



# Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors

Andy J. Minn,<sup>1,2</sup> Yibin Kang,<sup>1</sup> Inna Serganova,<sup>3</sup> Gaorav P. Gupta,<sup>1</sup> Dilip D. Giri,<sup>4</sup> Mikhail Doubrovin,<sup>3</sup> Vladimir Ponomarev,<sup>3</sup> William L. Gerald,<sup>4</sup> Ronald Blasberg,<sup>3</sup> and Joan Massagué<sup>1,5</sup>

<sup>1</sup>Cancer Biology and Genetics Program, <sup>2</sup>Department of Radiation Oncology, <sup>3</sup>Department of Neurology, <sup>4</sup>Department of Pathology, and <sup>5</sup>Howard Hughes Medical Institute, Memorial Sloan-Kettering Cancer Center, New York, New York, USA.

**We used bioluminescence imaging to reveal patterns of metastasis formation by human breast cancer cells in immunodeficient mice. Individual cells from a population established in culture from the pleural effusion of a breast cancer patient showed distinct patterns of organ-specific metastasis. Single-cell progenies derived from this population exhibited markedly different abilities to metastasize to the bone, lung, or adrenal medulla, which suggests that metastases to different organs have different requirements. Transcriptomic profiling revealed that these different single-cell progenies similarly express a previously described “poor-prognosis” gene expression signature. Unsupervised classification using the transcriptomic data set supported the hypothesis that organ-specific metastasis by breast cancer cells is controlled by metastasis-specific genes that are separate from a general poor-prognosis gene expression signature. Furthermore, by using a gene expression signature associated with the ability of these cells to metastasize to bone, we were able to distinguish primary breast carcinomas that preferentially metastasized to bone from those that preferentially metastasized elsewhere. These results suggest that the bone-specific metastatic phenotypes and gene expression signature identified in a mouse model may be clinically relevant.**

## Introduction

Cancer metastases are responsible for the majority of cancer-related deaths. A widely held hypothesis is that cancer metastasis arises from rare cells in the primary tumor that acquire the ability to progress through sequential steps necessary to grow at a distant site (1, 2). Some of these sequential steps include invasion through extracellular matrix, intravasation, survival in the circulation, extravasation into a distant site, and progressive growth at that site. Consistent with the multistep nature, there is experimental and clinical evidence to suggest that metastasis is an inefficient process whereby the vast majority of circulating tumor cells are not able to progressively grow at distant sites (3–6). Related to this is the observation that metastatic cells exhibit tissue tropism, preferring to grow in certain organs in a way that cannot be explained by circulatory patterns alone. In breast cancer, for example, metastasis affects the bone and the lung, and less frequently the liver, brain, and adrenal medulla. Although the genetic basis of these metastatic properties is poorly understood, acquisition of the ability to complete each step involved in metastasis is thought to be driven by the accumulation of genetic mutations that may result in a rare cell's acquisition of a full complement of these mutations relatively late during the evolution of the primary tumor (1).

Recently, the development of DNA microarray technology, which allows for genome-wide transcriptomic profiling, has provided new insight into the genetic basis of metastasis. Studies using primary

tumor material have identified a gene expression signature for breast cancer metastasis consisting of a set of 70 genes (7, 8). The presence of this “poor-prognosis” signature in the primary tumor from early stage breast cancer patients is highly prognostic for the development of distant metastasis and overall survival. Work using adenocarcinoma metastases and unmatched primary tumors from breast and other tumor types has revealed similar findings (9).

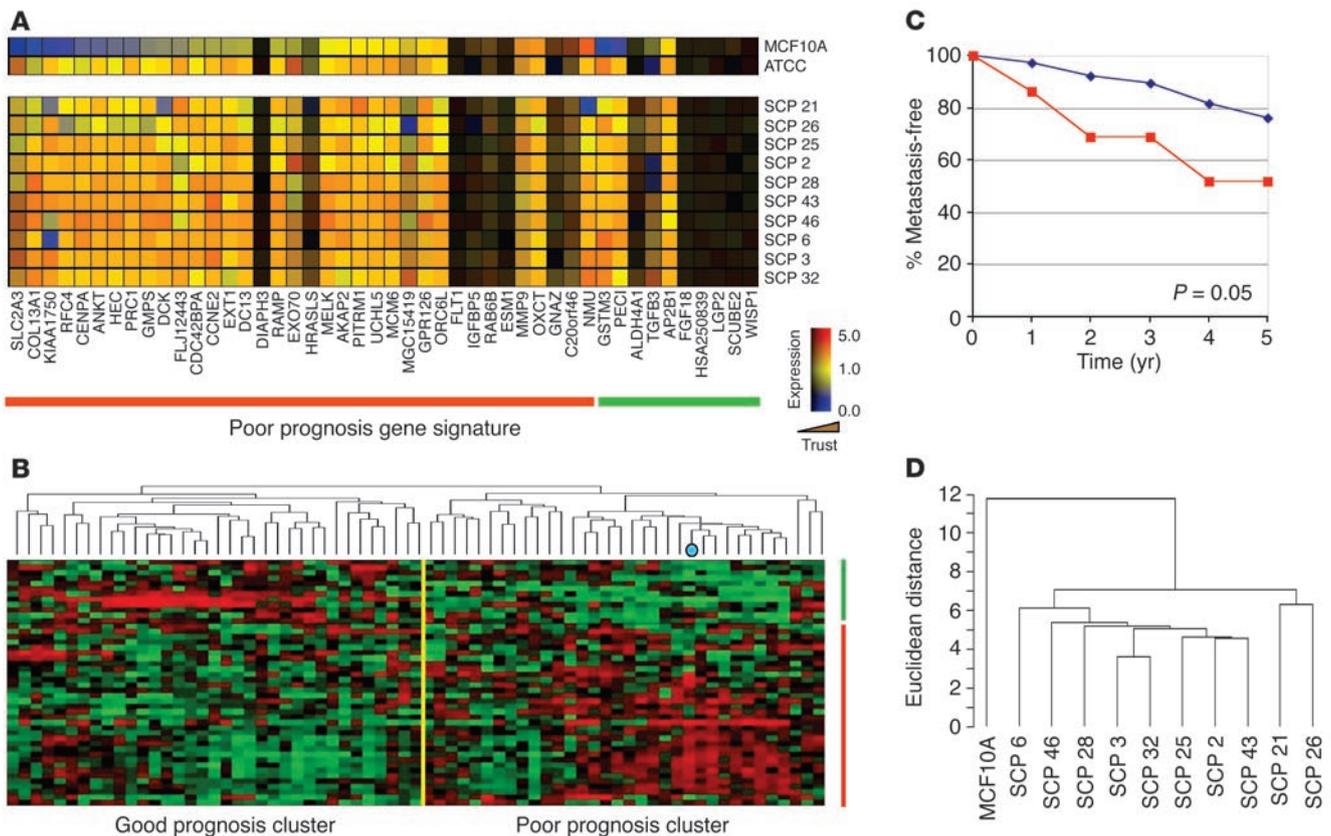
The fact that the poor-prognosis signature from early-stage primary cancers can be used to predict the development of distant metastasis has been interpreted as challenging the traditional model of metastasis because it suggests that metastatic cells may result from many of the early oncogenic events that drive primary tumor growth rather than developing from late-arising, rare cells that accumulate genomic alterations specific for metastasis (10). Other researchers have maintained the existence of distinct metastasis genes and have argued that a poor-prognosis signature may result from the aggregate contribution of these genes by subpopulations of cells that aberrantly express some but not all of the multiple genes required to complete metastasis (11). Thus, the cell that contains the full complement of metastasis-enabling genes still may be rare. Regardless, the ability of the poor-prognosis genes to directly mediate metastasis remains unknown.

Using *in vivo* selection of organ-specific metastatic cells from the human breast cancer cell line MDA-MB-231, we recently identified and functionally validated a set of genes that specifically mediate osteolytic bone metastasis in the mouse (12). Cells that express these genes and that are capable of bone metastasis pre-exist within the MDA-MB-231 parent line, which as a population already carries the poor-prognosis signature. This cell line was originally established as the total outgrowth of cells derived from a pleural effusion of a patient who relapsed years after removal of the primary tumor (13). In the present study, we investigate the

**Nonstandard abbreviations used:** ATCC, American Type Culture Collection; BLI, bioluminescence imaging; eGFP, enhanced green fluorescent protein; HSV1-TK, herpes simplex virus 1 thymidine kinase; R index, reproducibility index; ROI, region of interest; SCP, single cell-derived progeny; TGL, thymidine kinase, GFP, luciferase.

