



OPEN

Relevance of circulating hybrid cells as a non-invasive biomarker for myriad solid tumors

Matthew S. Dietz^{1,14,16}, Thomas L. Sutton^{2,16}, Brett S. Walker^{2,16}, Charles E. Gast³, Luai Zarour^{2,15}, Sidharth K. Sengupta³, John R. Swain³, Jennifer Eng⁴, Michael Parappilly³, Kristen Limbach², Ariana Sattler³, Erik Burlingame^{4,5}, Yuki Chin³, Austin Gower⁶, Jose L. Montoya Mira^{4,6}, Ajay Sapre⁶, Yu-Jui Chiu⁶, Daniel R. Clayburgh^{7,8,9}, SuEllen J. Pommier², Jeremy P. Cetnar^{9,10}, Jared M. Fischer^{6,9,11}, Jerry J. Jaboin^{9,12}, Rodney F. Pommier^{2,9}, Brett C. Sheppard^{2,9}, V. Liana Tsikitis², Alison H. Skalet^{9,13}, Skye C. Mayo^{2,9}, Charles D. Lopez^{9,10}, Joe W. Gray^{4,9}, Gordon B. Mills^{3,9}, Zahi Mitri^{9,10}, Young Hwan Chang^{4,5,9}, Koei Chin^{4,9} & Melissa H. Wong^{3,9}✉

Metastatic progression defines the final stages of tumor evolution and underlies the majority of cancer-related deaths. The heterogeneity in disseminated tumor cell populations capable of seeding and growing in distant organ sites contributes to the development of treatment resistant disease. We recently reported the identification of a novel tumor-derived cell population, circulating hybrid cells (CHCs), harboring attributes from both macrophages and neoplastic cells, including functional characteristics important to metastatic spread. These disseminated hybrids outnumber conventionally defined circulating tumor cells (CTCs) in cancer patients. It is unknown if CHCs represent a generalized cancer mechanism for cell dissemination, or if this population is relevant to the metastatic cascade. Herein, we detect CHCs in the peripheral blood of patients with cancer in myriad disease sites encompassing epithelial and non-epithelial malignancies. Further, we demonstrate that *in vivo*-derived hybrid cells harbor tumor-initiating capacity in murine cancer models and that CHCs from human breast cancer patients express stem cell antigens, features consistent with the potential to seed and grow at metastatic sites. Finally, we reveal heterogeneity of CHC phenotypes reflect key tumor features, including oncogenic mutations and functional protein expression. Importantly, this novel population of disseminated neoplastic cells opens a new area in cancer biology and renewed opportunity for battling metastatic disease.

Metastatic disease contributes to over 90% of cancer-related deaths¹. The disparity in survival between early- and late-stage disease reflects a knowledge gap in the biology underlying metastasis, such as the identity of tumor cells prone to disseminate and functional insight into their seeding of metastatic tumors. To successfully navigate the metastatic cascade, tumorigenic cells must invade surrounding tissue, intravasate into circulation, seed distant sites, and establish a permissive microenvironment for colonization and growth^{2–4}. Among these events, the mechanisms that allow tumor cells to escape into circulation and survive are arguably the least understood. Direct

¹Department of Pediatrics, Oregon Health & Science University (OHSU), Portland, OR 97239, USA. ²Department of Surgery, OHSU, Portland, OR 97239, USA. ³Department of Cell, Developmental and Cancer Biology, Oregon Health & Science University, 2720 S. Moody Ave., Mailcode KC-CDCB, Portland, OR 97201, USA. ⁴Department of Biomedical Engineering, OHSU, Portland, OR 97239, USA. ⁵Computational Biology Program, OHSU, Portland, OR 97239, USA. ⁶Cancer Early Detection Advanced Research Center, OHSU, Portland, OR 97201, USA. ⁷Department of Otolaryngology, OHSU, Portland, OR 97239, USA. ⁸Operative Care Division, Portland Veterans Affairs Medical Center, Portland, OR 97239, USA. ⁹The Knight Cancer Institute, OHSU, Portland, OR 97201, USA. ¹⁰Department of Medicine, OHSU, Portland, OR 97239, USA. ¹¹Department of Molecule and Medical Genetics, OHSU, Portland, OR 97239, USA. ¹²Department of Radiation Medicine, OHSU, Portland, OR 97239, USA. ¹³Casey Eye Institute, OHSU, Portland, OR 97239, USA. ¹⁴Present address: Department of Pediatrics, University of Utah, Salt Lake City, UT 84113, USA. ¹⁵Present address: Department of General Surgery, Legacy Medical Group, Gresham, OR 97030, USA. ¹⁶These authors contributed equally: Matthew S. Dietz, Thomas L. Sutton and Brett S. Walker. ✉email: wongme@ohsu.edu

investigation of tumor cells in circulation is hampered by their rarity, however it remains clear they provide an opportunity to better understand metastatic progression and potentially impact patient outcomes.

Detection of conventionally-defined circulating tumor cells (CTCs) in metastatic cancer patients provided the first direct evidence for tumor cell dissemination⁵. Distinct from other cells in circulation, CTCs express tumor proteins (e.g. cytokeratin (CK), epithelial cell adhesion molecule (EPCAM), and E-cadherin (ECAD) in epithelial malignancies) but lack leukocyte epitopes (e.g. CD45). CTCs are reported in many disease sites, including breast, colorectal, lung, pancreatic, and prostate cancer^{6–10}, as well as non-epithelial cancers, such as glioblastoma¹¹ and melanoma¹². The tumor-based origin of CTCs is supported by their mutual expression of nucleic acids, protein and cellular features. However, initial excitement for the utility of CTCs has been tempered by their rarity, highlighted by just 5 CTCs/7.5 mLs of blood in the clinical presence of high tumor burden⁷. Moreover, CTC enumeration has failed to inform therapeutic decision making, especially in early-stage disease¹³.

Our group reported a novel tumor-derived cell population detected in peripheral blood of patients with pancreatic ductal adenocarcinoma (PDAC)¹⁴. Circulating hybrid cells (CHCs) express neoplastic and immune cell functional attributes and are readily identifiable by their co-expression of tumor and leukocyte proteins. Our previous studies identified cell fusion as a major pathway for hybrid generation^{14–17}, though other described mechanisms can also result in tumor cells with dual expression states, including immune mimicry¹⁸, developmental gene expression¹⁹ and exosome-mediated protein/mRNA expression²⁰. Focusing on cell fusion, our studies have rigorously demonstrated that neoplastic-immune cell hybrids retain functional genotypic and phenotypic properties from both parent cells based on culture assays, murine models of tumorigenesis, and human patient samples^{14–16}. In vitro- and in vivo-derived fusion hybrids displayed enhanced motility, invasiveness and growth at metastatic sites, indicating they play a central role in the progression of metastatic disease¹⁴. In patients with pancreatic cancer, CHCs significantly outnumbered CTCs at every stage of disease, with levels capable of providing a prognostic indicator for overall survival¹⁴. Further, detection of CHCs with ECAD expression correlates with PDAC nodal staging²¹, indicating subpopulations of CHCs with differential protein expression may provide a predictive biomarker. However, as a newly defined population, the functional tumor-initiating capacity and expression of stem cell attributes of CHCs is unexplored.

Tumors contain significant cellular heterogeneity and select populations exhibit key attributes needed to successfully spread to distant sites. Presumably, generation of neoplastic hybrid cells supports the evolution of cells with capabilities to both migrate and initiate tumors. The concept that neoplastic cells with stem cell properties represent a source for metastatic spread and treatment resistance is strongly supported by the literature^{22–24}. Not only do neoplastic cells expressing stem cell surface antigens harbor tumorigenic potential, their levels within a tumor negatively correlate with patient survival in multiple cancer types^{25–27}. In breast cancer, tumor-initiating cancer stem cells express CD44, a transmembrane receptor for hyaluronic acid that is a known effector of metastasis, and maintain low levels of CD24, a sialoglycoprotein that facilitates a number of important signaling networks in development^{28–32}. This functional cell surface identity extends to other solid organ malignancies including PDAC, prostate cancer, colon cancer, hepatocellular carcinoma, melanoma, and glioma^{33,34}. Though rare, CTCs with tumorigenic ability mirror the expression of cancer stem cells, CD44⁺/CD24^{lo}^{35,36}, and supports their prognostic value^{37–40}. Based upon the metastatic potential of disseminated tumor cells, we sought to specifically investigate the prevalence of CHCs across numerous solid organ malignancies, identify sub-populations with stem cell markers, and evaluate conserved phenotypes between CHCs and tumor tissue.

Herein, we establish that CHCs represent the principal tumor-derived cell in circulation across a wide spectrum of malignancies, including epithelial and non-epithelial cancers. Using murine models and human patient samples, we provide evidence that hybrid cells possess tumor-initiating behaviors and characteristics that indicate their potential importance in metastatic disease. Furthermore, we reveal CHCs share phenotypic similarities with tumor tissue and reflect the cellular heterogeneity of neoplastic disease. With this evidence, we postulate that CHCs are effectors of metastatic spread and have translational value as a non-invasive analyte across cancer types.

