Bone metastases in gastric cancer follow a RANKL-independent mechanism.

(Article begins on next page)

07 May 2024

UNIVERSITÀ DEGLI STUDI DI TORINO

D'Amico L1, Satolli MA, Mecca C, Castiglione A, Ceccarelli M, D'Amelio P, Garino M, De Giuli M, Sandrucci S, Ferracini R, Roato I.

Primary breast cancer stem-like cells metastasize to bone and switch phenotype acquiring a bone tropism signature

Lucia D'Amico [1], Salvatore Patanè [1], Cristina Grange [2], Benedetta Bussolati [2], Claudio Isella [3], Lara Fontani [4], Laura Godio [5], Michele Cilli [6], Patrizia D'Amelio [7], Giancarlo Isaia [7], Enzo Medico [3]*, Riccardo Ferracini [8], Ilaria Roato [1]*

[1] CeRMS (Center for Experimental Research and Medical Studies), A.O. Città della Salute e della Scienza di Torino, via Santena 5, 10126 Torino, Italy

[2] Department of Internal Medicine, Molecular Biotechnology Center, University of Turin, C.so Dogliotti 14, 10126 Torino, Italy

[3] Department of Oncological Sciences, University of Turin and Laboratory of Oncogenomics, Institute for Cancer Research and Treatment (IRCC), Strada Provinciale 142, km 3.95, 10060 Candiolo, Italy

[4] Laboratory for Gene Transfer and Therapy, Institute for Cancer Research and Treatment (IRCC), Strada Provinciale 142, km 3.95, 10060 Candiolo, Italy

[5] Department of Pathology, A.O. Città della Salute e della Scienza di Torino, via Santena 7, 10126 Torino, Italy

[6] Stabulation Facility, National Institute for Cancer Research, Genova, L.go R. Benzi 10, 16132, Genova, Italy

[7] Department of Surgical and Medical disciplines, University of Turin, C.so Bramante 88, 10126 Torino, Italy

[8] Department of Orthopaedics, A.O. Città della Salute e della Scienza di Torino, C.so Bramante 88, 10126 Torino, Italy

* Corresponding Authors: Ilaria Roato: CeRMS, A.O. Città della Salute e della Scienza, via Santena 5, 10126 Torino, Italy. Tel +39-011-6334672; e-mail: roato78@libero.it

Enzo Medico: Institute for Cancer Research and Treatment, S.P. 142, km 3,95 – 10060 Candiolo (TO, Italy). Tel: +39-011-9933234; e-mail: enzo.medico@ircc.it

Abstract

BACKGROUND: Bone metastases represent a common and severe complication in breast cancer and the implication of cancer stem cells (CSCs) in promoting bone metastasis is currently under discussion. Here we used a human-in-mice model to study bone metastases formation due to primary breast CSCs-like colonization. METHODS: Primary $CD44^+CD24^-$ breast CSCs-like were transduced by a luciferase-lentiviral vector and injected through subcutaneous and intracardiac route in non-obese/severe combined immunodeficient mice carrying subcutaneous human bone implants. The CSCs-like localization was monitored by *in vivo* luciferase imaging. Bone metastatic CSCs-like were analysed through immunohistochemistry and flow citometry, and gene expression analysis were performed by microarray techniques. RESULTS: Breast CSCs-like colonized the human implanted bone, who resulted remodelled. Bone metastatic lesions were histologically apparent, with tumor cells expressing epithelial markers and vimentin. The bone isolated CSCs-like were $CD44^-CD24^+$, and showed tumorigenic abilities after injection in secondary mice. $CD44^-CD24^+$ CSCs-like displayed a distinct bone tropism signature, enriched in genes discriminating bone metastases of breast cancer from metastases at other organs.

CONCLUSION: Breast CSCs-like promote bone metastasis and displayed a CSCs-like bone tropism signature. This signature shows a clinical prognostic relevance, indeed it efficiently discriminates osteotropic breast cancers from tumor metastasizing to other sites.

Key words: bone metastasis, breast cancer, cancer stem cell, human-in-mice model, osteotropism

Introduction

Breast tumors are characterised by a phenotypic heterogeneity. They are composed of different groups of cells, which display a remarkable variability in many phenotypic traits, including the different ability to metastasize and to survive after therapy. The cancer stem cell hypothesis may explain the intra-tumor heterogeneity, suggesting that the variety between functionally important properties of individual cells arises from differences in their differentiation status (Campbell LL *et al*, 2007; Shackleton M *et al*, 2009). Cancer stem cells (CSCs) are defined as a subset of tumor-initiating cells, sharing surface markers with somatic stem/progenitor cells of their tissue of origin. They are able to self renewal, to proliferate and to differentiate, giving rise to a heterogeneous tumor cell population (Clarke MF *et al*, 2006; Rosen JM *et al*, 2009). In breast cancer, CSCs are identified as a rare population of $CD44^+CD24^-$ ephitelial-specific antigen^{low} cells, showing *in vivo* tumorigenicity based on xenotrasplant assays in non-obese/severe combined immunodeficient (NOD/SCID) mice (Al-Hajj M *et al*, 2003; Shipitsin M *et al*, 2007). Breast cancer cells are known for a peculiar bone tropism: indeed, 70% of patients develop bone metastases (Mundy GR, 2002). The metastatic process sometimes occurs very early in the natural history of the disease, while in other cases at later stages. Anyway, once the tumor has spread to the bone, cancer is no longer curable (Mundy GR *et al*, 2008). Recent reports highlighted the molecular mechanisms underlying the dissemination in bone of breast cancer cells, and lead to the identification of discrete sets of osteotropic genes associated with bone tropism (Kang Y *et al*, 2003; Liu R *et al*, 2007; Minn AJ *et al*, 2005; van den Hoogen C *et al*, 2010). Evidence that CSCs are responsible for bone metastases is still not conclusive, due to the difficulties to isolate CSCs, and to a lack of a valid animal model (Buijs JT *et al*, 2011; Liu H *et al*, 2010; van den Hoogen C *et al*, 2010). This paper investigates the feature of a non-metastatic human breast CSCs to metastasize to bone. NOD/SCID-human-in-mice model was utilised to inject $CD44^+CD24^-$ CSCs-like,

isolated from a non-bone metastatic primary breast cancer (Kuperwasser C *et al*, 2005; Roato I *et al*, 2010). These mice were implanted subcutaneously with a fragment of human bone to establish a human microenvironment. These breast $CD44^+CD24^-$ CSCs-like acquired the ability to metastasize the implanted bone and showed a switched phenotype $CD44^-CD24^+$, maintaining tumorigenic abilities *in vivo*. $CD44^-CD24^+$CSCs-like isolated from bone expressed a signature which persisted at subsequent passages also in the absence of surrounding bone tissue and it was significantly enriched in sets of genes biologically relevant in the process of cancer dissemination (Subramanian A *et al*, 2005). This signature derived from CSCs showed a clinical prognostic relevance because it efficiently discriminated osteotropic breast cancers from tumor metastasizing to other sites.

Methods

Breast CSCs-like

Breast CSCs-like were previously isolated and characterized for the expression of stemness markers, showing a $CD44^+CD24^-$ phenotype, and for tumorigenic potential (Bussolati B *et al*, 2009). Tumor specimens were obtained from a consenting patient, according to the Ethics Commitee of the San Giovanni Battista Hospital of Torino, Italy. The histologic analysis showed a lobular-infiltrating carcinoma of the pleomorphic type expressing oestrogen receptor in about 60% of cells and HER2-negative. As previously described (Bussolati B *et al*, 2009), cell suspension obtained from tumor tissue was processed and mammospheres were derived. These mammospheres were studied for their ability to metastasize bone, since patients with breast cancer frequently develop bone metastases (Smid M *et al*, 2008). $CD44^+CD24^-$ CSCs-like were transduced with a luciferase-lentiviral expression vector, originating Luc-CSCs-like. Both breast CSCs-like and Luc-CSCs-like were injected in mice to rule out possible modifications of the tumorigenic phenotype due to lentiviral transduction.

Human-in-mice model of bone metastasis induced by breast CSCs-like

Experimental animals were treated in compliance with the actual national and international guidelines (the Italian legislative decree 116/92 and the European Community Directive 86/609 CEE) and in accordance to the authorization provided by Italian Ministry of Health (as of D.M. 44/1994-A and subsequent integrations). All of the orthotopic xenograft models were established in NOD/SCID mice as previously described (Roato I *et al*, 2010). A fresh fragment of human bone obtained from the discarded femoral head of an adult patient submitted to total joint replacement (after the patient's informed consent), was transplanted subcutaneously in the left flank of 15 NOD/SCID 5-week-old female mice (Charles River Laboratories Italia). In 6 mice, 5-weeks old, breast CSCs-like (3 mice) and Luc-CSCs-like (3 mice) were injected subcutaneously (SC) close to bone implant; whereas in other 6 mice CSCs-like (3 mice) and Luc-CSCs-like (3 mice) were injected by intracardiac (IC) route to demonstrate the actual capability of CSCs to metastasize to bone. For SC injections, $1.5x\,10^5$ breast CSCs-like were resuspended in PBS and Matrigel 1:3 (BD Biosciences) and injected in a volume of 40 µL, using a 25-gauge needle. $1x10^3$ CSCs-like were injected through the IC route, in the left ventricle. The development of tumor mass and metastases to bone and/or other organs were monitored by IVIS for 45 days after the injection. In order to evaluate second tumor generation, mammospheres obtained from the metastatic lesion in human-implanted bone were cultured for one week, controlled for the phenotype by flow citometry for the expression od CD44 and CD24, then re-injected in other 6 mice (3 SC and 3 IC) with the same experimental conditions indicated above. The presence of circulating human IgG in mice sera was demonstrated by human IgG ELISA, purchased by ICL Inc.

Lentivirus production and breast CSCs-like transduction

Vector stocks were produced by transient transfection of the Luciferase transfer plasmid, the packaging plasmids pMDLg/pRRE and pRSV.REV, and the vescicular stomatitis virus (VSV) envelope plasmid pMD2.VSV-G (15, 6.5, 2.5, and 3.5 µg, respectively, for 10-cm dishes) in 293T. After 12-14 h the transfection solution was

removed and the cells washed twice with PBS 1X. The media was replaced and after 24 h was collected. The viral supernatants were filtered using a 0.45-μm low-protein–binding filter flask (Millipore) and viral particles were concentrated by ultracentrifugation as described (Follenzi A *et al*, 2000). Determination of the viral p24 antigen concentration was done by HIV-1 p24 Core profile ELISA (Perkin-Elmer Life).

Breast CSCs-like were resuspended in Opti-MEM in the presence of 8 μg/mL polybrene (Sigma-Aldritch) and plated in 96-well round-bottom plates. Then high-titer lentiviruses (210ng of p24 gag equivalent particles) were added to the cells and incubated for 4 hours.

Breast Luc-CSCs-like were tested for the Luciferase expression by using Luciferase Assay System (Promega Corp.) as described in protocol kit. Briefly, cells were rinsed with PBS and lysed with RLB lysis buffer by performing a single freeze-thaw cycle. 20μl of cell lysate was added to a luminometer tube containing the Luciferase Assay Reagent and read by performing a 2-second measurement delay followed by a 10-second measurement read for luciferase activity. Importantly, we did not observe significant changes in tumor phenotypes associated with lentiviral transduction, as demonstrated by histological appearance, flow cytometry analysis (see Fig. S2A-B) and micro-array analysis of parental breast CSCs-like and Luc-CSCs-like.

In vivo Bioluminescence imaging

For the bioluminescence imaging, mice were anesthetized with isoflurane inhalation, and were subsequently intraperitoneally injected with 200 μl of 15 mg/mL D-luciferin (Caliper Life Science). The bioluminescence signals were monitored using the IVIS system 2000 series (Xenogen Corp.) consisting of a highly sensitive cooled CCD camera. Two kinetic bioluminescent acquisitions were collected between 0 and 20 min after D-luciferin injection to confirm the peak photon emission recorded as maximum photon efflux per second; imaging times ranged from 1 to 60 s, depending on the amount of luciferase activity. Data were analyzed using the total photon flux emission (photons/second) in the regions of interest (ROI) defined manually. At various time

points after tumor implantation (25, 35 and 45 days), the mice were imaged using the IVIS 2000 system.

