## CANCER

# Tumor-derived circulating endothelial cell clusters in colorectal cancer

Igor Cima,[1,2] Say Li Kong,[3] Debarka Sengupta,[3] Iain B. Tan,[3,4] Wai Min Phyo,[1] Daniel Lee,[1] Min Hu,[1] Ciprian Iliescu,[1] Irina Alexander,[2,5] Wei Lin Goh,[1,6] Mehran Rahmani,[3] Nur-Afidah Mohamed Suhaimi,[1] Jess H. Vo,[1] Joyce A. Tai,[3] Joanna H. Tan,[3] Clarinda Chua,[4] Rachel Ten,[4] Wan Jun Lim,[4] Min Hoe Chew,[7] Charlotte A.E. Hauser,[2] Rob M. van Dam,[8] Wei-Yen Lim,[8] Shyam Prabhakar,[3] Bing Lim,[3] Poh Koon Koh,[1,6] Paul Robson,[3,9] Jackie Y. Ying,[1] Axel M. Hillmer,[3] Min-Han Tan[1,4,8,10]*

Clusters of tumor cells are often observed in the blood of cancer patients. These structures have been described as malignant entities for more than 50 years, although their comprehensive characterization is lacking. Contrary to current consensus, we demonstrate that a discrete population of circulating cell clusters isolated from the blood of colorectal cancer patients are not cancerous but consist of tumor-derived endothelial cells. These clusters express both epithelial and mesenchymal markers, consistent with previous reports on circulating tumor cell (CTC) phenotyping. However, unlike CTCs, they do not mirror the genetic variations of matched tumors. Transcriptomic analysis of single clusters revealed that these structures exhibit an endothelial phenotype and can be traced back to the tumor endothelium. Further results show that tumor-derived endothelial clusters do not form by coagulation or by outgrowth of single circulating endothelial cells, supporting a direct release of clusters from the tumor vasculature. The isolation and enumeration of these benign clusters distinguished healthy volunteers from treatment-naïve as well as pathological early-stage (≤IIA) colorectal cancer patients with high accuracy, suggesting that tumor-derived circulating endothelial cell clusters could be used as a means of noninvasive screening for colorectal cancer. In contrast to CTCs, tumor-derived endothelial cell clusters may also provide important information about the underlying tumor vasculature at the time of diagnosis, during treatment, and throughout the course of the disease.

## INTRODUCTION

Circulating tumor cells (CTCs) comprise a rare population of cells often found in the blood of cancer patients. They continue to undergo extensive characterization with the hopes that they will offer insights into cancer metastasis and also provide better biomarkers of disease progression and therapeutic response (1). Enumeration of CTCs has revealed prognostic information for metastatic breast cancer, colorectal cancer (CRC), and prostate cancer patients (2), whereas the data supporting CTC burden for the prediction of treatment effectiveness are less established (3, 4). The molecular profile of CTCs, rather than their quantity, might be a better approach for predicting drug responses (5).

In addition to their recognition as individual cells, CTCs have been observed as clusters of multiple cells since at least 1960 (6). Such clusters expressed mesenchymal markers in a longitudinal study of metastatic breast cancer patients (7), particularly FOXC1, which was a marker of epithelial-to-mesenchymal transition (EMT) (7, 8). Recently, CTC clusters in blood from both breast and prostate cancer patients correlated with shorter progression-free survival (9). Aceto et al. (9) further studied the formation and dynamics of CTC clusters in preclinical models, demonstrating a causative role for CTC clusters in breast cancer progression. A prospective trial in lung cancer patients confirmed that CTC clusters [also called circulating tumor microemboli (CTM)] were less apoptotic than single CTCs, and their detection was associated with poorer overall and progression-free survival (10).

Despite the high interest associated with these cells, their characterization has been relatively limited, owing to technical challenges of isolation and single-cell analysis. The clinical relevance of CTC clusters is not well defined, especially compared with that of single CTCs. To overcome these limitations, we comprehensively studied CTC clusters at the single-cell scale from a cohort of unselected CRC patients, by integrating microfiltration techniques, high-throughput DNA sequencing and RNA sequencing (RNA-seq), and computational modeling applied to clinical specimens. In CRC patients, CTC clusters were not malignant but were instead tumor-derived endothelial cells (tECs) correlating with features of the underlying tumor vasculature. Notably, the isolation and count of these clusters could distinguish healthy volunteers from preoperative, early-stage (≤IIA) CRC patients with high accuracy, indicating a potential role for circulating tEC clusters as a screening biomarker for CRC.

## RESULTS

### Isolation and retrieval of single CD45− clusters using a silicon-based microfiltration device

We originally aimed to analyze the transcriptional and genetic profiles of single CTCs in CRC patients and so developed a label-free microfiltration device, which isolated cells on the basis of size (11–13) and enabled both retrieval and downstream analysis of CTCs (Fig. 1, A to

C, and fig. S1A). From the device, 98% of the captured cells could be retrieved for downstream molecular analysis (fig. S1B). We optimized CTC enrichment and retrieval efficiency by spiking 1 ml of donor blood with 30 SW620 cells, a CRC cell line with a median diameter similar to CTCs (~12 μm) (14) (fig. S1C). We obtained an optimal trade-off between retrieval efficiency and cell purity using a flow rate of 0.25 ml/min and microsieve pore sizes of 9 to 10 μm (Fig. 1D). This resulted in >90%

SW620 retrieval efficiency with >5000-fold depletion of white blood cells, which are considered impurities (fig. S1D). Similar retrieval efficiencies were obtained using three alternative CRC cell lines of varying sizes (fig. S1, C and E).

We next filtered CTCs from a series of five CRC patients and stained the isolated cells for epithelial cell adhesion molecule (EpCAM), a canonical CTC marker, and CD45, a marker for white blood cells. We did not detect EpCAM[+] cells (Fig. 1, E and F); however, we noted the presence of large (average diameter, ~40 μm) clusters of EpCAM[−]/CD45[−] cells (Fig. 1, F to H). The isolated clusters had a cytomorphology consistent with malignancy, such as atypical nuclei, prominent nucleoli, and high nuclear-to-cytoplasmic ratio (Fig. 1H).
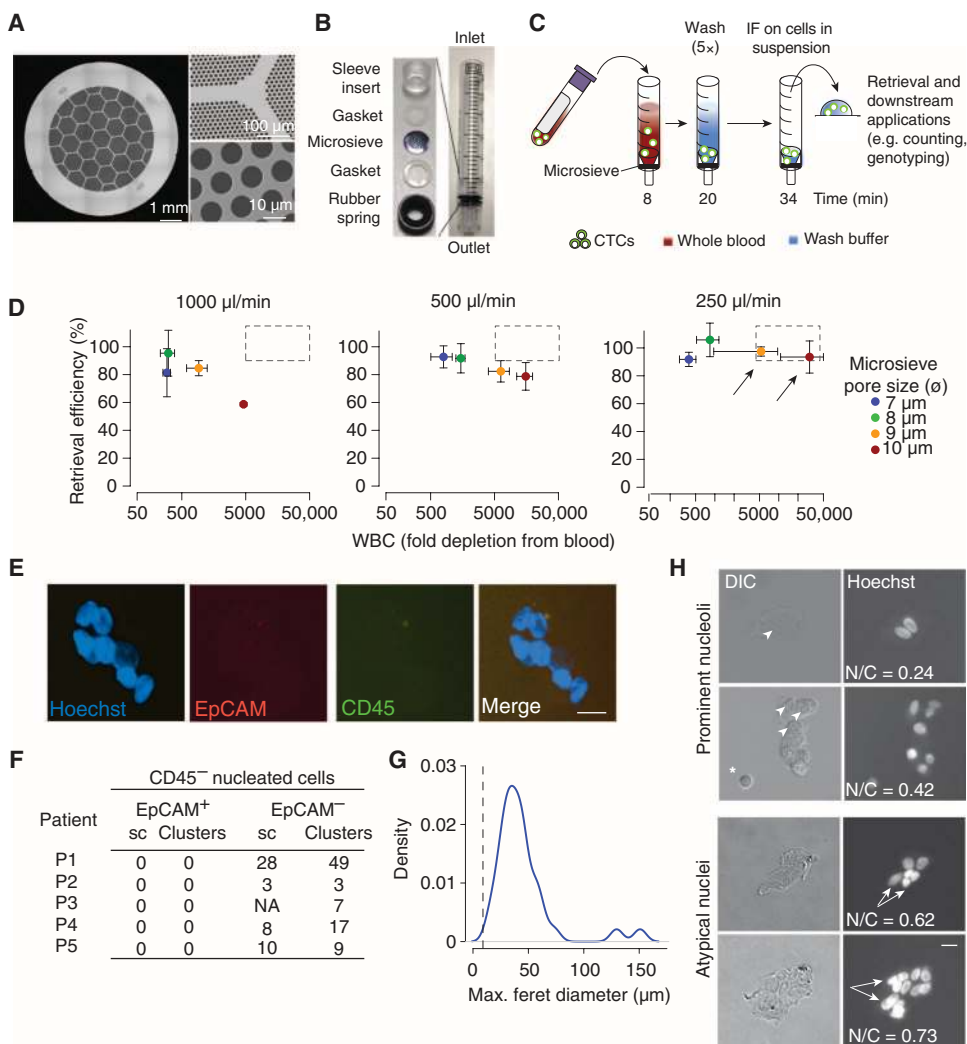