Results

CHCs are the dominant circulating tumor-derived cell in myriad cancers. Dissemination of neoplastic cells into circulation is a key step in metastatic progression. Therefore, to determine if hybrid formation and escape is generalizable across cancer types, we evaluated peripheral blood for CHCs from patients with 14 different epithelial or non-epithelial malignancies, including ampullary adenocarcinoma, breast adenocarcinoma, ovarian carcinoma, cholangiocarcinoma, colon adenocarcinoma, esophageal cancer, high grade glioma (pediatric and adult), head and neck squamous cell carcinoma, pancreatic ductal adenocarcinoma, pancreatic neuroendocrine tumor, prostate adenocarcinoma, rectal adenocarcinoma and uveal melanoma (Table 1). A standard ficoll-density gradient facilitated the isolation of peripheral blood mononuclear cells (PBMCs) and downstream CHC detection. Circulating tumor-derived cells were identified by protein expression of canonical tumor markers (Epithelial: CK⁺; Uveal melanoma: NKI beteb⁺; Glioma: glial fibrillary acidic protein (GFAP)⁺⁴¹; Pancreatic neuroendocrine tumor (PNET): chromogranin A (CHGA⁺), synaptophysin (SYP⁺)⁴²) using two different platforms, flow cytometry and fluorescence microscopy. Expression of the pan-leukocyte antigen CD45 provided the distinction between CTCs and CHCs, where CHCs were identified as cells co-positive for both tumor protein and CD45 (Fig. 1, Figure S1). Flow cytometry facilitated robust quantitative analyses (Fig. 1B,C), while microscopy provided visual confirmation of protein expression and their relevant cellular localization (Fig. 1E,F, Figure S2). Leveraging the collective power of these two platforms, we identified CHCs in every cancer we analyzed, including glioma, which is reported to seldomly disseminate outside of the central nervous system⁴³. In addition, we identified a significantly higher number of CHCs in each disease site relative to healthy subjects ($p < 0.00001$ for all, except PDAC, ECA, Adult Glioma $p < 0.0001$, and HNSCC $p < 0.05$), and found that CHCs outnumbered CTCs in all cancer types (Fig. 1, range of $p < 0.001–0.0001$). Despite differences in age,

Disease site	Number of patients	Age range	Disease burden			Treatment naïve (%)
			Local	Regional	Metastatic	
Ampullary adenocarcinoma	5	56–77	2	3	0	5 (100)
Breast adenocarcinoma	5	51–73	1	2	2	4 (80)
Cholangiocarcinoma	5	42–81	2	1	2	5 (100)
Colon adenocarcinoma	5	33–65	0	0	5	1 (20)
Esophageal cancer	5	52–75	3	2	0	0 (0)
Adult high-grade glioma	5	31–74	5	0	0	2 (40)
Pediatric high-grade glioma	5	4–15	5	0	0	1 (20)
Head and neck SCC	5	57–70	0	5	0	5 (100)
Non-small cell lung cancer	5	41–74	0	0	5	2 (50)
Ovarian carcinoma	5	60–76	0	4	1	1 (20)
Pancreatic adenocarcinoma	5	56–76	1	2	2	1 (20)
Pancreatic NET	5	55–88	4	1	0	5 (100)
Prostate adenocarcinoma	5	57–72	5	0	0	5 (100)
Rectal adenocarcinoma	5	47–67	0	0	5	0 (0)
Uveal melanoma	5	56–71	4	0	1	5 (100)
Total	75	4–88	32 (43%)	20 (27%)	23 (31%)	42 (56%)

Table 1. Patient clinicopathologic characteristics by disease site. *CHC* circulating hybrid cell, *SCC* squamous cell carcinoma, *NET* neuroendocrine tumor.

sex, tumor burden, and prior treatment status, across cancers the heterogeneous patient population (Table S1) harbored higher levels of CHCs than CTCs.

To further establish that CHCs derive from tumor tissue, we analyzed CHCs for the presence of known oncogenic mutations. Oncogenic *KRAS* mutations are implicated in the malignant transformation of pancreatic epithelia and are almost ubiquitously present in PDAC tumors^{44,45}. To define the oncogenic identity of CHCs at the genomic level, we focused our analysis on PDAC-derived CHCs. Using fluorescence-activated cell sorting (FACS), we isolated normal leukocytes (7000; CD45⁺/EpCAM⁻/ECAD⁻/CD49c⁻) and CHCs (580; CD45⁺/[EpCAM/ECAD/CD49c]⁺) from a patient with *KRAS*-G12D mutant PDAC. CTCs were too rare for capture by FACS. Using digital droplet polymerase chain reaction (ddPCR), we screened samples for wild type *KRAS* and seven common oncogenic *KRAS* mutations, including G12D^{45,46}. We identified that up to 9.1% of the isolated CHCs harbored an oncogenic *KRAS* allele while the leukocyte fraction only expressed wild type *KRAS* (Fig. 1D), indicating that CHCs derive from the primary tumor and retain key genomic drivers of cancer progression.

Hybrid cells harbor tumor-initiating properties. Retention of oncogenic driver mutations in disseminated tumor cell populations supports their relevance in metastatic progression. However, evidence of the capacity for CHCs to seed and grow metastatic tumors has not been previously explored. Our prior studies demonstrated that fusion-hybrids derived from allografted B16F10 melanoma cells are detectable in both primary and metastatic sites, with greater tumorigenicity than unfused cancer cells¹⁴. However, as they are derived from immortalized cell lines uniformly transformed and selected for their proliferative properties, this model is sub-optimal to assess tumor-initiating potential. In order to more accurately investigate the contribution of hybrid cells to metastatic progression, we generated a model for isolating fusion hybrids derived from autogenous malignancy. Here, we engaged the murine mammary tumor model (mouse mammary tumor virus-polyoma middle tumor-antigen; MMTV:PyMT), which develops mammary carcinoma⁴⁷. We genetically marked tumor cells with red fluorescent protein (RFP) by crossing the MMTV-PyMT mouse onto a CAG-RFP background, yielding MMTV-PyMT-RFP mice. Harvested RFP⁺ mammary tumor cells were dissociated into single cells and injected into the mammary fat pad of Actin-green fluorescent protein (GFP) transgenic mice, generating tumors designated as MMTV-PyMT-RFP into-GFP (Fig. 2A). In this model, co-expression of RFP and GFP identified tumor hybrid cells.

To assess the tumor-forming capacity of hybrid cells, we FACS-isolated hybrid cells (RFP⁺/GFP⁺) and unfused tumor cells (RFP⁺/GFP⁻) from MMTV-PyMT-RFP into-GFP tumors (Figure S3). We then independently injected each cell type into the mammary fat pad of secondary wild type recipient mice (2500 cells per animal; Fig. 2B). Neoplastic-derived hybrids supported rapid tumor growth, whereas unfused cancer cells did not generate tumors (Fig. 2C). To further investigate relative tumor-initiating capacity, we used a limiting-dilution assay. We found that to generate tumors in all mice, two orders of magnitudes more unfused cancer cells were required compared to the number of hybrids (Fig. 2D). The potent tumor-initiating capacity of hybrid cells is consistent with functional stem cell capacities. Cancer stem cell phenotypes, which are widely identified in human tumors, are associated with increased tumorigenic potential compared to their non-stem cell counter parts, have been linked to CTC phenotypes^{32,48–50}. The functional resemblance between fusion-hybrids and cancer stem cells necessitates investigation as whether human CHCs similarly harbor stem characteristics.