Isolation of bone metastatic cells

To isolate metastatic cells from the osteolytic lesions, human implanted bone was retrieved, finely minced and then digested by incubation for 30 min at 37°C in D-MEM containing collagenase I (Sigma-Aldrich). After collagenase neutralization, cells were washed by Hank's balanced saline solution (Lonza) and red blood cells were lysed with Red Blood cell Lysis Solution (Promega). Cell suspension was forced through a graded series of meshes to separate the cell components from stroma and aggregates. After filtration through 40 µL, single cells were plated in a selective medium serum-free DMEM-F12 (Gibco, Invitrogen), supplemented with 10ng/ml basic fibroblast growth factor (b-FGF), 20 ng/ml epidermal growth factor (EGF) (PeproTech), 5µg/ml insulin and 0,4% bovine serum albumine (Sigma-Aldrich), to avoid the presence of non-neoplastic contaminating cells. After one week, the appearance of non-adherent spherical clusters of cells, i.e. mammospheres, was observed. Mammospheres were collected on the bottom of a conical tube by spontaneous precipitation to remove non-living cells. After 5 days, mammospheres were collected by gentle centrifugation and disaggregated through a Non-enzymatic Cell Dissociation Solution (Sigma-Aldritch). In order to analyse the phenotype of these bone-derived mammospheres, the dissociated cells were counted and stained for flow cytometry analysis with anti-human CD44PE, CD24FITC (BD Pharmingen). For every antibody we used the relative isotypic control. Samples were analyzed in a FACs Calibur instrument and elaborated by Flowjo (Treestar).

Immunohistochemistry and histological analysis

Immunohistochemistry was performed on tissues fixed in 10% neutral buffered formalin and bone tissues were decalcified with EDTA treatment until soft. Tissues were embedded in paraffin, sections were deparaffinized, rehydrated through graded alcohols and subjected to antigen retrieval for immunohistochemistry. Sections were

stained for H&E to study the morphology. The presence of human vessels was demonstrated by staining for anti-CD34, clone QbndN/10 (Neomarkers) whereas tumors cells were stained by mouse monoclonal antibodies against low molecular weight cytokeratins (CK AE1/AE3, clone AE1/AE3/PCK26), epithelial membrane antigen (EMA, clone E29) and vimentin (clone R9) from DAKO; CD44 (clone C26) and CD24 (clone ML5) from BD Pharmingen. To identify the collagen fibers on the new bone, Trichrome stain was perfomed by Gomori's tricrhome stain kit (DAKO). TRAP staining was performed to show osteoclasts according to manufacturer's instruction (Roche).

CSC RNA extraction and microarray analysis

Total RNA was extracted with the miRNeasy Mini Kit (Qiagen). Cells were collected as a pellet, homogenized directly in QIAzol Lysis Reagent (a maximum of 10^7 cells in 0.7 ml of reagent) and processed according to the manufacturer's protocol. RNA quality was controlled on a Bioanalyzer 2100 (Agilent). Biotinylated antisense RNA (aRNA) was prepared from 500 ng of total RNA using the Illumina RNA Amplification Kit (Ambion) according to the manufacturer's directions, and quality controlled on a Bioanalyzer 2100 (Agilent). Labeled aRNA was hybridized on Illumina HumanHT-12 v4 expression BeadChips for 16 hours. Hybridized BeadChips were subsequently washed, stained with streptavidin-Cy3 and scanned on a BeadStation 500 (Illumina).

Microarray data analysis

Raw microarray data were processed and analyzed with the GenomeStudio software (Illumina) according to standard procedures. Briefly, bead-level data were summarized and cubic spline-normalized. Samples were subsequently organized in the following groups: (i) Breast CSCs-like, CD44$^+$CD24$^-$, (duplicate); (ii) Luc-CSCs-like, CD44$^+$CD24$^-$, (simplicate); (iii) Bone-isolated metastatic CSCs-like, CD44$^-$CD24$^+$, (duplicate) and subsequently (iv) grown *in vitro* as spheroids (simplicate) or (v) re-grown in subcutaneous implants (duplicate). The "Illumina Custom" variant of the T-test implemented in GenomeStudio was employed to select genes with differential

expression between Breast CSCs-like and bone-isolated metastatic CD44⁻CD24⁺ CSCs-like, (p-value < 0.001, fold-change >3), which resulted in 860 probes selected (of which 710 up-regulated and 150 downregulated in bone metastasis). Subsequently, probes were filtered to remove those showing differential expression (fold change >1.5) between breast CSCs-like and Luc-CSCs-like. This filter reduced up-regulated probes to 480 and down-regulated probes to 100. Finally, up- and down-regulated probes were tested for maintaining the change (>2-fold) compared with breast CSCs-like also in derivatives of bone-isolated metastatic CD44⁻CD24⁺ CSCs-like, i.e. spheroids and subcutaneous implants. Up-regulated and down-regulated probes passing this final filter were 96 and 23, respectively corresponding to 88 and 22 genes, collectively defining the breast CSCs-like bone tropism signature (see Table S3). Hierarchical clustering was performed using GEDAS Software (Fu L et al, 2007). The breast cell line microarray dataset was generated by Kang et al. (Kang Y et al, 2003) and obtained as supplementary material on the publisher's website. The breast cancer metestasis dataset generated by Zhang et al. (Zhang XH et al, 2009) was retrieved from the Gene Expression Omnibus database (GSE14020). GSE14020 is a composite of two sub-datasets obtained with two different Affymetrix array types. To homogenize them, data were cross-mapped based on probe-sets and scaled on median expression. GSEA (Mootha VK et al, 2003; Subramanian A et al, 2005) was performed online using the Broad Institute application (http://www.broadinstitute.org/gsea/msigdb/downloads.jsp) on our expression data filtered for detection. Genes represented by multiple probes were collapsed to the probe with the maximum value. From the Molecular Signatures Database of GSEA, we selected for the analysis the collections C2 (curated gene sets; file: c2.v3.symbols.gmt) and C5 (Gene Ontology; file: c5.v3.symbols.gmt). Enrichment statistics were performed with default settings and Signal_to_Noise metric with 1000 permutations. GSEA analysis considering the CSCs-like bone tropism signature as a gene set was run on published datasets (Kang Y et al, 2003; Zhang XH et al, 2009) to test for enrichment in genes with differential expression between bone-metastatic vs

non-metastatic cell or tissues. Gene set enrichment score was calculated on the "Ratio_of_classes" score for the genes in the set and FDR was evaluated with 1000 permutations.

Serum detection of IL-6, IL-8 and CCL20

To analyse the expression of human IL-6, IL-8 and CCL-20 in mice we utilised the multi analyte detection system Milliplex Map according to manufacturer' instructions. The molecule detection was performed with the instrument Luminex® 200™ (Luminex Corporation) and data were analysed through the MILLIPLEX analyst software.

Results

Breast CSCs-like show tumorigenic and metastatic characteristics in the human-in-mice model

To verify the ability of breast CSCs-like to metastasize bone, we utilized CSCs-like, previously isolated from a primary breast tumor. These cells were CD44$^+$CD24$^-$, grew as mammospheres and they expressed the stem cells markers as Oct-4 and nestin, but not the differentiation markers CK14, CK18 and α-SMA. They grew in SCID mice and were able to generate serial tumor, when re-injected in other mice (Bussolati B *et al*, 2009). After the isolation from the primary tumor these breast CSC-like cells were injected subcutaneously (SC) and in intracardiac route (IC) in NOD/SCID mice, carrying a small piece of human bone, previously implanted in a flank. The introduction of human bone in mice allowed us to study in a specie-specific manner the interaction between breast CSCs-like and bone microenvironment. Before the subcutaneous (SC) and intracardiac (IC) injection, we transduced these breast CSCs-like through Luciferase-lentiviral vector (Luc-CSCs-like) in order to monitor their localization *in vivo*. The number of mice for experimental group, the mice survival, the percentages of bone engraftment, bone and lung metastases are indicated in Table S1. To analyse the intensity of the Luciferase signal released from Luc-CSCs-like, increasing amounts of them were plated in a multiwell plate and subjected to IVIS imaging. We determined a

Luciferase signal starting from 5×10^2 cells and its intensity was roughly proportional to the different cell dilution (Figure S1). ~~After the~~ Luc-CSCs-like SC injection we ~~The tumor growth was~~ monitored the tumor growth at different time points for 45 days. After 20 days ~~from the Luc-CSCs-like injection,~~ the primary tumors were macroscopically evident and bone localizations were detectable (Figure 1A). The intensity of the signal in the bone increased progressively during the weeks, becoming stronger in the bone than in the tumor mass (Figure 1B-C). This result was confirmed by the quantification of the mean Luciferase intensity in bone and in SC tumor mass, showing a higher value in bone than in SC masses, particularly at day 45, $p < 0.05$ (Figure 1D). These data confirm our hypothesis that breast CSCs-like show osteotropism, moreover they find a particularly favourable soil in the human bone microenvironment because their number significantly increased in bone. In 5/12 mice, we detected lung metastases at day 20, then the animals were sacrificed before the occurence of the bone lesion formation (see also Figure S2 and Table S1). We also tested the bone metastatic ability of the subpopulation of CD44-/CD24+ cells, showing that they did not metastasize to bone (data not shown).

Breast CSCs-like metastasize human bone

To rule out the metastatisation of our breast CSCs-like to mice bone, we subjected mice to X-rays and we did not detect lesions in mice bone (Figure S3). To verify the validity of our human-in-mice model, we need to analyse the viability of the implanted bone in mice. The bone implants resulted viable because little to no necrosis was evident in the grafts, whereas bone marrow cells, mineralised area, stromal cells and neo-vascularization were present (Figure 2A). The engraftment of the human bone was also demonstrated by the newly synthesized bone present in the control (Figure 2B). Neo-vascularization was present and vessels expressed the human CD34, allowing to exclude the presence of murine vessels in the bone implants (Figure 2C). Breast CSCs-like colonized the human implanted bone: large area of invading tumor cells was evident and live bone resulted remodelled, both after IC and SC injection (Figure 2D

and 2E, respectively). The marked increase in bone remodelling due to bone metastases was evaluated through the Trichrome staining, which showed large area of new bone apposition (Figure 2F). In detail, the area of neo-apposition was lower (blue stain) in controls (Figure 2B) than in bone invaded by breast CSCs-like (Figure 2F). Moreover, we detected a lower number of TRAP positive osteoclasts in controls (Figure 2G) than in bone with metastatic lesion (Figure 2H). These results demonstrate an increase in bone remodelling due to the CSCs-like metastatisation. To demonstrate the human bone marrow activity in the implanted bone, we dosed human IgG in the serum of mice, detecting a medium level of 80 ug/ml (data not shown).

The bone-isolated metastatic cells show a phenotypic switch

To analyse the phenotype of breast CSCs-like injected in mice and retrieved from tumor masses and bone lesions, we utilised both FACS analysis and histological investigation for the expression of CD44 and CD24. The injected breast CSCs-like were CD44$^+$CD24$^-$ (96% $\pm$ 3,2%) (Figure 3B, see also Figure S4), but after their bone retrieval they mainly showed a CD44$^-$CD24$^+$ phenotype (40% $\pm$ 4,1%) (Figure 3C). The analysis of primary tumor masses demonstrated that they were mainly constituted of CD44$^+$CD24$^-$ cells (Figure 4A-D), whereas the bone lesions mainly showed CD44$^-$CD24$^+$ cells (Figure 4E-H). These data confirm the FACS results and suggest a phenotypic switch of the injected CD44$^+$CD24$^-$ cells, likely induced by the bone microenvironment. The CD44$^-$CD24$^+$ cells present in the bone lesions, also expressed low molecular weight CKs, EMA and vimentin (Figure 5A-C). This expression of both epithelial and mesenchymal markers suggest that CD44$^-$CD24$^+$ cells have the capability to perform epithelial mesenchymal transition (EMT) and viceversa (MET). To verify the tumorigenity of the CD44$^-$CD24$^+$ bone-isolated metastatic cells, we re-injected them in NOD/SCID mice. They showed a more aggressive growth SC than the parental CD44$^+$CD24$^-$ cells, because each mouse developed two or three tumor masses (Figure 6A, B), which were constituted by cellular heterogeneity as the original tumor. Cells in the secondary tumor mass expressed both CD44 and CD24

13

(Figure 6C,D), suggesting the ability of CD44$^-$CD24$^+$ cells to switch their phenotype, likely through an EMT since they acquired again the CD44 expression.

In order to rule out that this phenotypic change depends simply on the metastatic state of these cells rather than on the metastatic site, we also examined CD44 and CD24 expression in the lung metastases. Tumor cells in the lung lesions expressed both CD44 and CD24 (Figure 7), which indicates a different phenotype compared to the bone metastatic one. Thus the phenotypic switch observed in the CD44$^+$CD24$^-$ breast CSCs-like seems to depend on the contact with the human implanted bone, underlying the importance of the bone microenvironment in the metastatic process.

