A microfluidic device for label-free, physical capture of circulating tumor cell clusters

A Fatih Sarioglu[1–3,7,8], Nicola Aceto[2,4,8], Nikola Kojic[1,3], Maria C Donaldson[2], Mahnaz Zeinali[1,2], Bashar Hamza[1], Amanda Engstrom[2], Huili Zhu[2], Tilak K Sundaresan[2], David T Miyamoto[2,5], Xi Luo[2], Aditya Bardia[2,4], Ben S Wittner[2], Sridhar Ramaswamy[2,4], Toshi Shioda[2], David T Ting[2,4], Shannon L Stott[1,2,4], Ravi Kapur[1], Shyamala Maheswaran[2,3], Daniel A Haber[2,4,6] & Mehmet Toner[1,3]

Cancer cells metastasize through the bloodstream either as single migratory circulating tumor cells (CTCs) or as multicellular groupings (CTC clusters). Existing technologies for CTC enrichment are designed to isolate single CTCs, and although CTC clusters are detectable in some cases, their true prevalence and significance remain to be determined. Here we developed a microchip technology (the Cluster-Chip) to capture CTC clusters independently of tumor-specific markers from unprocessed blood. CTC clusters are isolated through specialized bifurcating traps under low–shear stress conditions that preserve their integrity, and even two-cell clusters are captured efficiently. Using the Cluster-Chip, we identified CTC clusters in 30–40% of patients with metastatic breast or prostate cancer or with melanoma. RNA sequencing of CTC clusters confirmed their tumor origin and identified tissue-derived macrophages within the clusters. Efficient capture of CTC clusters will enable the detailed characterization of their biological properties and role in metastasis.

The isolation and analysis of rare CTCs holds great promise for providing insight into blood-borne metastasis, as well as monitoring cancer response noninvasively following therapeutic interventions. Considerable progress has been made in the development of devices to capture one tumor cell among a billion normal blood cells[1,2]. Among these, microfluidic technologies have the important advantage of combining high-throughput processing with low-shear and efficient isolation[3–6] of unfixed cells, which are readily subjected to molecular and functional analyses[7].

Metastasis can result from single cancer cells that acquire a migratory epithelial-to-mesenchymal transition (EMT) phenotype, and as early as the 1950s, it was also suggested to be mediated by groupings of cells that appear to break off from primary tumors[8,9]. Such circulating tumor emboli have been reported in both mouse models[10] and human blood specimens[4,7,11–15], ranging from large thrombi or blood clots carrying tumor cells to clumps of tumor cells mixed with reactive stromal cells[16]. Moreover, intravenously injected tumor cell clusters show higher metastasis initiation capability in the mouse compared to otherwise identical single cells[8,9]. In patients with metastatic cancer, presence of CTC clusters has recently been associated with a poor prognosis[17,18].
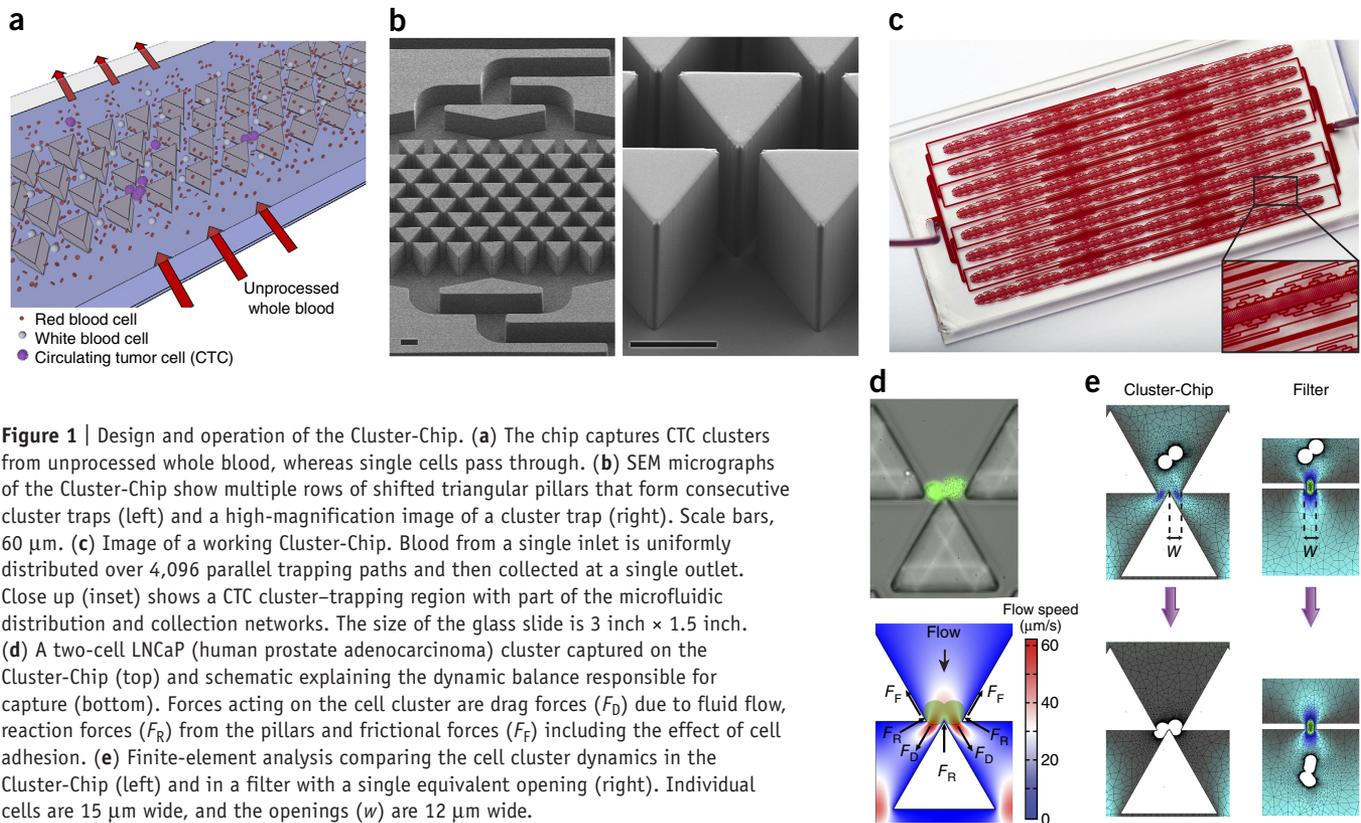
Existing isolation technologies are designed to capture single CTCs and may fail to preserve the integrity of CTC clusters or to sort them in a reliable manner. Here we introduce the Cluster-Chip, a microfluidic chip designed to specifically isolate CTC clusters from unprocessed patient blood samples with high sensitivity. The Cluster-Chip exploits the unique geometries of cellular aggregates to differentiate CTC clusters from single cells in blood. This chemistry-free approach enables specific and label-free isolation of CTC clusters from patients with different cancer types, as well as the release of CTC clusters following their capture, allowing for downstream molecular and functional assays.