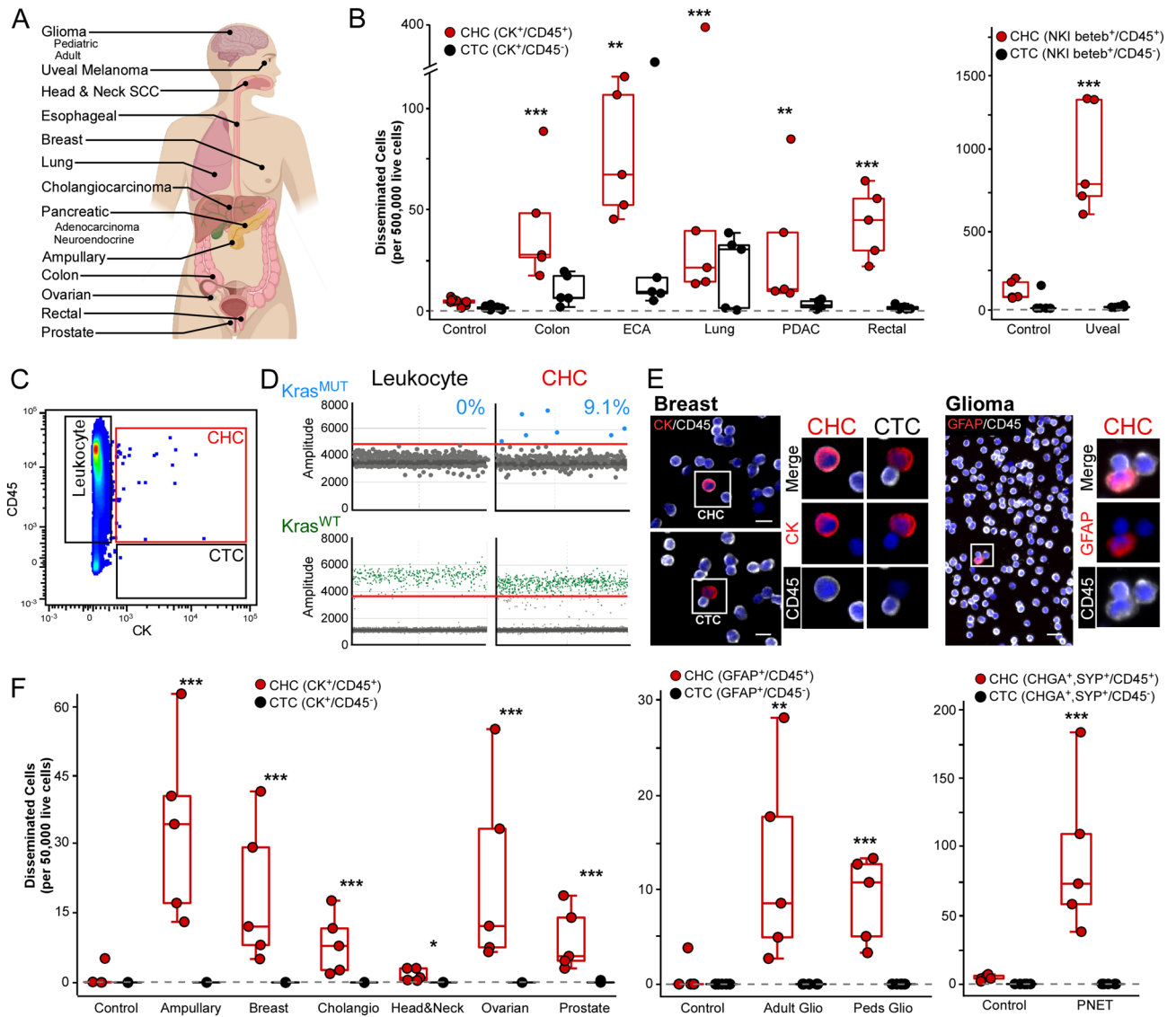


Figure 1. Circulating hybrid cells out number circulating tumor cells in myriad of human cancer disease sites. **(A)** Anatomical location of 14 cancers evaluated for disseminated CTCs and CHCs. **(B)** Flow cytometry analysis of peripheral blood from patients with advanced stage colon, esophageal, lung, pancreatic and rectal adenocarcinomas as well as uveal melanoma ($n = 5$ patients per disease site) demonstrates CHCs (mean value, red bar) present at significantly higher levels compared to CTCs (mean value, grey bar) and healthy controls (mean value, $n = 10$). **(C)** Representative flow cytometry scatter plot of cytokeratin vs CD45 expression, CHCs (red box), CTCs (gray box) full gating schema in Fig. S1. **(D)** Isolation of CHCs from pancreatic adenocarcinoma identifies KRAS mutation in 9.1% of CHCs using ddPCR. **(E)** In situ immunofluorescence microscopy analysis of peripheral blood cells analyzed for detection of glial fibrillary acid protein (GFAP; red) and CD45 (white) expression in pediatric high-grade glioma and cytokeratin (CK; red) and CD45 (white) in breast adenocarcinoma, identify CHCs and rare CTCs. Scale bars 10 and 20 μm , respectively. **(F)** Immunofluorescence microscopy analysis of peripheral blood from patients ($n = 5$ per disease site) with ampullary carcinoma, breast, cholangiocarcinoma, head and neck squamous cell carcinoma, ovarian, prostate, high grade glioma (adult and pediatric) and pancreatic neuroendocrine tumors reveals CHCs (mean value, red bar) are present at significantly higher levels compared to CTCs (mean value, grey bar) and healthy controls. *** $p < 0.00001$; ** $p < 0.0001$; * $p < 0.05$.

Human breast cancer CHCs display stem cell phenotypes. To investigate stem cell attributes of human CHCs, we leveraged the well-described immunophenotype of tumorigenic breast cancer stem cells identified by CD44⁺/CD24^{lo} expression^{28–32}. We evaluated peripheral blood specimens from a cohort of twenty-seven treatment-naïve patients with breast cancer representing all stages and subtypes (Table 2, Table S2). Using flow cytometry, PBMCs were interrogated for expression of EpCAM, CD45, CD31, CD44 and CD24 to identify CTC and CHCs with stem cell identity (Fig. 3). Consistent with the findings from our pan-cancer evaluation (Fig. 1), we identified greater numbers of CHCs than CTCs in all breast cancer patients (Fig. 3A, Figure S4). Further-

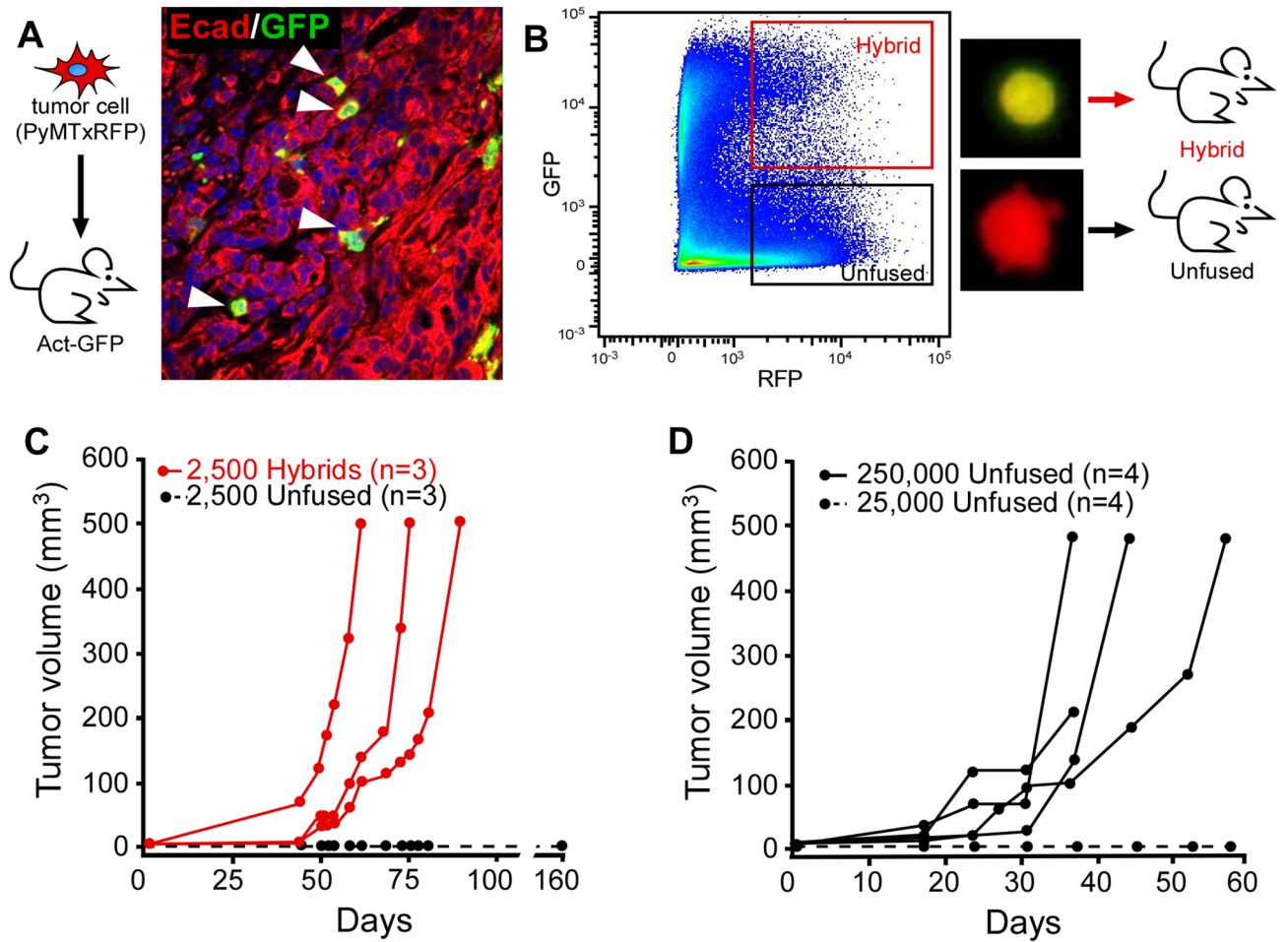


Figure 2. Murine hybrid cells harbor tumor-initiating capacity. (A) Mammary carcinoma cells from PyMT:RFP mice injected into GFP-expressing recipients. Resulting tumors express E-cadherin (red) and Actin:GFP (green) and were counterstained with Hoechst dye (blue). Arrowhead denotes co-positive hybrid cells. (B) Representative FACs plot of tumor-derived hybrid cells (GFP⁺/RFP⁺) and unfused tumor cells (RFP⁺/GFP⁻) from dissociated tumor and injected into syngeneic wt recipient mice. (C) Twenty-five hundred FACS-isolated cells were injected into recipient mice (hybrid cells, red line) and unfused cells, dashed black line) and temporally monitored for growth. (D) Injection of 250,000 unfused tumor cells developed tumors in recipient mice (black line), while 25,000 cells did not grow (dashed line).

Cancer subtype	Number of patients	Age range	Disease state			Treatment naïve (%)
			Local	Regional	Metastatic	
HR ⁺ /HER2 ⁻	15	28–80	8	5	2	15 (100)
HR ⁺ /HER2 ⁺	7	38–77	3	3	1	7 (100)
HR ⁻ /HER2 ⁻	3	26–58	1	2	0	3 (100)
HR ⁻ /HER2 ⁺	2	44–50	0	1	1	2 (100)
Total	27	26–80	11 (41%)	12 (44%)	4 (15%)	27 (100%)

Table 2. Clinicopathologic characteristics of breast cancer patients analyzed for chc stem cell properties. CHC circulating hybrid cell, HR hormone receptor, HER2 human epidermal growth factor receptor 2.

more, CHCs display high proportions of stem cell characteristics. All twenty-seven patient samples harbored CD44⁺/CD24^{lo} CHCs while only 74.1% (n=20) had detectable CD44⁺/CD24^{lo} CTCs. When present, CD44⁺/CD24^{lo} CTCs were found at significantly lower level than CD44⁺/CD24^{lo} CHCs, with means of 1.0 and 45.3 per 500,000 live cells, respectively (p<0.00001). Additionally, compared to the normal leukocyte population (CD45⁺/EpCAM⁻), there was a higher proportion of CHCs with CD44⁺/CD24^{lo} expression (p<0.01), indicating the high level of CHCs with stem marker expression does not simply reflect the normal distribution of leukocyte phenotypes. The presence of stem markers on CHCs suggests replicative potential and supports the relevance

