Molecular profiling of tumor cells in cerebrospinal fluid and matched primary tumors from metastatic breast cancer patients with leptomeningeal carcinomatosis

Authors: Mark Jesus M. Magbanua, Michelle Melisko, Ritu Roy, Eduardo V. Sosa, Louai Hauranieh, Andrea Kablanian, Lauren E. Eisenbud, Artem Ryazantsev, Alfred Au, Janet H. Scott, John W. Park

Affiliations:
1 Division of Hematology/Oncology, University of California San Francisco, San Francisco, California
2 Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, California
3 Helen Diller Family Comprehensive Cancer Center Computational Biology Core, University of California San Francisco, San Francisco, California

*Corresponding author: JWP: jpark@cc.ucsf.edu 1600 Divisadero St. San Francisco, California, 94115, Phone: 415-502-3844, Fax: 415-353-9592

Funding: This work was supported by grants from the UCSF Research Allocation Program awarded to Michelle Melisko.

Acknowledgments: We thank David Ginzinger and Christopher Corless for assistance in mutation analysis. We also thank the members of the UCSF Cancer Center Genome Core and the Laboratory for Cell Analysis.

Conflicts of Interest: JWP: Research Grant (Veridex). No potential conflicts of interest were disclosed by other authors.
Abstract

Although leptomeningeal carcinomatosis is a well-established clinical syndrome, virtually nothing is known about the tumor cells responsible for this particularly aggressive metastatic process. To isolate cerebrospinal fluid-derived tumor cells (“CSFTCs”) from 15 metastatic breast cancer patients diagnosed with leptomeningeal carcinomatosis, CSF samples were subjected to a two-step method involving immunomagnetic enrichment and fluorescence-activated cell sorting (IE/FACS), a technique previously used for isolating circulating tumor cells (CTCs) from blood. CSFTCs were subjected to genome-wide copy number analysis by array comparative genomic hybridization. Genomic profiling was successfully performed for 13 of the 15 patients (87%). Copy number analysis in CSFTCs revealed genomic alterations commonly observed in primary breast cancer and CTCs, indicating their malignant origin. Interestingly, 12 (92%) harbored high-level gains on the 8q24 locus, which includes the MYC oncogene. Comparison of CSFTCs against corresponding archival primary tumors in six patients revealed clonal relationships with some divergence. Good concordance among serial samples attested to the reproducibility of the assay. Our approach for isolation and molecular analysis of CSFTCs yielded new insights into the molecular nature of these cells. Further genomic and functional analyses may help elucidate mechanisms by which tumor cells metastasize to the central nervous system.
Introduction

Leptomeningeal carcinomatosis, in which tumor cells metastasize to the leptomeninges and cerebrospinal fluid (CSF), is a particularly virulent syndrome with extremely high morbidity and mortality (1-5). Furthermore, central nervous system (CNS) involvement of breast cancer, including leptomeningeal carcinomatosis, appears to be increasing due to longer survival times and more efficacious treatment against other sites of systemic disease (6). The particularly aggressive nature of leptomeningeal carcinomatosis and its increasing frequency highlight the importance of understanding the biology of CSFTCs. Despite efforts towards improving detection and diagnosis (7), almost no information exists regarding the underlying biology and molecular characteristics of these metastatic cells.

We have developed new methods for the isolation and molecular analysis of circulating tumor cells (CTCs) from the blood of metastatic cancer patients (8-10). This approach involves the efficient and complete isolation of CTCs without significant leukocyte or non-malignant epithelial cell contamination, followed by array comparative genomic hybridization (aCGH). It was therefore of interest to adapt this approach for isolation and profiling of CSFTCs, including gene expression and mutation analyses as well as aCGH. In a subset of patients, we compared the genomic profiles of CSFTCs to corresponding archival tumors to examine clonal relationships.
Methods

Patients and samples

Clinical samples were obtained from metastatic breast cancer patients who were diagnosed with leptomeningeal carcinomatosis and were treated at UCSF between September 2007 and February 2012 (Table 1). All patients gave informed consent for collection of additional CSF or use of leftover CSF for research purposes under a protocol approved by the UCSF Institutional Review Board. 4-10mL of CSF was collected via lumbar puncture or Ommaya reservoir. In a subset of patients, blood samples were also collected, and circulating tumor cells (CTCs) were enumerated in 7.5mL of blood via the CellSearch system following manufacturer’s instructions (11). Samples with $\geq$1CTC/mL were subjected to IE/FACS (see below) for isolation and genomic profiling of CTCs (8-10).

Primary and metastatic tumor samples

Archival formalin-fixed paraffin embedded (FFPE) primary tumor, nodal metastases and cell button from fine needle aspirate were processed for copy number analysis as previously described (8). Microdissected areas contained at least 70% tumor. Tumor DNA in whole cell lysates was subjected to whole genome amplification (see below). Clinical estrogen and progesterone receptor and HER2 statuses were obtained from the patients’ medical records.

Tumor cell purification

CSF samples were collected in tubes containing EDTA and were processed within 24 hr (Fig. S1). Tumor cells were purified from the CSF using the previously described IE/FACS method (8, 10). Briefly, CSFTCs were first enriched using magnetic beads coated with EPCAM mAb and then isolated via fluorescence activated cell sorting. CSFTCs were defined as nucleated, EPCAM-positive, and CD45-negative. Leukocytes serving as non-tumor controls
were also sorted and defined as: nucleated, EPCAM-negative, and CD45-positive. The cell inputs for genomic and gene expression profiling are listed in Table S1.