**Conflict of interest:** The authors have declared that no conflict of interest exists.

**Citation for this article:** *J. Clin. Invest.* 115:44–55 (2005).  
doi:10.1172/JCI200522320.



**Figure 1**

SCPs from MDA-MB-231 cells have a poor-prognosis gene expression signature. **(A)** Microarray expression data of 46 of the 70 poor-prognosis genes (7) that are present on the Affymetrix U133A GeneChip for the MCF10A normal breast epithelial cell line, parental MDA-MB-231 cell line, and various SCPs from MDA-MB-231. Each column represents a gene (denoted along the bottom) and each row represents a cell line (denoted along the right). Genes of the poor-prognosis signature that are expressed at higher levels in poor-prognosis tumors are above the red line, and those that are underexpressed are above the green line. Genes with low trust values due to low or absent expression are shaded in darker colors (Trust; wedge). **(B)** Microarray expression data of primary human breast carcinoma from 63 patients treated at our institution who had at least 5 years of clinical follow-up and/or developed metastatic disease. Hierarchical clustering of the patients' data was performed with the 46 poor-prognosis genes. Each column represents a patient and each row, a gene. The MDA-MB-231 cell line was included and is denoted by a blue dot in the dendrogram. Those patients in the good-prognosis versus the poor-prognosis cluster are separated by the yellow line. **(C)** Five-year metastasis-free survival data for the 63 patients classified according to the hierarchical clustering described in **B**. The *P* value shown in the graph was calculated by the  $\chi^2$  test. **(D)** Dendrogram showing hierarchical clustering of the SCPs and MCF10A using the poor-prognosis genes. A scale of the distance metric used is shown on the left.

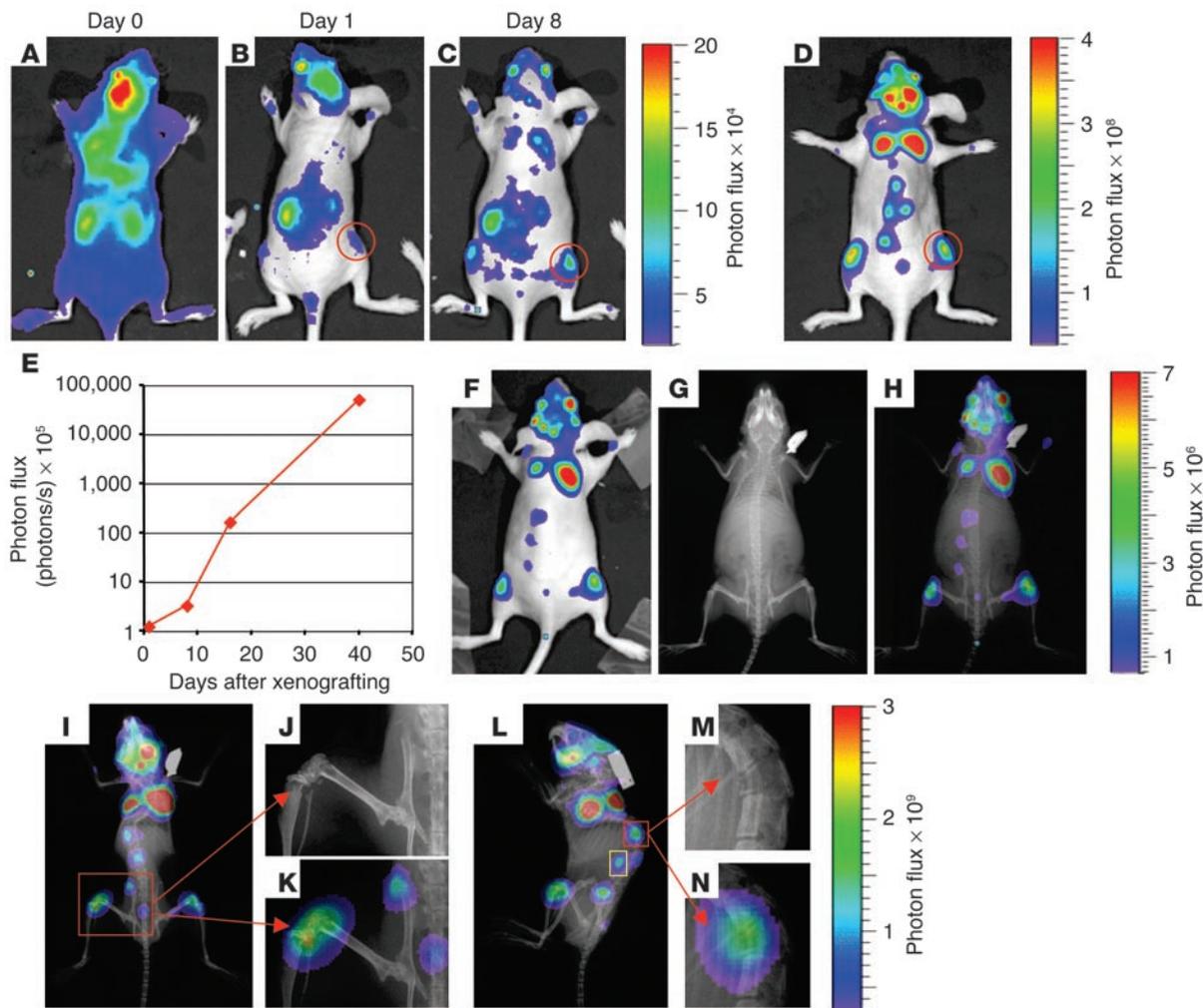
relationship between this bone metastasis signature, the general poor-prognosis signature, and the metastatic activity of individual cells from the parental population and of a cohort of metastatic human primary tumors.

**Results**

*Similar poor-prognosis gene expression signatures in different single cell-derived progenies.* The poor-prognosis gene expression signature for breast cancer, which can be used to predict the development of distant metastasis, consists of 70 genes, 58 of which are upregulated and 18 of which are downregulated, and correlates closely with negative estrogen receptor status (7). Most tumors in the poor-prognosis group have only a fraction (on average, approximately one third) of the 70 gene expression events that constitute the poor-prognosis signature. Furthermore, these gene expression events often show extensive variation among different tumors with a poor prognosis. We recently reported that MDA-MB-231 cells, as directly obtained from the American Type Culture Collection (ATCC), also have the

poor-prognosis signature. Of the 70 genes from this signature, 46 were present on the Affymetrix U133A GeneChip that we used for our microarray analysis (Figure 1A). Of the 58 upregulated genes of the poor-prognosis signature, 36 were present on this microarray. Compared with the MCF10A cell line derived from nonmalignant human breast epithelium, the majority of these 36 genes were upregulated in parental MDA-MB-231 cells. Of the 18 downregulated genes from the poor-prognosis signature, 10 were present on the U133A GeneChip. Consistent with downregulation in poor-prognosis tumors, 7 of the 10 had low trust values due to their low or absent expression.

To further confirm that MDA-MB-231 cells have a poor-prognosis gene expression signature, we compared the transcriptomic profile of these cells with that of a cohort of primary breast carcinomas from patients treated at the Memorial Sloan-Kettering Cancer Center. All of these patients had at least 5 years of clinical follow-up or had developed metastatic disease. Hierarchical clustering using the poor-prognosis gene expression signature (7) separated these tumors