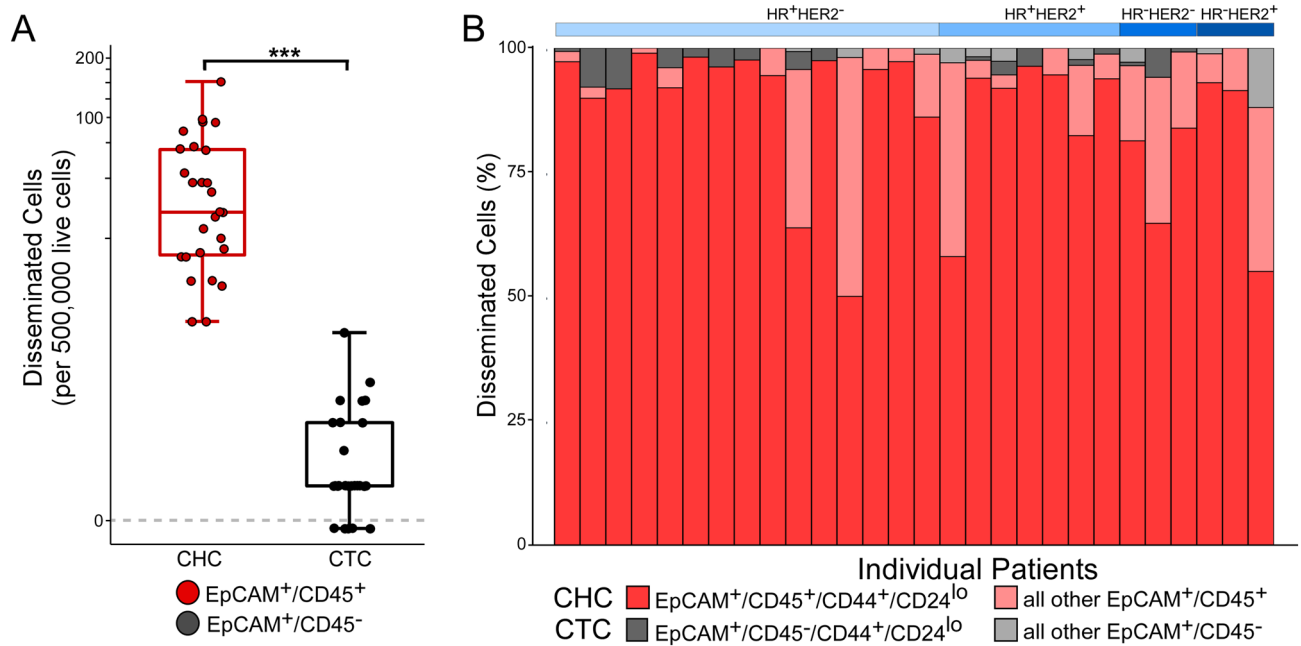


Figure 3. CTC and CHC stem cell antigen expression in breast cancer patients. **(A)** Flow cytometric analysis of peripheral blood from 27 untreated patients across breast cancer subtypes. Box whisker plot of CHCs (EpCAM⁺/CD45⁺; red) and CTCs (EpCAM⁺/CD45⁻; gray). CHCs levels are higher than CTCs at significant levels ($p < 0.0001$). **(B)** Depiction of percentages of CHC (red scale) and CTCs (gray scale) with stem cell identify (i.e. CD44⁺/CD24^{lo}).

of CHCs to metastatic progression. Coupled with their relative abundance compared to CTCs, these data suggest CHCs are a readily available circulating analyte with potential for liquid biopsy application. These results prompted our investigation into the extent to which CHCs reflect the discrete phenotypic features and diversity of neoplastic cells in cancer tissue.

Tumors disseminate a heterogeneous population of CHCs. The relative prevalence of CHCs and their suitability for liquid biopsy⁵¹ provides an exciting and unexplored opportunity to query their diagnostic, prognostic, or predictive utility for longitudinal care of cancer patients. One major challenge in the management of solid tumors is the vast heterogeneity that promotes variable or incomplete treatment response. Standard tissue biopsies may fail to capture the full spectrum of tumor heterogeneity and miss cells that could identify therapeutic vulnerabilities and resistance mechanisms. Further, repeat tumor sampling is fraught with logistical challenges and risks patient harm. CHCs are detectable in sufficient numbers to facilitate more complete tumor analyses (Fig. 1C), which could offer insight into tumor phenotyping and evolution over the course of a patient's treatment. To explore how CHCs reflect the phenotypic features and diversity within the tumor, we analyzed tumor biopsy specimens and corresponding CHCs from a patient with refractory metastatic breast cancer while receiving palliative chemotherapy (Fig. 4). The patient underwent an initial biopsy (BX1) at time of study enrollment and started monotherapy with the poly (ADP-ribose) polymerase (PARP) inhibitor, olaparib. A second biopsy (BX2) was obtained a month later. This evaluation aimed to provide feasibility data to determine if real time phenotypic changes within tumor cells from the biopsy reflected response to treatment.

To identify treatment resistant disease, we spatially defined the tumor microenvironment and interrogated disease heterogeneity with a multiplexed cyclic immunofluorescence (cyCIF) on longitudinal hilar lymph node biopsies from a patient with metastatic triple-negative breast cancer (TNBC). CyCIF utilizes iterative staining and imaging cycles to facilitate the spatial resolution of > 40 epitope-specific antibodies (Table S3) at the cellular level⁵² facilitating resolution of various tumor attributes e.g. stromal, immune, epithelial and vascular compartments. We performed cyCIF on formalin-fixed paraffin-embedded (FFPE) tumor biopsies to phenotypically identify viable treatment-resistant disease hot spots, and gain insight into cellular heterogeneity within the tumor ecosystem (Fig. 4A). Extraction of fluorescent intensity patterns from multiplexed, segmented images allowed for single-cell unsupervised learning to probe tumor heterogeneity⁵³. Using cellular features from BX1 and BX2, K-means clustering revealed discrete tumor cell populations based on clinically relevant protein expression (Fig. 4B). To interrogate intratumoral and inter-biopsy heterogeneity we focused on the neoplastic epithelial compartment (ECAD⁺ and/or CK⁺ cells) to quantitatively describe the phenotypes of viable, treatment-resistant disease (Fig. 4C). Within BX1 (pre-treatment), 42.5% of the epithelial compartment was Ki67⁺, with 14.9% of these cells also expressing CD44, suggestive of a proliferative stem cell phenotype. The distribution of hormone receptors, estrogen receptor (ER) and androgen receptor (AR), varied within this population, with 1.5% ER⁺/AR⁺, 1.4% ER⁺/AR⁻, and 6.5% ER⁻/AR⁺. After initiating therapy, only 20.6% of the malignant epithelia in BX2 were Ki67⁺ while the proportion of CD44⁺ cells remained similar at 14%, which may indicate a global tumor

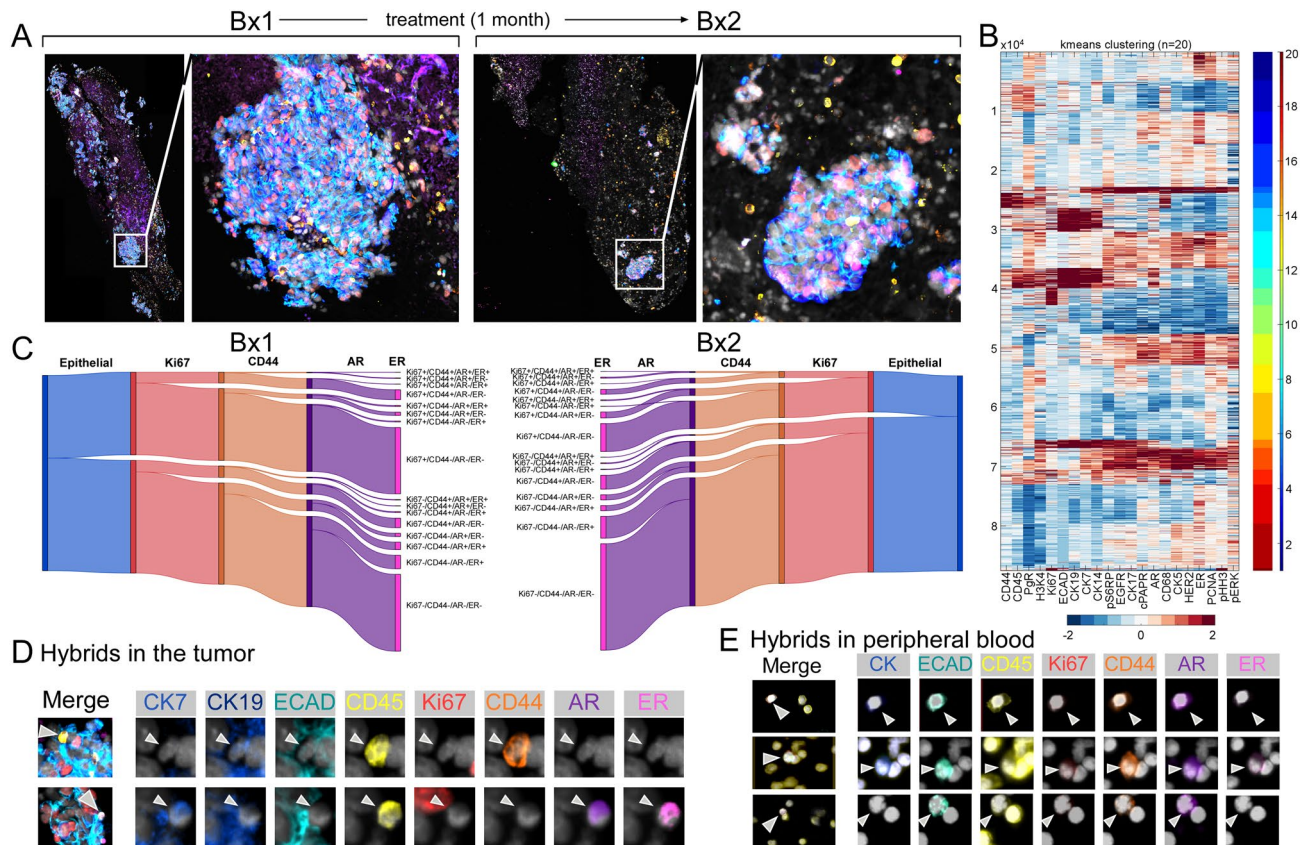


Figure 4. Multiplexed phenotypic analysis of tissue biopsy and CHCs from a subject with relapsed refractory breast cancer. **(A)** Cyclic immunofluorescence microscopy evaluation of tumor biopsy tissue before and during treatment reveals hot spots of viable disease. **(B)** Unsupervised clustering [k-means clustering with the number of clusters ($n = 20$)] of cellular immunophenotype identified distinct clusters and reveal tumor heterogeneity. **(C)** Sankey diagram of tumor phenotypes [proliferative (Ki67), stem (CD44), hormonal status (AR, HR)] quantifies heterogeneity of epithelial marker positive tumor cells. **(D)** Representative images of heterogeneous populations of hybrid cells expressing tumor epithelial (CK and/or ECAD) and immune (CD45) markers within biopsied tumor and **(E)** peripheral blood mononuclear cells. Heatmap created with MathWorks MATLAB ver. R2017b, www.mathworks.com.

response to PARP inhibition but less effect on the proliferative stem population. The ER⁺/AR⁺ and ER⁺/AR⁻ cells remained consistent (1.2% and 0.6%, respectively), however we observed an increase in ER⁻/AR⁺ cells to 20% between BX1 and BX2. Importantly, AR-expression is associated with cellular proliferation and metastatic spread in TNBC⁵⁴, therefore the observed threefold increase in concentration of Ki67⁺/ER⁻/AR⁺ neoplastic cells under PARP inhibition shows enrichment of a treatment-resistant proliferative population of cells. We identified hybrid cells within the tumor biopsy (Fig. 4D) by the co-expression of CD45 and epithelial tumor proteins, ECAD and cytokeratins, including those with the treatment-resistant tumor phenotype. Though only a qualitative analysis, hybrid cell presence in this metastatic patient supports their contribution to metastatic spread. With the evidence of proliferative, treatment-resistant disease within this patient's biopsy, we sought to investigate disseminated tumor cell populations from matched peripheral blood samples.