Identification of a transcriptional signature for CD44$^-$CD24$^+$ breast CSCs-like

Global gene expression profiling was carried out to compare parental breast CSCs-like and Luc-CSCs-like with bone-isolated CSCs-like. To determine whether transcriptional changes associated to the passage in bone were stable, we also profiled the purified CSCs-like bone-isolated metastatic cells, subsequently grown *in vitro* as mammospheres, or *in vivo* as subcutaneous secondary implants (See material and Methods for details on the analysis). Indeed, a robust 110-gene transcriptional signature (breast CSCs-like bone tropism signature) was found to distinguish bone-isolated CD44$^-$CD24$^+$ CSCs-like from parental CD44$^+$CD24$^-$ CSCs-like. This signature persisted after *in vitro* and *in vivo* expansion in the absence of surrounding bone tissue. This CSC-like bone tropism signature is composed of 88 up-regulated and 22 down-regulated genes (respectively 96 and 23 probes, see Table S2), modulated more than 2-fold compared with parental CSCs-like in both mammosphere- and subcutaneous derivatives of bone-isolated metastatic cells. 17 of the genes in the signature were overexpressed more than 10-fold, among them there were carbonic anhydrase IX (CA9), known to promote EMT and to be associated with high incidences of metastasis (Shin HJ *et al*, 2011); two chemokines CCL20 and IL-8 involved in the tumor metastatic process (Beider K *et al*, 2009; Li A *et al*, 2003); the insulin-like growth factor binding protein 1 (IGFBP1) linked to breast cancer outcome (Goodwin PJ *et al*, 2002); the

transcription factor early growth response-1 (EGR-1), which is involved in the prostate carcinoma progression and enhances IL-8 transcription and secretion (Ma J *et al*, 2009). In serum of mice we detected CCL20 at higher levels in sera of mice with bone metastases than in mice without metastases and controls (Fig. S5). IL-8 and IL-6 were detectable only in sera of mice with bone metastases (371 $\pm$ 65,2 pg/ml and 826 $\pm$ 597,1 pg/ml, respectively), while in the control mice they were not detectable. These chemokines and cytokines might be responsible for the attraction of CSCs-like in bone and might support the phenotypic switch.

Analysis of functional modules in the breast CSCs-like bone tropism signature

To identify functional gene modules ("gene sets") differentially expressed we carried out Gene Set Enrichment Analysis (GSEA) (Subramanian A *et al*, 2005), utilizing $CD44^{+}CD24^{-}$ parental cells (breast CSCs-like, Luc-CSCs-like) and all cell populations $CD44^{-}CD24^{+}$ derived after bone colonization, i.e. CSCs-like directly isolated from bone lesions (2 samples), subsequently grown as mammospheres (1 sample) or further implanted subcutaneously (2 samples). Thus GSEA allowed us to test for permanent transcriptional features acquired during bone colonization. In the "curated gene set" collection (C2), a total of 93 gene sets was significantly enriched in up-regulated genes (FDR < 0.25 and *p*-value < 0.01, see also Table S3A). In the Gene Ontology collection (C5) we also identified 12 up-regulated gene sets (FDR < 0.25 and p < 0.01). The most relevant up-regulated gene sets were condensed into a specific spectrum of functional modules, summarized in Table 1 (see also Table S3A-B). These modules include hypoxia response, metabolic stress, growth factors, cytokine stimulation, chemotaxis and acid-base balance, all known to be relevant to the biology of the bone metastatic process.

Among the down-regulated gene sets, 67 and 19 (respectively for the C2 and C5 collections) had a *p*-value < 0.01, but no one reached FDR <0.25. However, the C2 gene set with the highest down-regulation enrichment was composed of genes up-regulated in brain-metastatic breast cancer

(SMID_BREAST_CANCER_RELAPSE_IN_BRAIN_UP). The fact that such genes are preferentially down-regulated in the CSCs-like bone tropism signature supports organ specificity of the signature rather than a generic metastatic propension. The C5 gene set with the highest downregulation enrichment was ORGANELLE_ORGANIZATION and BIOGENESIS, with a core constituted by proteins modulating cytoskeleton.

The breast CSCs-like bone tropism signature discriminates osteotropic breast cancers from tumors metastasizing other sites

The ability of this CSCs-like to distinguish cell lines with different bone tropism was tested by hierarchical clustering of a dataset generated by Kang and colleagues, in which human breast cancer cells with high bone tropism were obtained by *in vivo* selection from a starting cell line with lower bone metastatic propensity (3). Expression data for 75 probes of this dataset, mapping to 51 of the 110 bone tropism genes, were converted into Log(2) ratio against mean expression in all samples. Our signature was able to segregate the strong bone metastatic cell lines from the two parental lines from which they were derived (Figure 8A). GSEA analysis confirmed that the CSCs-like bone tropism signature is significantly enriched in genes either up- or down-regulated in bone metastatic cells vs. their parental counterparts (Figure 8B, see also Table S2).

To verify whether the CSCs-like bone tropism signature actually discriminates bone metastatic breast cancers from cancers with different metastatic location, we mapped the signature on a published dataset including expression profiles of 65 breast cancer samples from metastatic lesions in the bone (n=18), brain (n=22), liver (n=5) and lung (n=20) (Zhang XH *et al*, 2009). Expression data for 128 probes of this dataset, mapping to 74 of the 110 bone tropism genes, were converted into Log2 ratio against mean expression in all samples and hierarchically clustered (Figure 8C). Interestingly, the CSCs-like bone tropism signature was able to aggregate 13 of 18 bone metastases within a cluster, where a subset of signature genes was highly expressed (top-right part of the heatmap). Metastases at other sites, such as liver, brain and lung were clustered in other groups. GSEA analysis on this dataset showed significant

enrichment of the CSCs-like bone tropism signature in genes discriminating bone metastases from metastases at other organs (Figure 8D, see also Table S2).

Discussion

This work analysed the ability of isolated CD44$^+$CD24$^-$ breast CSCs-like to metastasize to human bone implanted in immuno-compromised mice. This model was chosen because it provides a human microenvironment, which exerts a key role in the growth of metastatic cells and for their interaction with specialized niche in the bone marrow (Guise T, 2010; Tzukerman M *et al*, 2003).

The results of this paper indicate that breast CSCs-like had a mesenchymal and migratory CD44$^+$CD24$^-$ phenotype and are endowed with a metastatic potential to human implanted bone. The finding that breast CSCs-like grow in bone suggests that breast CSCs-like encountered a particularly favourable soil in the human implanted bone. In the tumor masses grown SC and in the bone lesions, the CSCs-like expressed both epithelial and mesenchymal markers. These observations support the existence of a migrating CSC, which possesses features both of stemness and mobility, may undergo an epithelial-to-mesenchymal transition (EMT) and migrate to bone. CSCs-like isolated from the bone lesion were CD44$^-$CD24$^+$, likely they performed a mesenchymal-to-epithelial transition (MET) after reaching the bone, with consequent phenotypical switch from CD44$^+$CD24$^-$ to CD44$^-$CD24$^+$, driven by the interaction with bone microenvironment. Moreover, after re-injection in mice, the CD44$^-$CD24$^+$ bone-isolated CSCs-like formed tumor masses, constituted by cells mainly CD44$^+$CD24$^-$, indicating another phenotypic transition. In different carcinomas, the importance of EMT for the induction of stem characteristics have been demonstrated (Kong D *et al*, 2010; Wellner U *et al*, 2009). The transitions between different states observed in our CSCs-like could be also explained by a quantitative Markov model of cell-state interconversion. According to this model, breast cancer include discrete populations, randomly perform transitions between states, without increasing their proliferation rate, to reach a

progressive equilibrium proportion (Gupta PB *et al*, 2011; Harris MA *et al*, 2008). Thus, the capability of the CD44⁻CD24⁺ bone-isolated CSCs-like to originate subcutaneous and heterogeneous tumor masses, with a high percentage of CD44⁺CD24⁻, could be consistent with the attempt of these CSCs-like to recapitulate the differentiation state heterogeneity present in the parental tumor from which they were derived.

The phenotypic switch of breast CSCs-like could be due also to their interaction with the haematopoietic stem cell niche. Supporting this view is the recent report that prostate cancer cells are able to occupy the haematopoietic stem cell niche in the bone marrow, which serves for tumor dissemination and plays a central role in bone metastases (Shiozawa Y *et al*, 2011).

For the first time, in this work we identified a CSCs-like bone tropism signature, able to distinguish CD44⁻CD24⁺ bone-isolated CSCs-like from parental breast CD44⁺CD24⁻ CSCs-like. This signature represents a characterization of a genetic/transcriptional program by which CSCs-like isolated from a primary breast tumor promote bone metastasis in a human implanted bone. Moreover, the signature can be a proof of concept of the phenotypic switch observed in CSCs-like after the passage in bone. The CSCs-like bone tropism signature is characterised by genes strongly up-regulated, such as CCL20, IL-8 and IL-6. These human products were also detected in sera of mice with bone metastases, whereas none of them were found in mice without bone colonization by CSC-like and in control mice. It is reasonable to assume that these chemokines and cytokines have a role in attracting CSCs-like in the bone microenvironment, and then in inducing the phenotypic switch. Indeed, they are up-regulated during the bone invasion and retained after *in vitro* and *in vivo* passages. The microarray analysis does highlight a strong cytokine/growth factor transcriptional module in the CSCs-like bone tropism signature. The up-regulation of a number of functional gene modules, involved in various aspect of the bone-metastatic process, was observed. A central role seems to be played by hypoxia, which promotes angiogenesis, bone metastatization (Dunn LK *et al*, 2009; Lu X *et al*, 2010), metabolic

stress adaptation (Winter SC *et al*, 2007), motility and invasiveness (Pennacchietti S *et al*, 2003). Another set of genes of potential interest includes telomere maintenance, reported as promoting cancer stemness through expression of EGFR (Beck S *et al*, 2011; Pantel K *et al*, 2004). In fact, the results showed up-regulation of multiple target gene sets and the Ras/ERK pathway, downstream from EGFR, which also contributes to the maintenance and acquisition of stem-like hallmarks (Tabu K *et al*, 2010). The significant up-regulation of ERK/MAP kinase targets suggests an activation of proliferation and it is likely flanked by concomitant up-regulation of genes of the MDM4 pathway, reflecting a possible functional block of p53 (Abdel-Fatah TM *et al*, 2010). The up-regulation of STAT3 targets has also been identified, supporting the role of STAT3 signalling in growth of breast CSCs-like growth (Marotta LL *et al*, 2011). Overall, these results suggest concomitant acquisition of proliferative and stemness potential by bone-isolated CSCs-like.

In conclusion, we demonstrated the ability of primary breast CSCs-like to metastasize to bone, showing their direct involvement in the pathogenesis of breast cancer bone metastasis. Anticipating a risk of bone metastasis is mandatory, because bone metastatic recurrence often follows a period of tumor dormancy. The identification of a CSCs-like bone tropism signature in breast cancer patients might be a relevant prognostic factor for clinical outcome, and it might allow to modify therapeutic regimens according to the breast cancer characteristics.

References

Abdel-Fatah TM, Powe DG, Agboola J, Adamowicz-Brice M, Blamey RW, Lopez-Garcia MA, Green AR, Reis-Filho JS, Ellis IO (2010) The biological, clinical and prognostic implications of p53 transcriptional pathways in breast cancers. *J Pathol* **220**(4): 419-434

Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF (2003) Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* **100**(7): 3983-3988

Beck S, Jin X, Sohn YW, Kim JK, Kim SH, Yin J, Pian X, Kim SC, Nam DH, Choi YJ, Kim H (2011) Telomerase activity-independent function of TERT allows glioma cells to attain cancer stem cell characteristics by inducing EGFR expression. *Mol Cells* **31**(1): 9-15

Beider K, Abraham M, Begin M, Wald H, Weiss ID, Wald O, Pikarsky E, Abramovitch R, Zeira E, Galun E, Nagler A, Peled A (2009) Interaction between CXCR4 and CCL20 pathways regulates tumor growth. *PloS one* **4**(4): e5125

Buijs JT, van der Horst G, van den Hoogen C, Cheung H, de Rooij B, Kroon J, Petersen M, van Overveld PG, Pelger RC, van der Pluijm G (2011) The BMP2/7 heterodimer inhibits the human breast cancer stem cell subpopulation and bone metastases formation. *Oncogene*:

Bussolati B, Grange C, Sapino A, Camussi G (2009) Endothelial cell differentiation of human breast tumour stem/progenitor cells. *J Cell Mol Med* **13**(2): 309-319

Campbell LL, Polyak K (2007) Breast tumor heterogeneity: cancer stem cells or clonal evolution? *Cell Cycle* **6**(19): 2332-2338

Clarke MF, Fuller M (2006) Stem cells and cancer: two faces of eve. *Cell* **124**(6): 1111-1115

Dunn LK, Mohammad KS, Fournier PG, McKenna CR, Davis HW, Niewolna M, Peng XH, Chirgwin JM, Guise TA (2009) Hypoxia and TGF-beta drive breast cancer bone

metastases through parallel signaling pathways in tumor cells and the bone microenvironment. *PloS one* **4**(9): e6896