**Figure 2**

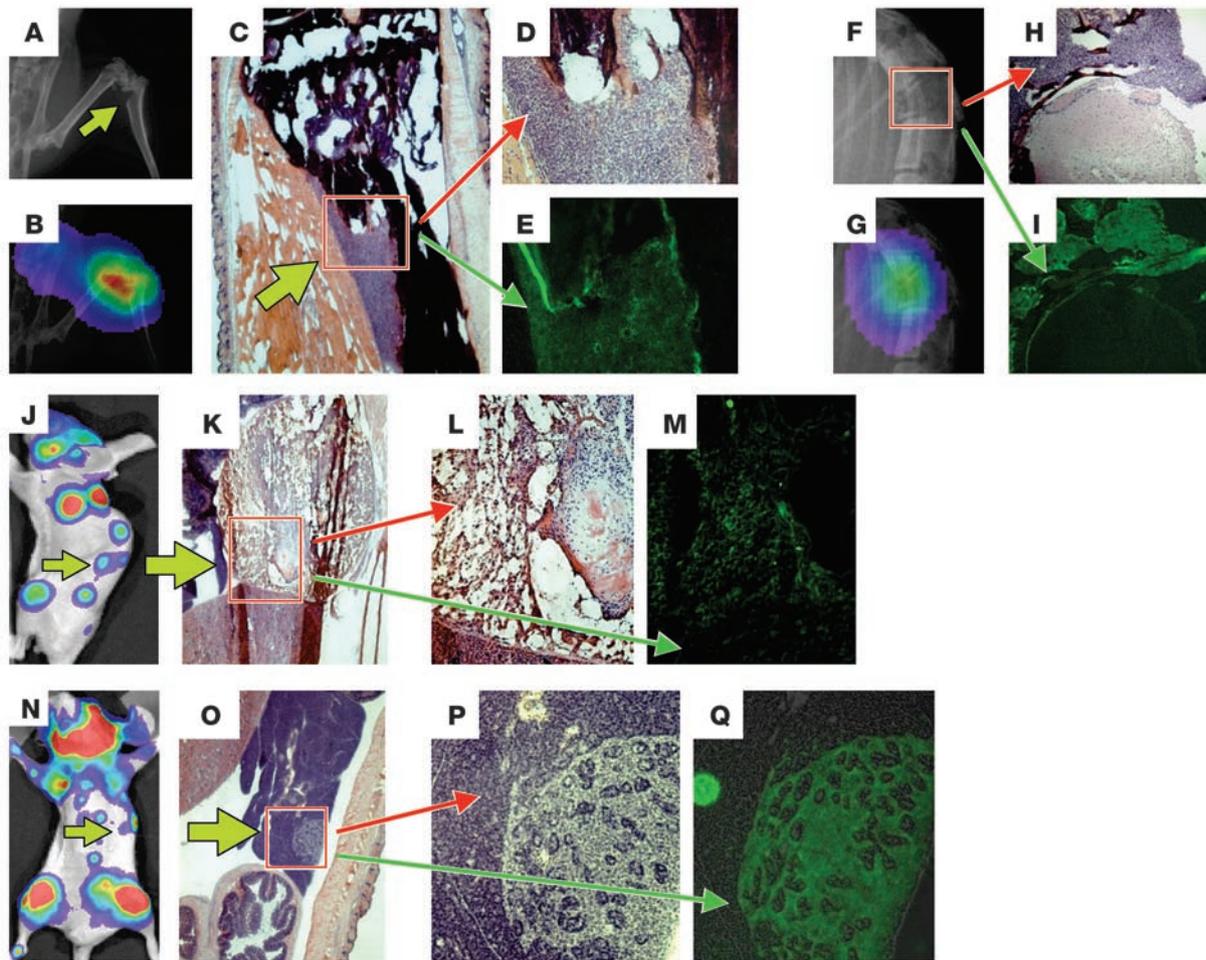
Noninvasive BLI to monitor the development of osteolytic metastases from the same mouse. (A–D) SCP2, a highly metastatic clone from MDA-MB-231, was transduced with the luciferase-containing TGL reporter gene and was injected into the left cardiac ventricle of an immunodeficient mouse. At the indicated times after xenografting, the bioluminescence signal was captured. The intensity of the signal, measured as photon flux, is shown as a color scale. Images for days 0, 1, and 8 are displayed on the same scale, while the day-35 image is shown on a different scale due to the exponential growth of the metastases. A metastasis to the right hindlimb is circled in red. (E) The growth kinetics of the right hindlimb metastasis outlined by the red circle shown in B–D was quantified by measurement of photon flux. (F–H) A bioluminescence image (F) and a skeletal x-ray image (G) were obtained on day 16 after xenografting. Images were superimposed (H) to demonstrate registration of the bioluminescence signals with skeletal anatomy. (I–N) A superimposed image from day 45 (I and L) reveals extensive areas of osteolytic destruction that correspond to bioluminescence signals. Magnification of regions outlined in red shows involvement of the femur/tibia, iliac crest of the pelvis, and the sacrum (J and K), in addition to the vertebrae (M and N). The bioluminescence signal from the region outlined in yellow on the left lateral projection (L) does not overlap with skeletal structures and originates from the adrenal gland (Figure 3, J–M).

into two major clusters, one cluster corresponding to patients with a poor-prognosis signature and the other representing those with a “good-prognosis” signature (Figure 1B). Consistent with previous reports, patients in our cohort with a poor-prognosis signature had a significantly worse 5-year metastasis-free survival than those with the good-prognosis signature (Figure 1C). MDA-MB-231 cells fall squarely within this poor-prognosis group (Figure 1B). Thus, MDA-MB-231 cells express a typical poor-prognosis tumor profile.

Among the questions raised by these observations is whether the particular set of poor-prognosis gene expression events presented by a poor-prognosis tumor reflects the presence of this particular pattern in the majority of malignant cells of the tumor or if it reflects contributions from different cells in the population.

To address this question in the MDA-MB-231 case, we used various single cell-derived progenies (SCPs) obtained from single-cell cloning and analyzed them for the presence of a poor-prognosis signature. Although there was some variation among the SCPs in the expression levels of the genes that comprised the signature, the SCPs maintained a set of poor-prognosis gene expression events similar to that found in the ATCC population from which they were derived (Figure 1A). A dendrogram of the SCPs using the poor-prognosis gene set confirmed that the distance metric between the SCPs was significantly less than the distance metric between the whole group of SCPs and MCF10A (Figure 1D).

Flow cytometry analysis of the parental MDA-MB-231 cell population indicated that approximately 10% of cells in this population



### Figure 3

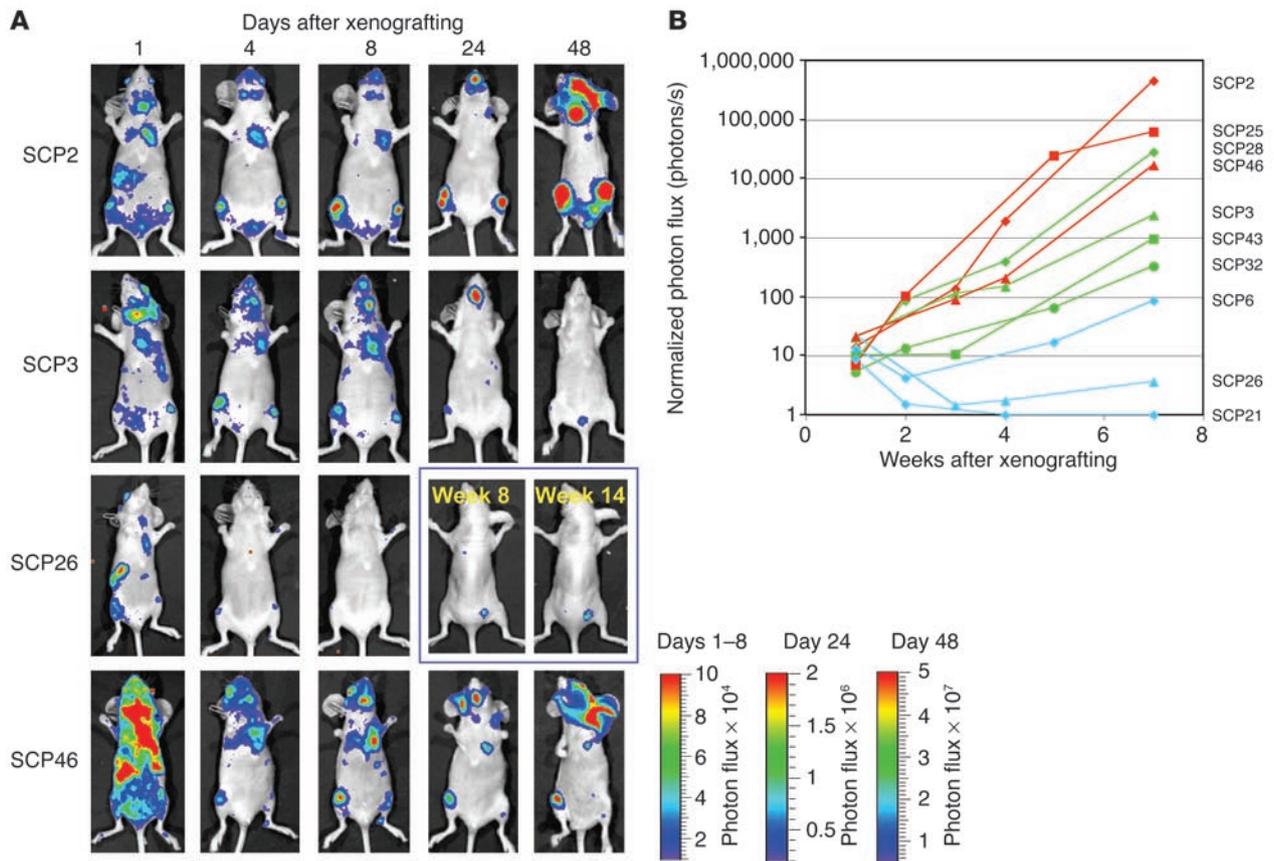
Verification of macroscopic and microscopic metastases by fluorescence histology. (A–I) A pathological fracture involving the proximal tibia (A–E) or vertebrae (F–I) is demonstrated by skeletal x-ray (A and B) and an overlay of this x-ray with BLI (B and G) from the same mouse as that described in Figure 2. To confirm metastases, we performed whole-mount frozen sectioning. Regions corresponding to the fractured tibia and vertebra were analyzed by H&E staining (C, D, and H) or unstained sections were analyzed for GFP fluorescence (E and I). (J–M) A lateral projection of a bioluminescence image from day 45 (J) corresponding to the same image as that in Figure 2L reveals a signal originating from the adrenal gland (green arrow), as shown by H&E staining (K). Magnification of the boxed region in K (L) and GFP fluorescence (M) of the left adrenal gland are shown. (N–Q) Inspection of organs in the left upper abdominal quadrant with areas of bioluminescence signal (N) reveals a focus of tumor growth in the pancreas (O). Magnification of the boxed region in O (P) and GFP fluorescence (Q) are shown.