To investigate heterogeneity among tumor cells in circulation, we utilized cyCIF to interrogate PBMCs sampled at the time of BX2. To maximize phenotypic profiling and minimize cell loss, we performed two iterative staining cycles using antibodies to CD45, ECAD, panCK, AR, ER, Ki67 and CD44. Our analysis revealed a heterogeneous population of CHCs, including those with CD44 and AR expression (Fig. 4E), which align with the phenotypes detected in both the proliferating tumor compartment (Fig. 4C) and that of tumor hybrid cells (Fig. 4D). These data indicate that tumors disseminate a heterogeneous population of CHCs reflective of cancer cell phenotypes, thus further exploration into the abundance and diversity of CHCs is warranted to define their capacity to anticipate disease evolution and treatment resistance.

Discussion

A cancer cell's successful navigation of the metastatic cascade drives cancer lethality, highlighting the importance of understanding the functional biology of disseminated tumor cells. Recently, we identified a novel disseminated tumor cell population harboring both neoplastic and immune cell identities, and now establish their conserved functional phenotypes. Here, we demonstrated the presence of CHCs in patients across many epithelial and non-epithelial cancers, indicating that the generation of hybrids and their escape into peripheral blood is a ubiquitous

and generalizable phenomenon of solid organ tumors, which warrant in-depth investigation. Although CHCs are more numerous than CTCs across all evaluated malignancies, the diversity and limited number of patients in each disease group prevents more robust determination of the significance of their overall levels. However, as our prior analysis revealed CHC burden correlates with PDAC stage and patient survival, studies with larger cohorts should be pursued across cancer types to determine how CHC quantification might translate to clinical practice. The data presented here also indicates that phenotypic analysis of CHCs provides insight into tumor heterogeneity, and therefore may hold potential as a predictive biomarker when queried for specific markers of interest (i.e. stem features and therapeutic targets). Beyond providing insight on tumor protein expression, our detection of mutant *KRAS* genes in CHCs isolated from a patient with PDAC highlights their promise as source of genomic material to facilitate tumor profiling for clinically actionable oncogenic alterations. Given that *KRAS* mutations were only identified in a subset of CHCs underscores there is much that remains to be understood about hybrid cell biology, including the degree to which hybrid cells retain each parent genome and the process of ploidy reduction⁵⁵. Additionally, evidence exists of allele recombination between parent genomes resulting in genetically distinct hybrid cells and contributes to the heterogeneity of hybrid cell populations⁵⁶. While it is worth noting that the frequency of oncogenic *KRAS* mutations we detected by ddPCR in CHCs is similar to what has been reported for PDAC-derived CTCs by the same method⁵⁷, it is unknown if this reflects conserved biology among disseminated tumor cell populations or is possibly related to the limitations of ddPCR analysis. Further, there is increasing evidence of the heterogeneity and distribution of *KRAS* mutations within individual tumors^{58,59}, which has been shown to influence PDAC biology^{60,61} and could possibly be reflected in the mutation burden of CHCs. Selection for cells with more aggressive phenotypes, predilection for dissemination, or fusogenic potential may also partially account for discrepancies seen between primary tumors and tumor-derived cells in circulation, and provide a rational basis for further investigation.

Disseminated neoplastic cells are effectors of tumor progression, yet the extent to which they reflect the spectrum of diversity among primary tumor cells remains unresolved. While it is possible that tumors randomly shed neoplastic cells into circulation, an active process may exist to mediate phenotype acquisition for escape from permissive tumor regions. To begin to unravel this question, we first employed a tractable murine system to collect hybrid cells and demonstrate their tumor-initiating capacity. We then translated these findings to human disease by evaluating disseminated tumor cells, including CHCs, for stem cell phenotypes. In a murine mammary tumor model, *autogenous* hybrid cells demonstrate the defining feature of stemness, namely growth in vivo, and with greater potency than unfused cancer cells. These findings suggest that hybrid identity independently supports in vivo replicative potential, reminiscent of stem cells, though the underlying mechanisms remains to be investigated. Our observation that classical breast cancer stem surface antigens (CD44⁺/CD24^{lo})^{32–34} are enriched on breast cancer CHCs, suggests their tumor initiating potential. Additionally, it is interesting to note that CD44 is a known effector of homotypic macrophage fusion in experimental models⁶², and its high expression in CHCs may indicate a previously unappreciated link between cancer stem cells and macrophage-cancer cell fusion. Beyond the ethos of CHC biology, focusing on the heterogeneity of disseminated tumor cells with stem phenotypes may expose cellular attributes necessary to navigate the metastatic cascade, such as homing to a premetastatic niche and the contributory features of proliferation, quiescence, senescence and cell cycle status. Moreover, interrogation of CHC and CTC phenotypes may illuminate evolving tumor heterogeneity and facilitate recognition of treatment resistant disease, and therefore be clinically leveraged as a liquid biopsy.

Cellular heterogeneity and the tumor microenvironment are at the forefront of translational research, as treatment failure is increasingly thought to result from evolution of resistant populations within the cellularly diverse neoplasm. However, the extent to which circulating tumor-derived cells recapitulate the heterogeneity within tumor tissue remains unclear. We demonstrate that tumor tissue is spatio-temporally variable with a heterogeneous collection of tumor cells and this can be appreciated in the phenotypes of CHCs. Our comparative evaluation of tumor tissue and CHC phenotypes was performed with a metastatic hilar lymph node, as prior treatment precluded our ability to obtain primary breast tumor tissue. While the lack of primary tumor tissue is a limitation to our study, our findings highlight that tumor heterogeneity extends to metastatic sites and traditional biopsy methods are unlikely to illustrate the complete tableau of a patient's disease. Additionally, our study showcases that neoplastic-immune cell hybrids are not limited to the primary tumor; however, we are cautious in generalizing the presence of hybrid cells within metastatic tumor tissue due to limited sampling, the rarity of tumor hybrids and the qualitative nature of the analysis. Importantly, these findings do support the potential power of CHCs to aid comprehensive tumor analysis and monitor changes in disease. Indeed, our highly multiplexed immunofluorescence analyses of tumor and peripheral blood samples provide an unprecedented spatial resolution of cellular heterogeneity that allowed for identification of diverse populations of hybrid cells in tumors and in circulation. These observations, taken with the ubiquitous nature of CHCs in cancer and the robust tumor-initiating capacity of hybrid cells in vivo, suggest that CHCs are a functioning population of cells relevant to the metastatic spread of cancer.

Only a minority of primary tumor cells functionally contribute to metastatic seeding. It is therefore unsurprising that phlebotomy may yield a more selective biopsy by enriching for biologically active tumor-derived cells. These advantages over standard tissue biopsy indicate an opportunity to monitor tumor evolution through serial blood draws. The abundance of CHCs relative to CTCs and the rarity of co-positive cells in healthy subjects supports their relevance in diagnostic and disease monitoring strategies. Further, the presence of oncogenic *KRAS* mutations in PDAC derived CHCs suggests their potential utility as a noninvasive analyte of disease biology. Finally, our data support the concept of the enriched liquid biopsy, as evidenced by our findings of robust stem signatures in breast cancer CHCs that far exceed previous reports of stem phenotypes within primary tumor tissue⁶³.

In conclusion, we demonstrate that immune-neoplastic cell hybrids represent a heterogeneous population of cells that disseminate into circulation as CHCs and is a generalizable phenomenon in human solid malignancies.

Neoplastic hybrids harbor genetic hallmarks of parental tumor cells, display stem markers, and recapitulate tumorigenesis to a greater degree than unfused cancer cells. Finally, we reveal that CHC phenotypes reflect the heterogeneity of functional protein expression from cancer tissue. This novel population is deserving of further study in an effort to understand the extent to which it can be leveraged as a liquid biomarker and in treatments to forestall metastatic progression.

Materials and methods

Human samples and ethics statement. All experimental protocols were approved by the Oregon Health & Science University (OHSU) Institutional Review Board and informed consent was obtained from all subjects. All experiments were performed in accordance with relevant guidelines and regulations. Peripheral blood was obtained from cancer patients at OHSU with non-small cell lung cancer, esophageal cancer (ECA), pancreatic ductal adenocarcinoma (PDAC), pancreatic neuroendocrine tumor (PNET), breast cancer, ovarian carcinoma, colon cancer, rectal cancer, adult and pediatric glioma, uveal melanoma, head and neck squamous cell carcinoma (HNSCC), cholangiocarcinoma, ampullary carcinoma, and healthy subjects (n=5 each site, Table 1, Table S1). Additional specimen from n=27 patients with untreated breast cancer and n=1 patient with relapsed refractory breast cancer were analyzed (Table 2, Table S2).