Follenzi A, Ailles LE, Bakovic S, Geuna M, Naldini L (2000) Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat Genet* **25**(2): 217-222

Fu L, Medico E (2007) FLAME, a novel fuzzy clustering method for the analysis of DNA microarray data. *BMC Bioinformatics* **8**: 3

Goodwin PJ, Ennis M, Pritchard KI, Trudeau ME, Koo J, Hartwick W, Hoffma B, Hood N (2002) Insulin-like growth factor binding proteins 1 and 3 and breast cancer outcomes. *Breast cancer research and treatment* **74**(1): 65-76

Guise T (2010) Examining the metastatic niche: targeting the microenvironment. *Semin Oncol* **37 Suppl 2**: S2-14

Gupta PB, Fillmore CM, Jiang G, Shapira SD, Tao K, Kuperwasser C, Lander ES (2011) Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**(4): 633-644

Harris MA, Yang H, Low BE, Mukherjee J, Guha A, Bronson RT, Shultz LD, Israel MA, Yun K (2008) Cancer stem cells are enriched in the side population cells in a mouse model of glioma. *Cancer research* **68**(24): 10051-10059

Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Cordon-Cardo C, Guise TA, Massague J (2003) A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* **3**(6): 537-549

Kong D, Banerjee S, Ahmad A, Li Y, Wang Z, Sethi S, Sarkar FH (2010) Epithelial to mesenchymal transition is mechanistically linked with stem cell signatures in prostate cancer cells. *PLoS One* **5**(8): e12445

Kuperwasser C, Dessain S, Bierbaum BE, Garnet D, Sperandio K, Gauvin GP, Naber SP, Weinberg RA, Rosenblatt M (2005) A mouse model of human breast cancer metastasis to human bone. *Cancer Res* **65**(14): 6130-6138

Li A, Dubey S, Varney ML, Dave BJ, Singh RK (2003) IL-8 directly enhanced endothelial cell survival, proliferation, and matrix metalloproteinases production and regulated angiogenesis. *J Immunol* **170**(6): 3369-3376

Liu H, Patel MR, Prescher JA, Patsialou A, Qian D, Lin J, Wen S, Chang YF, Bachmann MH, Shimono Y, Dalerba P, Adorno M, Lobo N, Bueno J, Dirbas FM, Goswami S, Somlo G, Condeelis J, Contag CH, Gambhir SS, Clarke MF (2010) Cancer stem cells from human breast tumors are involved in spontaneous metastases in orthotopic mouse models. *Proc Natl Acad Sci U S A* **107**(42): 18115-18120

Liu R, Wang X, Chen GY, Dalerba P, Gurney A, Hoey T, Sherlock G, Lewicki J, Shedden K, Clarke MF (2007) The prognostic role of a gene signature from tumorigenic breast-cancer cells. *N Engl J Med* **356**(3): 217-226

Lu X, Yan CH, Yuan M, Wei Y, Hu G, Kang Y (2010) In vivo dynamics and distinct functions of hypoxia in primary tumor growth and organotropic metastasis of breast cancer. *Cancer research* **70**(10): 3905-3914

Ma J, Ren Z, Ma Y, Xu L, Zhao Y, Zheng C, Fang Y, Xue T, Sun B, Xiao W (2009) Targeted knockdown of EGR-1 inhibits IL-8 production and IL-8-mediated invasion of prostate cancer cells through suppressing EGR-1/NF-kappaB synergy. *J Biol Chem* **284**(50): 34600-34606

Marotta LL, Almendro V, Marusyk A, Shipitsin M, Schemme J, Walker SR, Bloushtain-Qimron N, Kim JJ, Choudhury SA, Maruyama R, Wu Z, Gonen M, Mulvey LA, Bessarabova MO, Huh SJ, Silver SJ, Kim SY, Park SY, Lee HE, Anderson KS, Richardson AL, Nikolskaya T, Nikolsky Y, Liu XS, Root DE, Hahn WC, Frank DA, Polyak K (2011) The JAK2/STAT3 signaling pathway is required for growth of CD44CD24 stem cell-like breast cancer cells in human tumors. *The Journal of clinical investigation* **121**(7): 2723-2735

Minn AJ, Kang Y, Serganova I, Gupta GP, Giri DD, Doubrovin M, Ponomarev V, Gerald WL, Blasberg R, Massague J (2005) Distinct organ-specific metastatic potential of

individual breast cancer cells and primary tumors. *The Journal of clinical investigation* **115**(1): 44-55

Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC (2003) PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature genetics* **34**(3): 267-273

Mundy GR (2002) Metastasis to bone: causes, consequences and therapeutic opportunities. *Nat Rev Cancer* **2**(8): 584-593

Mundy GR, Sterling J (2008) Metastatic solid tumors to bone. Primer on the metabolic bone diseases and disorders of mineral metabolism. 7th ed. Vol. 79: pages pp. ASBMR: Washington

Pantel K, Brakenhoff RH (2004) Dissecting the metastatic cascade. *Nat Rev Cancer* **4**(6): 448-456

Pennacchietti S, Michieli P, Galluzzo M, Mazzone M, Giordano S, Comoglio PM (2003) Hypoxia promotes invasive growth by transcriptional activation of the met protooncogene. *Cancer cell* **3**(4): 347-361

Roato I, Caldo D, Godio L, D'Amico L, Giannoni P, Morello E, Quarto R, Molfetta L, Buracco P, Mussa A, Ferracini R (2010) Bone invading NSCLC cells produce IL-7: mice model and human histologic data. *BMC cancer* **10**: 12

Rosen JM, Jordan CT (2009) The increasing complexity of the cancer stem cell paradigm. *Science* **324**(5935): 1670-1673

Shackleton M, Quintana E, Fearon ER, Morrison SJ (2009) Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell* **138**(5): 822-829

Shin HJ, Rho SB, Jung DC, Han IO, Oh ES, Kim JY (2011) Carbonic anhydrase IX (CA9) modulates tumor-associated cell migration and invasion. *Journal of cell science* **124**(Pt 7): 1077-1087

Shiozawa Y, Pedersen EA, Havens AM, Jung Y, Mishra A, Joseph J, Kim JK, Patel LR, Ying C, Ziegler AM, Pienta MJ, Song J, Wang J, Loberg RD, Krebsbach PH, Pienta KJ, Taichman RS (2011) Human prostate cancer metastases target the hematopoietic stem cell niche to establish footholds in mouse bone marrow. *The Journal of clinical investigation* **121**(4): 1298-1312

Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J, Nikolskaya T, Serebryiskaya T, Beroukhim R, Hu M, Halushka MK, Sukumar S, Parker LM, Anderson KS, Harris LN, Garber JE, Richardson AL, Schnitt SJ, Nikolsky Y, Gelman RS, Polyak K (2007) Molecular definition of breast tumor heterogeneity. *Cancer Cell* **11**(3): 259-273

Smid M, Wang Y, Zhang Y, Sieuwerts AM, Yu J, Klijn JG, Foekens JA, Martens JW (2008) Subtypes of breast cancer show preferential site of relapse. *Cancer Res* **68**(9): 3108-3114

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**(43): 15545-15550

Tabu K, Kimura T, Sasai K, Wang L, Bizen N, Nishihara H, Taga T, Tanaka S (2010) Analysis of an alternative human CD133 promoter reveals the implication of Ras/ERK pathway in tumor stem-like hallmarks. *Mol Cancer* **9**: 39

Tzukerman M, Rosenberg T, Ravel Y, Reiter I, Coleman R, Skorecki K (2003) An experimental platform for studying growth and invasiveness of tumor cells within teratomas derived from human embryonic stem cells. *Proc Natl Acad Sci U S A* **100**(23): 13507-13512

van den Hoogen C, van der Horst G, Cheung H, Buijs JT, Lippitt JM, Guzman-Ramirez N, Hamdy FC, Eaton CL, Thalmann GN, Cecchini MG, Pelger RC, van der Pluijm G

(2010) High aldehyde dehydrogenase activity identifies tumor-initiating and metastasis-initiating cells in human prostate cancer. *Cancer Res* **70**(12): 5163-5173

Wellner U, Schubert J, Burk UC, Schmalhofer O, Zhu F, Sonntag A, Waldvogel B, Vannier C, Darling D, zur Hausen A, Brunton VG, Morton J, Sansom O, Schuler J, Stemmler MP, Herzberger C, Hopt U, Keck T, Brabletz S, Brabletz T (2009) The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat Cell Biol* **11**(12): 1487-1495

Winter SC, Buffa FM, Silva P, Miller C, Valentine HR, Turley H, Shah KA, Cox GJ, Corbridge RJ, Homer JJ, Musgrove B, Slevin N, Sloan P, Price P, West CM, Harris AL (2007) Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. *Cancer research* **67**(7): 3441-3449

Zhang XH, Wang Q, Gerald W, Hudis CA, Norton L, Smid M, Foekens JA, Massague J (2009) Latent bone metastasis in breast cancer tied to Src-dependent survival signals. *Cancer Cell* **16**(1): 67-78

Captions

Figure 1. Bioluminescence imaging. (A) Tumor growth was examined by BLI. Luc-CSCs-like were injected subcutaneously into flank of SCID mice. Representative images of 2 animals after different time points (20, 35 and 45 days post SC injection) were shown. (B) At day 20, the SC tumor mass was clearly evident, whereas the bone localization was detectable only in one mice (arrow). (C) At day 35, an increased tumor volume and invasion to human bone becomes particularly evident (arrow). (D) At 45 days, the signal derived from bone localization of the CSCs-like is stronger in bone than in SC tumor mass. (E) Increase in tumor volume and human bone colonization was indicated by quantification of the mean $\pm$ SD of photon count in SC (subcutaneous tumor) and bone (human bone).

Figure 2. Histological analysis of the implanted bone. (A) H&E-stained section in control mice shows the presence of human live bone, blood vessels, newly synthesized bone with osteoblasts lining cells, as indicated by the arrows (Magnification X 40). (B) Trichrome staining shows in blue the new collagen fibres in bone, osteoid originated after the bone implanted in SC of the control mice. (C) Human endothelial cells in the blood vessels are stained for anti-human CD34, as indicated by the arrows. (D and E, respectively) Breast Luc-CSCs-like metastasize the human implanted bone after IC and SC injection, in both cases area of pathological bone resorption are evident, as indicated by the arrows. (F) A marked neo-bone apposition is evident (osteoid is indicated by the blue staining) in bone invaded by breast CSCs-like. Osteoclasts stained for TRAP in bone of control mice (G) and mice injected with breast CSCs-like (H).

Figure. 2
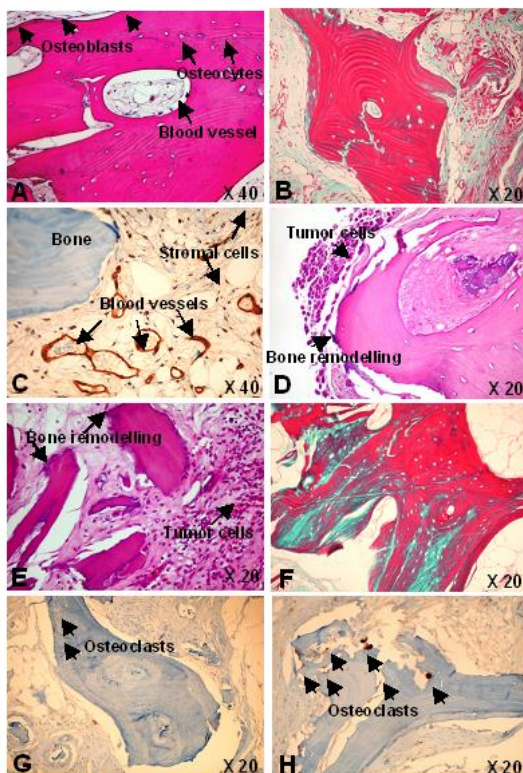
Figure 3. Phenotype of parental breast CSCs-like and bone-isolated CScs-like.