expressed CXCR4 (data not shown), a product representative of the bone metastasis gene expression signature (12). A similar percentage of SCPs were found to overexpress CXCR4 (12). Thus, based on these criteria at least, our single-cell cloning process did not introduce bias in the selection of cell clones representing the parental population.

**Noninvasive bioluminescence imaging of metastases.** After intracardiac injection of parental MDA-MB-231 cells into immune-deficient mice, approximately 30% will develop osteolytic bone metastasis that is evident by skeletal x-ray imaging (12). Subpopulations that are more osteolytic than the parental population have been obtained through a process of *in vivo* selection for bone metastasis or by isolation of SCPs from parental MDA-MB-231 cells. However, the sensitivity of skeletal x-ray in detecting nonosseous metastasis is poor. Likewise, findings at necropsy may also fail to reveal small and/or anatomically inconspicuous lesions. Indeed, at necropsy, MDA-MB-231 cells are infrequently found to have metastasized to nonosseous organs such as the adrenal medulla.

In order to better characterize the overall metastatic properties of MDA-MB-231 SCPs and their relationships to both the poor-prognosis and the bone metastasis gene sets, we used luciferase-based, noninvasive bioluminescence imaging (BLI) and fluorescence microscopy using a novel triple-modality reporter gene, thymidine kinase, GFP, luciferase (TGL) (14). This artificial gene encodes a triple fusion protein with herpes simplex virus 1 thymidine kinase (HSV1-TK) fused to the N terminus of enhanced green fluorescent protein (eGFP) and firefly luciferase fused to the C terminus of eGFP. When transduced into cells, HSV1-TK allows for nuclear imaging, eGFP can be utilized for fluorescence, and luciferase allows for BLI.

SCP2 is a single cell-derived population of MDA-MB-231 cells that produces aggressive osteolytic lesions by 8 weeks after left ventricular cardiac injection into immunodeficient mice. As a test of the sensitivity and resolution of the TGL reporter gene, we transduced SCP2 with the TGL reporter and monitored the development



**Figure 4** SCPs exhibit different abilities to metastasize to bone. (A and B) Each of the SCPs was labeled with the TGL reporter, and  $1 \times 10^5$  cells were injected into the left cardiac ventricle. At the indicated days after xenografting, bioluminescence images were acquired. (A) Representative mice injected with a representative set of SCPs are shown in the supine position. The intensity of the signal from days 1, 4, and 8 are on equivalent scales, while day 24 and day 48 are each on separate scales due to increasing signal strength and to avoid signal saturation. (B) The normalized photon flux from the dominant signal originating from the hindlimbs, forelimbs, or pelvis of all the SCPs studied was measured over the indicated time course. SCPs were ranked according to their growth kinetics in either bone or lung. SCPs with a higher rank order for bone are shown in red, and those with a higher rank order for lung are shown in green. The bottom three SCPs for both bone and lung are classified as being the least metastatic and are shown in blue.

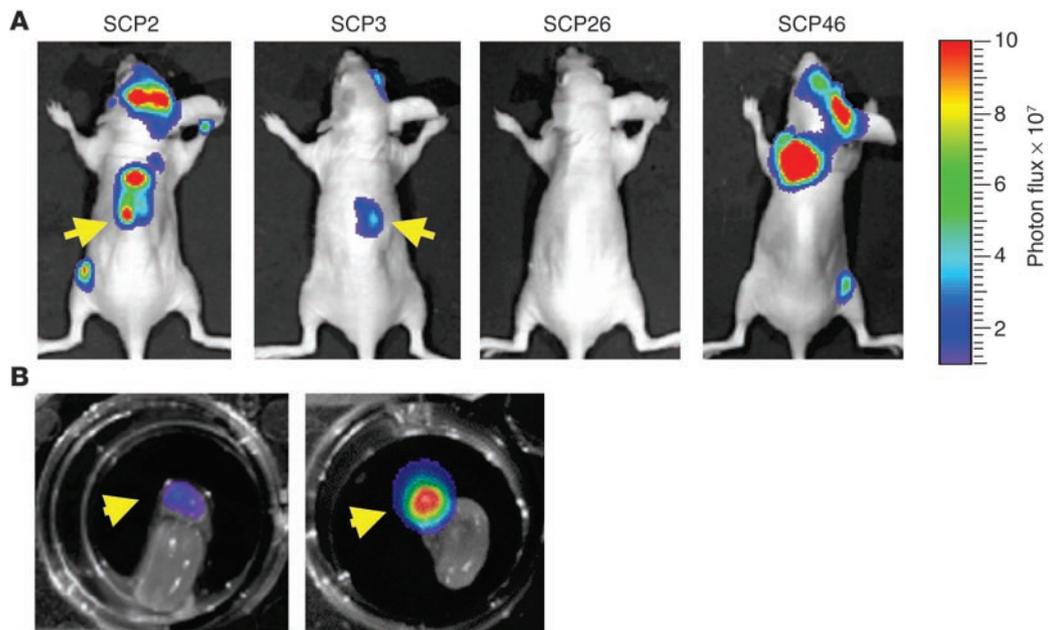
of osteolytic metastases. Shortly after the injection of  $1 \times 10^5$  cells into the left cardiac ventricle, a diffuse whole-body bioluminescence signal was detected (Figure 2A). This signal followed systemic blood flow patterns, with areas of strongest signal probably corresponding to organs receiving the highest percentage of cardiac output; namely, kidney, liver, and brain. At day 1 after injection, much of the diffuse signal disappeared; however, foci of arrested tumor cells could be seen. These foci increased in number and intensity through the first week (Figure 2, A–C). In particular, an increasing signal could be detected in the hindlimbs that corresponded to primary areas for the development of osteolytic metastasis (Figure 2, B–D, red circles). This major hindlimb signal was quantified by measurement of photon flux and demonstrated logarithmic growth (Figure 2E).

Because bioluminescence signals could be correlated only with surface anatomy, we sought a way to assign major areas of bioluminescence to anatomical structures. At day 16 after injection, we overlaid the bioluminescence signal with skeletal x-ray images in order to analyze the correlation between areas of signal with skeletal anatomy. The majority of the signal overlapped well with bony structures, including the distal femur/proximal tibia, bony pelvis, scapula, vertebra, distal ulna, and skull (Figure 2, F–H). Although inspection of the

x-ray images at day 16 did not reveal evidence of osteolytic destruction at the sites of overlap, skeletal x-ray imaging of the same animal at day 45 demonstrated overlapping areas with extensive osteolytic destruction involving the distal femur/proximal tibia, iliac crest, sacrum, and vertebral body (Figure 2, I–N). Thus, these data suggest that BLI can be significantly more sensitive in detecting bone metastasis than x-ray imaging, as it allows monitoring of the development of bone metastasis from initial arrest to osseous destruction.

*Verification of BLI by fluorescence histology.* In order to examine the regions of osteolytic metastasis histologically and to search for other, less obvious sites of occult metastases, we used whole-mount frozen sectioning to look for tumor-derived GFP fluorescence by microscopy. Skeletal x-ray and BLI identified a pathological fracture of the tibia (Figure 3, A and B). H&E staining of sections corresponding to this region revealed tumor cells eroding through the cortex of the tibia (Figure 3, C and D), and GFP fluorescence of a serial section confirmed the metastasis (Figure 3E). Similarly, a collapsed vertebral body was also demonstrated to be due to growth of tumor cells through the bone and into the spinal canal (Figure 3, F–I).