Flow cytometric analyses of CHCs and CTCs in human peripheral blood. Patient peripheral blood was collected in heparinized vacutainer tubes and diluted 1:2 with phosphate buffered saline (PBS). Peripheral blood mononuclear cells (PBMCs) were isolated either using RBC lysis or using density centrifugation with Ficoll-Paque PLUS. RBC lysis was performed with a 45-min incubation with Dextran T500 (Pharmacosmos, Denmark, PBS, 3% Dextran 0.1% Sodium Azide), then the top fraction was centrifuged, and the pellet subjected to a 1-min incubation in 0.2% NaCl followed by addition of the equivalent volume of 1.6% NaCl). Density centrifugation was performed by adding 12 mL Ficoll at the bottom of a conical tube containing 10 mLs of blood and PBS and centrifuging for 20 min at 800g with no brake. Isolated PBMCs were counted on a Countess Automated Cell Counter and resuspended in FACS Buffer (PBS, 1.0 mM EDTA, 2% FBS) to a concentration of 5×10^7 cells/mL and 10^7 cells were then prepared for antibody staining. All staining for flow cytometry was completed on ice, and cells were pelleted with centrifugation at 300g for 5 min after each step. Cells were incubated in PBS containing Live Dead Aqua with Fc Receptor Binding Inhibitor for 20 min. Cells were then incubated in FACS buffer for 30 min on ice with CD45, EPCAM, or NKI-Beteb. For evaluation of stem properties, cells were incubated with EpCAM, CD45, CD31-FITC, CD44 and CD24 or EpCAM, CD44, CD24, and CD45. To prepare cells for intracellular staining, they were incubated with eBioscience Fixation/Permeabilization solution for 30 min and washed with eBioscience Permeabilization Buffer. Cells were incubated in 200 μ L Permeabilization Buffer with pan-Cytokeratin for 30 min (Table S3). A BD LSRFortessa and Aria Fusion (Becton Dickinson, NJ, USA) FACS machine was used for sample analyses. The gates were established with single color and unstained controls (Figures S1, S3, S4).

ddPCR detection of KRAS mutations. After PBMCs were isolated from the peripheral blood of a patient with PDAC, circulating cell populations were collected by FACS. Utilizing the same methodology described for flow cytometry, PBMCs were stained with CD45, EpCAM, ECAD, and ITGA3. Using a BD FACSAria™ Fusion (BD Biosciences, CA, USA) cell sorter, 580 CHCs and 7000 normal leukocytes were isolated into FACS buffer. CHCs were defined by CD45 positivity and staining of any tumor marker (EpCAM, ECAD, ITGA3) either in singularly or in combination, while normal leukocytes were defined as CD45⁺/ECAD⁻/EpCAM⁻/ITGA3⁻. To achieve minimal cell requirements for DNA extraction, 2000 normal leukocytes were spiked into the CHC sample, leaving 5000 cells in the normal leukocyte sample. All samples were equilibrated to a total volume of 50 μ L and then DNA was extracted using Zymo Quick-DNA Microprep Kit (Zymo Research, CA). Procedure was performed using modifications to the manufacturer's protocol. Briefly, samples were incubated in proteinase K for a minimum of 30 min, 15 μ Ls of the sample were heated to 65 °C and applied to the extraction column, followed by a 10-min incubation, repeated twice. ddPCR was performed using the ddPCR™ KRAS G12/G13 Screening Kit #1863506 (Bio-Rad Laboratories, CA) which detects 7 KRAS mutations (KRAS p.G12A, p.G12C, p.G12D, p.G12R, p.G12S, p.G12V, and p.G13D), as well as wild type KRAS. Droplets were generated with the Auto Droplet Generator (Bio-Rad Laboratories, CA) and measured on the QX200™ Droplet Reader (Bio-Rad Laboratories, CA). Manufacturer recommended parameters for PCR were followed (95 °C for 10 min, followed by 40 cycles of 94 °C for 30 s and 55 °C for 1 min, followed by a final 98 °C heat treatment of 10 min for enzyme deactivation). Mutant and wild type KRAS thresholds were set to $\geq 99\%$ of mutant (A549 cell line) and wild type (A375 cell line) controls were positive and 100% of buffer alone specimens were negative (Table S4).

In situ detection and quantification of CHCs and CTCs from human peripheral blood. PBMCs were isolated from peripheral blood using density centrifugation with Ficoll-Paque PLUS as previously described and resuspended in FACS Buffer. Cells were then adhered to poly-D-lysine-coated slides through incubation at 37 °C for 15 min, permeabilized with Triton-X, and fixed with 4% PFA. Slides were stained with antibodies directed at cancer-specific antigens and CD45 (Table S3) and with DAPI. Slides were imaged using a Zeiss AxioObserver.Z1 light microscope, digitally scanned with a Zeiss AxioScanner.Z1, and analyzed using Zeiss Zen blue software (Carl Zeiss AG, Germany).

Manual quantification was performed for randomly selected slide regions containing > 50,000 nuclei by individuals blinded to the clinical status of the patients or healthy controls. Thresholds for positivity were set off histograms of the unstained portions of the slides. Cells with DAPI nuclear staining were evaluated for CD45, CK, GFAP, CHGA/SYP status. At least 50,000 cells per patient were enumerated. CTCs were defined as tumor-marker

(CK, GFAP, CHGA/SYP) positive, CD45 negative. CHCs were defined as cells with both tumor-marker and CD45 staining. Enumerated cells were normalized to 50,000 nuclei.

Statistical analysis of enumerated CHCs and CTCs. Fisher's exact test was performed for comparisons of CHC and CTC proportions for each specific disease site and between patients and healthy controls. Additional analyses were performed using Fisher's one-tailed *t* test, using CHC and CTC levels normalized to 500,000 live cells (for flow cytometry) and 50,000 nuclei (for immunohistochemistry) as numbers for each patient or control. Fisher's exact test was chosen as the primary test because values of zero do not change with normalization. All analyses were performed using SPSS 26 (IBM, New York).

Mice. All mouse experimental protocols were approved by the OHSU Institutional Animal Care and Use Committee and performed in accordance with relevant guidelines and regulations. All methods are reported in accordance with the ARRIVE Guidelines 2.0. Mice were housed in a specific pathogen-free environment under strictly controlled light cycle conditions, fed a standard rodent Lab Chow (#5001 PMI Nutrition International, St. Louis, MO), and provided water *ad libitum*. The following strains were used in the described studies: C57BL/6J (JAX #000664), FVB/N-Tg(MMTV-PyVT)634Mul/J (MMTV:PyMT; JAX #022974)⁶⁴ backcrossed onto C57BL/6J for 10 generations, B6.Cg-Tg(CAG-mRFP)1F1Hadj/J (Cag-RFP; JAX #005884)⁶⁵ and Tg(act-EGFP)Y01Osb (Act-GFP; JAX #006567)⁶⁶. Female mice were exclusively used for this study.

Generation of in vivo-derived hybrid cells. For isolation of in vivo-derived hybrids for assessment of hybrid cell tumor-initiating capacity, MMTV:PyMT mice were crossed onto a Cag-RFP background to yield PyMT:RFP mice. Female PyMT:RFP mice were allowed to age until mammary tumors measured 1 cm³, approximately 100 days. Tumors were resected, diced, and digested for 30 min at 37 °C in DMEM + 2 mg/mL Collagenase A + DNase under stirring conditions. Digested tumor was filtered through a 40 µm filter and washed with PBS. Dissociated single cells (1 × 10⁶) were injected into the mammary fat pad of recipient Act-GFP mice (n = 9). Tumors were allowed to grow until they reached a range of 1–2 cm³ in volume. Tumors were dissected and processed for FACS-isolation of hybrid cells, or a small specimen of each tumor processed for immunohistochemical analyses, as previously described. Tissue sections were stained with antibodies to E-cadherin to identify neoplastic hybrids within the tumor.

Analyses of tumor-initiating capacity of hybrids and unfused tumor cells. For FACS-isolation of hybrid cells and unfused tumor cells, tumors were harvested as described above, dissociated to single cells, and subjected to FACS-isolation by direct fluorescence on a Becton Dickinson InFlux sorter. To assess tumor-initiating capacity, 2500 double-positive RFP⁺/GFP⁺ cells (hybrids), or singly-positive RFP⁺/GFP⁻ cells (unfused tumor cells) were injected into the mammary fat pad of wild type recipient mice (technical replicates, n = 3, each). Tumor growth was monitored and measured when tumors became palpable. Mice were sacrificed when tumor reached 2 cm³ in diameter. A second round of tumor cell injections were conducted with 25,000 and 250,000 unfused tumor cells (n = 3–4).

Immunohistochemical analyses of tumor tissue. Formalin-fixed paraffin-embedded (FFPE) human tissues were sectioned at 4–5 microns and mounted on charged slides (Tanner Adhesive Slides, Mercedes Medical, TNR WHT45AD). The slides were baked overnight in a Robbin Scientific oven at 55 °C and an additional 30 min at 65 °C. Tissues were deparaffinized with xylene and rehydrated with graded ethanol baths. Two step antigen retrieval was performed in the Biocare Medical Decloaking Chamber Pro using the following settings: set point 1 (SP1), 125 °C, 30 s; SP2: 90 °C, 30 s; SP limit: 10 °C. Slides were further incubated in hot pH 9 buffer for 15 min. Slides were then washed in two brief changes of diH₂O (~ 2 s) and once for 5 min in 1 × PBS, pH 7.4. Sections were blocked in 10% normal goat serum (NGS), 1% bovine serum albumin (BSA) in PBS for 30 min at 20 °C in a humid chamber, followed by PBS washes.

Primary antibodies (supplemental table S3) were diluted in 5% NGS, 1% BSA in 1 × PBS and applied overnight at 4 °C in a humidity chamber, covered with plastic coverslips (IHC World, IW-2601). Following overnight incubation, tissues were washed 3 × 10 min in 1 × PBS. Coverslips were mounted in Slowfade Gold plus DAPI mounting media.

Fluorescence microscopy. Fluorescently stained slides were scanned on the Zeiss AxioScan.Z1 (Zeiss, Germany) with a Colibri 7 light source (Zeiss). The filter cubes used for image collection were DAPI (Semrock, LED-DAPI-A-000), AF488 (Zeiss 38 HE), AF555 (Zeiss 43 HE), AF647 (Zeiss 50) and Alexa Fluor 750 (AF750, Chroma 49007 ET Cy7). The exposure time was determined individually for each slide and stain to ensure good dynamic range but not saturation. Full tissue scans were taken with the 20 × objective (Plan-Apochromat 0.8NA WD = 0.55, Zeiss) and stitching was performed in Zen Blue image acquisition software (Zeiss).

Quenching fluorescence signal. After successful scanning, slides were soaked in 1 × PBS for 10–30 min in a glass Coplin jar, waiting until glass coverslip slid off without agitation. Quenching solution containing 20 mM sodium hydroxide (NaOH) and 3% hydrogen peroxide (H₂O₂) in 1 × PBS was freshly prepared from stock solutions of 5 M NaOH and 30% H₂O₂, and each slide placed in 10 mL quenching solution. Slides were quenched under incandescent light, for 30 min for FFPE tissue slides and 20 min for PBMCs adhered to glass slides. Slides were then removed from chamber with forceps and washed three times for two minutes in 1 × PBS. The next round of primary antibodies was applied, diluted in blocking buffer as previously described, and imaged.

ing and quenching were repeated over ten rounds for FFPE tissue slides, and two rounds for PBMCs adhered to glass slides.