RESULTS

Design of the Cluster-Chip

The Cluster-Chip captures CTC clusters by relying on the strength of cell-cell junctions as clusters flow at physiological speed through a set of triangular pillars (**Fig. 1a,b**). Three pillars make up the basic unit of the chip; two form a narrowing channel that funnels the cells into an opening, where the edge of the third pillar is positioned to bifurcate the laminar flow. As blood flows, single blood and tumor cells divert to one of the two streamlines at the bifurcation and pass through the 12 μm × 100 μm opening (**Fig. 1a**). In contrast, CTC clusters are held at the edge of the bifurcating pillar, even if they are deformable enough to squeeze through either one of the openings (**Fig. 1a,d**). This is because the bifurcating edge retains the captured cluster in both streamlines simultaneously under a dynamic force balance—cell-cell junctions support a stable equilibrium (not possible for a single cell) across the bifurcating

Figure 1 | Design and operation of the Cluster-Chip. (**a**) The chip captures CTC clusters from unprocessed whole blood, whereas single cells pass through. (**b**) SEM micrographs of the Cluster-Chip show multiple rows of shifted triangular pillars that form consecutive cluster traps (left) and a high-magnification image of a cluster trap (right). Scale bars, 60 μm. (**c**) Image of a working Cluster-Chip. Blood from a single inlet is uniformly distributed over 4,096 parallel trapping paths and then collected at a single outlet. Close up (inset) shows a CTC cluster–trapping region with part of the microfluidic distribution and collection networks. The size of the glass slide is 3 inch × 1.5 inch. (**d**) A two-cell LNCaP (human prostate adenocarcinoma) cluster captured on the Cluster-Chip (top) and schematic explaining the dynamic balance responsible for capture (bottom). Forces acting on the cell cluster are drag forces (F_D) due to fluid flow, reaction forces (F_R) from the pillars and frictional forces (F_F) including the effect of cell adhesion. (**e**) Finite-element analysis comparing the cell cluster dynamics in the Cluster-Chip (left) and in a filter with a single equivalent opening (right). Individual cells are 15 μm wide, and the openings (w) are 12 μm wide.

pillar, which serves as a fulcrum (**Fig. 1d,e**). The three-pillar unit is repeated in multiple rows for redundancy (**Fig. 1b,c**).

To ensure that CTC clusters do not dissociate, the chip is optimized to handle cellular aggregates with flow speeds (and therefore shear forces) well below those found in human capillaries. The peak flow speed of ~70 μm/s at the bifurcation (**Fig. 1d**) is much lower than that of existing microfluidic and filter-based CTC isolation technologies[19,20]. Yet the chip can interrogate clinical blood specimens at a rate of 2.5 ml/h owing to its highly parallel architecture (**Fig. 1c**).

The Cluster-Chip is also purposely designed to capture two-cell clusters. Our previous analysis showed that ~92% of CTC clusters are oligoclonal, including the majority of two-cell clusters, and characterized by an elevated metastatic potential compared to that of single CTCs[18].

The Cluster-Chip capture principle is fundamentally different from that of filter-based technologies such as porous membranes or microfluidic traps: it captures cell clusters regardless of their deformability, thus allowing capture of CTC clusters that might otherwise squeeze through a smaller pore (**Fig. 1e** and **Supplementary Fig. 1**). Moreover, captured clusters are retained under a dynamic force balance, unlike with filters where cells are retained primarily owing to their surface tension and are exposed to damaging stresses[20]. Finally, because single cells are not trapped by the Cluster-Chip, this technology enables clog-free processing of whole-blood samples.

Chip characterization and optimization using cell lines

To characterize the device, we first spiked artificially formed clusters of fluorescently labeled MDA-MB-231 human breast cancer cells into unlabeled blood samples from healthy donors and

processed the mixture using the Cluster-Chip (Online Methods). To test for chip-mediated cell aggregation, we introduced a 1:1 mixture of GFP-tagged single cells and mCherry-tagged clusters (2–30 cells) into whole blood and performed Cluster-Chip capture (**Fig. 2a**). No GFP-tagged cells were retained or adhered to clusters, a result consistent with our probabilistic analysis based on the rarity of CTCs in blood and the large number of uniformly distributed traps (**Supplementary Fig. 2**).

To quantify capture efficiency, we imaged and counted clusters captured on the Cluster-Chip as well as those that flowed through undetected (**Supplementary Fig. 3**). The efficiency increased with the size of cluster and decreased with increasing flow rate (**Fig. 2b**). Reduced capture at high flow rates may result from the failure of the dynamic force to balance smaller clusters and from the breaking up of captured clusters under higher shear stress. We therefore selected 2.5 ml/h as the optimal flow rate providing both high capture efficiency and high throughput (**Supplementary Video 1**). At 2.5 ml/h, the Cluster-Chip captured 169/171 (99%) MDA-MB-231 clusters of at least four cells, 28/40 (70%) of three-cell clusters and 48/117 (41%) of two-cell clusters (**Fig. 2b**). Distribution of captured two-cell and three-cell clusters to subsequent rows indicates that the capture efficiency for small clusters can further be improved by adding more rows (**Supplementary Fig. 4**). Also, the effectiveness of the CTC-cluster trap in capturing larger clusters (four cells or more) at 2.5 ml/h flow rate was confirmed, as captured clusters were predominantly (>80%) found in the first row.

To investigate possible damage to MDA-MB-231 clusters at this flow rate, we used microscopy to compare the cluster size distribution in the spiked population before processing with that of the captured population (**Supplementary Fig. 5** and Online

Figure 2 | Characterization of the Cluster-Chip using cell lines spiked in whole blood. (**a**) Procedure used to test for aggregation of single GFP-labeled cells (green) on the chip. Scale bar, 60 μm. (**b**) Capture efficiency at different flow speeds, measured at 4 °C to minimize cell adhesion, using artificial clusters of MDA-MB-231 cells spiked in whole blood. 100% capture efficiency corresponds to cases in which no clusters were detected in the waste (Online Methods). (**c**) Comparison of cell cluster size distribution in unprocessed (input) and captured populations. The experiment was performed at room temperature to maximize the capture efficiency of small clusters. Error bars, s.e.m. (*n* = 3 independent experiments). (**d**) Comparison of MDA-MB-231 cluster capture efficiency using the Cluster-Chip or membrane filters operated under different pressures. Effective whole-blood processing rate for each condition is noted in parentheses. Error bars, s.e.m. (*n* = 3 independent experiments). (**e**) Comparison of cluster capture efficiency of Cluster-Chip with immunoaffinity-based HB-Chip for three human breast cancer cell lines. Surface EpCAM expression is highest in MCF7 and lowest in LBX1. Error bars, s.e.m. (*n* = 4 (MCF7); *n* = 3 (MDA-MB-231 and LBX1)). The capture efficiencies for each cell type are normalized by the mean Cluster-Chip capture efficiency.
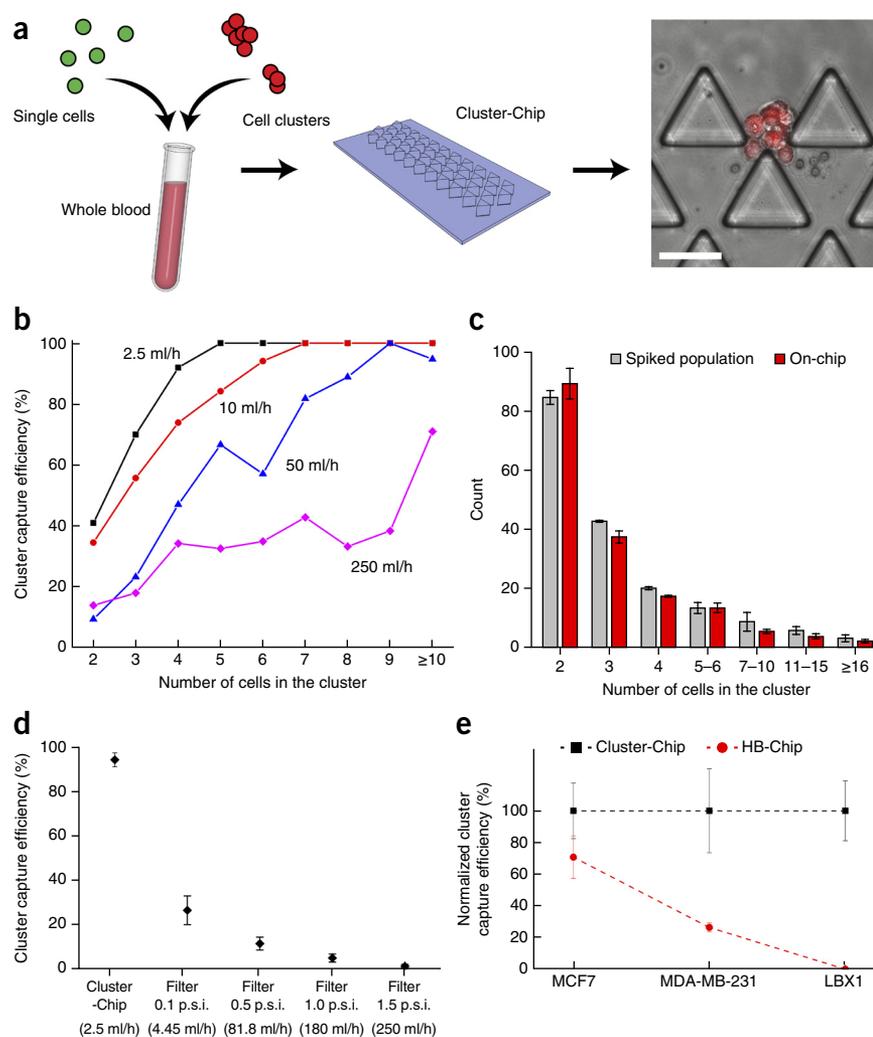
Methods). We found that the Cluster-Chip preserved the integrity of captured clusters (**Fig. 2c**). Moreover, more than 95% of captured clusters were in the first row, thereby excluding the possibility of damage due to bifurcating cluster traps.