Genomic profiling

Genome-wide copy number analysis of IE/FACS-isolated cells was performed as previously described (8-10). Briefly, genomic DNA was subjected to whole genome amplification using the GenomePlex® Single Cell Whole Genome Amplification Kit (WGA4) and closely adhering to manufacturer’s instructions. The resulting amplified DNA was then analyzed for copy number aberrations by bacterial artificial chromosome (BAC) array comparative genomic hybridization (CGH) analysis.

Somatic mutation analysis

WGA4-amplified DNA samples were screened for mutations using the Sequenom™ MassARRAY System (Sequenom, San Diego, CA). The MassARRAY platform involves the PCR amplification of the region containing the mutation, followed by a single base extension using mass-tagged primers and mass spectrometry (MALDI-TOF MS) of extended primers. The analysis was performed as a contract service at the Pathology Translational Research Laboratory, Oregon Health & Science University following manufacturer’s instructions and also as previously described (12). The Sequenom OncoCarta Panel version 2.0 assays, designed to detect 152 somatic mutations across 18 oncogenes and tumor suppressors, were used for the analysis. Candidate mutations were identified using an automated calling algorithm (Sequenom Typer software). Mass spectra were visually inspected to confirm putative mutations. A proof-of-concept experiment performed on non-amplified and WGA4-amplified DNA from cell culture (50ng) and on 20 MCF7 cells isolated by IE/FACS previously spiked into healthy blood
consistently detected the known mutation PIK3CA E545K in this cell line as listed in the Catalogue of Somatic Mutations in Cancer (COSMIC) database (Table S2).

**Gene expression analysis**

QPCR analysis was performed on 64 cancer-related genes chosen from a previously reported expression profiling of CTCs (13) (Table S3). In addition to CSFTCs, leukocytes defined as: nucleated, EpCAM-negative, and CD45-positive were also sorted by FACS from the same enriched CSF sample and served as non-tumor controls. EPCAM and PTPRC (encodes CD45), two genes included in the list, are markers specific for epithelial cells and hematopoietic cells, respectively. Also included are 6 candidate reference genes for normalization of gene expression data (see Statistical Analysis section). A custom Taqman® Low Density Array (TLD, Applied Biosystems) microfluidic card (384-well format) was designed to contain two sets of the 64 Taqman® gene expression assays printed in triplicate. Cell lysis, reverse transcription (RT), and preamplification were performed using Taqman® PreAmp Cells-to-CT™ kit (Ambion, Texas) following manufacturer’s instructions. Amplified cDNA was subjected to QPCR analysis using the ABI PRISM 7900HT to obtain cycle threshold (Ct) values.

**Quality control**

Quality control (QC) of DNA and RNA from CSFTC samples and resulting data (see Table S1) was performed as previously described (9, 13). For DNA profiling, amplified DNA was subjected to a multiplex PCR analysis to detect the presence of the housekeeping gene, GAPDH. Samples containing the amplifiable sequences (100, 200, 300 and 400bp) of GAPDH were considered for CGH analysis. The median absolute deviation (MAD) estimates were used as a measure of quality of the array data. Array data with MAD estimate ≥0.25 were considered noisy and were excluded from the analysis (9). For RNA profiling, pre-amplified cDNA was
analyzed via RT-PCR (20ul reactions in duplicate) at the UCSF Cancer Center Genome Core to determine the Ct values for ACTB, GAPDH, and RPS18. Initial testing revealed that samples with RPS18 Ct ≥26 and/or ACTB and GAPDH with Ct >36 resulted in a failed TLDA RT-PCR analysis, suggesting low quantity and low quality of RNA. Therefore, samples meeting these criteria were excluded from further analysis. After TLDA RT-PCR analysis, samples with less than 20% detection, i.e., detection of ≤12 of the 64 genes, were excluded from the analysis. In addition, since GAPDH and ACTB were chosen as reference genes to calculate for ΔCts, samples with no detection (missing values) for either of these genes were also excluded from the analysis.

Statistical Analysis

Array CGH microarray data was analyzed using Nexus 6.1 software (Biodiscovery) as previously described (8). The thresholds of log2 ratio values for single copy number gains and losses were 0.20 and −0.20, respectively; the thresholds for high-level number gains and homozygous deletions were 0.6 and −0.6, respectively. The rank segmentation algorithm was used to estimate copy number using the significance threshold of p-value <0.001. Regions of gains and losses present in ≥ 50% of each sample were considered recurrent. Results of the copy number analysis was compared to previously published aCGH data from primary breast tumors (14) and CTCs (9) that were processed in a manner similar to that of the samples in this dataset. All three datasets were obtained from hybridization using similar BAC arrays printed by the UCSF Array Core. We performed a Fisher’s Exact test in Nexus 6.1 to determine regions that were differentially gained or lost between cohorts. For higher stringency, we increased the cut-off to 35% as compared to 30% in our previous study (9). Therefore, gains and losses with an absolute difference of ≥35% between datasets with a p-value less than 0.05 were considered
statistically significant. Of note, the sequential position of the clones was considered during segmentation to minimize possible batch effects between datasets.

