Microfluidic CTC enrichment

The CTC-iChip technology for enrichment of CTC from whole blood has been described elsewhere [23]. In short, 8-20ml of whole blood is incubated with biotinylated antibodies against the WBC markers, CD45 (R&D Systems, clone 2D1), CD66b (AbD Serotec, clone 80H3), and CD16 (Janssen Diagnostics). Dynabeads MyOne Streptavidin T1 (Invitrogen) are then added to tag the WBCs. The blood is subsequently processed through the CTC-iChip, where RBC and platelets are removed through size-based separation, while WBCs are magnetically depleted. The CTC-enriched product is centrifuged, preserved in RNA-later (Ambion) and flash-frozen for long-term storage.

RNA extraction, amplification and detection

For all samples, RNA was extracted using RNeasy Micro Kit (Qiagen) and a quarter of the output was used for cDNA synthesis and 8 cycles of whole transcriptom amplification (WTA) using the SMART-Seq v4 Ultra Low Input RNA Kit (Clontech). 1% of the WTA product was used for each ddPCR reaction. WTA has been shown before to result in a reproducible amplification bias for any specific gene, which, while precluding the comparison of expression levels of different markers within the same sample, allows for highly qunatitative comparisons between the same marker in different samples[27]. ddPCR analysis was performed on the Biorad ddPCR system using predesigned Taqman-based qPCR assays (Invitrogen) and the ddPCR Supermix for Probes (No dUTP)(Biorad). A list of the Taqman assay probes used in

the final assay is available in Supplementary Table S7. For markers detected with multiple probes, the average transcript number was used.

Single cell and bulk RNA seq

RNA sequencing of single CTCs and bulk cell lines has been described previously[24, 47]. In brief, for single cell sequencing, CTCs were enriched to 0.1-10% purity through the CTC-iChip, then stained without fixation using a cocktail of anti-EpCAM, anti-EGFR, and anti-HER2 fluorescein-conjugated antibodies, (counterstaining with anti-CD45 as a negative marker), followed by manual picking of individual cells under fluorescence imaging using a micromanipulator. Single cell (low template protocol) RNA-Sequencing of individual breast CTCs is then performed, using a subset of transcripts to confirm breast lineage origin. Bulk RNAseq of cultured BRx-142, BRx-68 and MDA-231 is performed usng standard methods. These data are deposited and available online.