The Cluster-Chip outperforms other methods

To compare the Cluster-Chip with filter-based isolation, we processed whole blood spiked with MDA-MB-231 clusters using the Cluster-Chip and membrane filters with 5-μm pore diameter. The blood was processed using the Cluster-Chip at 2.5 ml/h at room temperature. For filtration, samples were diluted 1:10 (v/v), as undiluted whole blood led to clogging. In direct comparisons, the Cluster-Chip had higher cluster capture efficiency than membrane filters operated at different filtration pressures (0.1–1.5 p.s.i.) (**Fig. 2d**). At 1.5 p.s.i., the pressure typically used for filtration, we found that cluster capture was inefficient. At 0.1 p.s.i., which is substantially lower than pressures typically used for filtration, cluster capture efficiency increased to ~26%. At 0.1 p.s.i., the effective whole-blood processing rate of the filter was comparable to that of the Cluster-Chip, so further reduction was not practical (**Fig. 2d**).

Single-cell capture efficiencies reported for membrane filters[19,20] are higher than the cluster capture efficiency in our experiments, which may be for two reasons. First, clusters are more likely to be lost during the wash step than are single cells, which are partially or fully squeezed in pores. Second, we counted only intact clusters, excluding some that were damaged (**Supplementary Fig. 6**). Nevertheless, the operational conditions for filters were not optimized for clusters in our

experiments, and we cannot exclude the possibility that by changing some of the parameters, capture rate of viable clusters could increase.

To compare the Cluster-Chip with another microfluidic platform, we processed matched specimens through the herringbone-chip (HB-Chip)[4], an antibody-based capture chamber that we had first employed in detecting clusters[4]. We used three cell lines representing varying EpCAM expression levels observable across CTCs: MCF-7, representing high EpCAM-expressing epithelial breast cancer cells; MDA-MB-231, mesenchymal 'triple-negative' breast cancer cells with low EpCAM expression; and MCF10A-LBX1 (ref. 21), an EMT-induced breast cell line with virtually absent EpCAM expression. In direct comparisons, the Cluster-Chip had 50% and 400% higher efficiency than the anti-EpCAM–coated HB-Chip in capturing MCF7 and MDA-MB-231 clusters, respectively (**Fig. 2e**). MCF10A-LBX1 clusters were not captured by the HB-Chip but were readily identified in the Cluster-Chip (1,000-fold differential capture) (**Fig. 2e**). The results affirm that the reliance on physical properties rather than on specific cancer cell surface epitopes to isolate clusters makes the Cluster-Chip uniquely suited to study a range of cancer types. These include epithelial cancers in which activation of EMT during cancer cell invasion triggers loss of epithelial markers, as well as non-epithelial cancers such as melanoma.

Figure 3 | Release of captured clusters from the Cluster-Chip. (**a**) Experimental setup (top) and steps of the CTC cluster release process (bottom). The bulk of the blood sample is continuously rocked at room temperature and is cooled only when processed by the Cluster-Chip. (**b**) Release efficiency of MDA-MB-231 clusters from the chip as a function of the reverse flow rate and the processing temperature. (**c**) Nonspecific binding of leukocytes on the Cluster-Chip when the sample is processed at room temperature (top) and at 4 °C (bottom). Leukocytes were fixed with 4% PFA and stained with DAPI. (**d**) Left, images of the product released in solution from the Cluster-Chip operated at room temperature (top) and at 4 °C (bottom). Right, relative purity of released cell clusters against contaminating blood cells when Cluster-Chip is operated at room temperature and 4 °C. Scale bars, 60 μm.
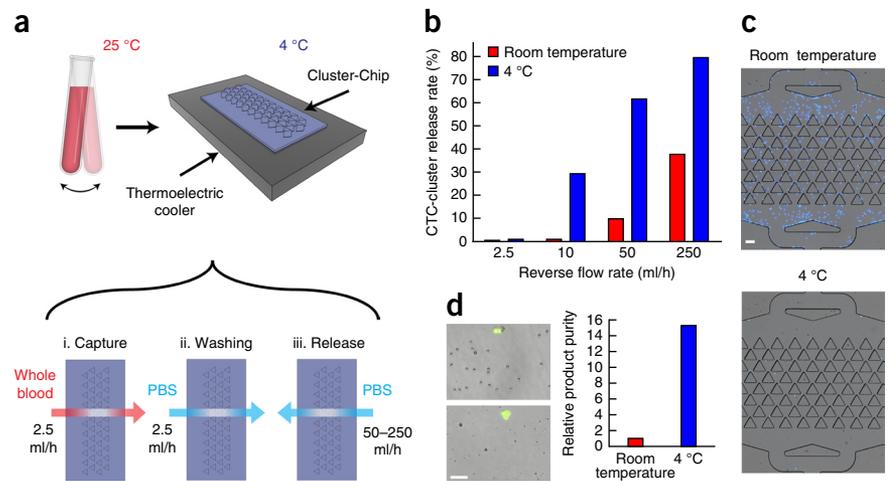
## Cluster release from the Cluster-Chip

In addition to sensitive capture, the viable release of captured CTC clusters is critically important for downstream molecular and functional assays. Despite the fact that clusters are not tethered to the Cluster-Chip through antibody-mediated interactions, we observed that release of captured CTC clusters was incomplete, especially for large clusters, after reversal of the flow. Using reverse flow rates of 2.5–250 ml/h, we succeeded in releasing only 114/308 (37%) of captured clusters. To address this problem, we tested sample processing at lower temperatures, which are known to reduce nonspecific cell adhesion[22]. We operated the Cluster-Chip on a thermoelectric cooler at 4 °C (**Fig. 3a**) so as to cool the samples transiently, thereby avoiding prolonged cold exposure—which is known to activate platelets[23]. Processing samples at 4 °C substantially improved the cluster release efficiency: 188/236 (80%) of captured clusters were released under 250 ml/h reverse flow. The improvement was particularly evident at low reverse flow speeds, which enabled the use of low shear forces, thereby enhancing release of viable cells (**Fig. 3b**). Indeed, processing at 4 °C and release of clusters under 250 ml/h reverse

flow had no notable effect on cell viability (**Supplementary Fig. 7**). We note that the Cluster-Chip technology is also compatible with nonadherent coating materials and sacrificial-layer approaches[24].