Fig. 1. **Retrieval of circulating CD45[−] cell clusters using a microfiltration device.** (**A**) Bright-field composite image (left panel) and scanning electron micrographs (right panels) of silicon microsieve. (**B**) Microfiltration device with components. (**C**) CTC microfiltration procedure for the downstream analyses. Numbers indicate the duration of each step. IF, immunofluorescence. (**D**) Optimization of retrieval efficiency and purity for downstream single-cell micromanipulation using 30 SW620 cells spiked in 1 ml of whole blood. Individual scatter-plots represent experiments using various flow rates and microsieve pore diameters. The black dashed rectangle indicates the target area of >90% retrieval efficiency and >5 × 10³ white blood cell (WBC) depletion for optimal downstream handling of retrieved cells. Data are means ± SEM (n = 3 independent experiments for each condition). (**E**) Representative immunofluorescence of CD45 and EpCAM of retrieved clusters from a CRC patient. Scale bar, 10 μm. (**F**) CD45[−] cell cluster counts for each staining and patient in the preliminary experiment. sc, single cells. (**G**) Size (diameter) distribution of CD45[−] cell clusters (n = 34 clusters). The dashed line indicates 9 μm, representing the minimal microsieve pore size used to isolate CD45[−] cell clusters. (**H**) Representative cytomorphology of CD45[−] cell clusters indicating prominent nucleoli (arrowheads), atypical nuclei (arrows), and cells with high nuclear/cytoplasmic ratio (N/C > 0.5). *White blood cell. Scale bar, 10 μm. DIC, differential interference contrast.
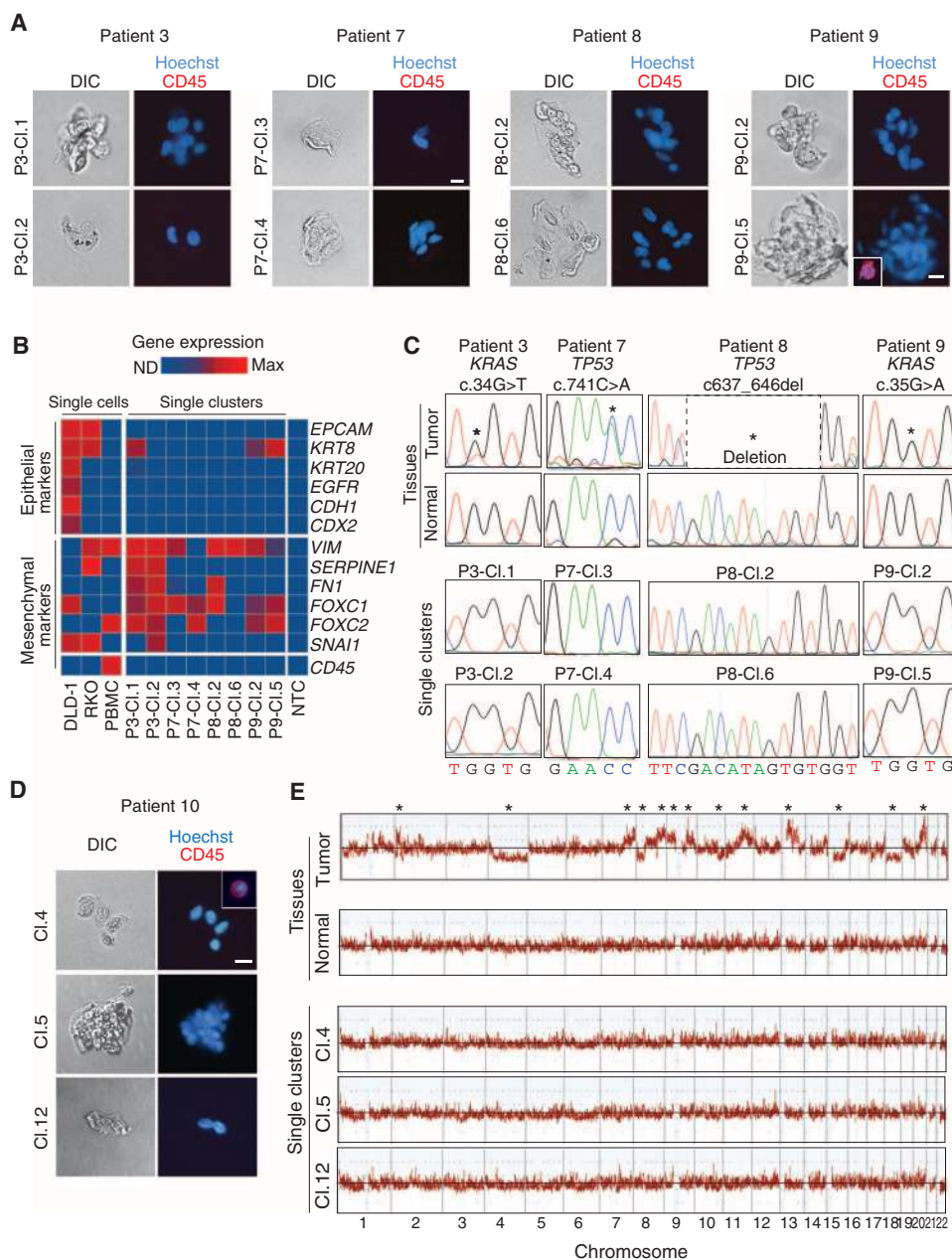
## The CD45[−] clusters express keratins and mesenchymal markers but do not mirror primary tumor mutations

CTCs have been shown to lose EpCAM expression in circulation, suggesting a drift from the epithelial phenotype (7, 8, 15). Because of their malignant cytomorphology and loss of EpCAM expression, which might indicate ongoing EMT of tumor cells (7), we hypothesized that these CD45[−] clusters harbored genetic alterations mirroring those occurring in the primary tumor. To identify DNA mutations in single cells or clusters undergoing EMT, we developed a polymerase chain reaction (PCR)–based method to simultaneously quantify RNA transcripts and detect DNA mutations, which is called single-cell RNA and mutational analysis PCR (scrmPCR) (fig. S2 and table S1). scrmPCR was first tested on known mutant alleles in an epithelial-like (DLD-1) and a mesenchymal-like (RKO) CRC cell line. The efficiency of targeted DNA amplification was about 70% for single cells and >90% for at least three cells for each amplicon (fig. S2). Eight clusters derived from four patients had mostly mesenchymal markers, including *FN1*, *SERPINE1*, and *FOXC1*, but three of the clusters also expressed the epithelial marker *KRT8* (Fig. 2, A and B). EMT marker profiles in the clusters were confirmed by staining for pan-cytokeratin (panCK) and vimentin (fig. S3A).

We sequenced the same single clusters for mutations present in the corresponding primary tumors because mutations matching the primary tumor are often found in CTCs (16). Surprisingly, all tested DNA sequence hotspots matched the wild-type alleles (Fig. 2C). We further applied targeted

high-throughput DNA sequencing to eight commonly mutated genes in DNA amplified from 15 single CTC clusters ($n = 6$ patients) and their corresponding primary tumor tissues. To avoid false-positive mutations introduced by the amplification procedure, we only considered mutations with >10% variant allele frequency (VAF) (table S2). In agreement with the scrmPCR results, mutations found in the tumor tissues were not detected in the CD45⁻ clusters (table S3). Similarly, mutations found in the clusters were not detected in the tumor tissues (table S4 and fig. S3, B and C). However, we detected multiple matching germline variants using the same method (0 of 58 matching mutations versus 12 of 28 matching germline variants; $P = 2 \times 10^{-7}$, power = 0.997, two-tailed Fisher's exact test) (table S5). We next performed array comparative genomic hybridization (aCGH) to detect chromosomal abnormalities using amplified DNA from 12 single clusters ($n = 4$ patients) because copy number variations for CTCs from lung cancer patients have been shown to mirror those of primary tumor tissue (18). However, unlike matched primary tumors, our clusters had normal cytogenetic profiles (Fig. 2, D and E, and fig. S3, D and E).