Genomic plots were generated as previously described (9). Briefly, microarray data was subjected to circular binary segmentation (CBS) (15), as implemented in the DNAcopy package from Bioconductor (16), to translate intensity measurements into regions of equal copy numbers and to make gain, loss, and amplification calls. Enrichment tests were done at the arm level to identify significantly gained and lost chromosome arms (9). The extent of the genomic instability, defined as the fraction of genome altered (FGA), was calculated as previously described (17). Briefly, the FGA for each sample was calculated by assigning each clone a distance equal to the sum of one half of the distance between its own center and that of its neighboring clones.

Concordance between two aCGH profiles was calculated using the weighted Pearson correlation coefficient ($r_w$) as previously described (9). Correlation coefficients falling in the corresponding intervals, 0 to <0.36, 0.36 to <0.68, 0.68 to <0.90, 0.90 to 1.0 were said to have weak, moderate, high, and very high correlation, respectively (18). The overall concordance of gains and losses between copy number datasets [this study vs CTCs (9) vs primary breast tumors (14)] was measured by estimating the concordance correlation coefficient ($r_c$) of the proportion of gains and losses, as described previously (9).

RealTime StatMiner® version 4.2 was used to analyze QPCR data. Genes with Cts $\geq 36$ were considered unreliable and were flagged as “not detected”. To select the appropriate reference gene(s), we used the geNorm algorithm within RealTime StatMiner® to calculate the gene stability measure (M) for all 6 candidate genes ($ACTB$, $GAPDH$, $GUSB$, $RPLP0$, $TFRC$, and $RPS18$). $ACTB$ and $GAPDH$ showed the lowest M values, indicating the most stable
expression across all samples, were chosen as references genes. The mean Ct for ACTB and GAPDH was used to calculate the ΔCts. Unsupervised complete linkage hierarchical clustering analysis was performed using Euclidean distance as a similarity measure. Differentially expressed genes between CSFTCs and leukocyte (CD45) were assessed using paired data available from the first time point of each patient. A paired t-test was performed with correction for multiple comparisons using the Benjamini-Hochberg method. An adjusted p-value <0.05 was considered statistically significant. Relative quantification (RQ) was reported in the logarithmic scale ($\log_{10} RQ = \log_{10} 2^{-\Delta\Delta Ct}$). For example, a $\log_{10} RQ = 0$ indicates no differential expression between test (CSFTC) versus calibrator (CD45) samples while a $\log_{10} RQ = 1$ and -1 indicates a ten-fold increase or decrease, respectively, in test versus calibrator samples.

MIAME compliant copy number and gene expression data have been deposited into GEO under accession # GSE46068.

Results

Patient characteristics

This study included 15 metastatic breast cancer patients diagnosed with leptomeningeal carcinomatosis by standard cytological detection of malignant cells in the CSF (14 patients) or atypical cells in the CSF with positive MRI showing leptomeningeal disease (1 patient). The median age was 51 years (Table 1). Eleven (73%) patients were ER positive, seven (47%) were HER2 positive, and two (13%) were triple negative. Ten (67%) patients were also diagnosed with brain metastasis.
**Isolation and genomic profiling of CSFTCs**

We evaluated the feasibility of applying our approach for CTC profiling (IE/FACS isolation followed by WGA/CGH) to genome-wide copy number analysis of CSFTCs. Nineteen CSF samples from 15 patients were obtained via lumbar puncture (n=11) or via an Ommaya reservoir (n=8). The samples were then subjected to IE/FACS for CSFTC isolation; genomic DNA was then amplified by WGA and analyzed by aCGH (Fig. S1). Seventeen of the 19 samples (89%) from 13 patients passed quality control (Fig. S2, Table S1).

The resulting CSFTC profiles demonstrated a wide range of genomic aberrations, confirming that the isolated cells were indeed cancer cells (Fig. 1). This is important since it was not clear whether nonmalignant EPCAM-expressing cells, which would be isolated by IE/FACS, might be present at low levels within CSF. Common aberrations (present in ≥50% of the samples) included gains on 1q, 8q and 20q, as well as losses on 1p, 3p, 3q, 4p, 8p, and 11p (Table 2). Notably, 12 out of 13 CSFTC samples (92%) harbored gains on 8q24, including the MYC oncogene (Table 2).

We compared the CSFTC profiles with those from a series of primary breast tumors, as reported previously (14). CSFTCs and primary tumors showed many concordant copy number alterations across the genome (rc gain=0.75 and rc loss=0.59) (Fig. S3A). We next compared CSFTC profiles with those obtained from a series of breast CTCs, as reported previously by us (9). CSFTCs and CTCs also showed many concordant copy number alterations across the genome (rc gain=0.72 and rc loss=0.60) (Fig. S3A). We then performed differential copy number analysis to explore potential regions of divergence among CSFTCs, primary tumors and CTCs. Results of this exploratory analysis revealed several alterations that were observed more frequently in CSFTCs than in primary tumors, such as gains in 8q24 (including MYC) (Table S4,
Fig. S3B). We have previously reported high-level gains in 8q24 (including MYC) in CTCs (9); however, 8q24 gains appeared even more frequently in CSFTCs than in CTCs (Table S5, Fig. S3C).