Processing blood samples at 4 °C also proved to substantially reduce nonspecific binding by contaminating blood cells. In experiments with blood samples from healthy donors, the temperature-dependent reduction in on-chip leukocyte contamination was as high as 50-fold (**Fig. 3c**). Moreover, less nonspecific binding of leukocytes to the chip translated into a 15-fold higher product purity for cold-processed samples (**Fig. 3d**), which is particularly important for downstream molecular applications.

Capture of CTC clusters from blood samples of patients

We applied the technology to blood samples collected from patients with metastatic cancer. Patients with breast cancer ($n = 27$), melanoma ($n = 20$) or prostate cancer ($n = 13$) (**Supplementary Table 1**) provided consent according to an IRB-approved protocol at Massachusetts General Hospital, and 4 ml of blood were directly processed through the Cluster-Chip, which was followed by immunofluorescence
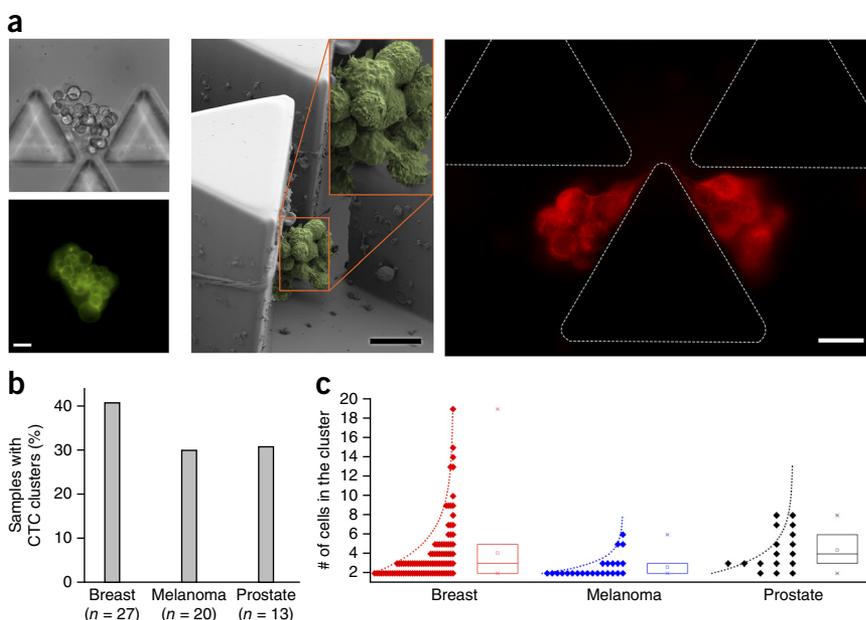

Figure 4 | Capture of CTC clusters from blood samples of patients with metastatic cancer. (**a**) Representative images of three CTC clusters isolated from patients with metastatic breast cancer. Left, bright-field and fluorescence images of a live CTC cluster stained for common breast cancer surface markers. Center, SEM micrograph of a fixed CTC cluster. Inset is ~2.1× magnified. Right, fluorescence image of a highly deformable CTC cluster stained for cytokeratin. Note that this CTC cluster is not split but is highly strained even under slow flow in the Cluster-Chip. (**b**) Percentage of patients with CTC clusters in breast, melanoma and prostate cancer. (**c**) Size distribution of CTC clusters isolated from breast, melanoma and prostate cancer patients. The box plots show the 25th, 50th and 75th percentiles for each disease type. Scale bars, 20 μm.

Figure 5 | Immunocytochemical and molecular characterization of patient CTC clusters. (**a**) Images of a Ki67⁻ and a Ki67⁺ CTC cluster stained with cytokeratin (red), Ki67 (yellow), CD45 (green) and DAPI (nuclei, blue). Arrows indicate Ki67⁺ cells within the CTC cluster. The bar graphs show the percentage of Ki67⁺ CTC clusters in this patient (left) and the percentage of Ki67⁺ cells within CTC clusters (right). (**b**) Image of a CTC cluster associated with a white blood cell (WBC). Cells were stained with cytokeratin (red), CD45 (green) and DAPI (nuclei, blue). (**c**) Images of WBC⁻ (top) and WBC⁺ (bottom) CTC clusters released from the Cluster-Chip and live stained with Texas red–conjugated antibodies against CD45, CD14 and CD16 (red). (**d**) Heat map showing expression levels of transcripts associated with CTCs, macrophages and monocytes, T cells, B cells, natural killer (NK) cells, hematopoietic stem cells (HSCs), granulocytes and platelets in 15 CTC clusters isolated at a single time point from a patient with metastatic breast cancer. RPM, reads per million. Scale bars, 20 μm.
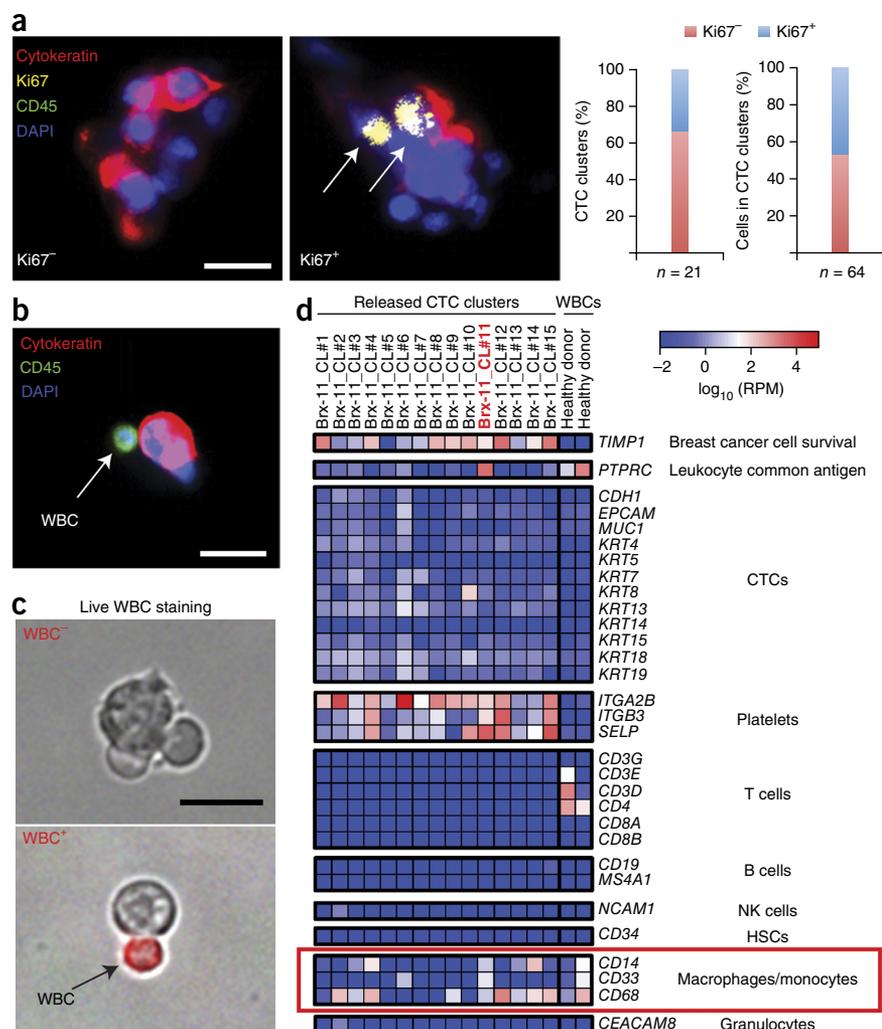
staining (Online Methods). We scored as positive those clusters that were (i) positive for specific and well-established cancer-associated markers for the disease type and (ii) negative for CD45 (a leukocyte marker). Previous studies indicate that those specific markers are very accurate for identifying bona fide cancer cells in circulation[4,7,18,25,26]. A representative CTC cluster from a patient with metastatic breast cancer is shown with on-chip capture (**Fig. 4a**) and subsequent release in solution (**Supplementary Fig. 8**). In addition to fluorescence microscopic imaging of live cells, we used scanning electron microscopy (SEM) to image fixed CTC clusters on the chip from the same patient (**Fig. 4a**). In one case, we observed a CTC cluster highly strained under even the very low flow speed in the Cluster-Chip. This case shows the extent that a CTC cluster can deform and also points to the need for the elasticity-independent capture mechanism of the Cluster-Chip, which the filter-based technologies do not have.