In summary, although analysis of individual CTC clusters revealed cytomorphologies and epithelial-mesenchymal marker expression similar to those of a previous report on CTCs (7), 100% (27 of 27) of single CD45⁻ clusters from nine patients did not share the genetic anomalies found in matching primary tumor tissues. In addition, DNA sequencing of 15 single CD45⁻ clusters from six patients revealed the presence of mutations that were not found in the tumor tissue (table S3 and fig. S4). These observations suggest that the source of the CD45⁻ clusters is not the tumor epithelium.
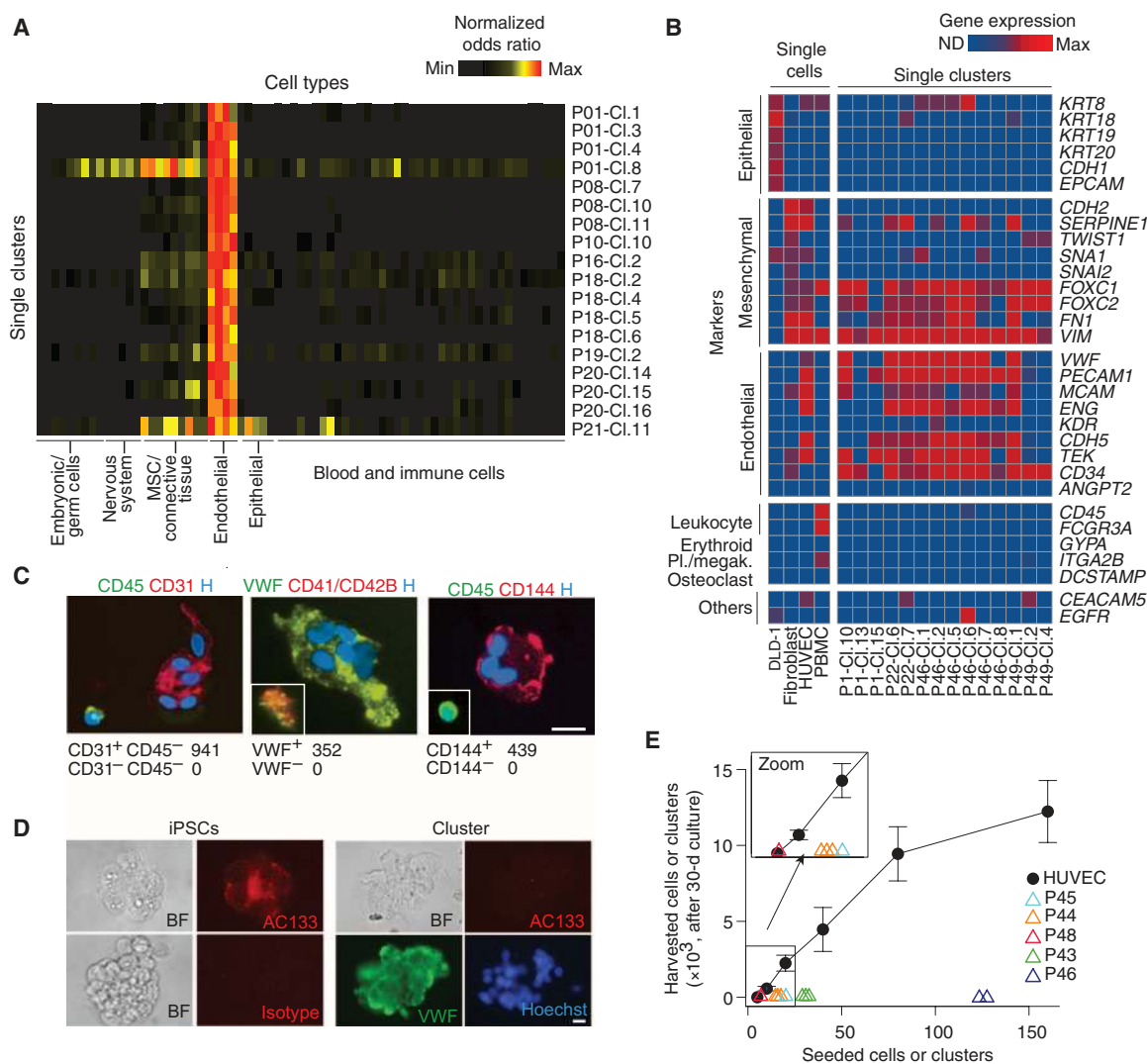
## Circulating CD45⁻ clusters are endothelial cells

We next profiled RNA expression in 18 single CD45⁻ clusters from eight patients and matched normal colon and tumor tissues using high-throughput RNA-seq (fig. S4 and tables S6 and S7). We also developed a workflow to infer 36 different cell types from RNA-seq data (fig. S5 and table S8). To this end, we first determined genes specific for each cell type (table S9 and fig. S6, A and B) and analyzed their presence in a data set made of published RNA-seq data (used as a positive control) (table S10). Using this approach, we obtained specific signals for each cell type (fig. S6, C and D). By applying this inference method to our RNA-seq data set of single CD45⁻ clusters, we found that all transcriptomes were associated with cell types of the endothelial lineage (Fig. 3A). We confirmed this



**Fig. 2. Circulating CD45⁻ cell clusters express EMT markers but do not harbor the mutations or chromosome abnormalities of matching primary tumors.** (**A**) Images of eight CD45⁻ cell clusters from four CRC patients with known primary tumor mutations before downstream scrmPCR. (**B**) Gene expression heat map of CD45⁻ cell clusters shown in (A), as well as control single cells for analysis of epithelial markers, mesenchymal markers, and *CD45*. ND, not detected; NTC, no-template control. (**C**) Chromatograms of hotspot gene sequences derived from the same single clusters, as shown in (A) and (B). Matching primary tumor and normal colon tissues (top panel**s**) were used to compare gene mutations to single cells (bottom panels). *Mutations found in tumor tissues. (**D**) Images of three clusters from a representative CRC patient with known chromosomal abnormalities. (**E**) aCGH analysis of CD45⁻ clusters shown in (D) with matching healthy and tumor tissues. Red lines indicate smoothed data calculated using Affymetrix ChAS software. *Large chromosomal abnormalities detected in the tumor sample.

result by scrmPCR in an additional 14 clusters from four patients (Fig. 3B). These results are in agreement with our findings from Fig. 2 because endothelial cells are known to express both mesenchymal markers and various keratins (classical epithelial markers) [19]. Without exception, all clusters, including those displaying a malignant cytomorphology, stained for endothelial markers such as CD31, VWF, and CD144 (Fig. 3C) but were negative for CD45 or markers of megakaryocytic lineages such as CD41 and CD42B. Thus, in our CRC patients,

all circulating CTC clusters were of endothelial origin. In addition, we did not detect single tumor cells within our clusters.

Our findings are in line with those of El-Heliebi et al. [20], reporting CD31 expression on circulating nonhematological cells (CNHCs) from kidney cancer patients, but are not consistent with a recent report describing CTC clusters of malignant origins [9]. Cell-type inference based on the single-cell scale RNA-seq data for CTC clusters in the study of Aceto et al. [9] indicated, in contrast to our data, the presence



Fig. 3. Circulating CD45⁻ cell clusters from CRC patients are endothelial cells. (A) Cell-type inference from RNA-seq data for single CD45⁻ cell clusters. Normalized odds ratios comparing number of genes enriched for each sample (rows) and cell type (columns) with random enrichment. The complete list of cell types is provided in the Supplementary Materials. MSC, mesenchymal stem cell. (B) scrmPCR gene expression in control single cells and 14 single CD45⁻ cell clusters (n = 4 patients). PBMC, peripheral blood mononuclear cell. (C) Immunofluorescence studies confirm endothelial lineage of CD45⁻ cell clusters. Representative clusters stained with the antibodies indicated and internal controls for each staining. (Inset, central panel) A CD41⁺/CD42B⁺ platelet aggregate. (Inset, right panel) A CD45⁺ white blood cell. Numbers indicate CD45⁻ cell cluster counts for the indicated immunofluorescence (n = 68, 8, and 4 patients, respectively). (D) Endothelial cell clusters do not contain EPCs. Representative image of a CD133 (AC133)–positive human induced pluripotent stem cell (iPSC) used as a positive control (left panel) and an endothelial cluster stained using the same conditions and VWF antibodies. Scale bar, 10 μm. BF, bright field. (E) EPC clonogenic assay for endothelial clusters. Scatterplot represents initial (x axis) and final (y axis) number of clusters (open colored triangles) or human umbilical vein endothelial cells (HUVECs) (black circles). Same-colored triangles are technical replicates (n = 5 patients). HUVEC data are presented as means ± SEM of eight technical replicates. One representative experiment of the two that were run is shown.

of epithelial-derived cells (fig. S7). Further experiments indicated that our circulating endothelial cell clusters were negative for CD133, an established marker for endothelial progenitor cells (EPCs) (Fig. 3D) (21). In addition, an EPC clonogenic assay (22) revealed that the endothelial clusters did not contain EPCs because live clusters failed to proliferate on a fibronectin substrate over a 30-day period (Fig. 3E). Therefore, the cellular structures characterized in our study represent a distinct population of nonproliferating endothelial cell clusters circulating in the blood of CRC patients (table S11).

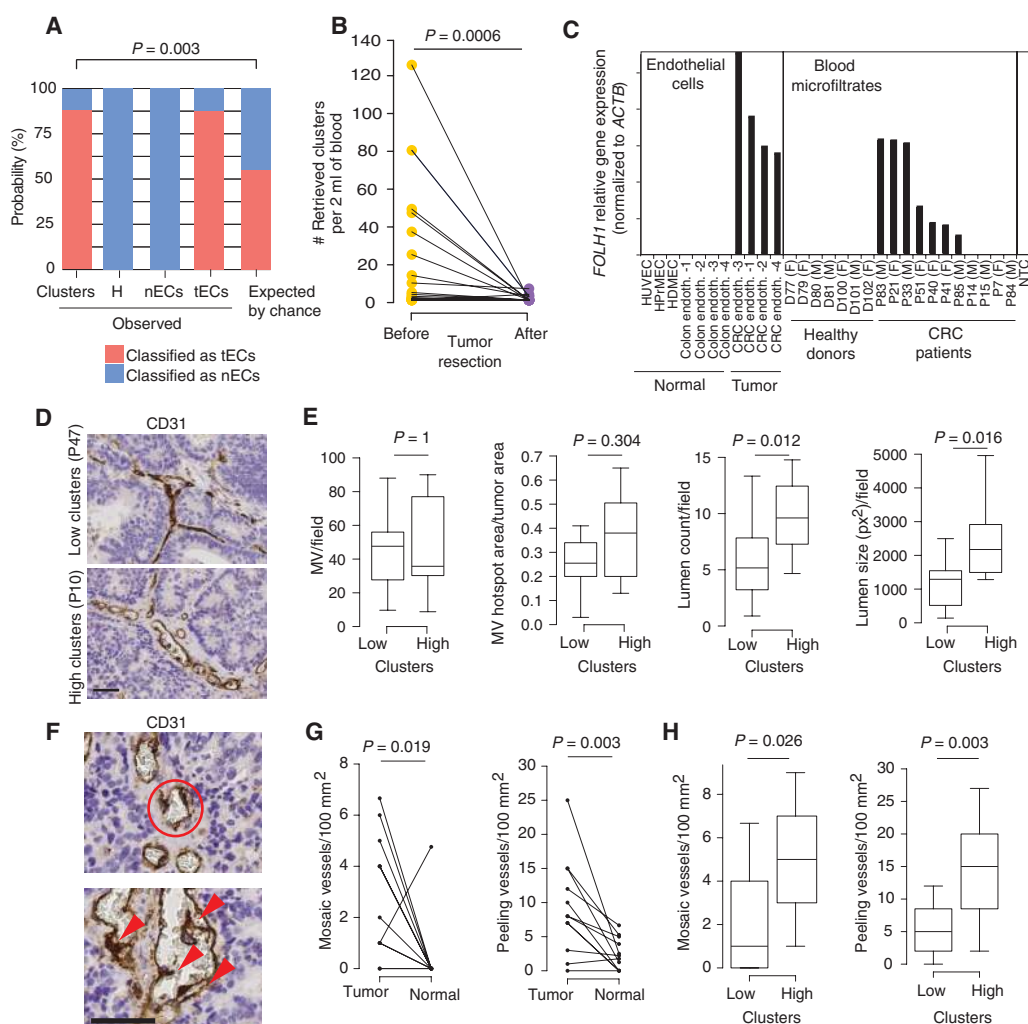## Circulating endothelial clusters are tumor-derived