Enumeration, isolation and genomic profiling of CTCs

Blood samples were collected from 6 of the 15 patients. CTCs were enumerated in 7.5mL of blood using the CellSearch protocol. Three patients had ≥5 CTCs in 7.5mL blood, while 2 patients had undetectable CTCs in blood despite having CSFTCs (Table 1). Samples from patients 108 and 4015 (with 12 and 37 CTC per 7.5mL of blood, respectively) underwent IE/FACS for CTC isolation and profiling. Comparison of the profiles between CTCs and their corresponding CSFTCs are discussed in the section below (Patient 4015) and in the supplementary information (Patient 108).

Genomic profiles of CSFTCs vs. matched primary tumors

In a subset of patient samples (n = 6) for which CSFTC isolation and profiling were performed, samples from the corresponding primary tumors were also obtained and successfully analyzed by CGH.

Overall, CSFTC samples showed many conserved alterations when compared to their respective primary tumor (r_c gain=0.76 and r_c loss=0.43) (Fig. S4A). However, high-level gains centered on 8q24 (including the MYC locus) were frequently seen in CSFTCs but were not observed in archival tumors (Fig. S4B). Furthermore, genomic changes in the CSFTC set were generally more extensive (i.e., new alterations or higher levels of gains/losses in the same loci) than in the primary tumor set. To assess this quantitatively, we compared the fraction of genome altered (FGA) of CSFTCs versus matched primary tumors. CSFTCs showed a significantly higher FGA at 23% as compared to primary tumors at 14% (p=0.0277, sign test), suggesting that
CSFTCs were more genomically aberrant than their corresponding primary tumor samples. Alternatively, it is possible that CSFTC samples contained a higher proportion of true tumor cells, due to the IE/FACS isolation procedure, than the primary tumor samples; greater purity would reduce any signal dilution from normal DNA associated with stromal or other non-tumor cell components.

Details of the 6 CSFTC cases as compared to their corresponding primary tumors, and in some cases to additional local regional or metastatic tumor tissue, are discussed below. A complete list of gained and lost arms for each sample can be found in Table S6. Seven additional cases are discussed in Supplementary Information (also see Fig. S5).

Patient 4037 is a 50-year-old woman with ER positive, PR positive and HER2 negative \textit{de novo} metastatic breast cancer (MBC). Genomic profiling of 20 CSFTCs isolated via IE/FACS revealed multiple genomic alterations (FGA=0.18) (Fig. 2A). Fine needle aspirate of the intact primary tumor collected a day prior to CSF collection was also profiled, and overall revealed similar genomic alterations (FGA=0.14). Comparison of CSFTCs to the primary tumor revealed moderate concordance ($r_w=0.54$), including shared gains in 8q and 9p. The primary tumor contained a single copy loss in 3p, which appeared as a homozygous deletion in CSFTCs, and a gain in 5q, which appeared to be deleted in CSFTCs.

Patient 4039 is a 51-year-old woman with ER positive, PR positive, and HER2 negative MBC. Seven CSFTCs were isolated and profiled by aCGH (Fig. 2B). The corresponding archival primary tumor from 7.4 years prior to CSF collection was also obtained and profiled. CSFTCs showed greater genomic aberrancy (FGA=0.36) than did the primary tumor (FGA=0.07). The patient subsequently expired, and metastatic tissue from a mesenteric lymph node was collected post mortem 20 days following CSF collection. Interestingly, the metastatic
lesion showed genomic aberrancy (FGA=0.03), which was more similar to the primary tumor than to the CSFTCs. Direct comparison similarly indicated that the metastatic tissue was more closely related to the primary tumor ($r_w=0.76$) than to the CSFTCs ($r_w=0.37$). For example, only CSFTCs showed focal amplifications in 1p36 and 11p12-13, which were not observed in either the primary tumor or mesenteric node metastasis.

Patient 4038 is a 40 year-old woman with ER positive, PR positive, and HER2 positive MBC. Genomic profiling of 20 CSFTCs (FGA=0.15) and the primary tumor (FGA=0.14) removed 2.4 years prior to CSF collection revealed a narrow focal amplification on 17q12 (HER2; arrow) in agreement with the HER2 clinical status (Fig. 2C, Table 1). High concordance was observed between CSFTCs and primary tumor ($r_w=0.80$). In addition to HER2 amplification at 17q12, gains in 1q, 4q, and 8q were observed in both CSFTCs and primary tumor.

Patient CSF6 is a 39 year-old woman with ER negative, PR negative, and HER2 positive de novo MBC. Twenty CSFTCs isolated by IE/FACS and profiled by aCGH showed multiple genomic alterations (FGA=0.26), including narrow focal amplification on 17q12 (HER2; arrow) in agreement with the HER2 clinical status (Fig. 2D). Primary tumor specimen from a core biopsy performed approximately 3 months prior to CSF collection was retrieved and profiled by aCGH. The primary tumor contained somewhat fewer genomic alterations (FGA=0.18), but generally showed moderate concordance with CSFTCs ($r_w=0.50$). Shared alterations included gains in 1q, 8q, 10p, 11p, and HER2 amplification on 17q21 and losses on 8p and 12p. However, the focal amplification on 1p32-34 seen in CSFTCs was not observed in the primary tumor.