Representative dot plots show the isotopic control (A), the CD44$^+$CD24$^-$ breast CSCs-like injected in mice (B) and the bone-isolated CSCs-like, expressing a CD44$^-$CD24$^+$ phenotype (C).
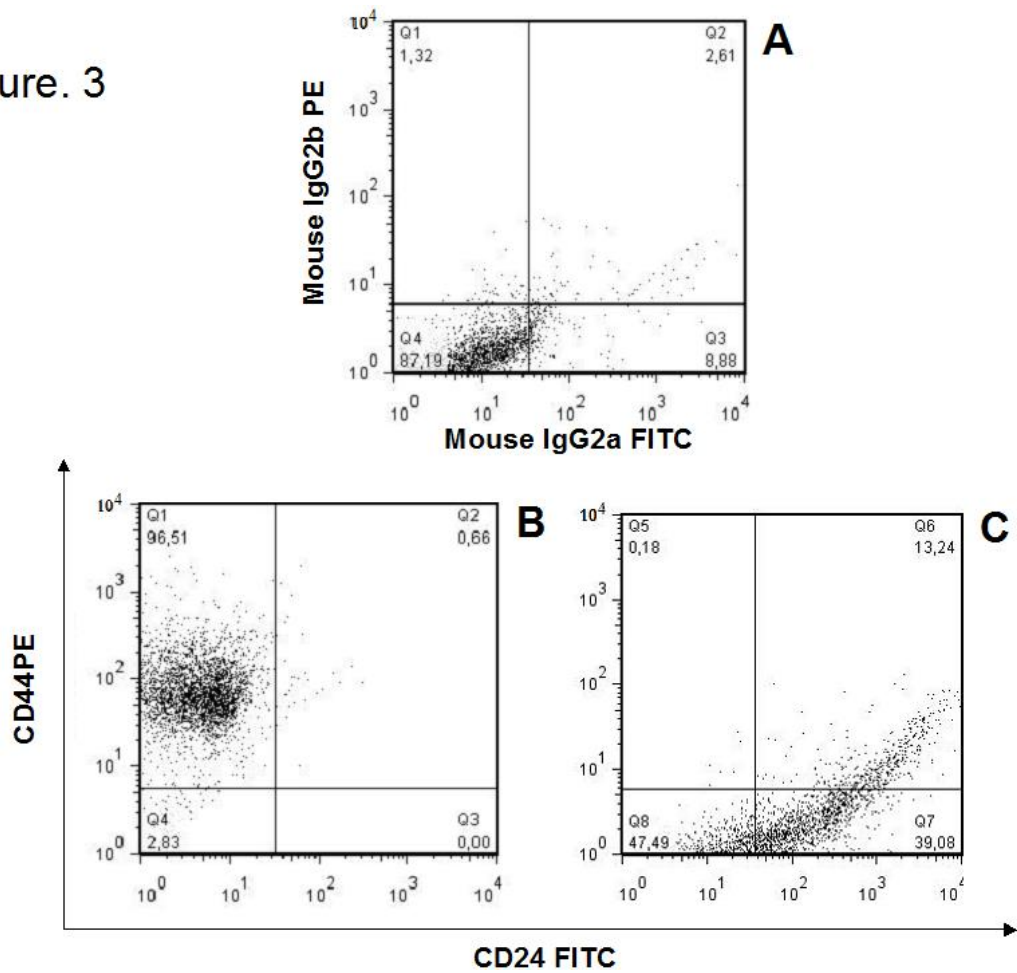
Figure 4. CD44 and CD24 expression in tumor mass and human implanted bone.

IHC staining of tumor masses shows tumor cells with a CD44$^+$CD24$^-$ phenotype (A-D). The bone-isolated CSCs-like are mainly CD44$^-$CD24$^+$ (E-H; Magnification X 20 and X 60).
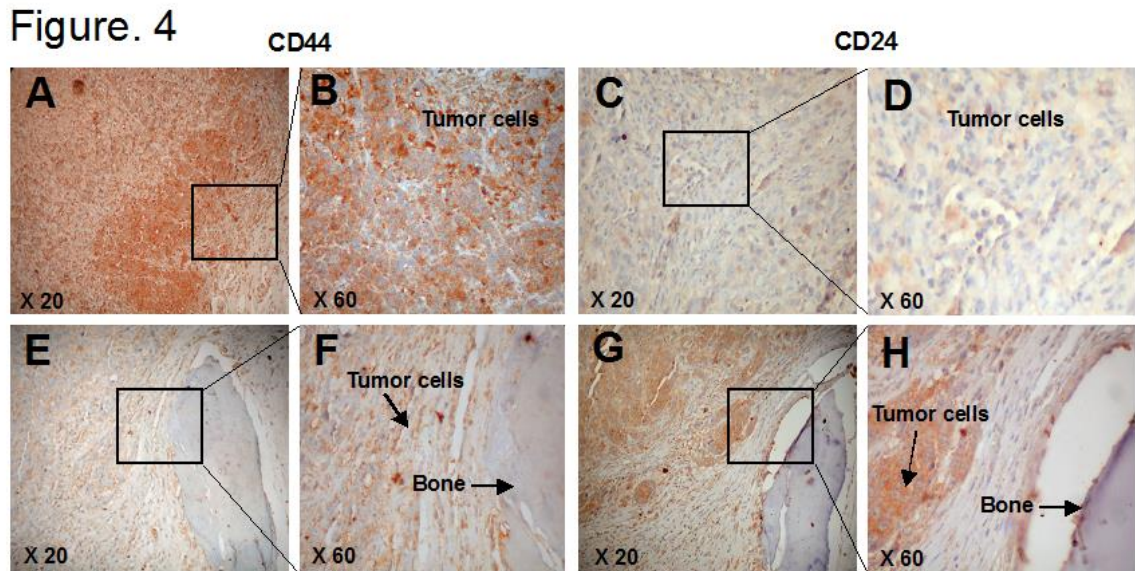
Figure 5. IHC analysis of CD44$^-$CD24$^+$ bone-isolated CSCs-like. CSCs-like isolated

from bone express CKs (A), EMA (B) and vimentin (C). (Magnification X 20).

Figure. 5
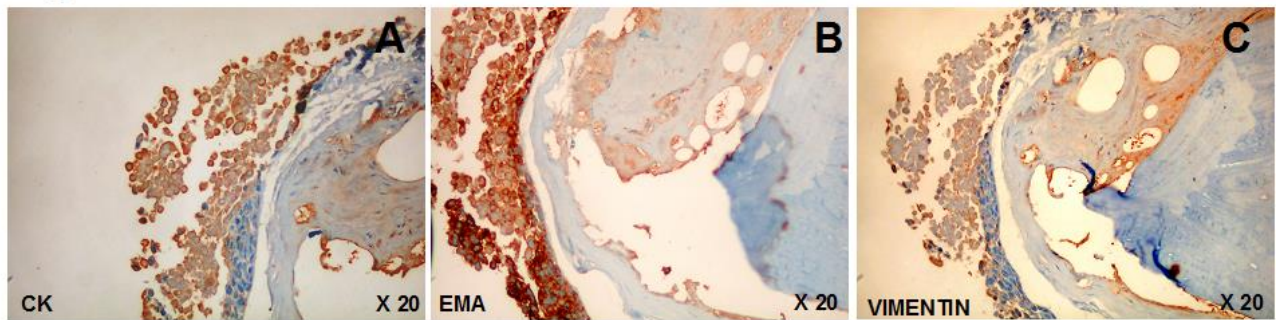
Figure 6. Analysis of secondary tumor induced by CD44⁻CD24⁺ bone-isolated CSCs-like. SC injection of CD44⁻CD24⁺ cells results in tumor masses formation as indicated by the arrows in A. H/E staining of tumor mass (B), CD44 (C) and CD24 (D).
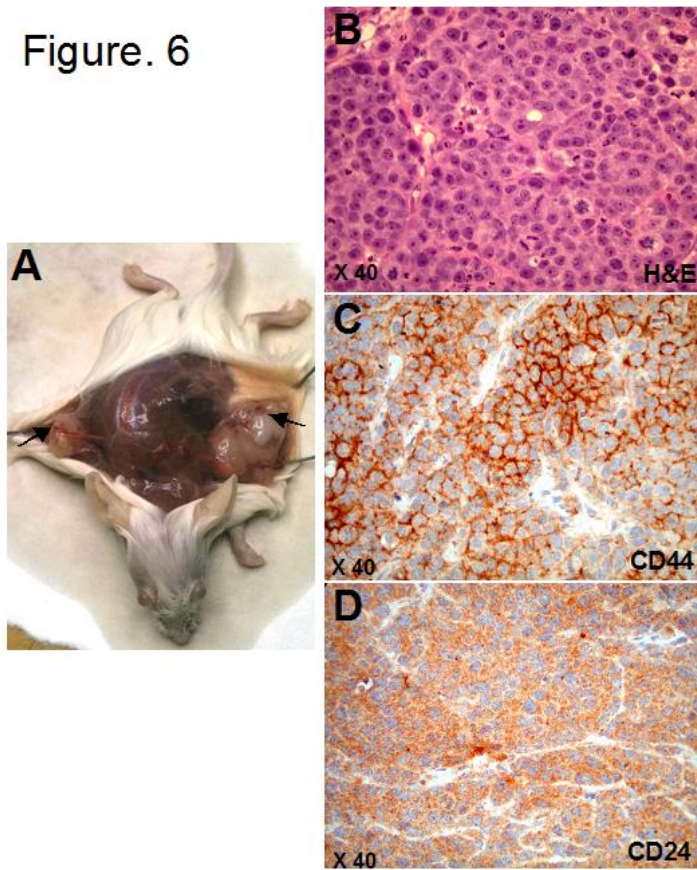
Figure 7. **Analysis of CSCs-like lung metastases.** (A) H&E-stained sections of a mouse lung shows the presence of a metastatic lesion, (Magnification X 20). (B, C) IHC staining for CD44 and CD24 shows both the markers expressed by the tumor cells. (Magnification X 60).

Figure. 7
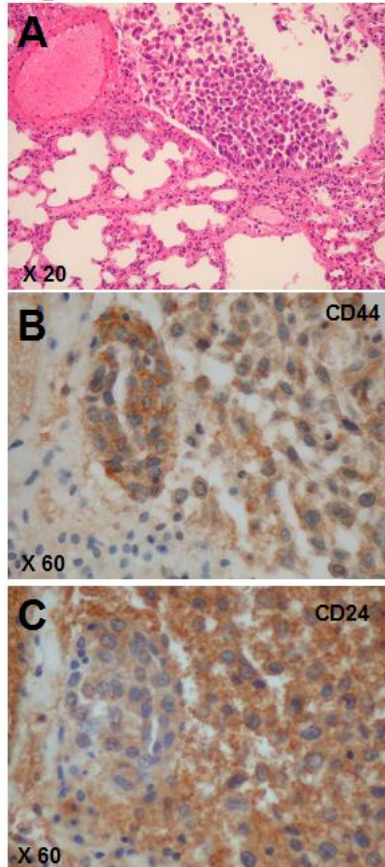
Figure 8. Microarray analysis of breast CSCs-like bone tropism signature. (A) Hierarchical clustering of human breast cancer cell lines with strong bone metastatic potential derived by parental lines in Kang and colleagues data set, using our CSCs-like bone tropism signature composed of 110 genes (119 probes). (B) GSEA analysis of Kang's samples for enrichment of the CSCs-like bone tropism signature in genes discriminating bone-metastatic from parental lines. Upper panel: positive signature genes are significantly high in cells with strong metastatic potential. Lower panel: negative signature genes are significantly low in strong metastatic potential. NES, normalized Enrichment Score; FDR, False Discovery Rate. (C) Hierarchical clustering of 65 breast cancer samples from metastatic lesions at different

sites in a published dataset, using our CSCs-like bone tropism signature. For convenience, bone metastasis samples are highlighted by light red boxes. A subset of signature genes with greatly higher expression in bone metastases is evident in the top right corner. (D) GSEA analysis on the metastasis samples for enrichment of the breast CSCs-like bone tropism signature in genes discriminating bone metastases from lesions at other sites. Left panel: positive signature genes are significantly high in bone metastases. Right panel: negative signature genes are significantly lo

