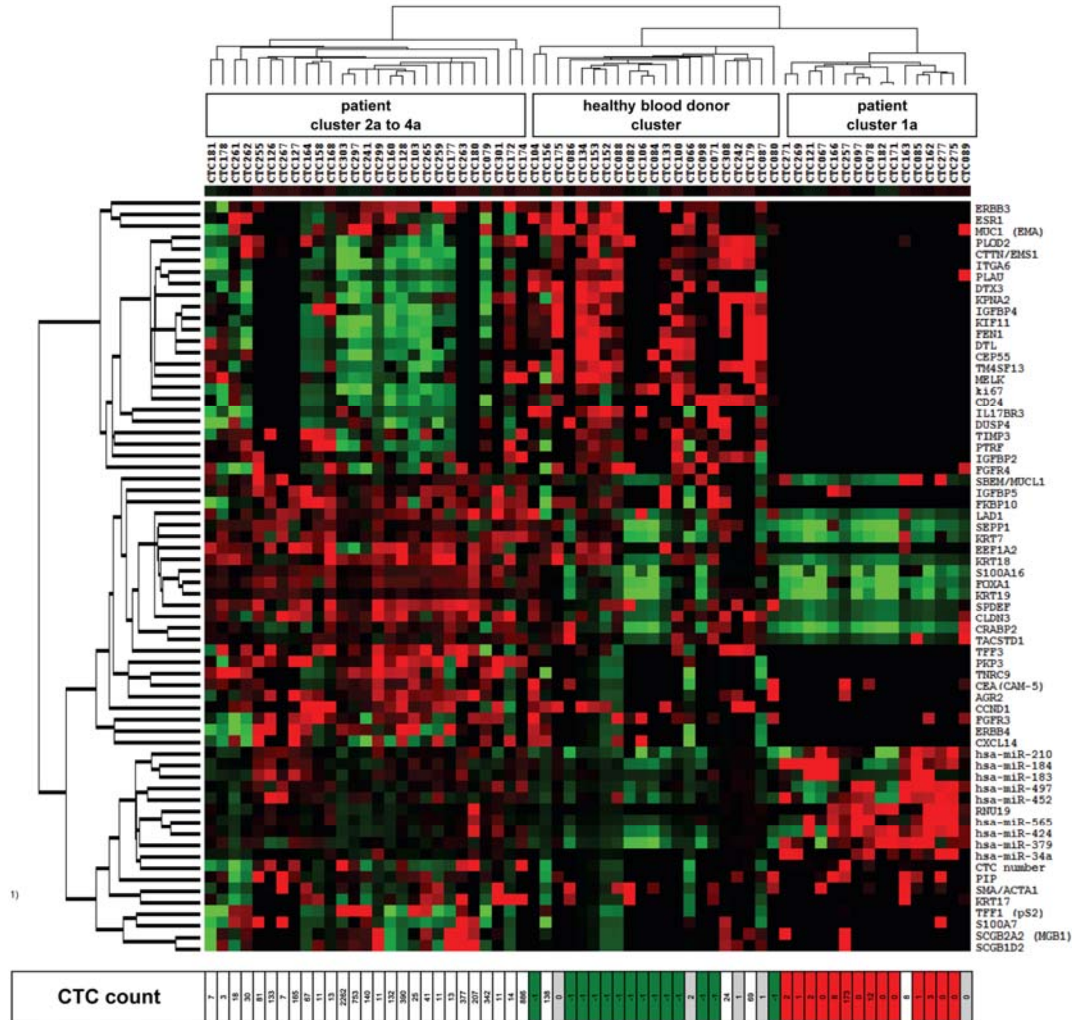
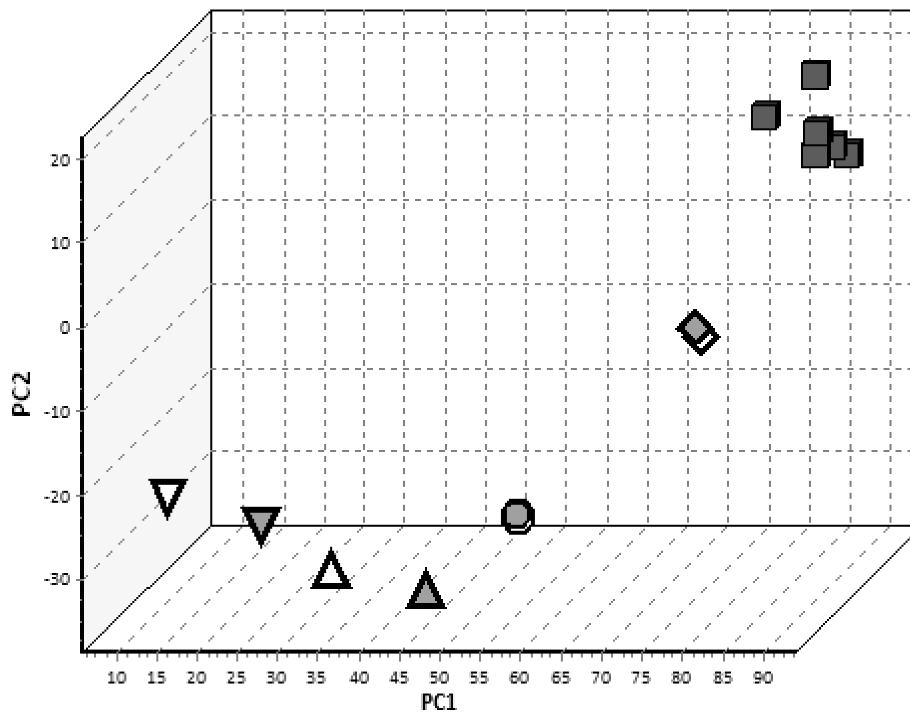


Supplemental Figure 1, on line only:

Unsupervised hierarchical cluster analysis comparing gene expression profiles in healthy blood donors and CTC-enriched blood samples of metastatic breast cancer patients.