Patient 4015 is a 54-year-old woman with ER negative, PR negative, and HER2 negative (triple negative) MBC. Twenty CSFTCs were isolated by IE/FACS for molecular analyses. aCGH revealed amplification of 8q24 (includes MYC) as well as gains in 19q12-13 (includes
CNNE1 and HPN) (Fig. 3A). FISH analysis confirmed gene amplification of the MYC locus as well as multiple copies of the centromeric region of chromosome 8 (Fig. 3B; Supplementary Methods (19)). Twenty CTCs were also isolated from the patient’s blood by IE/FACS one month after CSF collection. Comparison of CSFTC and CTC profiles revealed many shared alterations, including high-level gains in 8q24 and 19q12-13. The CTCs did contain additional copy number alterations not seen in CSFTCs. For example, a gain of whole chromosome 4 was observed in CTCs but not in CSFTCs. Archival primary tumor and axillary lymph node metastasis at her initial presentation 2.4 years prior to CSF collection were retrieved and subjected to aCGH analysis. The genomic profiles of these four samples (primary tumor, nodal metastasis, CTCs and CSFTCs) were compared. Overall, the profiles demonstrated many conserved alterations, including gains in 8q, 10p, and 19q, indicative of clonal relatedness. However, CSFTCs (FGA=0.22) and CTCs (FGA=0.51) exhibited more genomic aberrations than did the archival tumors (primary tumor, FGA=0.14; lymph nodes, mean FGA=0.09, s.d. 0.02). Genomic profiles of primary tumor and nodal metastasis showed high concordance (mean \( r_w = 0.88, \) sd 0.01), whereas CTCs and CSFTCs showed moderate concordance with each other \( (r_w=0.67) \).

Patient 4011 is a 51-year-old woman with triple negative MBC. To determine the feasibility and reproducibility of serial analysis, CSFTCs were isolated from 5 independent CSF collections obtained over 104 days. aCGH profiles of these samples revealed multiple genomic alterations \( (\text{mean FGA}=0.27, \text{s.d.} \ 0.03) \), including gains in 1q, 7q, 8q, 13q, and 14q and losses in 3p, 7q, 8p, 8q, 11q, and 13q (Fig. S6A). Over this period, 5 separate CSFTC samples showed consistent profiles and high overall concordance \( (\text{mean } r_w=0.89 \text{ s.d.} \ 0.04) \); these results attest to the reproducibility of the assay and suggest no major genomic changes during the time period...
studied (Fig. S6B). The aCGH profile of the corresponding primary tumor specimen from 9 months prior to initial CSF sampling showed multiple genomic alterations (FGA=0.24), including shared copy number changes and overall high concordance with CSFTCs (mean $r_w=0.87$, s.d. 0.04).

**HER2 status**

The HER2 status of CSFTCs was assessed via aCGH analysis and compared with clinical primary tumor HER2 results determined by IHC/FISH. All 13 patients showed concordance between CSFTC HER2 and the clinical HER2 status (Table 1). Patient 4011, from whom 5 serial CSFTC profiles were obtained, yielded a complex set of results. Her initial CSFTC sample showed no HER2 copy number gains by aCGH, and her primary tumor was similarly HER2 negative by clinical IHC. However, her four subsequent CSFTC samples showed HER2 copy gains by aCGH, as did her primary tumor. These results suggest that the patient had low level HER2 copy gain in both primary tumor and CSFTCs, detectable by aCGH.

For the 6 cases with paired CSFTC and primary tumor profiles, 5 (83%) exhibited concordant HER2 copy number status using the same aCGH technique. The only case not clearly concordant was patient 4011, who displayed mixed HER2 results as discussed above.

**Somatic mutation profiling**

In addition to aCGH analysis, we explored the feasibility of mutation screening in CSFTCs. After CSFTC isolation by IE/FACS and WGA, 3 of the CSFTC cases with corresponding primary tumor (patients 4015, 4038, and 4037) were assayed for 152 candidate mutations using the Sequenom™ MassARRAY System (Table S2). The $PI3KCA$ H1047R mutation was detected in patient 4015’s CSFTCs, as well as in her primary tumor, axillary lymph node metastasis, and CTCs. No mutations, including the $PI3KCA$ H1047R mutation, were
detected in matched CD45+ leukocytes obtained from the same CSF sample. In patient 4038, the
*PI3KCA* H1047R mutation was detected in both CSFTCs and the matched primary tumor. In
patient 4037, no mutations were observed in either the CSFTCs or the primary tumor.