Not all areas of bioluminescence signal could be overlaid with skeletal structures. For example, as shown on day 35 after xenografting,

**Figure 5**

Differential ability among SCPs to metastasize to the adrenal gland. **(A)** After intracardiac injection of individual SCPs, bioluminescence images were acquired and analyzed for signals originating from regions consistent with adrenal metastasis (arrows). Shown are representative mice at 7 weeks after injection with SCPs that show varying abilities to give rise to adrenal metastasis. **(B)** At necropsy, left and right adrenal glands (with the kidneys) were removed and were imaged ex vivo for bioluminescence. Arrows show the locations of the left and right adrenal glands, respectively, from a representative mouse with adrenal metastasis.

bioluminescence signals on both sides lateral to the vertebral column could be detected (Figure 2D). On a lateral projection, these signals lay anterior to the vertebrae (Figures 2L and 3J). At necropsy, enlarged and necrotic adrenal glands were noted (Figure 3, K and L), and fluorescence microscopy confirmed that this was due to metastasis (Figure 3M). In addition, careful analysis of whole-mount frozen sections also identified other nonosseous sites of microscopic metastases corresponding to weak regions of bioluminescence signal. For example, small foci of signal were noted in the upper left quadrant of the abdomen (Figure 3N). This signal was confirmed to be due to microscopic metastasis involving the pancreas (Figure 3, O–Q).

In total, these data demonstrate that the TGL reporter gene enables the use of a noninvasive method for tracking metastases from the initial arrest in distant organs to the development of gross lesions. The growth of these lesions can be quantified by measuring photon flux and confirmed by fluorescence microscopy. The sensitivity of the system is exemplified by the ability to detect and confirm microscopic metastases that would otherwise be overlooked by routine necropsy.

*Differential bone-metastatic activity with a similar poor-prognosis signature.* Empowered by the sensitivity of the TGL reporter system, we sought to fully characterize the metastatic phenotypes of the SCPs. To assess the metastatic activity that develops after hematogenous spread, we introduced each of the SCPs into the arterial circulation of immunodeficient mice by injection into the left cardiac ventricle. The major site of colonization and growth among the SCPs is the bone (hindlimbs, ribs, pelvis/sacrum, and skull/mandible) (Figure 4A). However, the SCPs displayed significant variation in their ability to grow in bone, even though the various SCPs proliferated in culture at comparable rates (data not shown). The dominant signals on the supine projections came from the hindlimbs and the

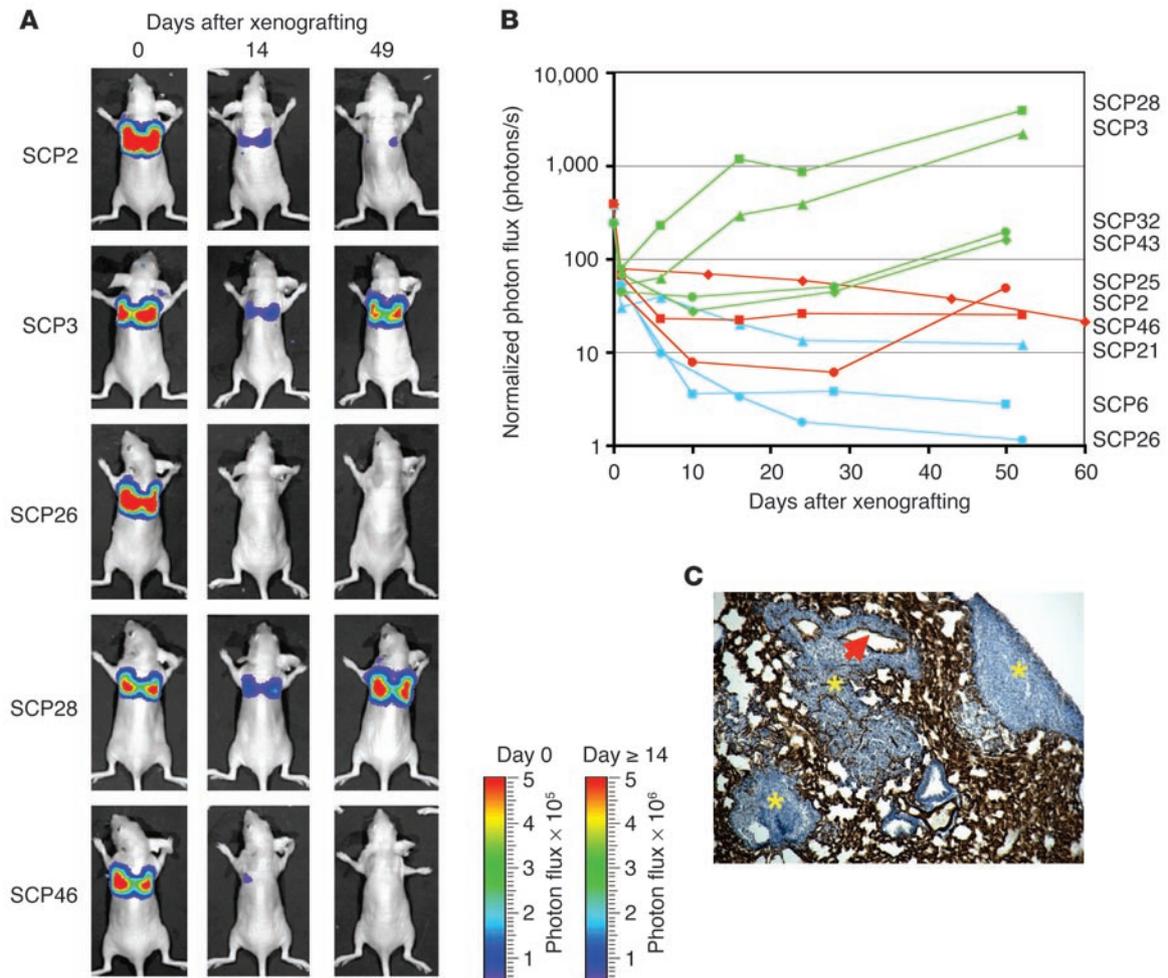
bones of the skull. For presentation purposes, the bioluminescence data from days 1–8 are displayed on the same scale and day 24 and day 48 are each displayed on a different scale. Comparisons within these groups across SCPs demonstrated that SCP2 and SCP46 were more metastatic to bone than are SCP3 and SCP26.

The dominant hindlimb lesion from the complete set of SCPs was quantified by measurement of photon flux, and the kinetics of growth are shown in Figure 4B. The aggressiveness of SCP2, SCP25, SCP28, and SCP46 in forming bone metastasis was shown by a 3- to 4-log growth of the dominant hindlimb lesion over the course of 7 weeks. Most of these mice became cachectic and were sacrificed. The aggressive nature of these SCPs is consistent with their expression of a previously described bone metastasis gene expression signature

**Table 1**  
Adrenal metastases and SCPs

| Progeny | Number of mice analyzed | Number of mice with adrenal metastases (%) |
|---------|-------------------------|--|
| SCP2    | 4                       | 2 (50)                                     |
| SCP3    | 9                       | 7 (78)                                     |
| SCP25   | 4                       | 1 (25)                                     |
| SCP6    | 5                       | 0  |
| SCP32   | 4                       | 0  |
| SCP43   | 5                       | 0  |
| SCP21   | 5                       | 0  |
| SCP26   | 4                       | 0  |
| SCP28   | 5                       | 0  |
| SCP46   | 4                       | 0  |

The presence of adrenal metastasis was determined for the entire cohort of SCPs.