Principal component analysis (PCA) revealed a stronger association of the endothelial cell cluster transcriptomes with the tumor tissues than with the normal tissue samples (fig. S8A), indicating that the clusters might be tECs. To test this hypothesis, we isolated normal endothelial cells (nECs) and

tECs from matched frozen tissue samples and profiled them by RNA-seq (fig. S8B and table S7). We also profiled the RNA expression of three primary endothelial cell types derived from the umbilical vein, normal skin, and prostate (HUVECs, HDMECs, and HPrMEC, respectively). The average expression of the endothelial cell cluster genes associated significantly with the intestinal-derived endothelial cells, but not with other nECs (fig. S8C). We subsequently classified the endothelial cell clusters on the basis of differential gene expression derived by comparing nEC with tEC transcriptomes (fig. S8D). Using a linear discriminant analysis (LDA) model derived from this comparison, we could classify 16 of 18 endothelial cell clusters as tECs and the remaining 2 clusters as nECs (Fig. 4A). This enrichment was significantly different from random classification and indicated that most circulating endothelial cell clusters belonged to the population of tECs.

If the endothelial cell clusters originated from the tumor vasculature, we would expect a drop of cluster count shortly after tumor resection.



**Fig. 4. Circulating endothelial cell clusters are tumor-derived and are associated with features of the underlying tumor vasculature.**
(**A**) Column chart stacked to 100% indicating classification of endothelial cell clusters as tECs and nECs. The left column indicates the observed probabilities for the single-cell clusters (Clusters, n = 18); the second column represents the normal endothelial samples HPrMECs, HDMECs, and HUVECs (H, n = 3); the third and fourth columns indicate the classification of nECs (n = 7) and tECs (n = 8) using leave-one-out cross-validation to assess the linear discriminant model performance (accuracy, 0.933); the right column indicates the mean probabilities expected by chance, obtained by 1000 random signatures. P value determined by two-tailed exact binomial test. (**B**) Ladder plot showing endothelial cluster counts 0 to 24 hours before and 24 to 72 hours after surgery. Lines connect data from the same patient (n = 34 paired samples from 17 patients). P value determined by two-tailed Wilcoxon signed-rank test; effect size r = 0.54. (**C**) FOLH1 expression for the indicated samples of endothelial cells and for the blood microfiltrates for the indicated donors and patients. F, female; M, male. (**D**) Representative images of CD31+ microvessels in patients with low endothelial cluster counts (≤10 clusters per 2 ml, n = 16 patients) and high endothelial cluster counts (>10 clusters per 2 ml blood, n = 12 patients). (**E**) Association of endothelial cluster counts with underlying vasculature features of matched tumors. Box plots represent the patients with low and high endothelial cluster count as described in (D). Effect sizes, r = 0.566 for lumen counts; r = 0.542 for lumen size. P values determined by two-tailed Wilcoxon-Mann-Whitney U test with Bonferroni correction. MV, microvessel; px², squared pixels. (**F**) Representative images of mosaic vessel (open circle) and peeling microvessels (arrows) stained with CD31 antibodies. (**G**) Ladder plots represent mosaic and peeling vessel counts in areas of normal and tumor tissues, respectively, on the same slide (n = 24 paired data from 12 patients). P values determined by two-tailed Wilcoxon signed-rank test. (**H**) Box plots represent mosaic and peeling vessels in the low and high endothelial cluster patients in (F). P values determined by two-tailed Wilcoxon-Mann-Whitney U test with Bonferroni correction.

To test this hypothesis, we collected paired samples from 17 CRC patients 0 to 24 hours before and 24 to 72 hours after surgical tumor resection (n = 34 samples in total). Tumor removal caused a sharp, significant decline in the numbers of circulating endothelial cell clusters, supporting the hypothesis that these structures were tumor-derived (Fig. 4B and table S12). In addition, this experiment indicates that circulating endothelial cell clusters do not originate from the peripheral circulation as outgrowth of single circulating endothelial cells but instead might be directly released as clusters from the tumor vasculature.

Folate hydrolase (FOLH1), the gene coding for prostate-specific membrane antigen (PSMA), is specifically expressed in the tumor vasculature of various cancer types but is absent in normal vasculature and peripheral blood (23, 24). FOLH1 expression was detected in CD31+CD45− cells isolated from fresh CRC tissues and in microfiltrates from the blood of 7 of 11 CRC patients, but not in endothelial cells isolated from normal tissues or in microfiltrates derived from healthy donors (Fig. 4C), providing further support for the tumor origin of our endothelial cell clusters. In addition, RNA-seq on the single circulating endothelial cell clusters indicated the expression of several published tumor endothelial markers (fig. S8E) (23, 25).

Using a CRC xenograft mouse model, we tested whether clusters of tECs were released in the circulation by transplanting tumors harboring fluorescent stroma into normal, wild-type mice and then analyzing the presence of single and multinucleated fluorescent endothelial cells in the blood by flow cytometry (fig. S9, A to D). Circulating, multinucleated green fluorescent protein (GFP)–positive endothelial cells were present in three of four mice with transplanted GFP+ xenografts but not in the control group (fig. S9E), indicating that tEC clusters were in the circulation of mice with growing tumors.

Finally, we investigated whether the number of circulating endothelial clusters correlated with features of the underlying tumor vasculature in CRC patients by analyzing blood vessels in tumor tissues isolated from patients with a low (≤10 per 2 ml of blood) or high (>10 per 2 ml of blood) endothelial cluster count. Although the median number of microvessel units did not differ between patients, the median number of lumens and their sizes were significantly higher in patients with high numbers of circulating endothelial cell clusters (Fig. 4, D and E). In addition, we noted an increase of mosaic and damaged vessels in tumor tissues compared to matched normal tissues and in patients

with high endothelial cluster counts compared to patients with low cluster counts (Fig. 4, F and G). Together, our results demonstrate that circulating endothelial clusters in CRC patients are not malignant entities but are clusters of tumor-derived, CD133−, nonproliferating endothelial cells, correlating with features of the underlying tumor vasculature.

## Tumor-derived endothelial cell clusters do not form by coagulation
CTC clusters are sometimes referred to as CTM, implying that clustering of these cells may be the result of coagulation events rather
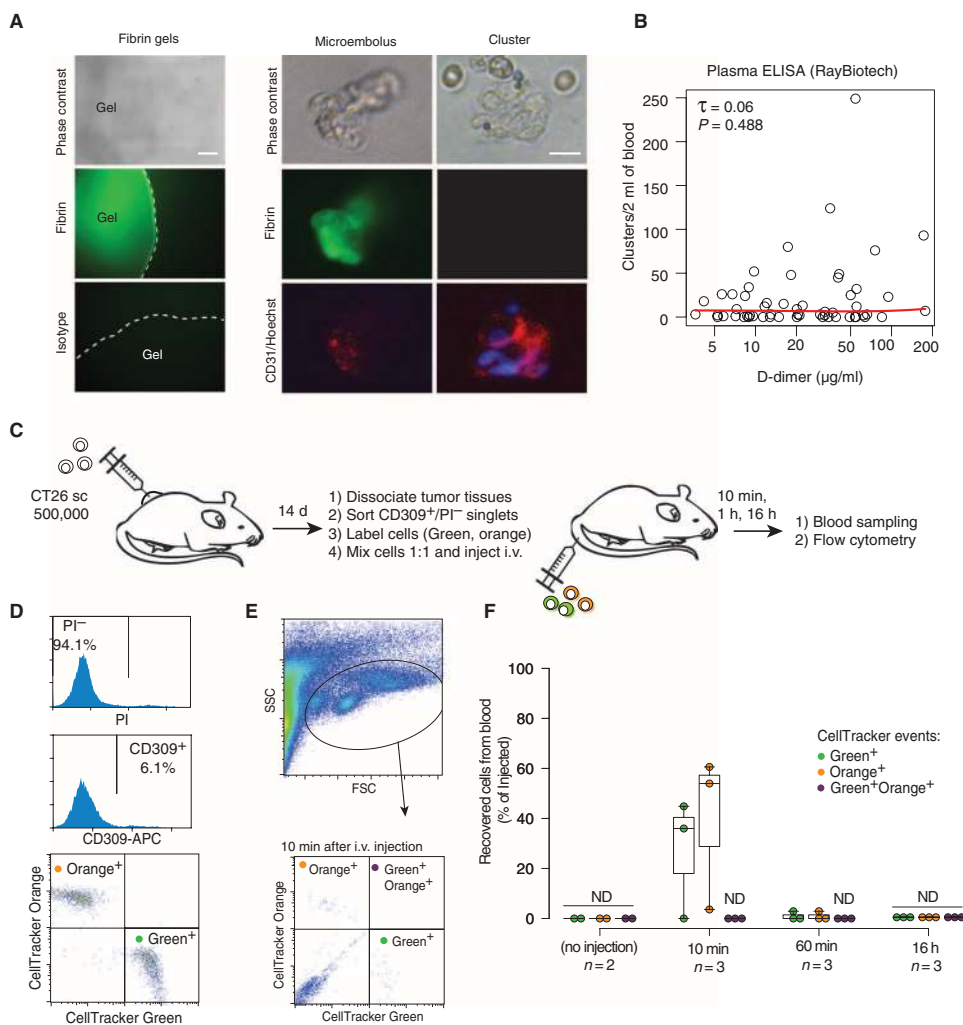