**Gene expression profiling**

To complement these DNA based analyses, we also performed an exploratory gene
expression analysis of CSFTCs (*Fig. S1*). We isolated CSFTCs in 24 samples from 6 patients
obtained via lumbar puncture (n=5) or Ommaya reservoir (n=19). Gene expression profiles of
CSFTCs were not compared with those of primary tumors in this study. Only archival tissue was
obtainable in select patients, which was amenable to aCGH analysis but not expression profiling.
After QC testing, 18 CSFTC samples (75%) from 5 patients were successfully analyzed for the
expression of 64 genes (*Fig. S2, Table S1, Table S3*). As a negative control, matching leukocyte
(CD45+) samples were isolated from the same CSF samples and were successfully profiled.
Unsupervised hierarchical clustering analysis revealed 3 distinct clusters: 2 clusters containing
only CSFTCs and one cluster containing the CD45+ samples (*Fig. 4A*). In general, serial
samples isolated from the same patient clustered together. Differential expression analysis of
CSFTCs vs. matching leukocytes showed, as expected, that CSFTCs have high mRNA levels of
*EPCAM* (*Log*$_{10}$ RQ = 3.14, adj p = 0.01) and low levels of *PTPRC/CD45* (*Log*$_{10}$ RQ = -4.08, adj
p = 0.005) (*Fig. 4B*). These results confirm the validity of our CSFTC isolation strategy, which
requires surface EPCAM protein in conjunction with absence of detectable CD45 protein (*Fig.
4B*). We also observed significant up-regulation of other genes in CSFTCs, such as *AGR2* (*Log*$_{10}$
RQ =3.53, adj p = 0.0007), *TFF3* (*Log*$_{10}$ RQ = 3.21, adj p = 0.0007), and *GRB7* (*Log*$_{10}$ RQ =
2.86, adj p = 0.01) (*Table S3*). Overexpression of these genes has been correlated with
aggressive cancer biology (20-23). Although statistical significance was not achieved, *MYC*
(Log10 RQ=0.85) was up-regulated in CSFTCs as compared to leukocytes (Table S3).

Additionally, ERBB2 (HER2) expression was almost 3 orders of magnitude higher in CSFTCs than in leukocytes, which is consistent with the known minimal expression of HER2 mRNA in hematopoietic cells (24, 25) (Table S3, Fig. 4C).

**Discussion**

The spread of tumor cells to the CSF [leptomeningeal carcinomatosis or leptomeningeal disease] is a devastating complication of breast cancer, as well as other solid tumor types including lung cancer and melanoma. Median survival after the diagnosis of leptomeningeal carcinomatosis is about 18 weeks (4). As with brain metastasis, there are formidable challenges to treating leptomeningeal carcinomatosis, including poor penetration of most anticancer agents into the CNS and poorly understood biology. Disease sites in the CNS are typically difficult to access and not frequently biopsied. Efforts to better understand the molecular basis of leptomeningeal carcinomatosis have been severely limited by the difficulties in isolating and characterizing CSF tumor cells. Consequently, almost nothing is known about the biology of metastatic tumor cells in the CNS, including in CSF. It is hypothesized that these tumor cells possess or acquire particular molecular features that enable them to colonize and proliferate in this space.

Our report is, to our knowledge, the first to provide detailed profiling of tumor cells isolated from the CSF, including genome-wide copy number, gene expression and mutation analyses. Although we previously described our approach for isolation and aCGH profiling of CTCs (8-10), the significance of the present work is the demonstration that similar methods can be used to study CSF tumor cells, which are directly responsible for the notably aggressive clinical syndrome of leptomeningeal carcinomatosis. Unlike CTC detection, which is neither
necessary nor sufficient for the diagnosis of breast cancer progression, the finding of tumor cells in the CSF is pathognomonic for leptomeningeal disease.

In this study, we demonstrated the feasibility of isolating CSFTCs using IE/FACS. Isolated CSFTCs were subjected to extensive molecular analyses, which confirmed their malignant origin. Copy number analysis of CSFTCs demonstrated a wide range of genomic aberrations frequently found in primary breast tumors. Furthermore, gene expression profiling confirmed the epithelial nature of the CSFTCs and demonstrated that they were readily distinguishable from the leukocyte population present in the CSF.

It is possible that our anti-EPCAM-based IE strategy may have missed tumor cells in the CSF with low EPCAM expression, which has been proposed in the case of CTC detection (26, 27). However, in this study of 15 patients with leptomeningeal disease, all were positive for CSFTCs using anti-EPCAM-based enrichment.

The evolutionary relationship between these metastatic cells in the CSF and the originating primary tumor is unknown. A recent study focused on HER2 FISH on CSFTCs, and suggested high concordance (14 out of 16) of HER2 copy status between CSFTCs and the primary tumor (28). We also observed high concordance between HER2 status of the CSFTCs and clinical HER2 of the corresponding primary tumor.

Our approach has yielded new insights into CSFTCs and their relationship with primary tumor tissue at the genome-wide level. Direct comparative analysis of CSFTCs and matched tumor samples, including primary tumors from the same patient, revealed conserved copy number alterations and clear evidence of clonality. We also observed divergence, including significantly more overall genomic aberrancy in CSFTCs than in the corresponding primary tumor, as well as specific genomic regions frequently altered in CSFTCs but not primary tumors.
Given the limited sample size of this pilot study, testing in a larger cohort is warranted, especially in patients with HER2 positive disease where the incidence of CNS metastasis and leptomeningeal spread is more common (29).

Molecular profiling of CSFTCs may lead to the discovery of candidate biomarkers and therapeutic targets relevant to CNS metastasis. For example, alterations including 8q24 (MYC) gain was observed in our cohort, suggesting opportunities for new targeted therapeutic strategies. We also demonstrated the feasibility and reproducibility of profiling CSFTCs from serial samples. Given the potential for serial analysis, further applications in real-time assessments of tumor evolution, monitoring of therapeutic effects, and personalization of therapy can be envisioned.
References


Figure Legends

Fig. 1. Genomic aberrations in CSFTCs. Frequency plot of copy number alterations in 13 CSFTC samples. Copy number gains and losses are indicated as blue and red, respectively. The thresholds of log2 ratio values for single copy number gains and losses were 0.20 and −0.20, respectively; the thresholds for high-level number gains and losses were 0.6 and −0.6, respectively.