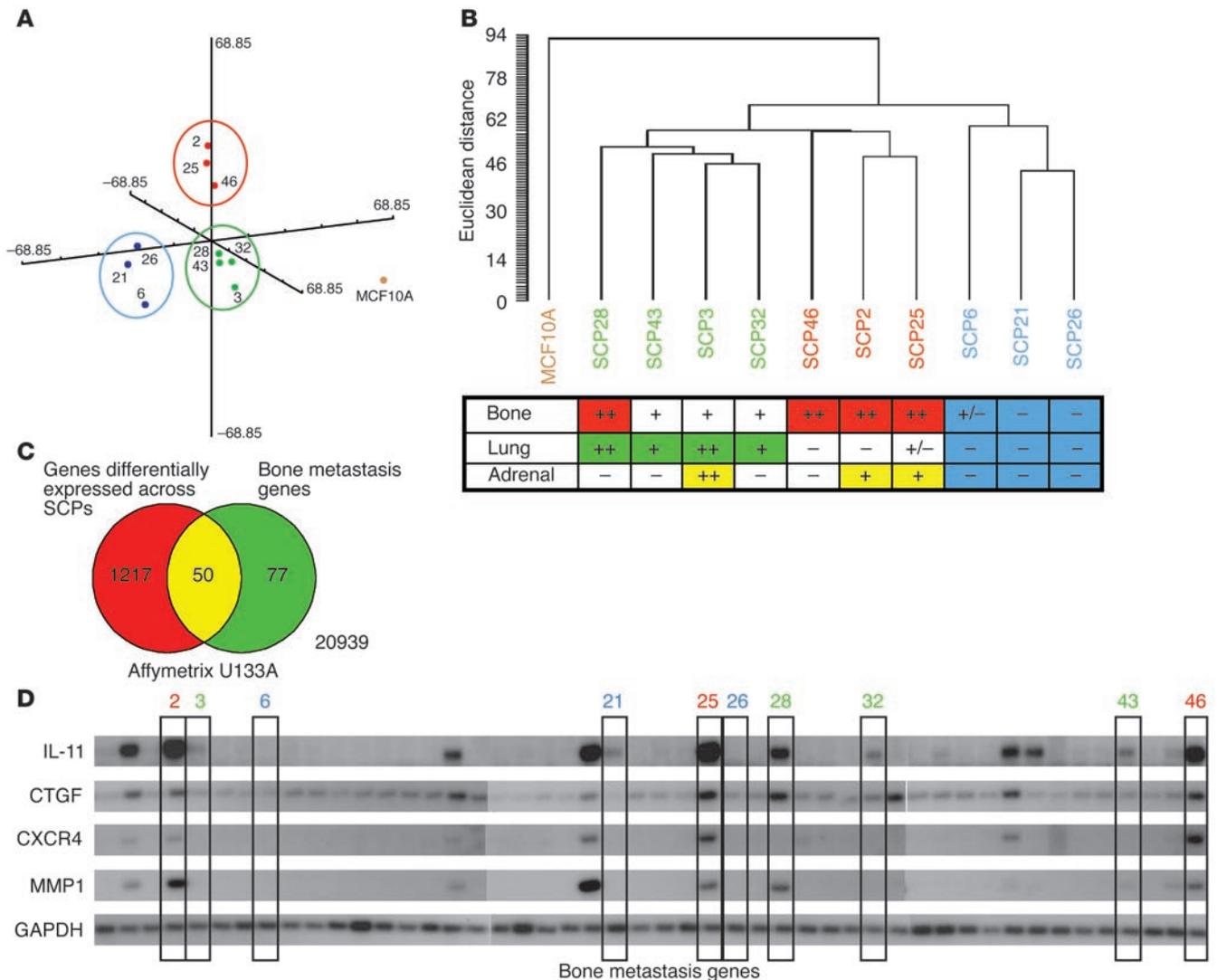
**Figure 6** SCPs demonstrate different abilities to metastasize to the lung. (A–C) Each of the SCPs was labeled with the TGL reporter, and  $2 \times 10^5$  cells were injected into the tail vein. At the indicated day after xenografting, bioluminescence images were acquired. (A) Representative mice injected with a representative set of SCPs are shown in the supine position. The intensity of the signal from day 0 is displayed on one scale, while that of days 14 and 49 (Day  $\geq 14$ ) are on a different scale due to increasing signal strength and to avoid signal saturation. (B) The normalized photon flux from the lung of all the SCPs studied was measured over the indicated time course. SCPs are color-coded as described in Figure 4B. (C) The lungs of SCPs that show growth in lung were analyzed histologically. A lung section from a representative SCP is shown stained for CD31, a marker for vascular endothelial cells, and counterstained with eosin. Asterisks mark regions of parenchymal tumor growth. The red arrow shows a CD31-positive blood vessel with an associated perivascular tumor growth pattern.

(12). However, SCP43, SCP3, and SCP32 were weaker in their metastatic growth to bone, while SCP6, SCP26, and SCP21 were the most weakly metastatic to bone. This reduction in bone metastasis ability correlated with the attenuation in expression of the bone metastasis genes (see Figure 7D). Interestingly, even among the weakest populations, we were able to detect the presence of bone metastasis. For example, at 14 weeks after xenografting of SCP26, a dormant metastatic focus within the hindlimbs was detectable in half of the mice (Figure 4A and Supplemental Figure 1; supplemental material available online with this article; doi:10.1172/JCI200522320DS1). Thus, these data demonstrate that the bone-metastatic activity of MDA-MB-231 cells does not correlate with the expression of their poor-prognosis signature but instead with the expression of our previously described bone metastasis gene set.

*Different organ specificity of metastasis by different cells from the same population.* After extensive analysis of metastatic growth by BLI,

whole-mount fluorescence microscopy, and micro-positron emission tomography (data not shown), we found bone to be the major site of tumor growth after arterial inoculation. In general, growth in other organs was rare, making comparable analysis unfeasible. However, one exception was metastatic growth in the adrenal gland, which occurred at an appreciable frequency. We were able to detect adrenal metastases in a minority of the SCPs by looking for dorsally located signals on either or both sides of the vertebral column that were suspicious for adrenal metastases (Figure 5A). These “suspicious” signals were confirmed at necropsy by gross inspection and/or ex vivo BLI (Figure 5B). Of the SCPs, SCP3 was the most consistent in producing adrenal metastasis (Table 1).

Due to size restrictions imposed by murine capillaries, human tumor cells are rarely able to pass from the arterial to the venous system (or vice versa) by way of the lungs (2). Therefore, we injected the SCPs into the tail vein in order to study the ability of SCPs to



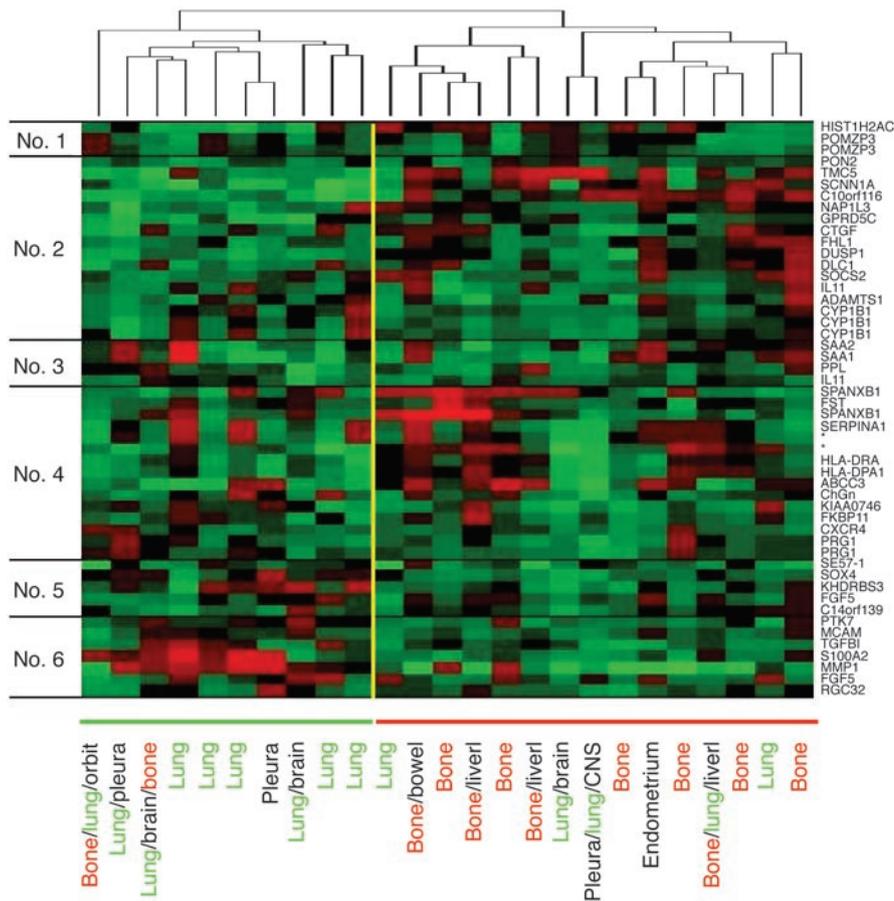
**Figure 7**

Genome-wide “unsupervised” classification of the SCPs correlates with metastatic phenotype. **(A)** A multidimensional scaling plot illustrates the relationship between the various SCPs and their primary metastatic tropism based on genes that are differentially expressed across the SCPs starting from the more than 22,000 present on the Affymetrix U133A GeneChip. SCPs are color-coded according to their primary metastatic tropism (green for lung, red for bone, and blue for weakly metastatic). The plot demonstrates that SCPs with the same primary metastatic tropism group together in 3-dimensional space. Each group is each enclosed in a circle. MCF10A is shown by itself (gold dot). **(B)** Hierarchical clustering of the SCPs based on genes differentially expressed reveals similar relationships and a similar association with metastatic tropism, as summarized in the table below the dendrogram. **(C)** A Venn diagram demonstrates the relationship between the genes differentially expressed across the SCPs and a previously described bone metastasis gene set. Of 1,267 differentially expressed genes, 50 of the 127 bone metastasis genes (102 are unique) overlap. **(D)** A Northern blot showing the expression levels of 4 of the bone metastasis genes among the SCPs used in this study (boxed and labeled by SCP, with the color of the label corresponding to tissue tropism). GAPDH, loading control.