Fig. 5. Tumor-derived endothelial cell clusters do not form by coagulation. (A) Immunofluorescence staining for fibrin and CD31 in a representative endothelial cluster using fibrin gels and a microembolus as positive controls. Scale bars, 10 μm. (B) Scatterplot indicating D-dimer concentration of plasma samples (x axis) from patients with matched endothelial cluster count (y axis). Smooth curve was computed by loess using scatter.smooth from the R package stats. P value determined by Kendall's τ correlation. ELISA, enzyme-linked immunosorbent assay. (C) Experimental approach to test whether single circulating tECs aggregate to form clusters in circulation. Formation of multicolored endothelial clusters in circulation was analyzed after injecting single tumor-derived CD309+ cells intravenously (i.v.) in Balb/c mice. (D) Sorting and labeling of single tumor-derived CD309+ cells. Cells were labeled using CellTracker Green (Green+) or Orange (Orange+). APC, allophycocyanin. (E) Gating strategy for detecting labeled populations of cells injected intravenously in Balb/c mice, as described in (C) and (D). FSC, forward scatter; SSC, side scatter. (F) Quantification of single and dual-labeled endothelial cells (clusters) in mouse blood at the indicated time points after injection (n = 11 mice total).

than being shed as a cluster from the tumor. However, the circulating endothelial clusters that we isolated from CRC patients stained negative for fibrin, the main component of the coagulation cascade (Fig. 5A), and did not contain platelets (Fig. 3C and table S11). In addition, in CRC patients, the number of endothelial clusters did not correlate with the concentration of plasma D dimer, a marker for intravascular coagulation (Fig. 5B).

To test whether circulating endothelial cells coagulate to form clusters in circulation, we injected fluorescently labeled single (dissociated) murine colorectal tumor (CT26)–derived endothelial cells into the tail vein of Balb/c mice and analyzed the formation of multinucleated events at different time points after injection (Fig. 5, C to E). Single endothelial cells did not aggregate to form multicolored clusters and were cleared within 60 min upon injection (Fig. 5F). These experiments provide additional support for the hypothesis that tEC clusters circulating in the blood of CRC patients may be directly shed from the tumor vasculature.

## Tumor-derived endothelial cell clusters are prevalent in CRC patients

The associations between circulating endothelial cell clusters and the primary tumor prompted us to investigate whether these clusters were indicators of CRC. We counted endothelial cell clusters from a test set composed of 141 clinical specimens taken from 125 subjects (45 control healthy volunteers and a consecutive series of 80 CRC patients, including the above-mentioned patients) (table S13). At least one cluster was detected in 76.2% (61 of 80) of CRC patients, compared to only 2.2% (1 of 45) of healthy individuals (Fig. 6A). Treatment-naïve CRC patients ($n = 52$) presented significantly higher endothelial cluster counts than patients who had already received any treatment for CRC ($n = 28$) (Fig. 6B). However, endothelial cell cluster count was not associated with clinical parameters such as tumor stage, grade, or presence of distant metastasis (Fig. 6B and fig. S10A).

Analysis of endothelial cell cluster distributions over time indicated that surgical resection of the primary tumor had the strongest effect on cluster counts (Fig. 6C and table S14), confirming the results in Fig. 4B that show a sharp decline in the numbers of circulating endothelial cell clusters before and after surgical resection of the tumor and providing further support for the association of circulating endothelial cell clusters with the presence of a

primary tumor. The presence of these cells in 86.5% of treatment-naïve patients (45 of 52) but only in 2.2% of healthy controls (1 of 45) indicated that the endothelial cell cluster count might be useful in assisting CRC diagnosis. Treatment-naïve patients were distinguished from healthy controls on the basis of their endothelial cell cluster burden [area under the curve (AUC), 0.930] (Fig. 6D). CRC patients with early-stage pathological tumors (stage ≤IIA) were also positive for endothelial cell clusters in 86.4% of cases (19 of 22), with an AUC of 0.922 (Fig. 6D). A second independent set of 100 samples (17 treatment-naïve CRC
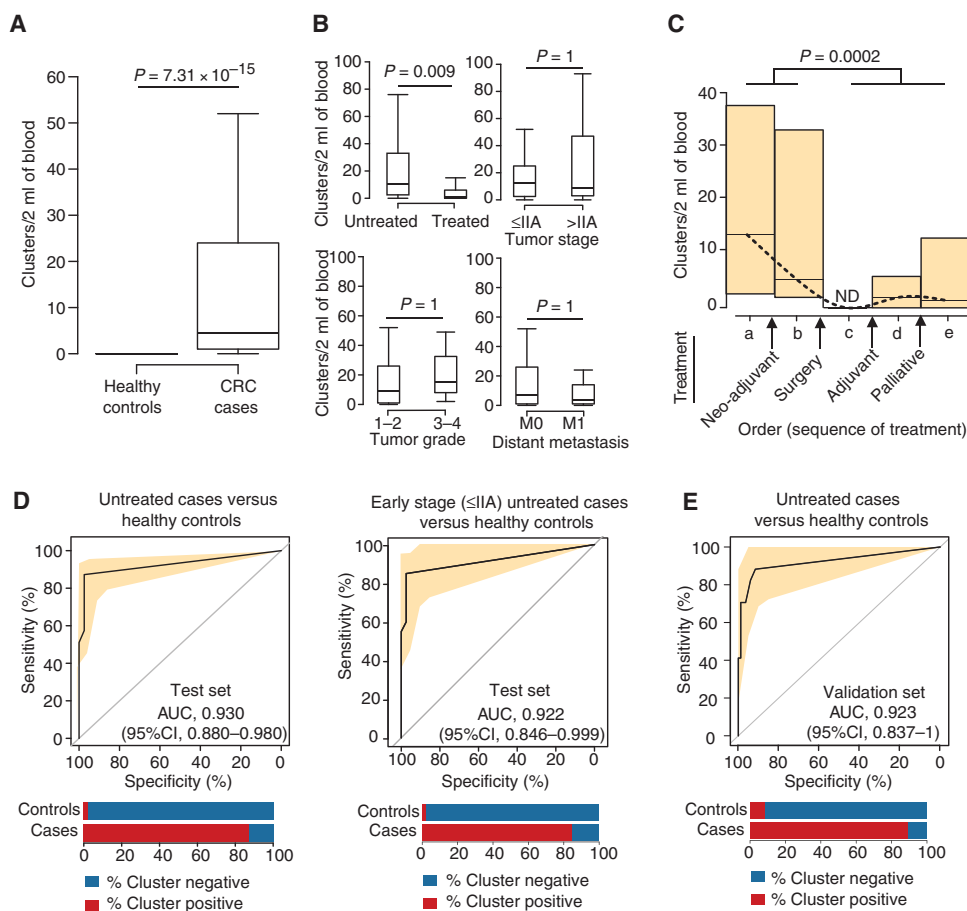


**Fig. 6. Tumor-derived endothelial cell clusters are prevalent in early-stage and treatment-naïve CRC patients and are affected by therapeutic intervention.** (**A**) Endothelial cell cluster count for healthy controls ($n = 45$) and CRC patients ($n = 80$). $P$ value determined by two-tailed Wilcoxon-Mann-Whitney $U$ test; effect size $r = 0.65$. (**B**) Association of endothelial cell cluster count with treatment ($n = 80$ CRC cases). $P$ value determined by two-tailed Wilcoxon-Mann-Whitney $U$ test with Bonferroni correction; effect size $r = 0.34$. (**C**) Trend of endothelial cell cluster count during sequence of treatment for CRC. Blood samples were collected independently at the indicated discrete time points: (a) treatment-naïve, (b) after neoadjuvant therapy, (c) after surgery, (d) during and after adjuvant therapy, and (e) during palliative therapy. Boxes indicate the interquartile range (IQR), line across boxes indicates the median, and the dashed line indicates spline interpolation of medians ($n = 80$ CRC cases). Postoperative samples in Fig. 4B were not included. Arrows indicate treatment events start. $P$ value determined by two-tailed Wilcoxon-Mann-Whitney $U$ test; effect size $r = 0.41$. (**D**) Test set comparing untreated cases versus healthy controls. (Left) Receiver operating characteristic (ROC) curve comparing treatment-naïve CRC patients with healthy controls (total $n = 89$ individuals). Yellow area represents the bootstrapped 95% confidence interval (CI); effect size $r = 0.716$. (Right) ROC curve comparing treatment-naïve, early-stage CRC patients (≤IIA) versus healthy controls (total $n = 61$ individuals); effect size $r = 0.706$. (**E**) Validation set comparing treatment-naïve CRC patients with healthy controls (total $n = 100$ individuals); effect size $r = 0.706$. In (D) and (E), 100% stacked bar charts indicate the percentage of circulating endothelial cell cluster–positive and endothelial cell cluster–negative samples for both healthy controls and CRC cases.

patients and 83 healthy individuals) validated the test set results in Fig. 6D, showing identical expected diagnostic performance (AUC, 0.923) (Fig. 6E and table S13). Furthermore, the number of circulating endothelial cell clusters did not correlate with potentially confounding variables, such as patient age, systemic inflammatory markers, or hypertension, cardiovascular diseases, or diabetes (fig. S10, B and C). Together, the widespread presence of circulating endothelial cell clusters in treatment-naïve patients, but not in healthy individuals, indicates the potential usefulness of endothelial cell cluster count as a diagnostic adjunct for CRC, even early in disease (stage ≤IIA).