Fig. 2. Copy number alterations in CSFTCs and corresponding archival tumors. Copy number profiling from a series of CSFTCs and corresponding archival primary tumors in patients, A) 4037, B) 4039, C) 4038, and D) CSF6. The log2 ratio value for each BAC clone is plotted on the y-axis. The x-axis represents the genomic position of each BAC clone on the array, with odd-numbered chromosomes shown below the plot and even-numbered chromosomes shown above the plot. Vertical solid lines indicate chromosome boundaries, and vertical red dashed lines represent the centromeric region dividing each chromosome into the p- or short arm (to the left of centromere) and the q- or long arm (to the right of the centromere). Green indicates copy number gain, red indicates copy number loss, black indicates no change, and blue indicates focal amplification. In the profiles for patients 4038 and CSF6, the arrows indicate HER2/ERBB2 focal amplification on chromosome 17.

Fig. 3. Copy number and FISH analysis in CSFTCs and corresponding archival tumors. A) Copy number profiling of CSFTCs and corresponding archival tumors from patient 4015. The arrow indicates focal amplification of MYC on chromosome 8. B) Fluorescence in situ hybridization (FISH) analysis of CSFTCs confirmed MYC amplification.

Fig. 4. Gene expression profiling of CSFTCs and matched leukocyte controls. A) Unsupervised hierarchical clustering analysis of CSFTCs (18 samples from 5 patients), including
serial analysis (indicated as T1, T2, T3 etc.) and matched leukocytes (n=9) from a subset of patients serving as non-tumor controls. The heat map displays normalized gene expression (ΔCt) of a panel of cancer-related genes including EPCAM and leukocyte-specific marker PTPRC (CD45). B) Differential expression analysis between paired CSFTCs and leukocytes obtained from the same enriched sample from 5 patients. Leukocytes (CD45) were treated as calibrator samples. Analysis was performed using a paired t-test. Genes with an adjusted p-value of <0.05 were considered statistically significant. Relative quantification (RQ) is reported in the logarithmic scale \(\log_{10} RQ = \log_{10} 2^{-\Delta\Delta CT}\) plotted on the y-axis. A \(\log_{10} RQ = 1\) or -1 means a gene is expressed 10 times or 1/10 as much, respectively, in CSFTC relative to leukocyte samples. C) ERBB2 (HER2) expression in CSFTCs relative to leukocytes in 5 patients. Clinical HER2 status are indicated on the x-axis.
Figure 2

a) CSFTC

b) CSFTC

c) CSFTC

d) CSFTC

Primary Tumor

Primary Tumor

Primary Tumor

Primary Tumor

Mesenteric Node Metastasis (autopsy)
Figure 3

a

Patient #4015

CSFTC

CTC

Primary Tumor

Axillary Node Metastases

b

FISH of CSFTCs

MYC - orange
CEP8 - aqua
Nucleus - blue
Figure 4

(a) Heatmap showing gene expression levels for PTPRC and EPCAM. The expression levels are color-coded from low to high.

(b) Bar graph showing Log2 RQ values for AGR2, EPCAM, ORBT, PTPRC, and TFF3. The expression levels are indicated as up or down-regulated in CSFTC.

(c) Bar graph showing FOSZ expression as a ratio relative to housekeeping genes. The expression levels are indicated as negative or positive.
Table 1. Patient and tumor characteristics. Clinical data and relevant information on CSFTC and corresponding tumor samples from patients in the study.

<table>
<thead>
<tr>
<th>Patient ID</th>
<th>Age</th>
<th>CTC(^a) per 7.5mL blood</th>
<th>aCGH(^a) data</th>
<th>Matched PT(^a) aCGH data</th>
<th>Matched mutation data</th>
<th>Expression data</th>
<th>Clinical status</th>
<th>Status by aCGH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ER(^a) PR(^a) HER2</td>
<td>CSFTC(^a) PT</td>
</tr>
<tr>
<td>108</td>
<td>41</td>
<td>12</td>
<td>yes</td>
<td>.</td>
<td>.</td>
<td>pos</td>
<td>pos</td>
<td>neg</td>
</tr>
<tr>
<td>112</td>
<td>39</td>
<td>0</td>
<td>yes</td>
<td>.</td>
<td>.</td>
<td>pos</td>
<td>pos</td>
<td>pos</td>
</tr>
<tr>
<td>4011(^b)</td>
<td>51</td>
<td>0</td>
<td>yes</td>
<td>yes</td>
<td>.</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
</tr>
<tr>
<td>4014</td>
<td>63</td>
<td>6</td>
<td>yes</td>
<td>.</td>
<td>.</td>
<td>pos</td>
<td>pos</td>
<td>pos</td>
</tr>
<tr>
<td>4015</td>
<td>54</td>
<td>37</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>neg</td>
<td>neg</td>
<td>neg</td>
</tr>
<tr>
<td>4037</td>
<td>50</td>
<td>.</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>pos</td>
<td>pos</td>
<td>pos</td>
</tr>
<tr>
<td>4038</td>
<td>40</td>
<td>.</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>pos</td>
<td>pos</td>
<td>pos</td>
</tr>
<tr>
<td>4039</td>
<td>51</td>
<td>.</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>pos</td>
<td>pos</td>
<td>neg</td>
</tr>
<tr>
<td>CSF1</td>
<td>67</td>
<td>.</td>
<td>yes</td>
<td>.</td>
<td>.</td>
<td>pos</td>
<td>neg</td>
<td>pos</td>
</tr>
<tr>
<td>CSF2</td>
<td>39</td>
<td>5</td>
<td>yes</td>
<td>.</td>
<td>.</td>
<td>pos</td>
<td>pos</td>
<td>neg</td>
</tr>
<tr>
<td>CSF3</td>
<td>68</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>neg</td>
<td>neg</td>
<td>pos</td>
</tr>
<tr>
<td>CSF4</td>
<td>55</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>pos</td>
<td>neg</td>
<td>pos</td>
</tr>
<tr>
<td>CSF5</td>
<td>51</td>
<td>.</td>
<td>yes</td>
<td>.</td>
<td>.</td>
<td>pos</td>
<td>neg</td>
<td>neg</td>
</tr>
<tr>
<td>CSF6</td>
<td>39</td>
<td>.</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>neg</td>
<td>neg</td>
<td>pos</td>
</tr>
<tr>
<td>CSF7</td>
<td>79</td>
<td>.</td>
<td>yes</td>
<td>.</td>
<td>.</td>
<td>pos</td>
<td>pos</td>
<td>neg</td>
</tr>
</tbody>
</table>