A

STRONG-2274
STRONG-2269
STRONG-2268
STRONG-1833
STRONG-2271
STRONG-2287
PARENTAL
PARENTAL

-1.0   0   +1.0

B

Positive Bone Tropism Genes

Enrichment Score

NES = 2.85
p-value < 0.01
FDR < 0.01

Hits

Negative Bone Tropism Genes

Enrichment Score

NES = -1.83
p-value = 0.016
FDR = 0.016

Hits

C

-1.0   0   +1.0

D

Positive Bone Tropism Genes

Enrichment Score

NES = -1.68
p-value = 0.033
FDR = 0.032

Hits

Negative Bone Tropism Genes

Enrichment Score

NES = -1.60
p-value = 0.035
FDR = 0.10

Hits

Figure S2

20 days

Figure S3
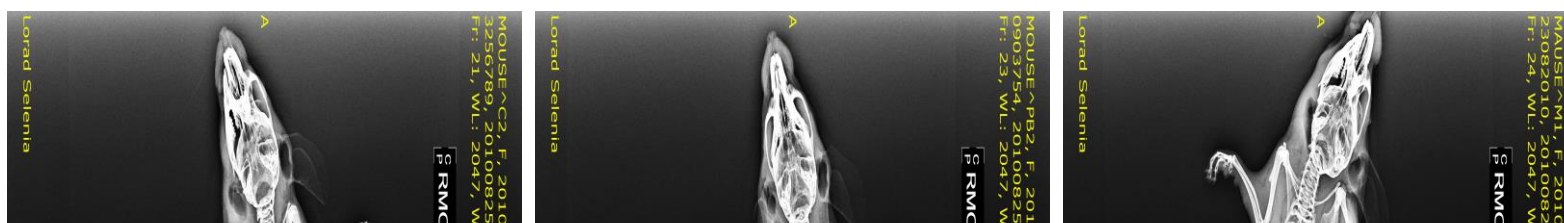
Figure S4
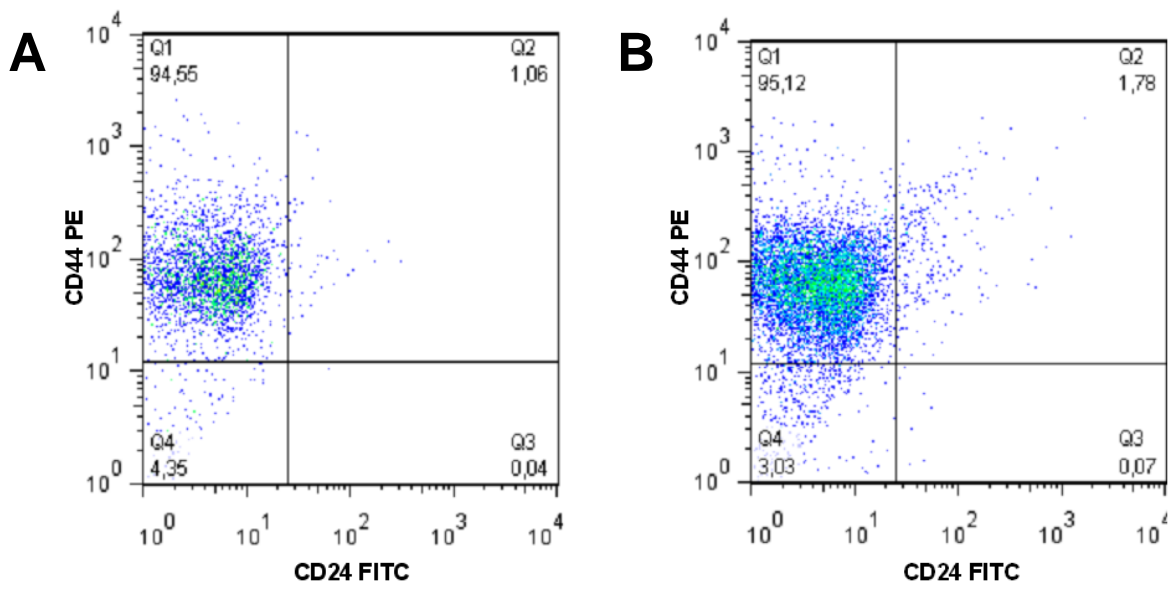
Figure S5
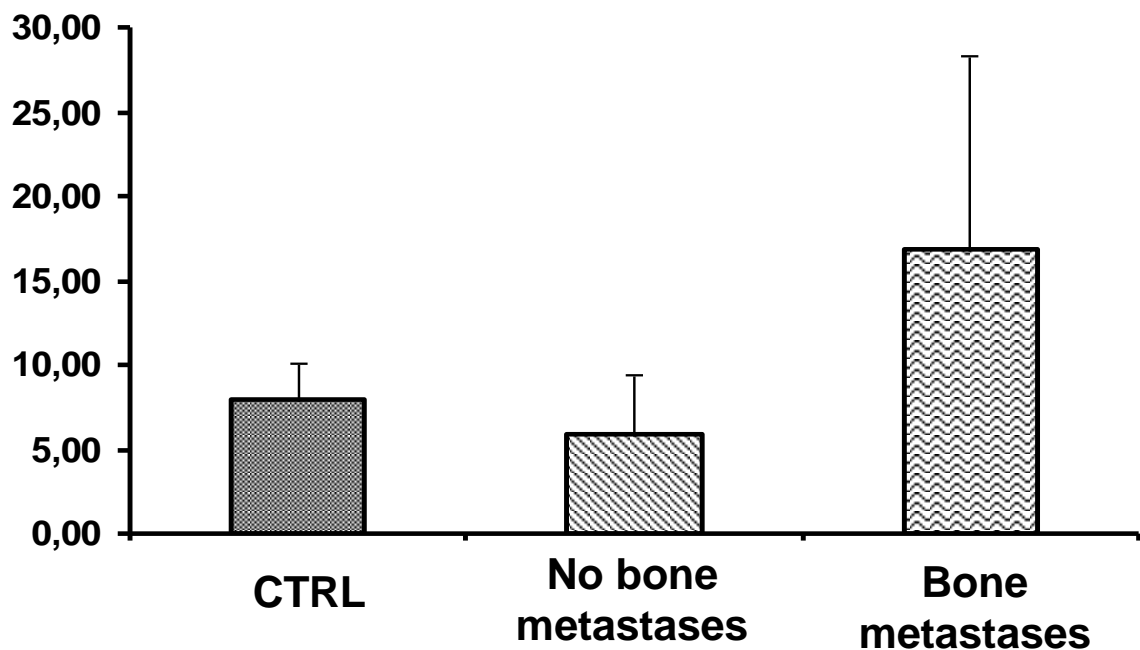
Figure S1. Analysis of the Luciferase signal intensity

Luminescence of breast Luc-CSCs-like (total flux expressed as photon per sec) was expressed as a function of the cells numbers. Cells were seeded at densities of 5×10^2 to 3×10^4. Number of cells correlated with amount of emitted light, after D-luciferin addition to all wells.

Figure S2. Bioluminescence imaging of lung metastases. Tumor growth was examined by BLI. Breast Luc-CSCs-like injected IC induce metastases in lung. Representative image of a lung tumor after 20 days post injection.

Figure S3. X-Ray imaging of mice with the human implanted bone. Human implanted bone in the left flank (arrow) of mice is detectable. Images show (A) control mouse, (B) mouse with Luc-CSCs-like injected through IC route, (C) mouse with Luc-CSCs-like injected SC, and with a soft tissue breast cancer mass (arrow head) detectable as subcutaneous growth.

Figure S4. Phenotype of parental CSCs-like and luciferase transduced CSCs-like.

The flow citometry analysis shows the same phenotype, CD44$^+$CD24$^-$ for both parental breast CSCs-like (A) and CSCs-like transduced by Luc-lentiviral vector (B).

Figure S5. Human CCL20 is present in mice sera. The level of human serum CCL20 (mean $\pm$ SD) results higher in mice with bone invasion by CSCs-like than in control mice and in mice without bone metastases.

Table S1. Mice groups and summary results

mice groups	injection site	survival	bone engrafment	bone metastasis	lung metastasis
CTRL		5/5 (100%)	5/5 (100%)	0/5 (0%)	0/5 (0%)
CSCs-like	SC	6/6 (100%)	6/6 (100%)	6/6(100%)	2/6 (33%)
CSCs-like	IC	5/6 (83%)	6/6 (100%)	3/6 (50%)	3/6 (50%)

CTRL indicates the control mice subjected only to bone implant. CSCs-Luc were injected in two groups of mice: subcutaneously (SC) and through intracardiac route (IC).

Table S2. Genes of the breast CSCs-like bone tropism signature
(A) Genes up-regulated in breast CSCs-like bone tropism signature

Symbol	DEFINITION	Illumina PROBE_ID	Accession N.
ADM	Homo sapiens adrenomedullin (ADM), mRNA.	ILMN_1708934	NM_001124.1
ANGPTL4	Homo sapiens angiopoietin-like 4 (ANGPTL4), transcript variant 1, mRNA.	ILMN_1707727	NM_139314.1
ANKRD37	Homo sapiens ankyrin repeat domain 37 (ANKRD37), mRNA.	ILMN_1756417	NM_181726.1
AXUD1	Homo sapiens AXIN1 up-regulated 1 (AXUD1), mRNA.	ILMN_1703123	NM_033027.2
BAIAP2L2	Homo sapiens BAI1-associated protein 2-like 2 (BAIAP2L2), mRNA.	ILMN_1771652	NM_025045.3
BTG1	Homo sapiens B-cell translocation gene 1, anti-proliferative (BTG1), mRNA.	ILMN_1775743	NM_001731.1
C13ORF15	Homo sapiens chromosome 13 open reading frame 15 (C13orf15), mRNA.	ILMN_1658494	NM_014059.1
C8ORF13	Homo sapiens chromosome 8 open reading frame 13 (C8orf13), mRNA.	ILMN_1687213	NM_053279.1
CA12	Homo sapiens carbonic anhydrase XII (CA12), transcript variant 1, mRNA.	ILMN_1720998	NM_001218.3
CA12	Homo sapiens carbonic anhydrase XII (CA12), transcript variant 1, mRNA.	ILMN_2382942	NM_001218.3
CA2	Homo sapiens carbonic anhydrase II (CA2), mRNA.	ILMN_2199439	NM_000067.1
CA9	Homo sapiens carbonic anhydrase IX (CA9), mRNA.	ILMN_1725139	NM_001216.1
CATSPER2	Homo sapiens cation channel, sperm associated 2 (CATSPER2), transcript variant 4, mRNA.	ILMN_2286014	NM_172097.1
CCL20	Homo sapiens chemokine (C-C motif) ligand 20 (CCL20), mRNA.	ILMN_1657234	NM_004591.1
CD68	Homo sapiens CD68 molecule (CD68), transcript variant 1,	ILMN_2359907	NM_001251.2

	mRNA.		
COL8A1	Homo sapiens collagen, type VIII, alpha 1 (COL8A1), transcript variant 2, mRNA.	ILMN_1685433	NM_020351.2
CXCL16	Homo sapiens chemokine (C-X-C motif) ligand 16 (CXCL16), mRNA.	ILMN_1728478	NM_022059.1
CYFIP2	Homo sapiens cytoplasmic FMR1 interacting protein 2 (CYFIP2), transcript variant 3, mRNA.	ILMN_1677200	NM_014376.1
DMBT1	Homo sapiens deleted in malignant brain tumors 1 (DMBT1), transcript variant 3, mRNA.	ILMN_2410612	NM_017579.1
DNER	Homo sapiens delta/notch-like EGF repeat containing (DNER), mRNA.	ILMN_1791679	NM_139072.2
DUSP4	Homo sapiens dual specificity phosphatase 4 (DUSP4), transcript variant 1, mRNA.	ILMN_1808391	NM_001394.5
DUSP5	Homo sapiens dual specificity phosphatase 5 (DUSP5), mRNA.	ILMN_1656501	NM_004419.3
DUSP6	Homo sapiens dual specificity phosphatase 6 (DUSP6), transcript variant 2, mRNA.	ILMN_2396020	NM_022652.2
EGR1	Homo sapiens early growth response 1 (EGR1), mRNA.	ILMN_1762899	NM_001964.2
EID3	Homo sapiens EP300 interacting inhibitor of differentiation 3 (EID3), mRNA.	ILMN_1776000	NM_001008394.1
FAM153B	Homo sapiens family with sequence similarity 153, member B (FAM153B), mRNA.	ILMN_2205470	NM_001079529.2
FAM167A	Homo sapiens family with sequence similarity 167, member A (FAM167A), mRNA.	ILMN_3248511	NM_053279.2
GCNT3	Homo sapiens glucosaminyl (N-acetyl) transferase 3, mucin type (GCNT3), mRNA.	ILMN_1712082	NM_004751.1
HIST1H1C	Homo sapiens histone cluster 1, H1c (HIST1H1C), mRNA.	ILMN_1757406	NM_005319.3
HIST1H2BG	Homo sapiens histone cluster 1, H2bg (HIST1H2BG), mRNA.	ILMN_1716195	NM_003518.3
HIST2H2AA3	Homo sapiens histone cluster 2, H2aa3 (HIST2H2AA3), mRNA.	ILMN_1659047	NM_003516.2
HIST2H2AC	Homo sapiens histone cluster 2, H2ac (HIST2H2AC), mRNA.	ILMN_1768973	NM_003517.2
IER3	Homo sapiens immediate early response 3 (IER3), mRNA.	ILMN_1682717	NM_052815.1
IGFBP1	Homo sapiens insulin-like growth factor binding protein 1 (IGFBP1), transcript variant 2, mRNA.	ILMN_2387385	NM_001013029.1
IGFBP1	Homo sapiens insulin-like growth factor binding protein 1 (IGFBP1), mRNA.	ILMN_1728445	NM_000596.2
IL6	Homo sapiens interleukin 6 (interferon, beta 2) (IL6), mRNA.	ILMN_1699651	NM_000600.1
IL8	Homo sapiens interleukin 8 (IL8), mRNA.	ILMN_2184373	NM_000584.2
ISG20	Homo sapiens interferon stimulated exonuclease gene 20kDa (ISG20), mRNA.	ILMN_1659913	NM_002201.4
JUN	Homo sapiens jun oncogene (JUN), mRNA.	ILMN_1806023	NM_002228.3
KLF2	Homo sapiens Kruppel-like factor 2 (lung) (KLF2), mRNA.	ILMN_1735930	NM_016270.2
KLF4	Homo sapiens Kruppel-like factor 4 (gut) (KLF4), mRNA.	ILMN_2137789	NM_004235.3
LDLR	Homo sapiens low density lipoprotein receptor (familial hypercholesterolemia) (LDLR), mRNA.	ILMN_2053415	NM_000527.2
LOC100132377	PREDICTED: Homo sapiens hypothetical protein LOC100132377 (LOC100132377), mRNA.	ILMN_3248260	XM_001722278.1
LOC338758	PREDICTED: Homo sapiens hypothetical protein LOC338758 (LOC338758), mRNA.	ILMN_1795835	XM_931359.1
LOC391764	PREDICTED: Homo sapiens similar to hCG1989915 (LOC391764), mRNA.	ILMN_3289679	XM_373076.3
LOC642127	PREDICTED: Homo sapiens similar to TBP-associated factor 11 (LOC642127), mRNA.	ILMN_3286154	XM_936272.2
LOC729731	PREDICTED: Homo sapiens similar to hCG1809904 (LOC729731), mRNA.	ILMN_3245028	XM_001131140.1
LOX	Homo sapiens lysyl oxidase (LOX), mRNA.	ILMN_1695880	NM_002317.3
MAFB	Homo sapiens v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian) (MAFB), mRNA.	ILMN_1764709	NM_005461.3
MT2A	Homo sapiens metallothionein 2A (MT2A), mRNA.	ILMN_1686664	NM_005953.2