We identified CTC clusters in 11/27 patients with breast cancer (41%, ~0.5 clusters/ml); 6/20 patients with melanoma (30%, ~0.15 clusters/ml) and 4/13 patients with prostate cancer (31%, ~0.28 clusters/ml) (**Fig. 4b**). For some patients, multiple (two or three) blood samples were obtained, and patients scored were as positive when a CTC cluster was observed during at least one time point. The number of cells within a CTC cluster ranged from 2 to 19 cells and followed a trend toward exponential distribution (**Fig. 4c**). In a subset of our patient cohort (n = 19), we compared the number of CTC clusters captured using the Cluster-Chip with the number of single CTCs simultaneously identified using CTC-iChip[5] and found no correlation (**Supplementary Fig. 9**). When processing patient samples, we observed cell debris and fibrins, which did not interfere with chip operation owing to the large number of traps working in parallel.

Immunocytochemical and molecular analysis of patient CTC clusters

To test the versatility of the Cluster-Chip to address the biology of CTC clusters, we first measured tumor cell proliferation markers to explore intratumor cell heterogeneity and then analyzed nontumor cells that were adherent to tumor cells within the clusters. Staining for the proliferation marker Ki67 correlates well with invasiveness and disease progression in comparison with mitotic activity index or phospho-histone H3 staining[27]. In a patient with metastatic breast cancer with large numbers of single and clustered CTCs, costaining of CTC clusters for Ki67, cytokeratin (a tumor marker) and CD45 showed no notable difference between the proliferative indices of these two cell populations: 34/64 (53%) CTCs within clusters were Ki67⁺, compared with 162/439 (37%) single CTCs (**Fig. 5a** and **Supplementary Fig. 10**). Of note, 66% of CTC clusters had at least one Ki67⁺ cell (**Fig. 5a**).

The low shear stress of the Cluster-Chip also facilitates the identification of heterologous cell types[4,16,28] that may be attached to tumor cells. Given recent progress in cancer immunotherapy, the identification of leukocyte populations adherent to tumor cells in the bloodstream is of particular interest. Overall, we found nontumor cells to be rare among clusters captured using the Cluster-Chip (<5% of 60 patients). Although these cells consistently expressed the pan-leukocyte marker CD45 (**Fig. 5b**), their

precise identity is unknown. To address this question, we used the Cluster-Chip to capture CTC clusters from the blood of a breast cancer patient, released the clusters in solution, stained for leukocyte cell surface markers, and then isolated intact CTC clusters individually using a micromanipulator (**Fig. 5c**).

From a single time point, we retrieved 15 CTC clusters and performed individual RNA sequencing analysis on each (Online Methods). Expression analysis revealed that (i) all Brx-11 CTC clusters expressed low but detectable levels of transcripts encoding CTC markers such as keratins, MUC1, EpCAM and/or CDH1; (ii) 14/15 CTC clusters expressed high levels of *TIMP1*, encoding a matrix metalloproteinase widely associated with breast cancer cell survival and absent in white blood cells (WBCs)[29–32]; and (iii) all Brx-11 CTC clusters were associated with platelet transcripts, whereas control WBCs were not, a result consistent with previous reports[18,28] (**Fig. 5d**). Most CTC clusters appeared to exist in a hybrid epithelial-mesenchymal state (**Supplementary Fig. 11**), a phenotype observed in instances of breast CTCs[7] and consistent with the possibility of grouped migration. One CTC cluster associated with a WBC (cluster #11, **Fig. 5c**) expressed transcriptional signatures of tissue-derived macrophages (for example, high *CD14*, *CD33* and *CD68* expression) as well as transcripts encoding CD45, keratins, TIMP1 and platelet-associated proteins (**Fig. 5d**). No enrichment for other leukocyte subclasses—including T cells, B cells, natural killer (NK) cells, hematopoietic stem cells and granulocytes—were observed (**Fig. 5d**). Additional transcripts for platelets and epithelial cells were noted, consistent with previous reports[28].

DISCUSSION

We introduce a microfluidic technology that specifically isolates CTC clusters from unprocessed blood samples of patients with cancer. The dynamic capture of multicellular structures as they are balanced on triangular structures under low flow offers important capabilities that are not readily achieved with current CTC isolation strategies. Existing technologies primarily target single CTCs and exhibit lower sensitivity and specificity for capturing clusters. Widely used batch purification techniques[13,14,19,20,33] involve multiple processing steps that are likely to disrupt CTC clusters. High-speed fluorescence imaging of minimally enriched blood samples is efficient for CTC cluster detection[13], but the low purity complicates downstream molecular analysis. Microfluidic devices optimized to isolate single CTCs can also isolate CTC clusters[4,15]; however, substantial losses of clusters may be associated with the optimization of flow conditions for single CTC capture. Filtering blood samples through membranes with small pores[14,19,20] may be effective, but these approaches employ high flow rates, which results in extremely high shear forces. As such, CTC clusters are likely damaged or even squeeze through relatively smaller pores, as modeled in our computer simulation.

Using the Cluster-Chip, we determined that CTC clusters are heterogeneous, including both actively proliferating cells and apparently quiescent cells. Occasionally, CTC clusters are also found associated with cells of the immune system, and our RNA sequencing data revealed that these are most likely tumor-associated macrophages. This finding supports the increasingly appreciated role played by tumor-associated macrophages in cancer progression[34,35]. The fact that such tissue-derived macrophages travel with CTC clusters in the bloodstream has implications for the ability to noninvasively monitor tumor–immune cell interactions, a potentially important benefit given the increasing use of immune checkpoint blockade in the treatment of multiple types of cancer[36].

The Cluster-Chip enables label-free isolation of unfixed CTC clusters from unprocessed whole-blood specimens from patients with cancer. The reliance on structural properties of CTC clusters is particularly important given the variation in tumor epitope expression and the ability of highly flexible clusters to pass through simple pores. The ability to capture CTC clusters at relatively high frequency in patients with metastatic cancer and to release them for biological studies will enable detailed analyses of the physiological role of these clusters in the progression and metastasis of human cancer.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. NCBI Gene Expression Omnibus: GSE67939.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

AUTHOR CONTRIBUTIONS
A.F.S., N.A., S.M., D.A.H. and M.T. designed the research, analyzed the data and prepared the manuscript. N.K. performed computer simulations. M.C.D., M.Z. and A.E. processed clinical samples, performed immunofluorescence staining and scanning. B.H. manufactured devices for clinical studies. H.Z. and T.S. performed amplification and RNA sequencing. T.K.S., D.T.M., X.L. and A.B. provided clinical samples. B.S.W. performed statistical analysis on the RNA sequencing data. S.R., D.T.T., S.L.S. and R.K. commented on the manuscript.

