gammon, L., Liang, X., Costea, D. E. & Mackenzie, I. C. Phenotypic Plasticity Determines Cancer Stem Cell Therapeutic Resistance in Oral Squamous Cell Carcinoma. *EBioMedicine* **4**, 138–145 (2016).
75. Thankamony, A. P., Saxena, K., Murali, R., Jolly, M. K. & Nair, R. Cancer Stem Cell Plasticity - A Deadly Deal. *Front Mol Biosci* **7**, 79 (2020).
76. Warriar, N. M. *et al.* Understanding cancer stem cells and plasticity: Towards better therapeutics. *Eur J Cell Biol* **102**, 151321 (2023).
77. Cabrera, M. C. Cancer stem cell plasticity and tumor hierarchy. *World J Stem Cells* **7**, 27 (2015).

78. Das, P. K. *et al.* Plasticity of Cancer Stem Cell: Origin and Role in Disease Progression and Therapy Resistance. *Stem Cell Rev Rep* **16**, 397–412 (2020).
79. Paul, R., Dorsey, J. F. & Fan, Y. Cell plasticity, senescence, and quiescence in cancer stem cells: Biological and therapeutic implications. *Pharmacol Ther* **231**, 107985 (2022).
80. Doherty, M., Smigiel, J., Junk, D. & Jackson, M. Cancer Stem Cell Plasticity Drives Therapeutic Resistance. *Cancers (Basel)* **8**, 8 (2016).
81. Hirata, A., Hatano, Y., Niwa, M., Hara, A. & Tomita, H. Heterogeneity in Colorectal Cancer Stem Cells. *Cancer Prevention Research* **12**, 413–420 (2019).
82. Magee, J. A., Piskounova, E. & Morrison, S. J. Cancer Stem Cells: Impact, Heterogeneity, and Uncertainty. *Cancer Cell* **21**, 283–296 (2012).
83. Eun, K., Ham, S. W. & Kim, H. Cancer stem cell heterogeneity: origin and new perspectives on CSC targeting. *BMB Rep* **50**, 117–125 (2017).
84. Tang, D. G. Understanding cancer stem cell heterogeneity and plasticity. *Cell Res* **22**, 457–472 (2012).
85. Heumos, L. *et al.* Best practices for single-cell analysis across modalities. *Nat Rev Genet* **24**, 550–572 (2023).
86. Spitzer, M. H. & Nolan, G. P. Mass Cytometry: Single Cells, Many Features. *Cell* **165**, 780–791 (2016).
87. Porter, R. L. *et al.* Epithelial to mesenchymal plasticity and differential response to therapies in pancreatic ductal adenocarcinoma. *Proceedings of the National Academy of Sciences* **116**, 26835–26845 (2019).
88. Collisson, E. A. *et al.* Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat Med* **17**, 500–503 (2011).
89. McConkey, D. J. *et al.* Molecular Characterization of Pancreatic Cancer Cell Lines. in *Pancreatic Cancer* 457–469 (Springer New York, New York, NY, 2010). doi:10.1007/978-0-387-77498-5\_19.
90. Ferrara, B. *et al.* A Comprehensive Characterization of Stemness in Cell Lines and Primary Cells of Pancreatic Ductal Adenocarcinoma. *Int J Mol Sci* **23**, 10663 (2022).
91. Minami, F. *et al.* Morphofunctional analysis of human pancreatic cancer cell lines in 2- and 3-dimensional cultures. *Sci Rep* **11**, 6775 (2021).
92. Deer, E. L. *et al.* Phenotype and Genotype of Pancreatic Cancer Cell Lines. *Pancreas* **39**, 425–435 (2010).
93. Gil Vazquez, E. *et al.* Dynamic and adaptive cancer stem cell population admixture in colorectal neoplasia. *Cell Stem Cell* **29**, 1213–1228.e8 (2022).
94. Sheng, W. *et al.* Musashi2 promotes the development and progression of pancreatic cancer by down-regulating Numb protein. *Oncotarget* **8**, 14359–14373 (2017).
95. Zhou, L. *et al.* Musashi2 promotes the progression of pancreatic cancer through a novel ISYNA1-p21/ZEB-1 pathway. *J Cell Mol Med* **24**, 10560–10572 (2020).