metastasize to the lung. Shortly after tail vein injection, all detectable cells became trapped in the lung (Figure 6A). Within the first few days, there was a substantial attenuation of this signal. In SCP6 and SCP26, this attenuation continued over the ensuing weeks, suggesting that as in the bone, these SCPs were unable to efficiently survive and grow in the lung. The highly bone-metastatic populations SCP2 and SCP46 were also unable to grow in the lung but were able to survive over the course of several weeks, as shown by their persistent bioluminescence signal. In contrast, SCP3 and SCP28, and to a lesser extent SCP32 and SCP43, were able to grow in the lung. To confirm the presence of lung metastases, we per-

formed histological analysis. Immunohistochemistry with CD31, which is a marker for vascular endothelial cells, revealed multiple areas of perivascular tumor growth and growth within the capillary-rich lung parenchyma (Figure 6C).

It is hypothesized that growth at metastatic sites is enhanced by genes that confer productive tumor-stroma interaction. Thus, metastatic cells that grow well at one site may not grow well at another. Based on the metastatic tropisms of each SCP defined by BLI, we ranked SCPs according to their growth kinetics in either bone or lung. As shown in Figures 4B and 6B, SCPs with a higher rank order for bone were color-coded in red, and those with a higher rank order



**Figure 8**

Segregation of primary breast carcinomas using the bone metastasis gene signature. The microarray data for primary breast tumors from patients that developed distant metastasis were used in hierarchical clustering using the 50 bone metastasis genes described in Figure 7C. Both the patient samples (columns) and the genes (rows) were clustered. The patient samples were classified into two major clusters with an overall R index of 0.90 (a robustness index; see Methods). The site(s) of distant recurrence for each patient is (are) listed along the bottom, with the site of first recurrence listed first. The genes were clustered into six groups (labeled along the left), with the gene symbol of each gene shown (on the right). The asterisks indicate genes without symbols (from top to bottom, Affymetrix probe set identifiers 211429\_s\_at and 211796\_s\_at). Orbit, orbit of the eye.

for lung were coded in green. The bottom three SCPs for both bone and lung growth were classified as least metastatic and were coded in blue. Consistent with the concept of metastatic tissue tropism, the SCPs that were the best at growing in bone were not the best at growing in lung, and the most lung-metastatic SCPs generally were not the most metastatic to bone (Figures 4B and 6B).

In summary, extensive analysis of the metastatic activity of the various SCPs derived from the same cancer cell line has revealed significant variability in their metastatic activity. This variability is seen in cell survival, organ-specific colonization, and organ-specific growth.

*Genome-wide variation correlates with organ-specific metastatic phenotype.*

Because the presence of the poor-prognosis signature of the various SCPs does not strongly correlate with any recognizable aspects of their metastatic activity, this supports the hypothesis that many characteristics of metastatic activity are governed by a different set(s) of genes. To test this idea, we first analyzed the SCP microarray data to estimate the amount of variation in the expression levels of the more than 22,000 genes represented on the Affymetrix U133A GeneChip. After filtering out genes in which more than half of the SCPs showed less than 1.5-fold change in expression level and by eliminating genes that were absent in all of the datasets, 1,267 differentially expressed genes remained. A higher-stringency filter that required a minimum twofold change in expression reduced this list further to 286 genes. Multidimensional scaling was then used to determine the relatedness of the different SCPs based on the 1,267 broadly differentially expressed genes. SCP2, SCP25, and SCP46 formed one distinct group in three-dimensional space, while SCP28, SCP3, SCP32, and SCP43 formed another (Figure 7A). A

third group was formed by SCP6, SCP21, and SCP26. Although distinct, these three groups of SCPs were significantly closer to each other than they were to MCF10A. As expected, hierarchical clustering also revealed similar relationships (Figure 7B). Both the multi-dimensional scaling and the hierarchical clustering were repeated with the more stringent 286 gene list and numerous other filtered lists and gave similar results (data not shown). Interestingly, both of these unsupervised methods (i.e., methods wherein knowledge of class assignments are not used in the analysis) defined groups that reflected the BLI-assigned primary metastatic tropisms of the SCPs, as shown by the color coding. The group formed by SCP28, SCP43, SCP3, and SCP32 was mainly metastatic to the lung (green), while the group formed by SCP46, SCP2, and SCP25 exhibited aggressive metastatic growth in the bone (red). The least metastatic of the SCPs, SCP6, SCP21, and SCP26, formed the third group (blue). Some SCPs showed significant multi-tropic properties (Figure 7B). Thus, the “unsupervised” separation of the SCPs into broad groups that correlate with primary properties of their metastatic phenotypes supports the notion that distinct gene expression patterns are responsible for the variability seen in their metastatic activities.

To validate that metastasis-specific genes were among the 1,267 differentially expressed genes, we determined how many of the 102 unique genes from our previously described (12) and independently derived bone metastasis gene set (represented on the U133A GeneChip by 127 probe sets) were among the 1,267 genes. As seen in the Venn diagram in Figure 7C, 50 of the 127 bone metastasis genes were overlapping. This set of 50 included *IL11*, *CTGF*, and *CXCR4*, three genes that were determined to specifi-



cally cause bone metastasis (12). Accordingly, Figure 7D demonstrates that the expression of these genes strictly correlated with bone-specific growth (Figure 7D).

*Segregation of primary tumors using a bone metastasis gene expression signature.* The existence of a poor-prognosis gene expression signature from the bulk expression data of primary breast cancers suggests that the emergence of cells that express metastasis genes may occur early during tumorigenesis. Therefore, we wanted to determine whether the bone metastasis genes that we identified in our MDA-MB-231 model system in the mouse (12) could be detectable within primary breast carcinomas. To this end, we used the 50 bone metastasis genes expressed among the bone-metastatic SCPs. Hierarchical clustering of all 63 primary breast tumors in our cohort did not robustly distinguish those tumors that gave rise to bone metastasis from those that did not (data not shown). This suggests that either our bone metastasis signature carries little predictive value or our genes are expressed only by an undetectable subpopulation of tumor cells.

To help distinguish between these two possibilities, we restricted our analysis to those primary tumors that gave rise to distant metastasis (mainly to bone and/or to lung) (Figure 8). Under these conditions, the 50 bone metastasis genes could be used to divide the primary breast carcinoma groups into two major clusters with an overall reproducibility index (R index) of 0.90, which is indicative of the robustness of this cluster. The primary breast carcinomas that gave rise to bone metastasis were predominantly associated with the second cluster. In contrast, those samples that produced lung metastasis were mainly grouped together by the first cluster. The 50 bone metastasis genes were also clustered together into six groups based on similarity in their expression pattern. Gene cluster 2 represented genes that were generally upregulated in the primary tumors that developed bone metastasis. Genes in this cluster included *CTGF* and *IL11*, in addition to other genes that are upregulated in the bone-metastatic SCPs, including *NAPIIL3*, *DUSP1*, *ADAMTS1*, and *SOCS2* (Supplemental Table 1). Some genes that are upregulated in the bone-metastatic SCPs, such as *MMP1*, are not selectively upregulated in the breast carcinoma primary tumors that develop bone metastasis; for example, *MMP1* is also involved in lung metastasis (our unpublished observations). The failure of other genes to display concordant expression patterns in the SCPs and the breast primary tumors may be because they are not biologically relevant and/or because of unknown peculiarities of the clinical data set or the MDA-MB-231 model system. Nonetheless, these data suggest that the development of distant sites of metastasis in breast cancer patients is related to differences in the gene expression pattern that is discernible by our bone metastasis gene expression signature.

## Discussion

In this study, we have demonstrated that SCPs from a metastatic parental breast cancer population carry a poor-prognosis signature. This signature varied little from SCP to SCP; however, the metastatic activity of different SCPs varied significantly. With the sensitivity afforded by noninvasive BLI coupled with fluorescence microscopy, we were able to fully characterize the metastatic activities of individual SCPs by evaluating tissue tropism and growth kinetics. We determined that some SCPs were capable of efficient metastasis to bone, others metastasized better to lung, and a minority were also able to colonize and grow within the adrenal gland and/or other sites. This activity resembles the typical distribution of breast cancer metastases observed in patients. Some

SCPs exhibit multiple tropisms, while others, in contrast, are only weakly metastatic and/or give rise to dormant lesions. The presence of cells with different metastatic properties from the same pleural effusion-derived cell line may reflect an accumulation of circulating tumor cells from multiple metastatic sites within the pleural fluid of the patient from which the cells were derived.