NAMPT	Homo sapiens nicotinamide phosphoribosyltransferase (NAMPT), mRNA.	ILMN_1653871	NM_005746.1
NDRG1	Homo sapiens N-myc downstream regulated gene 1 (NDRG1), mRNA.	ILMN_1809931	NM_006096.2
NKD2	Homo sapiens naked cuticle homolog 2 (Drosophila) (NKD2), mRNA.	ILMN_1731206	NM_033120.2
NOTCH3	Homo sapiens Notch homolog 3 (Drosophila) (NOTCH3), mRNA.	ILMN_1658926	NM_000435.1
NT5E	Homo sapiens 5'-nucleotidase, ecto (CD73) (NT5E), mRNA.	ILMN_1697220	NM_002526.1
OLFM1	Homo sapiens olfactomedin 1 (OLFM1), transcript variant 1, mRNA.	ILMN_1742025	NM_014279.2
OLFM1	Homo sapiens olfactomedin 1 (OLFM1), transcript variant 2, mRNA.	ILMN_1714709	NM_006334.2
P4HA1	Homo sapiens procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase), alpha polypeptide I (P4HA1), transcript variant 1, mRNA.	ILMN_1693334	NM_000917.2
PHLDA2	Homo sapiens pleckstrin homology-like domain, family A, member 2 (PHLDA2), mRNA.	ILMN_1671557	NM_003311.3
PI3	Homo sapiens peptidase inhibitor 3, skin-derived (SKALP) (PI3), mRNA.	ILMN_1693192	NM_002638.2
PID1	Homo sapiens phosphotyrosine interaction domain containing 1 (PID1), mRNA.	ILMN_1671891	NM_017933.3
PPP1R15A	Homo sapiens protein phosphatase 1, regulatory (inhibitor) subunit 15A (PPP1R15A), mRNA.	ILMN_1659936	NM_014330.2
PTGS2	Homo sapiens prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase) (PTGS2), mRNA.	ILMN_1677511	NM_000963.1
PTGS2	Homo sapiens prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase) (PTGS2), mRNA.	ILMN_2054297	NM_000963.1
QSOX1	Homo sapiens quiescin Q6 sulfhydryl oxidase 1 (QSOX1), transcript variant 2, mRNA.	ILMN_2411282	NM_001004128.2
RASD1	Homo sapiens RAS, dexamethasone-induced 1 (RASD1), mRNA.	ILMN_1740426	NM_016084.3
RN5S9	Homo sapiens RNA, 5S ribosomal 9 (RN5S9), ribosomal RNA.	ILMN_3234762	NR_023371.1
RN7SK	Homo sapiens RNA, 7SK small nuclear (RN7SK), non-coding RNA.	ILMN_1739423	NR_001445.1
RNU1-3	Homo sapiens RNA, U1 small nuclear 3 (RNU1-3), small nuclear RNA.	ILMN_3246273	NR_004408.1
RNU1-5	Homo sapiens RNA, U1 small nuclear 5 (RNU1-5), small nuclear RNA.	ILMN_3236653	NR_004400.1
RNU1A3	Homo sapiens RNA, U1A3 small nuclear (RNU1A3), small nuclear RNA.	ILMN_3245678	NR_004430.1
RNU1F1	Homo sapiens RNA, U1F1 small nuclear (RNU1F1), small nuclear RNA.	ILMN_3240220	NR_004402.1
RNU4-1	Homo sapiens RNA, U4 small nuclear 1 (RNU4-1), small nuclear RNA.	ILMN_3309453	NR_003925.1
RNU4-2	Homo sapiens RNA, U4 small nuclear 2 (RNU4-2), small nuclear RNA.	ILMN_3308138	NR_003137.2
RNU4ATAC	Homo sapiens RNA, U4atac small nuclear (U12-dependent splicing) (RNU4ATAC), small nuclear RNA.	ILMN_3240594	NR_023343.1
RNY4	Homo sapiens RNA, Ro-associated Y4 (RNY4), small cytoplasmic RNA.	ILMN_3241021	NR_004393.1
RPPH1	Homo sapiens ribonuclease P RNA component H1 (RPPH1), RNase P RNA.	ILMN_1704056	NR_002312.1
SERPINB3	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 3 (SERPINB3), mRNA.	ILMN_1703855	NM_006919.1
SERPINB3	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 3 (SERPINB3), mRNA.	ILMN_2120222	NM_006919.1
SLC35F3	Homo sapiens solute carrier family 35, member F3 (SLC35F3),	ILMN_1794959	NM_173508.1

Symbol	DEFINITION	Illumina PROBE_ID	Accession N.
	mRNA.		
SNAP25	Homo sapiens synaptosomal-associated protein, 25kDa (SNAP25), transcript variant 2, mRNA.	ILMN_1740555	NM_003081.2
SNORD3A	Homo sapiens small nucleolar RNA, C/D box 3A (SNORD3A), small nucleolar RNA.	ILMN_3239574	NR_006880.1
SNORD3C	Homo sapiens small nucleolar RNA, C/D box 3C (SNORD3C), small nucleolar RNA.	ILMN_3241034	NR_006881.1
SNORD3D	Homo sapiens small nucleolar RNA, C/D box 3D (SNORD3D), small nucleolar RNA.	ILMN_3242315	NR_006882.1
SOD2	Homo sapiens superoxide dismutase 2, mitochondrial (SOD2), nuclear gene encoding mitochondrial protein, transcript variant 3, mRNA.	ILMN_2406501	NM_001024466.1
SPHK1	Homo sapiens sphingosine kinase 1 (SPHK1), transcript variant 1, mRNA.	ILMN_2357134	NM_021972.2
SULF2	Homo sapiens sulfatase 2 (SULF2), transcript variant 1, mRNA.	ILMN_1686981	NM_018837.2
SULF2	Homo sapiens sulfatase 2 (SULF2), transcript variant 1, mRNA.	ILMN_1667460	NM_018837.2
TM4SF5	Homo sapiens transmembrane 4 L six family member 5 (TM4SF5), mRNA.	ILMN_2167808	NM_003963.2
TMEM45A	Homo sapiens transmembrane protein 45A (TMEM45A), mRNA.	ILMN_2148913	NM_018004.1
TNFAIP6	Homo sapiens tumor necrosis factor, alpha-induced protein 6 (TNFAIP6), mRNA.	ILMN_1785732	NM_007115.2
TRIM15	Homo sapiens tripartite motif-containing 15 (TRIM15), transcript variant 1, mRNA.	ILMN_1660904	NM_033229.1
TRIM15	Homo sapiens tripartite motif-containing 15 (TRIM15), mRNA.	ILMN_1754476	NM_052812.1
TSPAN3	Homo sapiens tetraspanin 3 (TSPAN3), transcript variant 1, mRNA.	ILMN_1790549	NM_198902.1
TSPAN3	Homo sapiens tetraspanin 3 (TSPAN3), transcript variant 1, mRNA.	ILMN_1655469	NM_005724.4
USH1C	Homo sapiens Usher syndrome 1C (autosomal recessive, severe) (USH1C), transcript variant b3, mRNA.	ILMN_1784283	NM_005709.2

(B) Genes down-regulated in breast CSCs-like bone tropism signature

Symbol	DEFINITION	Illumina PROBE_ID	Accession N.
ADA	Homo sapiens adenosine deaminase (ADA), mRNA.	ILMN_1803686	NM_000022.2
AGXT2L1	Homo sapiens alanine-glyoxylate aminotransferase 2-like 1 (AGXT2L1), mRNA.	ILMN_1757807	NM_031279.2
CCDC109A	Homo sapiens coiled-coil domain containing 109A (CCDC109A), mRNA.	ILMN_1672759	NM_138357.1
CD24	Homo sapiens CD24 molecule (CD24), mRNA.	ILMN_2060413	NM_013230.2
FAM105A	Homo sapiens family with sequence similarity 105, member A (FAM105A), mRNA.	ILMN_1695711	NM_019018.1
FAM5C	Homo sapiens family with sequence similarity 5, member C (FAM5C), mRNA.	ILMN_2092589	NM_199051.1
FER1L3	Homo sapiens fer-1-like 3, myoferlin (C. elegans) (FER1L3), transcript variant 1, mRNA.	ILMN_2370976	NM_013451.2
GNG11	Homo sapiens guanine nucleotide binding protein (G protein), gamma 11 (GNG11), mRNA.	ILMN_1782419	NM_004126.2
HIF1A	Homo sapiens hypoxia-inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor) (HIF1A), transcript variant 1, mRNA.	ILMN_1763260	NM_001530.2
KLHL5	Homo sapiens kelch-like 5 (Drosophila) (KLHL5), transcript variant 3, mRNA.	ILMN_1706687	NM_001007075.1
LITAF	Homo sapiens lipopolysaccharide-induced TNF factor (LITAF), mRNA.	ILMN_1713934	NM_004862.2
LXN	Homo sapiens latexin (LXN), mRNA.	ILMN_1723962	NM_020169.2
MB	Homo sapiens myoglobin (MB), transcript variant 1, mRNA.	ILMN_1666109	NM_005368.2

MX1	Homo sapiens myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse) (MX1), mRNA.	ILMN_1662358	NM_002462.2
MYH10	Homo sapiens myosin, heavy chain 10, non-muscle (MYH10), mRNA.	ILMN_1815154	NM_005964.1
PRSS23	Homo sapiens protease, serine, 23 (PRSS23), mRNA.	ILMN_1797776	NM_007173.3
PSMB9	Homo sapiens proteasome (prosome, macropain) subunit, beta type, 9 (large multifunctional peptidase 2) (PSMB9), transcript variant 1, mRNA.	ILMN_2376108	NM_002800.4
PTPLB	Homo sapiens protein tyrosine phosphatase-like (proline instead of catalytic arginine), member b (PTPLB), mRNA.	ILMN_1690806	NM_198402.2
RARRES 1	Homo sapiens retinoic acid receptor responder (tazarotene induced) 1 (RARRES1), transcript variant 1, mRNA.	ILMN_1800091	NM_206963.1
SEPT3	Homo sapiens septin 3 (SEPT3), transcript variant B, mRNA.	ILMN_1746673	NM_019106.4
SEPT3	Homo sapiens septin 3 (SEPT3), transcript variant B, mRNA.	ILMN_1659953	NM_019106.4
SLC22A5	Homo sapiens solute carrier family 22 (organic cation transporter), member 5 (SLC22A5), mRNA.	ILMN_1699357	NM_003060.2
SLC27A2	Homo sapiens solute carrier family 27 (fatty acid transporter), member 2 (SLC27A2), mRNA.	ILMN_1700831	NM_003645.2

Table S3. GSEA results

[1]NES: Normalised Enrichment Score; [2]FDR: False Discovery Rate
A) C2 collection gene sets enriched in up-regulated genes