COMPETING FINANCIAL INTERESTS
The authors declare competing financial interests: details are available in the online version of the paper.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Pantel, K., Brakenhoff, R.H. & Brandt, B. Detection, clinical relevance and specific biological properties of disseminating tumour cells. *Nat. Rev. Cancer* **8**, 329–340 (2008).
2. Yu, M., Stott, S., Toner, M., Maheswaran, S. & Haber, D.A. Circulating tumor cells: approaches to isolation and characterization. *J. Cell Biol.* **192**, 373–382 (2011).
3. Nagrath, S. *et al.* Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* **450**, 1235–1239 (2007).
4. Stott, S.L. *et al.* Isolation of circulating tumor cells using a microvortex-generating herringbone-chip. *Proc. Natl. Acad. Sci. USA* **107**, 18392–18397 (2010).
5. Ozkumur, E. *et al.* Inertial focusing for tumor antigen-dependent and -independent sorting of rare circulating tumor cells. *Sci. Transl. Med.* **5**, 179ra47 (2013).

6. Yoon, H.J. *et al.* Sensitive capture of circulating tumour cells by functionalized graphene oxide nanosheets. *Nat. Nanotechnol.* **8**, 735–741 (2013).

7. Yu, M. *et al.* Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* **339**, 580–584 (2013).

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

9. Fidler, I.J. The relationship of embolic homogeneity, number, size and viability to the incidence of experimental metastasis. *Eur. J. Cancer* **9**, 223–227 (1973).

10. Yu, M. *et al.* RNA sequencing of pancreatic circulating tumour cells implicates WNT signalling in metastasis. *Nature* **487**, 510–513 (2012).

11. Brandt, B. *et al.* Isolation of prostate-derived single cells and cell clusters from human peripheral blood. *Cancer Res.* **56**, 4556–4561 (1996).

12. Molnar, B., Ladanyi, A., Tanko, L., Sréter, L. & Tulassay, Z. Circulating tumor cell clusters in the peripheral blood of colorectal cancer patients. *Clin. Cancer Res.* **7**, 4080–4085 (2001).

13. Cho, E.H. *et al.* Characterization of circulating tumor cell aggregates identified in patients with epithelial tumors. *Phys. Biol.* **9**, 016001 (2012).

14. Krebs, M.G. *et al.* Analysis of circulating tumor cells in patients with non-small cell lung cancer using epithelial marker-dependent and-independent approaches. *J. Thorac. Oncol.* **7**, 306–315 (2012).

15. Hou, H.W. *et al.* Isolation and retrieval of circulating tumor cells using centrifugal forces. *Sci. Rep.* **3**, 1259 (2013).

16. Duda, D.G. *et al.* Malignant cells facilitate lung metastasis by bringing their own soil. *Proc. Natl. Acad. Sci. USA* **107**, 21677–21682 (2010).

17. Hou, J.M. *et al.* Clinical significance and molecular characteristics of circulating tumor cells and circulating tumor microemboli in patients with small-cell lung cancer. *J. Clin. Oncol.* **30**, 525–532 (2012).

18. Aceto, N. *et al.* Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **158**, 1110–1122 (2014).

19. Vona, G. *et al.* Isolation by size of epithelial tumor cells: a new method for the immunomorphological and molecular characterization of circulating tumor cells. *Am. J. Pathol.* **156**, 57–63 (2000).

20. Coumans, F.A.W., van Dalum, G., Beck, M. & Terstappen, L.W.M.M. Filtration parameters influencing circulating tumor cell enrichment from whole blood. *PLoS ONE* **8**, e61774 (2013).

21. Yu, M. *et al.* A developmentally regulated inducer of EMT, LBX1, contributes to breast cancer progression. *Genes Dev.* **23**, 1737–1742 (2009).

22. Rico, F., Chu, C., Abdulreda, M.H., Qin, Y. & Moy, V.T. Temperature modulation of integrin-mediated cell adhesion. *Biophys. J.* **99**, 1387–1396 (2010).

23. Kattlove, H.E. & Alexander, B. The effect of cold on platelets. I. Cold-induced platelet aggregation. *Blood* **38**, 39–48 (1971).

24. Shah, A.M. *et al.* Biopolymer system for cell recovery from microfluidic cell capture devices. *Anal. Chem.* **84**, 3682–3688 (2012).

25. Luo, X. *et al.* Isolation and molecular characterization of circulating melanoma cells. *Cell Rep.* **7**, 645–653 (2014).

26. Miyamoto, D.T. *et al.* Androgen receptor signaling in circulating tumor cells as a marker of hormonally responsive prostate cancer. *Cancer Discov.* **2**, 995–1003 (2012).

27. Inwald, E.C. *et al.* Ki-67 is a prognostic parameter in breast cancer patients: results of a large population-based cohort of a cancer registry. *Breast Cancer Res. Treat.* **139**, 539–552 (2013).

28. Labelle, M., Begum, S. & Hynes, R.O. Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. *Cancer Cell* **20**, 576–590 (2011).

29. Kuvaja, P. *et al.* Tumor tissue inhibitor of metalloproteinases-1 (TIMP-1) in hormone-independent breast cancer might originate in stromal cells, and improves stratification of prognosis together with nodal status. *Exp. Cell Res.* **318**, 1094–1103 (2012).

30. Hekmat, O. *et al.* TIMP-1 increases expression and phosphorylation of proteins associated with drug resistance in breast cancer cells. *J. Proteome Res.* **12**, 4136–4151 (2013).

31. Würtz, S.O., Schrohl, A.-S., Mouridsen, H. & Brünner, N. TIMP-1 as a tumor marker in breast cancer–an update. *Acta Oncol.* **47**, 580–590 (2008).

32. Würtz, S.Ø. *et al.* Tissue inhibitor of metalloproteinases-1 in breast cancer. *Endocr. Relat. Cancer* **12**, 215–227 (2005).

33. Talasaz, A.H. *et al.* Isolating highly enriched populations of circulating epithelial cells and other rare cells from blood using a magnetic sweeper device. *Proc. Natl. Acad. Sci. USA* **106**, 3970–3975 (2009).

34. Qian, B.-Z. & Pollard, J.W. Macrophage diversity enhances tumor progression and metastasis. *Cell* **141**, 39–51 (2010).

35. Wyckoff, J.B. *et al.* Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. *Cancer Res.* **67**, 2649–2656 (2007).

36. Pardoll, D.M. The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* **12**, 252–264 (2012).

ONLINE METHODS

Microfluidic chip fabrication. Microfluidic devices were fabricated using soft lithography. For mold fabrication, SU-8 photoresist (MicroChem) spun on a silicon wafer was patterned in the form of microfluidic channels through a chrome photomask by conventional photolithography. Poly(dimethylsiloxane) (PDMS) prepolymer and cross-linker (Sylgard 184, Dow Corning) mixed at 10:1 ratio was poured on the mold, first degassed and then cured at 65 °C for at least 4 h. The final device was built by bonding the cured PDMS peeled off from the mold and a glass substrate after surface activation in oxygen plasma. The fabricated devices were primed by flushing ethanol through microfluidic channels and then washed using deionized water and PBS before use.

Cluster-Chip design. The Cluster-Chip is composed of 4,096 parallel tracks, each equipped with seven consecutive CTC-cluster traps, comprising a cumulative flow cross-section of ~10 mm^2 at bifurcation nodes. The blood flow is uniformly distributed across these trapping tracks that are designed to be equal in hydraulic resistance through microfluidic networks. The depth of the chip is 100 μm.