Digital quantification and analysis of FFPE tissue cyclic immunofluorescence. Each image acquired during the cyCIF assay was registered based on DAPI features acquired from each round of staining⁶⁷. In-house software⁶⁸ was used to generate nuclear, cell and membrane segmentation masks by classifying pixels on the basis of a combination of marker expression to identify cells and membranes, respectively. Extracted single-cell features included centroids and mean intensity of each marker from its biologically-relevant segmentation mask, e.g. Ecad_Ring, Ki67_Nuclei. The last round DAPI image was used to filter out cells lost during each round of cyCIF staining. For downstream analysis, first intensity normalization is performed for each biopsy sample based on RESTORE (robust intensity normalization method)⁶⁹ to minimize intensity variation across samples. Then, a heatmap was constructed using unsupervised clustering (k-means clustering with the number of clusters; n = 20).

Quantification and analysis of PBMC cyclic immunofluorescence. Digitally scanned slides as described above were processed using Zeiss blue software. Regions of interest were created and conserved between round of staining and quenching. Thresholds for positivity were set off histograms of the unstained portions of the slides. Region of interest were registered using the image correlation feature, registered cell overlays were used for phenotypic profiling at the single cell level. Cells with DAPI nuclear staining were evaluated for CD45, CK, ECAD, ER, AR, Ki67 and CD44 status. At least 50,000 cells were enumerated. CTCs were defined as cells with positive tumor-marker staining with no CD45 staining compared to other PBMCs. CHCs were defined as cells with both tumor-marker staining and strong CD45 staining.

Data availability

No datasets were generated or analyzed during the current study.

Received: 17 March 2021; Accepted: 9 June 2021

Published online: 01 July 2021

References

- Chaffer, C. L. & Weinberg, R. A. A perspective on cancer cell metastasis. *Science* **331**, 1559–1564. <https://doi.org/10.1126/science.1203543> (2011).
- Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* **144**, 646–674. <https://doi.org/10.1016/j.cell.2011.02.013> (2011).
- Valastyan, S. & Weinberg, R. A. Tumor metastasis: Molecular insights and evolving paradigms. *Cell* **147**, 275–292. <https://doi.org/10.1016/j.cell.2011.09.024> (2011).
- Pantel, K. & Brakenhoff, R. H. Dissecting the metastatic cascade. *Nat. Rev. Cancer* **4**, 448–456. <https://doi.org/10.1038/nrc1370> (2004).
- Ashworth, T. R. A case of cancer in which cells similar to those in the tumours were seen in the blood after death. *Aust. Med. J.* **14**, 146 (1869).
- Cohen, S. J. *et al.* Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer. *J. Clin. Oncol.* **26**, 3213–3221. <https://doi.org/10.1200/jco.2007.15.8923> (2008).
- Cristofanilli, M. *et al.* Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N. Engl. J. Med.* **351**, 781–791. <https://doi.org/10.1056/NEJMoa040766> (2004).
- Krebs, M. G. *et al.* Evaluation and prognostic significance of circulating tumor cells in patients with non-small-cell lung cancer. *J. Clin. Oncol.* **29**, 1556–1563. <https://doi.org/10.1200/jco.2010.28.7045> (2011).
- Court, C. M. *et al.* Circulating tumor cells predict occult metastatic disease and prognosis in pancreatic cancer. *Ann. Surg. Oncol.* **25**, 1000–1008. <https://doi.org/10.1245/s10434-017-6290-8> (2018).
- Danila, D. C. *et al.* Circulating tumor cell number and prognosis in progressive castration-resistant prostate cancer. *Clin. Cancer Res.* **13**, 7053–7058. <https://doi.org/10.1158/1078-0432.Ccr-07-1506> (2007).
- Sullivan, J. P. *et al.* Brain tumor cells in circulation are enriched for mesenchymal gene expression. *Cancer Discov.* **4**, 1299–1309. <https://doi.org/10.1158/2159-8290.Cd-14-0471> (2014).
- Palmieri, G. *et al.* Prognostic value of circulating melanoma cells detected by reverse transcriptase-polymerase chain reaction. *J. Clin. Oncol.* **21**, 767–773. <https://doi.org/10.1200/jco.2003.01.128> (2003).
- Smerage, J. B. *et al.* Circulating tumor cells and response to chemotherapy in metastatic breast cancer: SWOG S0500. *J. Clin. Oncol.* **32**, 3483–3489. <https://doi.org/10.1200/jco.2014.56.2561> (2014).
- Gast, C. E. *et al.* Cell fusion potentiates tumor heterogeneity and reveals circulating hybrid cells that correlate with stage and survival. *Sci. Adv.* **4**, 7828. <https://doi.org/10.1126/sciadv.aat7828> (2018).
- Rizvi, A. Z. *et al.* Bone marrow-derived cells fuse with normal and transformed intestinal stem cells. *Proc. Natl. Acad. Sci. USA* **103**, 6321–6325. <https://doi.org/10.1073/pnas.0508593103> (2006).
- Powell, A. E. *et al.* Fusion between Intestinal epithelial cells and macrophages in a cancer context results in nuclear reprogramming. *Can. Res.* **71**, 1497–1505. <https://doi.org/10.1158/0008-5472.CAN-10-3223> (2011).
- Davies, P. S., Powell, A. E., Swain, J. R. & Wong, M. H. Inflammation and proliferation act together to mediate intestinal cell fusion. *PLoS One* **4**, e6530. <https://doi.org/10.1371/journal.pone.0006530> (2009).
- Bakhoum, S. F. *et al.* Chromosomal instability drives metastasis through a cytosolic DNA response. *Nature* **553**, 467–472. <https://doi.org/10.1038/nature25432> (2018).
- Laughney, A. M. *et al.* Regenerative lineages and immune-mediated pruning in lung cancer metastasis. *Nat. Med.* **26**, 259–269. <https://doi.org/10.1038/s41591-019-0750-6> (2020).
- Whiteside, T. L. Tumor-derived exosomes and their role in cancer progression. *Adv. Clin. Chem.* **74**, 103–141. <https://doi.org/10.1016/bs.acc.2015.12.005> (2016).
- Cheng, K. S. *et al.* ALICE: A hybrid AI paradigm with enhanced connectivity and cybersecurity for a serendipitous encounter with circulating hybrid cells. *Theranostics* **10**, 11026–11048. <https://doi.org/10.7150/thno.44053> (2020).
- Peitzsch, C., Tyutyunnykova, A., Pantel, K. & Dubrovskaya, A. Cancer stem cells: The root of tumor recurrence and metastases. *Semin. Cancer Biol.* **44**, 10–24. <https://doi.org/10.1016/j.semcancer.2017.02.011> (2017).