## DISCUSSION

Here, we present a subpopulation of tumor-derived cells of endothelial origin that appear to challenge the consensus that all CTC clusters are malignant entities. Intending to isolate CTCs from CRC patients, we instead identified clusters with both epithelial and mesenchymal characteristics that expressed endothelial markers. Through a series of biochemical, genetic, and in vivo analyses, we determined that these cells are nonmalignant, are tumor-derived, are of endothelial origin, and reflect the properties of the primary tumor vasculature. Our study provides experimental evidence for the occurrence of a second tumor-derived cellular entity in the blood, alongside bona fide CTCs. In support of our findings, a recent study described a subpopulation of CD276[+] mature circulating endothelial cells, correlating with the presence of a primary tumor in esophageal and non–small cell lung cancer patients (26).

Tumor-derived endothelial cell clusters might be senescent or non-replicative endothelial cells shed as multiple-cell structures from the chaotic tumor vasculature as it undergoes pathological angiogenesis (27), a recognized early event in CRC tumor progression (28). It is not clear how large clusters of cells travel from the primary tumor to the peripheral blood vasculature. One of the possible routes to bypass the capillary circuit could be through arteriovenous anastomoses, which are large circulatory connections between arterial and venous circulation (29). Recently, Au et al. (30) showed that clusters of tumor cells, owing to their high deformability, can even traverse capillary-sized blood vessels.

Unlike Aceto et al. (9), we did not detect the presence of clusters of epithelial CTCs. This might be the result of different patient profiles. Aceto et al. (9) analyzed blood samples from terminal breast cancer patients, whereas our patients were mostly preoperative CRC patients. A second explanation may be that epithelial clusters are more deformable than endothelial clusters and consequently traversed the microsieve pores of our device during the isolation procedure (30). Nonetheless, we can conclude that CTC clusters do not constitute a homogeneous population and that at least two different tumor-derived population exist: an epithelial-derived population, characterized by malignant potential, and the endothelial population described in this study. Although the endothelial cluster cytomorphology is reminiscent of malignancy, with expression of keratins and mesenchymal markers, these cells should not be confused with malignant CTCs undergoing EMT. Research efforts interrogating the role of EMT in CTC clusters should consider the spurious relationships introduced by unintentional isolation of circulating endothelial cell clusters. For example, our endothelial clusters not only expressed vimentin but also were positive for mesenchymal genetic markers such as FN1, SERPINE1, and FOXC1. A few circulating endothelial clusters expressed transcripts for keratin 8 and 18 or stained positive

for the pan-keratin antibody C11, further suggesting a potential confounding role of tECs in CTC studies, especially those using label-free isolation techniques (17).

Our study has some limitations. For example, we could not study the number of single tECs in circulation because our device would not fully recover cells with a diameter of <10 μm. In addition, we could not determine whether mutations found in our endothelial clusters were technical artifacts, owing to the whole-genome amplification procedure (which are known to occur especially in PCR-based amplifications), or real sporadic mutations (table S4). The presence of potentially false-positive mutations in our samples, however, does not compromise our results ultimately because only matching mutations between circulating clusters and tumor tissue would be conclusive. Finally, we note that aberrant endothelial clusters similar to the ones we have characterized here have been reported in patients with increased risk for acute myocardial infarction (31). This indicates that the presence of endothelial clusters in the blood might not be specific to neoplastic diseases, but rather a general indicator for vascular fragility.

In conclusion, we describe circulating endothelial cell clusters as a tumor-derived entity widespread in all stages of CRC. This insight opens up multiple new opportunities for translational research in querying disease status, in CRC and other cancers. For example, endothelial cluster burden may be used for early CRC detection. Additional prospective trials, including patients with colorectal adenomas, would be required to validate such diagnostic utility. Profiling of circulating endothelial cell clusters also promises to reveal features of the underlying tumor vasculature, particularly on a dynamic basis. For example, drug responses to anti-angiogenic drugs could be tested using endothelial cluster counts. These cells may also be characterized to select patients who would respond to drugs targeting endothelial surface markers currently under clinical evaluation, such as PSMA (24). In addition, live circulating endothelial clusters could offer an intriguing in vitro model to test drug responses targeting the tumor vasculature.

## MATERIALS AND METHODS

### Study design

The aim of this study was to comprehensively characterize CTCs in unselected CRC patients as a noninvasive method to derive information on underlying tumor characteristics. Devices and methods were optimized using cell lines and healthy donor blood samples before the study of clinical samples. Initial observations of the presence of circulating CD45[−] clusters were recorded from a pilot series of five patients. Further molecular characterization studies were performed on circulating CD45[−] clusters in unselected patients as described in table S14. Concurrently, we designed a case-control study to determine whether CD45[−] clusters counts were associated with the presence of CRC. To determine the required minimal sample size, we first assumed that there was no association between samples positive for CD45[−] clusters (≥1 cluster per 2 ml of blood) and CRC (null hypothesis). For a target power of 0.95, this resulted in an estimated minimal sample size of $n = 72$ using the pwr.chisq.test function of the pwr package in the R environment (version 3.1.0) (32). An effect size of $w = 0.5$ was assumed at the two-sided significance level of 0.01. The effect size $w$ was chosen on the basis of the pilot test of five CRC patients (all positive for circulating CD45[−] clusters) (Fig. 1D) and information derived from four healthy controls with negative cluster counts. Moreover, we

considered a review of the literature that reported the rare presence of clusters in healthy individuals and the abundant presence of CTC clusters and CTM in various types of cancer. An independent set of 100 individuals was further used to validate initial results on CD45⁻ cluster counts. Replicates and repeats are indicated in figure captions, and the collection of clinical data and its blinding are described in the Supplementary Materials and Methods.

### Clinical samples collection

All samples were obtained according to protocols approved by the Institutional Review Boards of the National University of Singapore, Fortis Surgical Hospital (FSH), and Singapore Health Services. All 227 subjects gave informed written consent to participate. The clinical samples for the test set ($n = 127$) were obtained between July 2012 and April 2014. The clinical samples for the validation set ($n = 100$) were obtained between July 2014 and August 2015. Consecutive blood samples from 99 CRC patients were provided by FSH and the National Cancer Centre, Singapore. Blood samples from 128 healthy subjects were provided by the Singapore Consortium of Cohort Studies and by donors of the Institute of Bioengineering and Nanotechnology.

### Device fabrication

Silicon microsieves were fabricated as described by Lim *et al.* (*13*). The microsieve consisted of a silicon disk with an overall diameter (ø) of 7.3 mm and a 300-μm-thick support ring. The central capture region measured ø = 5.3 mm and was 60 μm thick; it contained 100,000 circular pores obtained by deep reactive ion etching. To embed the microsieve in a sterile 3-ml syringe, we designed an acrylic sleeve insert consisting of an inlet channel with ø = 8.58 mm tapering to a ø = 5.54-mm channel, which corresponded to the microsieve cell capture region. The sleeve insert housed the microsieve and silicone O-rings (0.5 mm thick) that ensured a good seal and provided cushioning, as shown in Fig. 1B. Details on device assembly and microfiltration optimization are provided in the Supplementary Materials and Methods.

### On-sieve immunofluorescence

Cell suspensions were stained for 30 min directly "on-sieve" after five washes in phosphate-buffered saline containing 0.5% bovine serum albumin, 2 mM EDTA, and human FcR Blocking Reagent (Miltenyi Biotec) using the antibodies and fluorescent dyes listed in the Supplementary Materials and Methods. After a washing step, cells were retrieved and visualized in suspension under an inverted fluorescence microscope (IX81, Olympus) for imaging, counting, and/or micromanipulation. Images were recorded using MetaMorph software (Molecular Devices) with a CoolSNAP HQ2 CCD camera (Photometrics).

### Circulating CD45⁻ clusters definition and count

Circulating CD45⁻ clusters were defined in this study as "any CD45⁻ cell or CD45⁻ cellular cluster with a major axis of >10 μm and having at least two clearly distinct nuclei." All identified clusters stained positive for endothelial markers CD31, CD144, or VWF (Fig. 3C). Therefore, in our study, CD45⁻ clusters and endothelial cell clusters refer to the same population. Cellular populations belonging to the megakaryocytic lineages, with large lobulated single nuclei or large round single nuclei, were excluded from the analysis. The characteristic cytomorphology of these cells was clearly different from that of CD45⁻ clusters with positive staining for CD41 and CD42B. Megakaryocytic cells were predominantly observed in CRC patients undergoing treatment. Single endothelial cells were also

excluded from the analysis. CD45⁻ clusters were counted by transferring the microfiltrate obtained from 2 ml of whole blood to the well of a 96-well plate. After brief centrifugation, the cell clusters were identified and counted by manually scanning the target well three times using a 20× objective. A positive sample was defined by the detection of at least one CD45⁻ cluster.

### Single-cell RNA and mutational analysis PCR

All primers were designed using Primer-BLAST (*33*). For each RNA transcript, primers were designed either to span exon-exon boundaries or to anneal to regions separated by at least one intron on the corresponding genomic DNA region. Primers for mutational analysis were designed to bind intronic regions of the target gene (table S1). The scrmPCR method is highlighted in fig. S2A and integrated several workflow steps derived from the CellsDirect One-Step qRT-PCR Kit (Invitrogen), the Biomark HD system (Fluidigm), and the work of Peixoto *et al.* (*34*). Further details are provided in the Supplementary Materials and Methods.