96. Yang, H. *et al.* RNA-binding protein Musashi2 regulates Hippo signaling via SAV1 and MOB1 in pancreatic cancer. *Medical Oncology* **37**, 84 (2020).
97. Sheng, W. *et al.* Cooperation of Musashi-2, Numb, MDM2, and P53 in drug resistance and malignant biology of pancreatic cancer. *The FASEB Journal* **31**, 2429–2438 (2017).
98. Sheng, W. *et al.* Musashi2 promotes EGF-induced EMT in pancreatic cancer via ZEB1-ERK/MAPK signaling. *Journal of Experimental & Clinical Cancer Research* **39**, 16 (2020).
99. Sleightholm, R. L. *et al.* Emerging roles of the CXCL12/CXCR4 axis in pancreatic cancer progression and therapy. *Pharmacol Ther* **179**, 158–170 (2017).
100. Singh, S., Srivastava, S. K., Bhardwaj, A., Owen, L. B. & Singh, A. P. CXCL12–CXCR4 signalling axis confers gemcitabine resistance to pancreatic cancer cells: a novel target for therapy. *Br J Cancer* **103**, 1671–1679 (2010).
101. Ding, Y. & Du, Y. Clinicopathological significance and prognostic role of chemokine receptor CXCR4 expression in pancreatic ductal adenocarcinoma, a meta-analysis and literature review. *International Journal of Surgery* **65**, 32–38 (2019).
102. Morita, T. *et al.* CXCR4 in Tumor Epithelial Cells Mediates Desmoplastic Reaction in Pancreatic Ductal Adenocarcinoma. *Cancer Res* **80**, 4058–4070 (2020).
103. Harper, M. M. *et al.* Interaction of immune checkpoint PD-1 and chemokine receptor 4 (CXCR4) promotes a malignant phenotype in pancreatic cancer cells. *PLoS One* **17**, e0270832 (2022).
104. Wang, Y., Yi, J. & Liu, X. Roles of Dclk1 in the pathogenesis, diagnosis, prognosis and treatment of pancreatic cancer: A review. *Expert Rev Gastroenterol Hepatol* **16**, 13–19 (2022).
105. Westphalen, C. B. *et al.* Dclk1 Defines Quiescent Pancreatic Progenitors that Promote Injury-Induced Regeneration and Tumorigenesis. *Cell Stem Cell* **18**, 441–55 (2016).
106. Maruno, T. *et al.* Visualization of stem cell activity in pancreatic cancer expansion by direct lineage tracing with live imaging. *Elife* **10**, (2021).
107. Schlesinger, Y. *et al.* Single-cell transcriptomes of pancreatic preinvasive lesions and cancer reveal acinar metaplastic cells' heterogeneity. *Nat Commun* **11**, 4516 (2020).
108. Ito, H. *et al.* Dominant Expression of DCLK1 in Human Pancreatic Cancer Stem Cells Accelerates Tumor Invasion and Metastasis. *PLoS One* **11**, e0146564 (2016).
109. Qu, D. *et al.* Doublecortin-Like Kinase 1 Is Elevated Serologically in Pancreatic Ductal Adenocarcinoma and Widely Expressed on Circulating Tumor Cells. *PLoS One* **10**, e0118933 (2015).
110. Ito, H. *et al.* Dominant Expression of DCLK1 in Human Pancreatic Cancer Stem Cells Accelerates Tumor Invasion and Metastasis. *PLoS One* **11**, e0146564 (2016).
111. da Silva-Diz, V., Lorenzo-Sanz, L., Bernat-Peguera, A., Lopez-Cerda, M. & Muñoz, P. Cancer cell plasticity: Impact on tumor progression and therapy response. *Semin Cancer Biol* **53**, 48–58 (2018).
112. Pitter, K. L. *et al.* Systematic comparison of pancreatic ductal adenocarcinoma models identifies a conserved highly plastic basal cell state. *Cancer Res* (2022) doi:10.1158/0008-5472.CAN-22-1742.

113. Lu, W. & Kang, Y. Epithelial-Mesenchymal Plasticity in Cancer Progression and Metastasis. *Dev Cell* **49**, 361–374 (2019).
114. Gil Vazquez, E. *et al.* Dynamic and adaptive cancer stem cell population admixture in colorectal neoplasia. *Cell Stem Cell* **29**, 1213-1228.e8 (2022).
115. R Core Team. R: A language and environment for statistical computing. . Preprint at (2023).
116. Schuyler, R. P. *et al.* Minimizing Batch Effects in Mass Cytometry Data. *Front Immunol* **10**, (2019).
117. Timothy Keyes. tidytof: Analyze Mass Cytometry Data Using Tidy Data Principles. Preprint at (2023).
118. Werba, G. *et al.* Single-cell RNA sequencing reveals the effects of chemotherapy on human pancreatic adenocarcinoma and its tumor microenvironment. *Nat Commun* **14**, 797 (2023).
119. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).
120. Kolberg, L. *et al.* g:Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res* **51**, W207–W212 (2023).
121. Chan-Seng-Yue, M. *et al.* Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. *Nat Genet* **52**, 231–240 (2020).
122. Puram, S. V. *et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* **171**, 1611-1624.e24 (2017).
123. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag, New York, 2016).
124. Raivo Kolde. pheatmap: Pretty Heatmaps. Preprint at (2019).

## Acknowledgements

The authors thank Christine Iacobuzio-Donahue, Catherine Hogan, Eric O’Neill, Ming Tsao, Claudia Monaco and Liye Chen for providing the cell lines, and Jonathan Williams and Sarah Pericli for their advice and assistance with clinical sample genotyping. The work in the laboratory of S.P. was supported by a Cancer Research UK Career Development Fellowship (Grant ID: C59392/A25064) and Pancreatic Cancer UK (Grant/Award Number: 2018RIF\_03). E-H.E. was supported by the Kristjan Jaak Scholarship Programme. U.O. was supported by Innovate UK, the National Institute for Health Research Oxford Biomedical Research Centre, Cancer Research UK, the Leducq Epigenetics of Atherosclerosis Network (LEAN) program grant from the Leducq Foundation and the Myeloma Single Cell Consortium.

## Author contributions

E-H.E., D.A., and S.P. designed the project and performed the experiments. E-H.E. and F.L. analysed the data and visualised the results. A.R. and Z.S. provided pathology support and clinical tissue samples. S.P. provided resources. S.P., P.H. and U.O. supervised the project and provided intellectual input. The original draft was written by E-H.E. and S.P. All authors reviewed the manuscript.

## Competing interests

All authors declare no competing interests.

## **Materials and correspondence**

Correspondence and material requests to Siim Pauklin.

## **Supplementary information**

Supplementary information in a separate file.