NAME	SIZE	NES[1]	NOM p-val	FDR[2] q-val
MARTORIATI_MDM4_TARGETS_FETAL_LIVER_UP	83	1.952	0.000	0.042
RASHI_RESPONSE_TO_IONIZING_RADIATION_1	31	1.897	0.000	0.074
AMIT_EGF_RESPONSE_40_HELA	40	1.879	0.000	0.064
WINTER_HYPOXIA_UP	80	1.813	0.000	0.117
REACTOME_RNA_POLYMERASE_I_PROMOTER_CLEARANCE	56	1.810	0.000	0.108
MAHAJAN_RESPONSE_TO_IL1A_UP	37	1.802	0.000	0.113
AMIT_EGF_RESPONSE_480_MCF10A	35	1.781	0.000	0.126
ST_ERK1_ERK2_MAPK_PATHWAY	26	1.775	0.000	0.121
MANALO_HYPOXIA_UP	142	1.751	0.000	0.124
AMIT_SERUM_RESPONSE_40_MCF10A	27	1.749	0.000	0.116
AMIT_EGF_RESPONSE_60_MCF10A	34	1.740	0.000	0.122
THEILGAARD_NEUTROPHIL_AT_SKIN_WOUND_UP	59	1.735	0.000	0.119
JIANG_HYPOXIA_NORMAL	184	1.734	0.000	0.107
HARRIS_HYPOXIA	61	1.731	0.000	0.106
REACTOME_MAPK_TARGETS_NUCLEAR_EVENTS_MEDIATED_BY_ MAP_KINASES	24	1.722	0.000	0.110
WINTER_HYPOXIA_METAGENE	172	1.718	0.000	0.112
DAUER_STAT3_TARGETS_UP	31	1.714	0.000	0.114
REACTOME_TELOMERE_MAINTENANCE	55	1.710	0.000	0.115
CHEN_LVAD_SUPPORT_OF_FAILING_HEART_UP	77	1.705	0.000	0.118
SEKI_INFLAMMATORY_RESPONSE_LPS_UP	51	1.704	0.000	0.112
KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY	54	1.700	0.000	0.110
CROONQUIST_STROMAL_STIMULATION_UP	31	1.681	0.000	0.122
MARTORIATI_MDM4_TARGETS_NEUROEPITHELIUM_UP	79	1.676	0.000	0.131
MCLACHLAN_DENTAL_CARIES_DN	124	1.668	0.000	0.140
REACTOME_PEPTIDE_LIGAND_BINDING_RECEPTORS	36	1.667	0.000	0.138

REACTOME_RNA_POLYMERASE_I_III_AND_MITOCHONDRIAL_TRANSCRIPTION	93	1.665	0.000	0.136
GERY_CEBP_TARGETS	85	1.657	0.000	0.143
MCLACHLAN_DENTAL_CARIES_UP	109	1.655	0.000	0.138
NAKAMURA_METASTASIS_MODEL_DN	37	1.654	0.000	0.137
VALK_AML_CLUSTER_5	15	1.651	0.000	0.138
BRUECKNER_TARGETS_OF_MIRLET7A3_UP	88	1.650	0.000	0.135
KANNAN_TP53_TARGETS_UP	38	1.646	0.000	0.133
ROSS_ACUTE_MYELOID_LEUKEMIA_CBF	51	1.642	0.000	0.136
MARCHINI_TRABECTEDIN_RESISTANCE_DN	36	1.638	0.000	0.139
GRAHAM_NORMAL_QUIESCENT_VS_NORMAL_DIVIDING_UP	45	1.638	0.000	0.137
DELLA_RESPONSE_TO_TSA_AND_BUTYRATE	19	1.636	0.000	0.137
KYNG_ENVIRONMENTAL_STRESS_RESPONSE_UP	23	1.635	0.000	0.134
NAGASHIMA_NRG1_SIGNALING_UP	150	1.634	0.000	0.132
REACTOME_RNA_POLYMERASE_I_PROMOTER_OPENING	35	1.629	0.000	0.139
MARKEY_RB1_ACUTE_LOF_UP	136	1.620	0.000	0.146
SHEPARD_CRUSH_AND_BURN_MUTANT_UP	116	1.619	0.000	0.145
REACTOME_CLASS_A1_RHODOPSIN_LIKE_RECEPTORS	52	1.619	0.000	0.144
MITSIADES_RESPONSE_TO_APLIDIN_UP	345	1.617	0.000	0.144
LEONARD_HYPOXIA	29	1.613	0.000	0.149
MISSIAGLIA_REGULATED_BY_METHYLATION_UP	83	1.611	0.000	0.145
URS_ADIPOCYTE_DIFFERENTIATION_DN	16	1.611	0.000	0.143
KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS	67	1.610	0.000	0.142
VALK_AML_CLUSTER_13	15	1.606	0.000	0.148
SEMENZA_HIF1_TARGETS	28	1.605	0.000	0.147
REACTOME_PACKAGING_OF_TELOMERE_ENDS	29	1.603	0.000	0.148
RUGO_ENVIRONMENTAL_STRESS_RESPONSE_UP	21	1.602	0.000	0.146
ST_TUMOR_NECROSIS_FACTOR_PATHWAY	27	1.602	0.000	0.144
HALMOS_CEBPA_TARGETS_UP	31	1.596	0.000	0.150
MENSE_HYPOXIA_UP	83	1.594	0.000	0.150
CHIANG_LIVER_CANCER_SUBCLASS_CTNNB1_DN	75	1.593	0.000	0.149
ELVIDGE_HYPOXIA_BY_DMOG_UP	110	1.589	0.000	0.153
SESTO_RESPONSE_TO_UV_C1	64	1.585	0.000	0.157
KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY	38	1.585	0.000	0.155
TSENG_ADIPOGENIC_POTENTIAL_DN	23	1.580	0.000	0.160
WANG_ESOPHAGUS_CANCER_VS_NORMAL_UP	73	1.575	0.000	0.165
YAO_TEMPORAL_RESPONSE_TO_PROGESTERONE_CLUSTER_5	20	1.573	0.000	0.163
VERHAAK_AML_WITH_NPM1_MUTATED_UP	107	1.573	0.000	0.161
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES	26	1.571	0.000	0.161
BOYLAN_MULTIPLE_MYELOMA_C_CLUSTER_DN	16	1.570	0.000	0.162
HOOI_ST7_TARGETS_DN	68	1.570	0.000	0.161
KEGG_LEISHMANIA_INFECTION	44	1.568	0.000	0.163
ELVIDGE_HYPOXIA_UP	138	1.568	0.000	0.162
REACTOME_INNATE_IMMUNITY_SIGNALING	69	1.566	0.000	0.160
ELVIDGE_HIF1A_TARGETS_DN	72	1.563	0.000	0.160
BROWNE_HCMV_INFECTION_30MIN_UP	30	1.559	0.000	0.162
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	89	1.547	0.000	0.174

NAME	SIZE	NES	NOM p-val	FDR q-val
GRAHAM_CML_DIVIDING_VS_NORMAL_QUIESCENT_DN	56	1.546	0.000	0.174
HAN_SATB1_TARGETS_DN	256	1.544	0.000	0.172
ROZANOV_MMP14_TARGETS_UP	162	1.543	0.000	0.171
SCHEIDEREIT_IKK_INTERACTING_PROTEINS	20	1.531	0.000	0.185
YAMASHITA_LIVER_CANCER_STEM_CELL_UP	31	1.527	0.000	0.188
KLEIN_PRIMARY_EFFUSION_LYMPHOMA_UP	38	1.525	0.000	0.189
REACTOME_GLYCOLYSIS	17	1.525	0.000	0.187
XU_HGF_SIGNALING_NOT_VIA_AKT1_6HR	20	1.525	0.000	0.185
DAZARD_UV_RESPONSE_CLUSTER_G2	17	1.522	0.000	0.183
RIZ_ERYTHROID_DIFFERENTIATION_12HR	19	1.521	0.000	0.181
CHIARADONNA_NEOPLASTIC_TRANSFORMATION_KRAS_CDC25_DN	27	1.517	0.000	0.186
PRAMOONJAGO_SOX4_TARGETS_UP	47	1.508	0.000	0.199
ELVIDGE_HIF1A_AND_HIF2A_TARGETS_DN	84	1.504	0.000	0.199
CROMER_TUMORIGENESIS_UP	22	1.497	0.000	0.207
HALMOS_CEBPA_TARGETS_DN	26	1.492	0.000	0.212
ONDER_CDH1_TARGETS_2_DN	302	1.491	0.000	0.214
RUGO_STRESS_RESPONSE_SUBSET_H	31	1.489	0.000	0.215
KYNG_DNA_DAMAGE_BY_GAMMA_AND_UV_RADIATION	31	1.489	0.000	0.214
VART_KSHV_INFECTION_ANGIOGENIC_MARKERS_UP	75	1.486	0.000	0.216
HORIUCHI_WTAP_TARGETS_UP	196	1.482	0.000	0.223
KOBAYASHI_EGFR_SIGNALING_6HR_DN	16	1.467	0.000	0.239
CASORELLI_ACUTE_PROMYELOCYTIC_LEUKEMIA_UP	122	1.461	0.000	0.244

B) C5 collection gene sets enriched in up-regulated genes

NAME	SIZE	NES	NOM p-val	FDR q-val
CYTOKINE_ACTIVITY	31	1.749	0.000	0.097
REGULATION_OF_ANGIOGENESIS	15	1.748	0.000	0.083
RESPONSE_TO_VIRUS	28	1.694	0.000	0.140
CYTOKINE_METABOLIC_PROCESS	17	1.693	0.000	0.121
CYTOKINE_BIOSYNTHETIC_PROCESS	16	1.686	0.000	0.112
ANGIOGENESIS	26	1.666	0.000	0.132
CARBON_OXYGEN_LYASE_ACTIVITY	22	1.647	0.000	0.147
ANATOMICAL_STRUCTURE_FORMATION	32	1.640	0.000	0.144
HYDRO_LYASE_ACTIVITY	18	1.637	0.000	0.136
BEHAVIOR	45	1.599	0.000	0.178
LOCOMOTORY_BEHAVIOR	31	1.595	0.000	0.176

Table 1. Relevant gene sets grouped in functional modules

FUNCTIONAL MODULE	GENE SET NAME	Collection	SIZE	NES	NOM p-val	FDR q-val
MDM4 p53 Target	MARTORIATI_MDM4_TARGETS_FETAL_LIVER_UP	C2	83	1.952	0.000	0.042
	MARTORIATI_MDM4_TARGETS_NEUROEPITHELIUM_UP	C2	79	1.676	0.000	0.131
Response to Growth Factor	AMIT_EGF_RESPONSE_40_HELA	C2	40	1.879	0.000	0.064
	AMIT_EGF_RESPONSE_480_MCF10A	C2	35	1.781	0.000	0.126
	ST_ERK1_ERK2_MAPK_PATHWAY	C2	26	1.775	0.000	0.121
	AMIT_SERUM_RESPONSE_40_MCF10A	C2	27	1.749	0.000	0.116
	AMIT_EGF_RESPONSE_60_MCF10A	C2	34	1.740	0.000	0.122
	REACTOME_MAPK_TARGETS_NUCLEAR_EVENTS_MEDIATED_BY_MAP_KINASES	C2	24	1.722	0.000	0.110
Cytokine stimulation	CYTOKINE_METABOLIC_PROCESS	C5	17	1.693	0.000	0.121
	CYTOKINE_BIOSYNTHETIC_PROCESS	C5	16	1.686	0.000	0.112
	CYTOKINE_ACTIVITY	C5	31	1.749	0.000	0.097
Angiogenesis	VASCULATURE_DEVELOPMENT	C5	31	1.813	0.000	0.079
	ANGIOGENESIS	C5	26	1.666	0.000	0.132
	REGULATION_OF_ANGIOGENESIS	C5	15	1.748	0.000	0.083
Hypoxia	WINTER_HYPOXIA_UP	C2	80	1.813	0.000	0.117
	MANALO_HYPOXIA_UP	C2	142	1.751	0.000	0.124
	JIANG_HYPOXIA_NORMAL	C2	184	1.734	0.000	0.107
	HARRIS_HYPOXIA	C2	61	1.731	0.000	0.106
	WINTER_HYPOXIA_METAGENE	C2	172	1.718	0.000	0.112
ATM Target	RASHI_RESPONSE_TO_IONIZING_RADIATION_1	C2	31	1.897	0.000	0.074
Histone	REACTOME_TELOMERE_MAINTENANCE	C2	55	1.710	0.000	0.115
Chemotaxis	LOCOMOTORY_BEHAVIOR	C5	31	1.595	0.000	0.176
	BEHAVIOR	C5	45	1.599	0.000	0.178
	THEILGAARD_NEUTROPHIL_AT_SKIN_WOUND_UP	C2	59	1.735	0.000	0.119
Stat3	DAUER_STAT3_TARGETS_UP	C2	31	1.714	0.000	0.114
Acid- base balance	CARBON_OXYGEN_LYASE_ACTIVITY	C5	22	1.647	0.000	0.147
	HYDRO_LYASE_ACTIVITY	C5	18	1.637	0.000	0.136