### Targeted resequencing and aCGH of circulating CD45⁻ clusters

Single CD45⁻ clusters ($n = 34$ from 12 patients) were subjected to whole-genome amplification using the GenomePlex Single Cell Whole Genome Amplification Kit (Sigma) according to the manufacturer's instructions. This method was chosen after comparison tests for efficiency of amplification with an alternative method based on multiple displacement amplification. For the aCGH experiment, 50 pg of tissue DNA samples was amplified using the same procedure. All amplified samples were tested for *ACTB* DNA amplification by quantitative PCR. Nine clusters were removed from the analysis because of poor amplification of the reference gene. For targeted resequencing, a custom gene panel was designed to target exons for *NRAS*, *CTNNB1*, *PIK3CA*, *EGFR*, *BRAF*, *PTEN*, *KRAS*, *AKT1*, and *TP53* genes (~6.1 kb). The libraries were constructed using Ion AmpliSeq Library Kits 2.0 (Life Technologies) with 10 ng of input DNA from the amplified clusters or the unamplified tissue DNA. Fifteen libraries were successfully generated from an initial number of 25 single CD45⁻ clusters. The targeted resequencing run was performed using an Ion Torrent Personal Genome Machine Sequencer (Life Technologies). Variants were called using the Ion Torrent Variant Caller Plugin with high-stringency settings. aCGH was performed by hybridizing 250 ng of DNA to CytoScan 750K arrays (Affymetrix) according to the manufacturer's instructions and using recommended reagents. Data were analyzed and visualized using ChAS software version 2.1 (Affymetrix).

### cDNA synthesis and RNA-seq

cDNA was synthesized from 30 single CD45⁻ clusters, 10 pg of whole tissue, 10 pg of laser-dissected endothelial cells, or 10 pg of primary endothelial cell RNA. To this end, we used the SMARTer Ultra Low RNA Kit (Clontech Laboratories) and long-distance PCR for 25 cycles for the clusters and 18 cycles for the remaining samples. At this stage, four clusters were excluded from the analysis because of non-amplification. For each sample, cDNA was sheared using the Adaptive Focused Acoustics system (Covaris). Libraries were constructed using the NEBNext DNA Library Prep Master Mix Kit (New England Biolabs). Eight clusters were further removed from the analysis because of poor library quality. All remaining libraries were barcoded using unique indices and pooled for RNA-seq run on an Illumina HiSeq 2000 platform. Data were mapped to human genome version 19 (hg19) using TopHat

(version 2). Cufflinks (version 2.2) was used to quantify fragments per kilobase of transcript per million mapped reads (FPKM) (*35*).

## RNA-seq data analysis

The cell-type inference workflow was implemented in a custom R script (available upon request). PCA, differential gene expression analysis, and LDA were performed in the R environment (version 3.1.0) (*32*). Further details are in the Supplementary Materials and Methods.

## Microvessel staining and analysis

Immunohistochemical staining of CD31 was performed on all available formalin-fixed and paraffin-embedded tissues ($n = 28$) with matching CD45$^-$ cluster count. Tissue sections (4 μm) were placed on poly-L-lysine slides and stained with anti-CD31 antibodies in a 1:50 dilution (JC70A, Dako) for 20 min following the manufacturer's instructions on a BOND Automated Stainer (Leica Biosystems). Microvessel density and lumen counts on all sections were determined similarly as described previously (*36*). See the Supplementary Materials and Methods for more details.

## Experimental mice

All animals were housed at the Biological Resource Centre, Agency for Science, Technology and Research (A*STAR). Animal procedures were approved by the Institutional Animal Care and Use Committee of A*STAR and performed in compliance with the National Advisory Committee for Laboratory Animal Research (NACLAR) Guidelines. Experimental procedures are described in Fig. 5C and fig. S9A, with details provided in the Supplementary Materials and Methods.

## Statistical analyses

All statistical analyses were performed in the R environment (version 3.1.0) (*32*). Unpaired samples were tested using a two-tailed Wilcoxon-Mann-Whitney *U* test with Bonferroni correction in the case of multiple comparisons. For paired samples, the two-tailed Wilcoxon signed-rank test was used. For each test, the exact *P* value was computed using the coin package. ROC curves with AUC and 95% CI were computed using the pROC package. Hierarchical clustering with bootstrapped *P* values in fig. S8C was computed using pvclust. To simplify the interpretation and comparison of effect size, the effect size *r* was derived for each statistical test as follows: $r = |Z| / \sqrt{n}$, where $Z$ is the $z$ score calculated from the Wilcoxon-Mann-Whitney $U$ or the Wilcoxon signed-rank test. *r* was converted from the resulting AUC and odds ratio as described in (*37*). Using Cohen's definition, $r = 0.1$ can be interpreted as a small effect, $r = 0.3$ can be interpreted as a medium effect, and $r = 0.5$ can be interpreted as a large effect (*37*). Box plots were drawn with boxes representing the IQR, a line across the box indicating the median, and whiskers indicating 1.5 × IQR. Outliers were not shown. Correlations were described using Kendall's tau (τ) coefficient and its derived *P* value. The significance threshold was set at 0.05.

## SUPPLEMENTARY MATERIALS

## REFERENCES AND NOTES

1. C. L. Chaffer, R. A. Weinberg, A perspective on cancer cell metastasis. *Science* **331**, 1559–1564 (2011).
2. M. C. Miller, G. V. Doyle, L. W. M. M. Terstappen, Significance of circulating tumor cells detected by the CellSearch system in patients with metastatic breast colorectal and prostate cancer. *J. Oncol.* **2010**, 617421 (2010).
3. C. Alix-Panabières, K. Pantel, Challenges in circulating tumour cell research. *Nat. Rev. Cancer* **14**, 623–631 (2014).
4. C. Raimondi, A. Gradilone, G. Naso, E. Cortesi, P. Gazzaniga, Clinical utility of circulating tumor cell counting through CellSearch: The dilemma of a concept suspended in limbo. *Onco. Targets Ther.* **7**, 619–625 (2014).
5. E. S. Antonarakis, C. Lu, H. Wang, B. Luber, M. Nakazawa, J. C. Roeser, Y. Chen, T. A. Mohammad, Y. Chen, H. L. Fedor, T. L. Lotan, Q. Zheng, A. M. De Marzo, J. T. Isaacs, W. B. Isaacs, R. Nadal, C. J. Paller, S. R. Denmeade, M. A. Carducci, M. A. Eisenberger, J. Luo, AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. *N. Engl. J. Med.* **371**, 1028–1038 (2014).
6. G. C. Finkel, G. H. Tishkoff, Malignant cells in a peripheral blood smear: Report of a case. *N. Engl. J. Med.* **262**, 187–188 (1960).
7. M. Yu, A. Bardia, B. S. Wittner, S. L. Stott, M. E. Smas, D. T. Ting, S. J. Isakoff, J. C. Ciciliano, M. N. Wells, A. M. Shah, K. F. Concannon, M. C. Donaldson, L. V. Sequist, E. Brachtel, D. Sgroi, J. Baselga, S. Ramaswamy, M. Toner, D. A. Haber, S. Maheswaran, Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* **339**, 580–584 (2013).
8. R. Kalluri, R. A. Weinberg, The basics of epithelial-mesenchymal transition. *J. Clin. Invest.* **119**, 1420–1428 (2009).
9. N. Aceto, A. Bardia, D. T. Miyamoto, M. C. Donaldson, B. S. Wittner, J. A. Spencer, M. Yu, A. Pely, A. Engstrom, H. Zhu, B. W. Brannigan, R. Kapur, S. L. Stott, T. Shioda, S. Ramaswamy, D. T. Ting, C. P. Lin, M. Toner, D. A. Haber, S. Maheswaran, Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **158**, 1110–1122 (2014).
10. J.-M. Hou, M. G. Krebs, L. Lancashire, R. Sloane, A. Backen, R. K. Swain, L. J. C. Priest, A. Greystoke, C. Zhou, K. Morris, T. Ward, F. H. Blackhall, C. Dive, 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).
11. S. H. Seal, A sieve for the isolation of cancer cells and other large cells from the blood. *Cancer* **17**, 637–642 (1964).
12. G. Vona, A. Sabile, M. Louha, V. Sitruk, S. Romana, K. Schütze, F. Capron, D. Franco, M. Pazzagli, M. Vekemans, B. Lacour, C. Bréchot, P. Paterlini-Bréchot, 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).
13. L. S. Lim, M. Hu, M. C. Huang, W. C. Cheong, A. T. L. Gan, X. L. Looi, S. M. Leong, E. S.-C. Koay, M.-H. Li, Microsieve lab-chip device for rapid enumeration and fluorescence in situ hybridization of circulating tumor cells. *Lab Chip* **12**, 4388–4396 (2012).

14. F. A. W. Coumans, G. van Dalum, M. Beck, L. W. M. M. Terstappen, Filter characteristics influencing circulating tumor cell enrichment from whole blood. *PLOS One* **8**, e61770 (2013).

15. A. J. Armstrong, M. S. Marengo, S. Oltean, G. Kemeny, R. L. Bitting, J. D. Turnbull, C. I. Herold, P. K. Marcom, D. J. George, M. A. Garcia-Blanco, Circulating tumor cells from patients with advanced prostate and breast cancer display both epithelial and mesenchymal markers. *Mol. Cancer Res.* **9**, 997–1007 (2011).