\(^a\)CTC= circulating tumor cells, aCGH = array comparative genomic hybridization, PT = primary tumor, CSFTC = tumor cells from CSF, ER = estrogen receptor, PR = progesterone receptor, pos = positive, neg = negative, pos\(^{\text{amp}}\) = positive with focal amplification, (.) = no data; \(^b\)Five serial CSF samples were collected from patient 4011.
Table 2. Recurrent copy number aberrations in CSFTCs. List of common gains and losses in the CSFTC samples (n=13). The rank segmentation algorithm was used to estimate copy number using the significance threshold of p-value <0.001. Regions of gains and losses present in ≥50% of the samples were considered recurrent. Chr- chromosome.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cytoband Location</th>
<th>Event</th>
<th>No. of genes in the region</th>
<th>Frequency %</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr1</td>
<td>chr1:85,166,259-94,234,657 p22.3 - p22.1</td>
<td>Loss</td>
<td>72</td>
<td>53.8</td>
</tr>
<tr>
<td>chr1</td>
<td>chr1:179,320,361-214,349,532 q25.3 - q41</td>
<td>Gain</td>
<td>251</td>
<td>76.9</td>
</tr>
<tr>
<td>chr3</td>
<td>chr3:58,536,745-91,700,000 p14.2 - q11.1</td>
<td>Loss</td>
<td>81</td>
<td>53.8</td>
</tr>
<tr>
<td>chr4</td>
<td>chr4:8,291,188-19,278,038 p16.1 - p15.31</td>
<td>Loss</td>
<td>206</td>
<td>53.8</td>
</tr>
<tr>
<td>chr8</td>
<td>chr8:2,218,233-8,270,256 p23.2 - p23.1</td>
<td>Loss</td>
<td>67</td>
<td>69.2</td>
</tr>
<tr>
<td>chr8</td>
<td>chr8:14,358,974-16,795,085 p22</td>
<td>Loss</td>
<td>4</td>
<td>53.8</td>
</tr>
<tr>
<td>chr8</td>
<td>chr8:23,103,650-30,031,602 p21.3 - p12</td>
<td>Loss</td>
<td>55</td>
<td>61.5</td>
</tr>
<tr>
<td>chr8</td>
<td>chr8:61,470,364-65,191,920 q12.1 - q12.3</td>
<td>Gain</td>
<td>12</td>
<td>61.5</td>
</tr>
<tr>
<td>chr8</td>
<td>chr8:77,495,948-78,897,673 q21.11 - q21.12</td>
<td>Gain</td>
<td>3</td>
<td>69.2</td>
</tr>
<tr>
<td>chr8</td>
<td>chr8:99,355,501-114,713,038 q22.2 - q23.3</td>
<td>Gain</td>
<td>59</td>
<td>84.6</td>
</tr>
<tr>
<td>chr8</td>
<td>chr8:125,617,193-129,974,351 q24.13 - q24.21</td>
<td>Gain</td>
<td>21</td>
<td>92.3</td>
</tr>
<tr>
<td>chr11</td>
<td>chr11:4,760,685-14,357,699 p15.4 - p15.2</td>
<td>Loss</td>
<td>158</td>
<td>53.8</td>
</tr>
<tr>
<td>chr11</td>
<td>chr11:14,656,897-17,718,436 p15.2 - p15.1</td>
<td>Loss</td>
<td>18</td>
<td>53.8</td>
</tr>
<tr>
<td>chr20</td>
<td>chr20:52,085,105-56,357,440 q13.2 - q13.32</td>
<td>Gain</td>
<td>30</td>
<td>53.8</td>
</tr>
</tbody>
</table>
Molecular profiling of tumor cells in cerebrospinal fluid and matched primary tumors from metastatic breast cancer patients with leptomeningeal carcinomatosis

Mark Jesus M. Magbanua, Michelle Melisko, Ritu Roy, et al.

Cancer Res  Published OnlineFirst October 18, 2013.