23. Steinbichler, T. B. *et al.* Cancer stem cells and their unique role in metastatic spread. *Semin. Cancer Biol.* **60**, 148–156. <https://doi.org/10.1016/j.semcancer.2019.09.007> (2020).
24. Li, F., Tiede, B., Massagué, J. & Kang, Y. Beyond tumorigenesis: Cancer stem cells in metastasis. *Cell Res.* **17**, 3–14. <https://doi.org/10.1038/sj.cr.7310118> (2007).
25. Miranda, A. *et al.* Cancer stemness, intratumoral heterogeneity, and immune response across cancers. *Proc. Natl. Acad. Sci. USA* **116**, 9020–9029. <https://doi.org/10.1073/pnas.1818210116> (2019).
26. Ng, S. W. *et al.* A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature* **540**, 433–437. <https://doi.org/10.1038/nature20598> (2016).
27. Smith, B. A. *et al.* A human adult stem cell signature marks aggressive variants across epithelial cancers. *Cell Rep.* **24**, 3353–3366. <https://doi.org/10.1016/j.celrep.2018.08.062> (2018).
28. Todaro, M. *et al.* CD44v6 is a marker of constitutive and reprogrammed cancer stem cells driving colon cancer metastasis. *Cell Stem Cell* **14**, 342–356. <https://doi.org/10.1016/j.stem.2014.01.009> (2014).
29. Wang, L., Zuo, X., Xie, K. & Wei, D. The role of CD44 and cancer stem cells. *Methods Mol. Biol.* **1692**, 31–42. https://doi.org/10.1007/978-1-4939-7401-6_3 (2018).
30. Yan, Y., Zuo, X. & Wei, D. Concise review: Emerging role of CD44 in cancer stem cells: A promising biomarker and therapeutic target. *Stem Cells Transl. Med.* **4**, 1033–1043. <https://doi.org/10.5966/sctm.2015-0048> (2015).
31. Basakran, N. S. CD44 as a potential diagnostic tumor marker. *Saudi Med. J.* **36**, 273–279. <https://doi.org/10.15537/smj.2015.3.9622> (2015).
32. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. USA* **100**, 3983–3988. <https://doi.org/10.1073/pnas.0530291100> (2003).
33. Dalerba, P., Cho, R. W. & Clarke, M. F. Cancer stem cells: Models and concepts. *Annu. Rev. Med.* **58**, 267–284. <https://doi.org/10.1146/annurev.med.58.062105.204854> (2007).
34. Wicha, M. S., Liu, S. & Dontu, G. Cancer stem cells: An old idea—a paradigm shift. *Cancer Res.* **66**, 1883–1890. <https://doi.org/10.1158/0008-5472.Can-05-3153> (2006) ((discussion 1895–1886)).
35. Kaigorodova, E. V. *et al.* Heterogeneity of circulating tumor cells in neoadjuvant chemotherapy of breast cancer. *Molecules* <https://doi.org/10.3390/molecules23040727> (2018).
36. Strati, A., Nikolaou, M., Georgoulas, V. & Lianidou, E. S. Prognostic significance of TWIST1, CD24, CD44, and ALDH1 transcript quantification in EpCAM-positive circulating tumor cells from early stage breast cancer patients. *Cells* <https://doi.org/10.3390/cells8070652> (2019).
37. Watanabe, T. *et al.* Circulating tumor cells expressing cancer stem cell marker CD44 as a diagnostic biomarker in patients with gastric cancer. *Oncol. Lett.* **13**, 281–288. <https://doi.org/10.3892/ol.2016.5432> (2017).
38. Okumura, T. *et al.* Circulating tumor cells detected by the expression of cancer stem cell marker CD90 and CD44 in patients with esophageal cancer. *Int. Surg.* <https://doi.org/10.9738/INTSURG-D-16-00022.1> (2017).
39. Lee, C. H. *et al.* Baseline circulating stem-like cells predict survival in patients with metastatic breast cancer. *BMC Cancer* **19**, 1167. <https://doi.org/10.1186/s12885-019-6370-1> (2019).
40. Blassl, C. *et al.* Gene expression profiling of single circulating tumor cells in ovarian cancer—establishment of a multi-marker gene panel. *Mol. Oncol.* **10**, 1030–1042. <https://doi.org/10.1016/j.molonc.2016.04.002> (2016).
41. Müller, C. *et al.* Hematogenous dissemination of glioblastoma multiforme. *J. Sci. Transl. Med.* **6**, 47ra101. <https://doi.org/10.1126/scitranslmed.3009095> (2014).
42. Kyriakopoulos, G., Mavroei, V., Chatzellis, E., Kaltsas, G. A. & Alexandraki, K. I. Histopathological, immunohistochemical, genetic and molecular markers of neuroendocrine neoplasms. *Ann. Transl. Med.* **6**, 252. <https://doi.org/10.21037/atm.2018.06.27> (2018).
43. Lun, M., Lok, E., Gautam, S., Wu, E. & Wong, E. T. The natural history of extracranial metastasis from glioblastoma multiforme. *J. Neurooncol.* **105**, 261–273. <https://doi.org/10.1007/s11060-011-0575-8> (2011).
44. Liu, J. *et al.* Critical role of oncogenic KRAS in pancreatic cancer (Review). *Mol. Med. Rep.* **13**, 4943–4949. <https://doi.org/10.3892/mmr.2016.5196> (2016).
45. Waters, A. M. & Der, C. J. KRAS: The critical driver and therapeutic target for pancreatic cancer. *Cold Spring Harb. Perspect. Med.* <https://doi.org/10.1101/cshperspect.a031435> (2018).
46. Khan, A. Q. *et al.* RAS-mediated oncogenic signaling pathways in human malignancies. *Semin. Cancer Biol.* **54**, 1–13. <https://doi.org/10.1016/j.semcancer.2018.03.001> (2019).
47. Gast, C. E., Shaw, A. K., Wong, M. H. & Coussens, L. M. Surgical procedures and methodology for a preclinical murine model of de novo mammary cancer metastasis. *J. Visual. Exp.* <https://doi.org/10.3791/54852> (2017).
48. Savelieva, O. E. *et al.* Heterogeneity of stemlike circulating tumor cells in invasive breast cancer. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms21082780> (2020).
49. Singh, S. K. *et al.* Identification of human brain tumour initiating cells. *Nature* **432**, 396–401. <https://doi.org/10.1038/nature03128> (2004).
50. Zhou, B. B. *et al.* Tumour-initiating cells: Challenges and opportunities for anticancer drug discovery. *Nat. Rev. Drug Discov.* **8**, 806–823. <https://doi.org/10.1038/nrd2137> (2009).
51. Sutton, T. L., Walker, B. S. & Wong, M. H. Circulating hybrid cells join the fray of circulating cellular biomarkers. *Cell. Mol. Gastroenterol. Hepatol.* **8**, 595–607. <https://doi.org/10.1016/j.jcmgh.2019.07.002> (2019).
52. Lin, J.-R. *et al.* Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *Elife* **7**, e31657. <https://doi.org/10.7554/eLife.31657> (2018).
53. Eng, J. *et al.* Cyclic Multiplexed-Immunofluorescence (cmIF), a Highly Multiplexed Method for Single-Cell Analysis 521–562 (Humana, 2020).
54. Giovannelli, P. *et al.* The androgen receptor in breast cancer. *Front. Endocrinol. (Lausanne)* **9**, 492–492. <https://doi.org/10.3389/fendo.2018.00492> (2018).
55. Dörnen, J., Sieler, M., Weiler, J., Keil, S. & Dittmar, T. Cell fusion-mediated tissue regeneration as an inducer of polyploidy and aneuploidy. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms21051811> (2020).
56. Miroshnychenko, D. *et al.* Spontaneous cell fusions as a mechanism of parasexual recombination in tumour cell populations. *Nat. Ecol. Evol.* <https://doi.org/10.1038/s41559-020-01367-y> (2021).
57. Kulemann, B. *et al.* Pancreatic cancer: Circulating tumor cells and primary tumors show heterogeneous KRAS mutations. *Sci. Rep.* **7**, 4510. <https://doi.org/10.1038/s41598-017-04601-z> (2017).
58. Biankin, A. V. *et al.* Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* **491**, 399–405. <https://doi.org/10.1038/nature11547> (2012).
59. Lennerz, J. K. & Stenzinger, A. Allelic ratio of KRAS mutations in pancreatic cancer. *Oncologist* **20**, e8–9. <https://doi.org/10.1634/theoncologist.2014-0408> (2015).
60. Muzumdar, M. D. *et al.* Survival of pancreatic cancer cells lacking KRAS function. *Nat. Commun.* **8**, 1090–1090. <https://doi.org/10.1038/s41467-017-00942-5> (2017).
61. Windon, A. L. *et al.* A KRAS wild type mutational status confers a survival advantage in pancreatic ductal adenocarcinoma. *J. Gastrointest. Oncol.* **9**, 1–10. <https://doi.org/10.21037/jgo.2017.10.14> (2018).

62. Vignery, A. Osteoclasts and giant cells: Macrophage–macrophage fusion mechanism. *Int. J. Exp. Pathol.* **81**, 291–304. <https://doi.org/10.1111/j.1365-2613.2000.00164.x> (2000).
63. Abraham, B. K. *et al.* Prevalence of CD44⁺/CD24⁻/low cells in breast cancer may not be associated with clinical outcome but may favor distant metastasis. *Clin. Cancer Res.* **11**, 1154–1159 (2005).
64. Guy, C. T., Cardiff, R. D. & Muller, W. J. Induction of mammary tumors by expression of polyomavirus middle T oncogene: A transgenic mouse model for metastatic disease. *Mol. Cell. Biol.* **12**, 954–961. <https://doi.org/10.1128/mcb.12.3.954> (1992).
65. Long, J. Z., Lackan, C. S. & Hadjantonakis, A. K. Genetic and spectrally distinct in vivo imaging: Embryonic stem cells and mice with widespread expression of a monomeric red fluorescent protein. *BMC Biotechnol.* **5**, 20. <https://doi.org/10.1186/1472-6750-5-20> (2005).
66. Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T. & Nishimune, Y. “Green mice” as a source of ubiquitous green cells. *FEBS Lett.* **407**, 313–319 (1997).
67. Young Hwan, C. *et al.* Deep learning based Nucleus Classification in pancreas histological images. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2017**, 672–675. <https://doi.org/10.1109/embc.2017.8036914> (2017).
68. Eng, J. *et al.* Cyclic multiplexed-immunofluorescence (cmIF), a highly multiplexed method for single-cell analysis. *Methods Mol. Biol.* **2055**, 521–562. https://doi.org/10.1007/978-1-4939-9773-2_24 (2020).
69. Chang, Y. H. *et al.* RESTORE: Robust intEnSiTy nORmalization mEthod for multiplexed imaging. *Commun. Biol.* **3**, 111. <https://doi.org/10.1038/s42003-020-0828-1> (2020).

Acknowledgements

We acknowledge members of the OHSU Flow Cytometry Shared Resource, and Advanced Light Microscopy Core at The Jungers Center for technical assistance. We also acknowledge support from the OHSU Cancer Early Detection and Research (CEDAR) department for patient identification. Figure illustrations created in part with Biorender.com.

Author contributions

Conceptualization, M.S.D., T.L.S., B.S.W., C.E.G., L.Z., and M.H.W.; methodology, M.S.D., T.L.S., B.S.W., C.E.G., J.E., K.C., E.B., and Y.H.C.; formal analysis, M.S.D., T.L.S., B.S.W., C.E.G., E.B., Y.H.C., J.E., K.C., and M.H.W.; investigation, M.S.D., T.L.S., B.S.W., C.E.G., L.Z., J.R.S., J.E., M.P., K.L., A.S., E.B., A.S., J.M.M., A.G., and Y.C.; resources, J.M.M., A.S., Y.C., D.R.C., S.J.P., J.P.C., J.M.F., K.T., J.J.J., R.F.P., B.C.S., A.H.S., V.L.T., S.C.M., C.D.L., J.W.G., G.M. and Z.M.; data curation, S.K.S.; writing—original draft preparation, M.S.D., T.L.S., B.S.W., and M.H.W.; writing—review and editing, A.H.S., M.S.D., T.L.S., B.S.W., and M.H.W.; visualization, B.S.W., S.K.S., M.H.W., and M.S.D.; all authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by National Center for Research Resources (NCRR), Grant number S10-RR023432. National Cancer Institute of the National Institutes of Health, U54CA112970, and P30 CA069533. Additional support to JWJG: NIH/NCI Grants U54CA209988 and U2CCA233280, to KC: Small Business Innovation Research (SBIR) Grant R44CA224994 and Prospect Creek Foundation, to CEG: NIH/T32GM071388 and NCI/T32CA106195, to TLS: the Medical Research Foundation, the Collins Medical Trust, to MSD: NIH/NCATS/TL1TR002371, Medical Research Foundation, and Umpqua Bank Innovation in Pediatric Cancer Award, to BSW the Collins Medical Trust and Medical Research Foundation, to AHS: Melanoma Research Foundation, American Association for Cancer Research, P30 EY010572, and unrestricted departmental funding from Research to Prevent Blindness, to MHW the OHSU Center for Women’s Health Circle of Giving Grant, the Michael J. Newton Esophageal Cancer Foundation, the Prospect Creek Foundation, the Brenden-Colson Center for Pancreatic Health, NIH/NCI grants R01CA118235, R21CA172334, and Department of Defense CA170613 W81XWH-18-1-0621.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-93053-7>.

Correspondence and requests for materials should be addressed to M.H.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021