16. S. Maheswaran, L. V. Sequist, S. Nagrath, L. Ulkus, B. Brannigan, C. V. Collura, E. Inserra, S. Diederichs, A. J. Iafrate, D. W. Bell, S. Digumarth, A. Muzikansky, D. Irimia, J. Settleman, R. G. Tompkins, T. J. Lynch, M. Toner, D. A. Haber, Detection of mutations in *EGFR* in circulating lung-cancer cells. *N. Engl. J. Med.* **359**, 366–377 (2008).

17. I. Cima, C. W. Yee, F. S. Iliescu, W. M. Phyo, K. H. Lim, C. Iliescu, M. H. Tan, Label-free isolation of circulating tumor cells in microfluidic devices: Current research and perspectives. *Biomicrofluidics* **7**, 011810 (2013).

18. X. Ni, M. Zhuo, Z. Su, J. Duan, Y. Gao, Z. Wang, C. Zong, H. Bai, A. R. Chapman, J. Zhao, L. Xu, T. An, Q. Ma, Y. Wang, M. Wu, Y. Sun, S. Wang, Z. Li, X. Yang, J. Yong, X.-D. Su, Y. Lu, F. Bai, X. S. Xie, J. Wang, Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 21083–21088 (2013).

19. M. Miettinen, J. F. Fetsch, Distribution of keratins in normal endothelial cells and a spectrum of vascular tumors: Implications in tumor diagnosis. *Hum. Pathol.* **31**, 1062–1067 (2000).

20. A. El-Heliebi, T. Kroneis, E. Zöhrer, J. Haybaeck, K. Fischereder, K. Kampel-Kettner, R. Zigeuner, H. Pock, R. Riedl, R. Stauber, J. B. Geigl, B. Huppertz, P. Sedlmayr, C. Lackner, Are morphological criteria sufficient for the identification of circulating tumor cells in renal cancer? *J. Transl. Med.* **11**, 214 (2013).

21. F. Bertolini, Y. Shaked, P. Mancuso, R. S. Kerbel, The multifaceted circulating endothelial cell in cancer: Towards marker and target identification. *Nat. Rev. Cancer* **6**, 835–845 (2006).

22. E. Colombo, F. Calcaterra, M. Cappelletti, D. Mavilio, S. Della Bella, Comparison of fibronectin and collagen in supporting the isolation and expansion of endothelial progenitor cells from human adult peripheral blood. *PLOS One* **8**, e66734 (2013).

23. S. S. Chang, D. S. O'Keefe, D. J. Bacich, V. E. Reuter, W. D. W. Heston, P. B. Gaudin, Prostate-specific membrane antigen is produced in tumor-associated neovasculature. *Clin. Cancer Res.* **5**, 2674–2681 (1999).

24. S. R. Denmeade, A. M. Mhaka, D. M. Rosen, W. N. Brennen, S. Dalrymple, I. Dach, C. Olesen, B. Gurel, A. M. DeMarzo, G. Wilding, M. A. Carducci, C. A. Dionne, J. V. Møller, P. Nissen, S. B. Christensen, J. T. Isaacs, Engineering a prostate-specific membrane antigen–activated tumor endothelial cell prodrug for cancer therapy. *Sci. Transl. Med.* **4**, 140ra86 (2012).

25. B. S. Croix, C. Rago, V. Velculescu, G. Traverso, K. E. Romans, E. Montgomery, A. Lal, G. J. Riggins, C. Lengauer, B. Vogelstein, K. W. Kinzler, Genes expressed in human tumor endothelium. *Science* **289**, 1197–1202 (2000).

26. R. Mehran, M. Nilsson, M. Khajavi, Z. Du, T. Cascone, H. K. Wu, A. Cortes, L. Xu, A. Zurita, R. Schier, B. Riedel, R. El-Zein, J. V. Heymach, Tumor endothelial markers define novel subsets of cancer-specific circulating endothelial cells associated with antitumor efficacy. *Cancer Res.* **74**, 2731–2741 (2014).

27. P. Carmeliet, R. K. Jain, Angiogenesis in cancer and other diseases. *Nature* **407**, 249–257 (2000).

28. K. Akagi, Y. Ikeda, Y. Sumiyoshi, Y. Kimura, J. Kinoshita, M. Miyazaki, T. Abe, Estimation of angiogenesis with anti-CD105 immunostaining in the process of colorectal cancer development. *Surgery* **131** (suppl. 1), S109–S113 (2002).

29. I. Zeidman, J. M. Buss, Transpulmonary passage of tumor cell emboli. *Cancer Res.* **12**, 731–733 (1952).

30. S. H. Au, B. D. Storey, J. C. Moore, Q. Tang, Y.-L. Chen, S. Javaid, A. F. Sarioglu, R. Sullivan, M. W. Madden, R. O'Keefe, D. A. Haber, S. Maheswaran, D. M. Langenau, S. L. Stott, M. Toner, Clusters of circulating tumor cells traverse capillary-sized vessels. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4947–4952 (2016).

31. S. Damani, A. Bacconi, O. Libiger, A. H. Chourasia, R. Serry, R. Gollapudi, R. Goldberg, K. Rapeport, S. Haaser, S. Topol, S. Knowlton, K. Bethel, P. Kuhn, M. Wood, B. Carragher, N. J. Schork, J. Jiang, C. Rao, M. Connelly, V. M. Fowler, E. J. Topol, Characterization of circulating endothelial cells in acute myocardial infarction. *Sci. Transl. Med.* **4**, 126ra33 (2012).

32. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, 2005).

33. J. Ye, G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, T. L. Madden, Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinf.* **13**, 134 (2012).

34. A. Peixoto, M. Monteiro, B. Rocha, H. Veiga-Fernandes, Quantification of multiple gene expression in individual cells. *Genome Res.* **14**, 1938–1947 (2004).

35. C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, L. Pachter, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).

36. G. P. Gupta, D. X. Nguyen, A. C. Chiang, P. D. Bos, J. Y. Kim, C. Nadal, R. R. Gomis, K. Manova-Todorova, J. Massagué, Mediators of vascular remodelling co-opted for sequential steps in lung metastasis. *Nature* **446**, 765–770 (2007).

37. M. Borenstein, L. V. Hedges, J. P. T. Higgins, H. R. Rothstein, *Introduction to Meta-Analysis* (John Wiley & Sons Ltd., Chichester, UK, 2009).

38. N. A. Mabbott, J. K. Baillie, H. Brown, T. C. Freeman, D. A. Hume, An expression atlas of human primary cells: Inference of gene function from coexpression networks. *BMC Genomics* **14**, 632 (2013).

39. J. Schug, W.-P. Schuller, C. Kappen, J. M. Salbaum, M. Bucan, C. J. Stoeckert Jr., Promoter features related to tissue specificity as measured by Shannon entropy. *Genome Biol.* **6**, R33 (2005).

40. C. Wu, C. Orozco, J. Boyer, M. Leglise, J. Goodale, S. Batalov, C. L. Hodge, J. Haase, J. Janes, J. W. Huss III, A. I. Su, BioGPS: An extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.* **10**, R130 (2009).

41. S. Tarazona, F. García-Alcalde, J. Dopazo, A. Ferrer, A. Conesa, Differential expression in RNA-seq: A matter of depth. *Genome Res.* **21**, 2213–2223 (2011).

42. J. M. J. Herbert, D. Stekel, S. Sanderson, V. L. Heath, R. Bicknell, A novel method of differential gene expression analysis using multiple cDNA libraries applied to the identification of tumour endothelial genes. *BMC Genomics* **9**, 153 (2008).

43. N. Kumar, J. Richter, J. Cutts, K. T. Bush, C. Trujillo, S. K. Nigam, T. Gaasterland, D. Brafman, K. Willert, Generation of an expandable intermediate mesoderm restricted progenitor cell line from human pluripotent stem cells. *eLife* **4**, e08413 (2015).

44. J. W. Rowley, A. J. Oler, N. D. Tolley, B. N. Hunter, E. N. Low, D. A. Nix, C. C. Yost, G. A. Zimmerman, A. S. Weyrich, Genome-wide RNA-seq analysis of human and mouse platelet transcriptomes. *Blood* **118**, e101–e111 (2011).

Editor's Summary

**A new type of circulating tumor cluster**

Setting out to isolate circulating tumor cells (CTCs) in the blood of cancer patients, Cima *et al.* instead noticed that clusters, rather than single cells, were present in their sorting device. However, these clusters were not like others reported previously: The cells did not express the cell marker epithelial cell adhesion molecule (EpCAM), and they did not have the same mutations seen in the primary tumor−−two features that are typical of CTCs. Instead, these clusters matched the expression profile and phenotype of endothelial cells, which are the cells lining the blood vessels. The authors first confirmed that these endothelial cells were shed by the tumor, through a series of experiments using patient-derived samples and in vivo studies in mice, and then demonstrated that the clusters did not form by coagulation after being shed; it was the tumor that was shedding intact clusters. Moving back to patients, Cima and colleagues also found that the presence of endothelial cell clusters correlated with early-stage disease, suggesting that these groups of cells might be a unique indicator of cancer, before treatment starts.

| | |
|---|---|
| **Article Tools** | Visit the online version of this article to access the personalization and article tools:<br>http://stm.sciencemag.org/content/8/345/345ra89 |
| **Supplemental Materials** | *"Supplementary Materials"*<br>http://stm.sciencemag.org/content/suppl/2016/06/27/8.345.345ra89.DC1 |
| **Related Content** | The editors suggest related resources on *Science*'s sites:<br>http://stm.sciencemag.org/content/scitransmed/5/179/179ra47.full<br>http://science.sciencemag.org/content/sci/339/6119/580.full<br>http://science.sciencemag.org/content/sci/313/5794/1785.full |