Finite-element analysis simulations. Modeling studies were performed using our finite-element software package, PAK (http://fempak.fink.rs/), and run on a desktop supercomputer consisting of 32 cores (SupermicroSuperServer: 4xEight-Core Intel Xeon Processor 2.70 GHz; 512 GB total memory). The 2D finite-element computational model was developed building on our previous numerical platform[37]. The model is based on strong solid-fluid coupling, which allows the inclusion of deformable solid viscoelastic bodies (i.e., two-cell clusters) in fluid-filled channels. Specifically, we examined how the change in geometry affected transit of two-cell viscoelastic clusters through the system. The following assumptions were incorporated into the model: (i) to prevent the cell from penetrating into the solid wall, a repulsive spring-like force takes effect when the cell is within 0.75 μm of the wall; (ii) the two-cell cluster geometry was constructed by merging two cells, each with a diameter of 15 μm; (iii) the two-cell cluster was treated as an incompressible viscoelastic solid with an elastic modulus of 8 Pa, a viscous damping coefficient of 2 Pa, and density equal to water; (iv) the fluid surrounding the two-cell cluster was assumed to be water; (v) the inlet velocity on the top surface was taken to be a constant 10 μm/s, and the outlet (bottom) surface was treated as a free surface, meaning the fluid can pass through freely.

Probabilistic analysis of on-chip formation of artifact clusters. Our model assumes a blood sample with 1,000 CTCs being processed using the Cluster-Chip. Each of the 4,096 parallel cluster traps is assumed to have equal probability of receiving CTCs. We assume that whenever multiple CTCs go into the same trap, a CTC cluster is formed. On the basis of these assumptions, the probability of forming a CTC cluster of k cells was calculated using Poisson approximation

$$p(k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

where $\lambda = 1,000/4,096$.

Cluster-Chip characterization and flow speed optimization. To characterize the capture sensitivity of the Cluster-Chip, we used an analytical version of the Cluster-Chip (with 512 cluster-trapping tracks). The chip was connected to a microfluidic waste chamber that is 50 μm in height to allow us to accurately identify and analyze noncaptured MDA-MB-231 clusters in whole blood (**Supplementary Fig. 3**). We processed 250 μl of whole-blood spiked with fluorescent MDA-MB-231 clusters in each chip at 4 °C under different flow speeds.

Characterization of cell clusters before spiking. Fluorescently labeled MDA-MB-231 cells were prepared as clusters (see separate section on cell culture and reagents), and then a 2.5-μl cluster suspension was deposited on an ultralow-attachment culture dish. The cluster population was characterized by acquiring a fluorescence microscope image. The cluster suspension was then repipetted, spiked into blood samples and processed using the Cluster-Chip at room temperature. The culture dish was reimaged to account for cell clusters that possible remain attached to the surface. The captured microscope images were post-processed, and cells within each cluster were counted.

Processing of samples with membrane filters. MDA-MB-231 samples were prepared as clusters (see separate section on cell culture and reagents) and were spiked into 1:10 (v/v)-diluted blood samples collected from healthy donors. The simulated samples run through 13-mm polycarbonate track-etched membrane filters with 5-μm pore diameter (GE Whatman Nucleopore) placed in a filter folder (EMD Millipore Swinnex) using a pneumatic blood processing system under controlled pressure. Subsequently, membrane filters were removed from the holder, gently washed with PBS and deposited on a glass slide for enumeration.

SEM sample preparation and imaging. After the blood sample was processed using the Cluster-Chip, the PDMS portion containing the captured CTC clusters was separated from the glass substrate using a razor blade. The CTC clusters were first fixed in 2% glutaraldehyde and then in 1% osmium tetroxide, both diluted in 0.1 M sodium cacodylate. After fixation, cells were dehydrated in 50%, 70%, 80% and 95% ethanol (200 proof) solutions in water and 100% ethanol successively for 15 min in each. The samples were dried using a critical-point dryer (Tousimis Autosamdri-815) and then coated with Pt/Pd using a sputtering system (Cressington 208 HR). Prepared samples were imaged using a Zeiss Supra 55VP field emission scanning electron microscope at the Harvard University Center for Nanoscale Systems.

Blood sample collection and processing. The blood samples were collected from patients with metastatic breast and prostate cancers and melanoma following the MGH Institutional Review Board (IRB)-approved protocols. Healthy donor blood samples were either ordered from Research Blood Components, LLC or collected at Massachusetts General Hospital under an IRB-approved protocol. Patients and healthy donors provided consent according to an IRB-approved protocol at Massachusetts General Hospital. To minimize coagulation in the samples, tubes for blood sample collection contained the anticoagulant EDTA, and samples were processed within 4 h of withdrawal. Blood samples were run through the Cluster-Chip using a pneumatic blood-processing system[3].

Blood samples (1–10 ml) in a sealed conical tube were continuously rocked as the sample was driven through the microfluidic chip at 2.5 ml/h under constant pressure. After blood flow, the chip was washed in forward direction with PBS at 2.5 ml/h for 1 h to remove the nonspecifically bound single cells. In experiments that required release of CTC clusters, the Cluster-Chip was taped on a thermoelectric cooler (AHP1200CPV, TECA Corp.) and kept at 4 °C during all steps in the capture and release processes. Following the blood flow and forward steps describe above, the CTC clusters were released from the chip under reverse flow (up to 250 ml/h) into a Petri dish for further studies. In experiments where precise flow rate was required (cluster capture/release efficiency vs. flow rate, **Figs. 2b** and **3b**), the samples were run through the chip using syringe pumps (Harvard Apparatus Infuse/Withdraw PHD Ultra) instead of the pneumatic system described above. To quantify the efficiency of cell release from Cluster-Chip, we counted clusters that were successfully released from the chip under reverse flow as well as the ones that remained trapped in the chip. Aliquots of healthy donor blood samples spiked with fluorescent MDA-MB-231 clusters were side-by-side processed using multiple Cluster-Chips, each subjected to a different combination of blood processing temperature and reverse flow speed.

Cell cultures and reagents. MCF7, LNCaP and H1650 cells were purchased from the American Type Culture Collection (ATCC) and propagated according to the manufacturer's instructions. MDA-MB-231 cells were obtained from J. Massague (MSKCC) and propagated in DMEM (Life Technologies) supplemented with 10% FBS (Life Technologies). All cell lines were tested and found negative for *Mycoplasma* contamination. To generate single cells or clusters, we incubated cells growing in monolayer at 80% confluence with trypsin (Life Technologies) for 1 min to generate floating clusters. Clusters were then distributed equally in two separate dishes. In one of the two dishes, clusters were mechanically dissociated by pipetting to generate a single-cell suspension. The plasmid expressing GFP was obtained from C. Ponzetto (University of Torino). The plasmid expressing mCherry was purchased from Addgene. Lentiviral packaging vectors (Addgene) were used to transfect 293T cells (ATCC) and produce lentiviral particles. Infections of target cells lines was performed overnight at an MOI of 10 in growth medium containing 8 µg/ml polybrene (Thermo Scientific).

Immunofluorescence staining. Cells captured on the Cluster-Chip were fixed with 4% paraformaldehyde (PFA) and washed with PBS. Fixed cells were then permeabilized with 1% NP40 in PBS, blocked with 3% goat serum/2% bovine serum albumin (BSA), and immunostained with antibodies against wide-spectrum cytokeratin (Abcam, ab9377) and Ki67 (Life Technologies, 18-0192Z) for breast CTCs; NG2 (Abcam, ab78284), CD146 (BioLegend, 342008), TYRP-1 (Abcam, 18801) and αSMA (R&D, MAB1420) for melanoma; and PSA (Cell Signaling Technology, 5365) and wide-spectrum cytokeratin (Abcam) for prostate cancer. Antibodies against CD45 (Abcam, ab30470) were used to detect leukocytes. 4′,6-diamidino-2-phenylindole (DAPI) was used to detect nuclei. Stain-positive cells were detected using the BioView Ltd. automated imaging system (Billerica). High-resolution pictures were obtained with an upright fluorescence microscope (Eclipse 90i, Nikon). For live staining upon release of CTC clusters in solution, unfixed cells were stained with Alexa 488–conjugated antibodies against EpCAM (Cell Signaling Technology, #5198), cadherin 11 (R&D Systems, FAB17901G) and HER2 (BioLegend, 324410) to identify CTCs, as well as Texas red–conjugated antibodies against CD45 (BD Biosciences, 562279), CD14 (BD Biosciences, 562320) and CD16 (BD Biosciences, 562334) to identify contaminating white blood cells.