Expression levels were analyzed with real-time RT-PCR with 65 TaqMan Gene Expression Assays in cDNA generated from RNA isolated from blood of 14 healthy blood donors and CTC-enriched fractions of 50 metastatic breast cancer patients. Sample loading and RNA integrity was controlled with 3 additional universal reference genes (*GUSB, HPRT1* and *HMBS*). Prior to real-time PCR, cDNA was pre-amplified in 15 cycles with the PreAmp method from ABI as described in the Materials and Methods section, using the same TaqMan Gene Expression Assays that were used for the real-time PCR.

Data shown have been subjected to median normalization of each individual gene across all samples followed by median normalization of each individual sample across all genes. Each horizontal row represents a gene, and each vertical column corresponds to a sample. Red color indicates a transcript level above the median level, black color indicates a median transcript level and green color indicates a transcript level below the median level of the particular assay as measured in all samples. The number of CTCs as established by the CellSearch Epithelial Kit is given right of the graph. Here, HBD samples are marked green and the in Figure 2 identified patients of cluster 1 are marked red. The grey marked cells indicate that the clustering of these CTC samples agrees with the expected absence of a epithelial-derived CTC signal.



Supplemental Figure 2: Gene expression profiles of cells from breast cancer cell lines compared with those of HBDs before and after mixing with breast cancer cells.

Gene expression profiles of cells from four different breast cancer cell lines were compared with those of HBD samples of 5 different healthy volunteers and a HBD sample of a healthy volunteer in which RNA of the 4 different tumor cell lines were spiked. For the first, RNA of cells from 4 different human breast cancer cell lines in a final quantity equivalent to approximately one CTC (~10 pg) was analysed. For the latter, 20% of the RNA isolated after CellSearch from 7.5 mL blood of five individual HBDs was analysed and RNA isolated from 7.5 mL blood of a sixth HBD was aliquoted in 5 equal parts, of which 4 aliquots were spiked with RNA from the 4 different human breast cancer cell lines in a final quantity equivalent to approximately one CTC (~10 pg) per 1.5 mL blood. Following pre-amplification and real-time PCR, the gene expression levels were compared with those of the separate cell lines and the unspiked HBDs. Gene expression levels using our CTC-specific set of 55 TaqMan Gene Expression Assays and 10 TaqMan Human MicroRNA Assays were normalized as described in the Materials and Methods section. Prior to comparison by principal component analysis (PCA), the delta Ct leukocyte cut-offs as identified in our clinical cohort of 50 breast cancer patients was applied.
Six different HBDs without CTCs after CellSearch enrichment, ***rectangles***.
RNA equivalent to 1 reference cell in the absence of HBD: CAMA-1, ***open circle***; SKBR-3, ***open up-ward triangle***; MDA-MB-231, ***open diamond***; ZR75.1, ***open down-ward triangle***.
RNA equivalent to 1 reference cell mixed with RNA obtained from 1.5 mL CellSearch-enriched blood from a HBD: CAMA-1, ***shaded circle***; SKBR-3, ***shaded up-ward triangle***; MDA-MB-231, ***shaded diamond***; ZR75.1, ***shaded down-ward triangle***.

Supplementary Materials and Methods

miRNA and mRNA isolation from CTCs, FF and FFPE

For gene expression studies, in parallel with the enumeration studies, 7.5 mL blood of the same healthy donors and patients was drawn in EDTA tubes and enriched for CTCs on the CellTracks™ AutoPrep System using the CellSearch Profile Kit (Veridex LLC). The cells in the enriched CTC fractions were lysed by adding 250 µL of Qiagen AllPrep DNA/RNA Micro Kit Lysis Buffer (RLT+ lysis buffer) (Qiagen BV, Venlo, The Netherlands) and stored immediately at -80°C until RNA isolation was performed with the AllPrep DNA/RNA Micro Kit (Qiagen) according to the manufacturer's instructions. In brief, by using a gradient of ethanol (Absolute Ethyl Alcohol (EtOH) Merck, Darmstadt, Germany) the larger RNAs (>200 nt) were first captured in a RNeasy Mini spin column in the presence of 35% EtOH and eluated separately from the small RNA molecules (≤200 nt) present in the flow through. These >200 nt aliquots were treated with DNAse I according the manufacturer's instruction. Next, the ≤200 nt molecules present in the flow through were captured in a new RNeasy Mini spin column in the presence of a final concentration of 60% EtOH and thus eluted separately from the >200 nt molecules. Using this approach consisting of two sequential filtrations with different ethanol concentrations, a 12 µL RNA fraction highly enriched in RNA species ≤200 nt and a 14 µL RNA fraction enriched in RNA species >200 nt could be obtained from the same sample.

Total RNA was isolated from FF tissue with RNA-Bee as described before(29) and from FFPE tissue with the column-based High Pure RNA Paraffin Kit (Roche Applied Science, Almere, The Netherlands) according to the manufacturer's instructions.