Micromanipulation of CTC clusters. The product released from the Cluster-Chip was collected in a 35-mm Petri dish, live stained with Texas red–labeled antibodies against the leukocyte markers CD45 (R&D Systems), CD14 (BD Biosciences) and CD16 (BD Biosciences) and then viewed using a Nikon Eclipse Ti inverted fluorescence microscope. CTC clusters were individually micromanipulated with a 10-µm transfer tip on an Eppendorf TransferMan NK 2 micromanipulator and ejected into PCR tubes containing RNA protective lysis buffer (10× PCR buffer II, 25 mM MgCl$_2$, 10% NP40, 0.1 M DTT, SUPERase-In, RNase Inhibitor, 0.5 µM UP1 primer, 10 mM dNTP and nuclease-free water) and immediately flash frozen in liquid nitrogen. Each CTC cluster was processed individually as a group of cells.

RNA amplification and sequencing. RNA samples extracted from CTC clusters were thawed on ice and incubated at 70 °C for 90 s. To generate cDNA, we treated samples with reverse transcription master mix (0.05 µl RNase inhibitor, 0.07 µl T4 gene 32 protein and 0.33 µl SuperScript III reverse transcriptase per 1× volume) and incubated them on thermocycler at 50 °C for 30 min and 70 °C for 15 min. To remove free primers, we added 1.0 µl of EXOSAP mix to each sample and then incubated the mixture at 37 °C for 30 min and inactivated at 80 °C for 25 min. Next, a 3′ poly(A) tail was added to the cDNA in each sample by incubating in master mix (0.6 µl 10× PCR buffer II, 0.36 µl 25 mM MgCl$_2$, 0.18 µl 100 mM dATP, 0.3 µl terminal transferase, 0.3 µl RNase H and 4.26 µl H$_2$O per 1× volume) at 37 °C for 15 min and inactivated at 70 °C for 10 min. A second strand of cDNA was synthesized by dividing each sample into 4 and incubating in master mix (2.2 µl 10× high-fidelity PCR buffer, 1.76 µl 2.5 mM each dNTP, 0.066 µl UP2 primer at 100 µM, 0.88 µl 50 mM MgSO4, 0.44 µl Platinum *Taq* DNA polymerase and 13.654 µl H$_2$O per 1× volume) at 95 °C for 3 min, 50 °C for 2 min and 72 °C for 10 min. PCR amplification (95 °C for 3 min and then 20 cycles of 95 °C for 30 s, 67 °C for 1 min and 72 °C for 6 min 6 s) was performed with master mix (4.1 µl 10× high-fidelity PCR buffer, 1.64 µl 50 mM MgSO$_4$, 4.1 µl 2.5 mM each dNTP, 0.82 µl AUP1 primer at 100 µM, 0.82 µl AUP2 primer at 100 µM, 0.82 µl Platinum *Taq* DNA polymerase and 6.7 µl H$_2$O per 1× volume). The four reactions of each sample were pooled and purified using the Qiagen PCR purification kit (cat. no 28106) and eluted in 50 µl EB buffer. Samples were selected by testing for genes *GAPDH*, *ACTB*, *PTPRC* (CD45), *KRT8*, *KRT18* and *KRT19* using qPCR. Each sample was again divided in 4 and a second round of PCR amplification (nine cycles of 98 °C for 3 min, 67 °C for 1 min and 72 °C for 6 min 6 s) was performed with master mix (9 µl 10× high-fidelity PCR buffer, 3.6 µl 50 mM MgSO$_4$, 13.5 µl 2.5 mM each dNTP, 0.9 µl AUP1 primer at 100 µM, 0.9 µl AUP2 primer at 100 µM, 1.8 µl Platinum *Taq* DNA polymerase and 59.1 µl H$_2$O per 1× volume). Samples were pooled and purified using Agencourt AMPure XP beads and eluted in 40 µl 1× low-TE buffer. The universal

primer 1 sequence was 5′-NH$_2$-ATATGGATCCGGCGCGCCGT CGACTTTTTTTTTTTTTTTTTTTTTTTTT-3′, and the universal primer 2 sequence was 5′-NH$_2$-ATATCTCGAGGGCGCGCC GGATCCTTTTTTTTTTTTTTTTTTTTTTTTT-3′.

Sequencing library construction. To shear the DNA using the Covaris S2 System, we added 1× low-TE buffer and 1.2 μl shear buffer to each sample. Conditions of the shearing program include: 6 cycles, 5 °C bath temperature, 15 °C bath temperature limit, 10% duty cycle, intensity of 5, 100 cycles/burst, and 60 s. Then samples were end polished at room temperature for 30 min with a master mix (40 μl 5× reaction buffer, 8 μl 10 mM dNTP, 8 μl end polish enzyme1, 10 μl end polish enzyme2 and 14 μl H$_2$O per 1× volume). DNA fragments larger than 500 bp were removed with 0.5× volumes of Agencourt AMPure XP beads. Supernatant was transferred to separate tubes. To size-select 200- to 500-bp DNA products, we added 0.3× volumes of beads, and samples were washed twice with 70% EtOH. The products were eluted in 36 μl low TE buffer. A dA tail was added to each size-selected DNA by treating with master mix (10 μl 5× reaction buffer, 1 μl 10 mM dATP and 5 μl A-tailing enzyme I per 1× volume) and incubated at 68 °C for 30 min and then cooled to room temperature. To label and distinguish each DNA sample for sequencing, we ligated barcode adaptors (5500 SOLiD 4464405) to DNA using the 5500 SOLiD Fragment Library Enzyme Module (4464413). After barcoding, samples were purified twice using the Agencourt

AMPure XP beads and eluted in 22 μl low-TE buffer. Following a round of PCR amplification (95 °C for 5 min, 12 cycles of 95 °C for 15 s, 62 °C for 15 s and 70 °C for 1 min, and finally 70 °C for 5 min), the libraries were purified with AMPure XP beads. Finally, to quantify the amount of ligated DNA, we used the SOLiD Library TaqMan Quantitation Kit to perform qPCR. Completed barcoded libraries were then subjected to emulsion PCR with template beads preparation and sequenced on the ABI 5500XL.

RNA sequencing data analysis. For determination of reads per million (RPM): color-space reads were aligned using TopHat and Bowtie1 with the no-novel-juncs argument set with human genome version hg19 and transcriptome defined by the hg19 knownGene table from genome.ucsc.edu. Reads that did not align or aligned to multiple locations in the genome were discarded. The hg19 table knownToLocusLink from genome.ucsc.edu was used to map, if possible, each aligned read to the gene whose exons the read had aligned to. The reads count for each gene was the number of reads that were so mapped to that gene. This count was divided by the total number of reads that were mapped to any gene and multiplied by 1 million to form the reads per million (RPM) count. We used RPM rather than RPKM because we noted a 3′ bias in the alignments.

37. Kojić, N. *et al.* A 3-D model of ligand transport in a deforming extracellular space. *Biophys. J.* **99**, 3517–3525 (2010).