Although we cannot rule out the possibility that minor variations in the poor-prognosis signature may contribute to these differences in metastatic phenotypes, hierarchical clustering based on the poor-prognosis genes does not clearly segregate the SCPs into different groups that correlate with particular aspects of metastatic activity such as colonization and growth within specific organs. This suggests that the genes that make up the poor-prognosis signature do not control these more specific metastatic properties. In contrast, hierarchical clustering based on the entire gene expression data set does segregate the SCPs into different groups with different organ tropisms. The poor-prognosis signature was defined in a way that does not take into account particular characteristics of metastasis such as tissue tropism and growth kinetics. In a recent report comparing human primary breast tumors to distant metastases from various organs, the primary tumor showed extensive genetic similarity to the distant metastasis from the same patient, and a “supervised” method was unable to generate a classifier to distinguish primary tumors from metastases (15). These results are in line with the concepts of a poor-prognosis signature; however, because the metastasis samples were from various organs, the presence of site-specific metastasis genes could not be determined. Thus, the poor-prognosis signature may be composed of gene expression events acquired early during primary tumor development that function to endow tumor cells with baseline metastatic properties or that mark a particular cell phenotype that is liable to express metastatic functions. Indeed, MBA-MD-231 cells are derived from the pleural effusion of a patient with widespread metastatic disease, and all of the individual clones from this population that we analyzed show at least some level of metastatic activity.

Based on the identification of metastasis genes associated with osteolytic bone metastasis, our previous study proposed that in addition to the poor-prognosis signature, metastatic cells need to acquire a genetic “tool box,” or a set of genes that confer the functions necessary for efficient tissue-specific growth. The genes that make up this “tool box” would be regarded as metastasis-specific genes that are acquired through mutation or epigenetic changes. However, the classification of genes into this category would require a level of specificity such as tissue tropism. Our current study provides support for this requisite, as the expression of these genes strictly correlated with efficient bone metastasis and not with other recognizable aspects of metastatic activity. In addition, multi-dimensional scaling of genes that are differentially expressed across SCPs defines groups that correlate with primary tissue tropism, and our bone metastasis gene set overlaps with these differentially expressed genes. We expect that within these differentially expressed genes, a lung metastasis gene set will also exist (our unpublished observations). Thus, SCPs with different genetic profiles can exhibit marked differences in their ability to colonize and to grow exponentially in various metastatic sites. These results support the idea of the importance of productive tumor-stroma interactions that foster metastatic growth, consistent with Paget’s “seed and soil” hypothesis (16), or interactions such as those between tumor and vasculature that result in differential tissue arrest.



Some of the SCPs that we analyzed demonstrated the ability to grow effectively at more than one metastatic site. For example, SCP28 grew well in both the bone and the lung, SCP3 was metastatic to both lung and adrenal, and SCP2 exhibited both bone and adrenal tropism. In contrast, SCP46 was metastatic only to the bone. The multi-tropic properties of metastatic cells raise the possibility that metastatic cells from one site may spawn metastasis to another site. Because there are limited clinical situations in which single metastasis or oligometastasis is effectively treated by surgical excision, knowledge of whether metastatic cells are single or multi-tropic may be of important clinical relevance.

A metastatic cell must complete a series of sequential steps in order to successfully colonize and grow at a distant site. Our data suggests that the expression of a poor-prognosis signature can mark only a baseline ability to accomplish some of these steps. The signature may comprise genes related to the early oncogenic changes that drive primary tumor formation, but is absent in genes that dictate organ-specific metastatic activity. These additional metastasis genes provide the capability to become fully metastatic and confer properties such as organotropism. It is unclear whether these metastasis genes are acquired during the growth of the primary tumor or during colonization at a distant site (17). Indeed, our hierarchical clustering of a mixed cohort of primary breast tumors with a bone metastasis gene expression signature (12) did not allow robust classification of those tumors that gave rise to bone metastasis versus those that did not. Nonetheless, this signature was able to distinguish between primary breast carcinomas that preferentially metastasized to bone from those that preferentially metastasized elsewhere. These results suggest that the development of distant sites of metastasis in breast cancer patients is related to differences in primary tumor gene expression pattern that are discernible by our bone metastasis gene expression signature. A further enrichment of the list of bone metastasis genes may allow in the future accurate prediction of the bone metastasis tropism of breast cancer primary tumors.

## Methods

**Cell culture and retroviral gene transfer.** MDA-MB-231 cells were obtained from ATCC and were cultured in Dulbecco's modified Eagle's, high glucose supplemented with 10% FBS. SCPs were derived from MDA-MB-231 cells as described previously (12). The construction and retroviral gene transfer of the triple-modality reporter gene TGL has been described previously (14). In brief, 20 µg of the TGL reporter plasmid SFG<sub>NES</sub>TGL was transfected into the GPG29 packaging cell line with Lipofectamine 2000 (Invitrogen). Virus-containing supernatants were harvested between 72 and 96 hours, were filtered with a 0.45-µm syringe filter, and were used to infect MDA-MB-231 SCPs for 12–24 hours in the presence of 8 µg/ml of polybrene (Sigma-Aldrich). At 72 hours after infection, successful gene transfer was confirmed by visualization of GFP by fluorescence microscopy. These cells were enriched by fluorescence-activated cell sorting (FACS-Vantage; Becton Dickinson). Luciferase activity was confirmed in vitro by seeding of  $1 \times 10^5$  cells into a 24-well plate followed by the addition of 0.03 mg of D-Luciferin (Xenogen). Luciferase activity was measured with the IVIS Imaging System (Xenogen).

**Mouse xenografting.** For intracardiac injections, subconfluent cells were harvested, washed in PBS, and resuspended at a concentration of  $1 \times 10^6$  cells/ml. BALB/c nude mice (NCI) were anesthetized by intraperitoneal injection of ketamine (100 mg/kg) and xylazine (10 mg/kg) and were placed in the supine position. With a 26-gauge needle,  $1 \times 10^5$  cells were injected into the left ventricle via the third intercostal space after visualization of arterial blood flow into the syringe. For tail vein injections, unanesthetized mice

were warmed with a heat lamp to allow for venous dilation. Mice were then placed into a plastic restraining apparatus, and  $2 \times 10^5$  cells were injected via the lateral tail vein. Successful injections were confirmed by immediate BLI. All animal studies were performed in accordance with an IACUC-approved protocol at the Memorial Sloan-Kettering Cancer Center.

**BLI and analysis.** Anesthetized mice were injected retro-orbitally with 75 mg/kg of D-Luciferin (Xenogen) in PBS. Bioluminescence images were acquired with the IVIS Imaging System (Xenogen) at 2–5 minutes after injection. Acquisition times at the beginning of the time course started at 60 seconds and were reduced in accordance with signal strength to avoid saturation. Analysis was performed using LivingImage software (Xenogen) by measurement of photon flux (measured in photons/s/cm<sup>2</sup>/steradian) with a region of interest (ROI) drawn around the bioluminescence signal to be measured. For bone metastasis, an ROI was drawn around the major bioluminescence signal from the hindlimb, forelimb, or pelvis/sacrum. For lung metastasis, an ROI was used that encompassed the thorax of the mouse. For determination of the “fold increase” above background, average background measurements were obtained using the same ROI on a corresponding region from control mice. Data were divided by the average background measurement and were normalized to the signal obtained immediately after xenografting (day 0).

**Histology.** For whole-mount analysis, sacrificed mice were frozen in liquid nitrogen and were stored at  $-80^\circ\text{C}$ . Prior to frozen sectioning, tissue was embedded in M1 embedding media (Shandon). Sections 20 µm in thickness were mounted on slides and were fixed with 100% methanol for 30 seconds. GFP was visualized in these mounted sections using a fluorescence microscope. H&E staining was then performed on serial sections of interest. For immunohistochemistry for CD31, lungs were fixed in 4% paraformaldehyde overnight and were incubated in 30% sucrose for an additional 12–24 hours prior to cryosectioning. CD31 staining was performed with the Discovery AutoStainer (Ventana Medical Systems) and anti-CD31 (sc-1506; Santa Cruz Biotechnology) at a concentration of 1 µg/ml.

**DNA microarray analysis.** Methods for RNA extraction, labeling, and hybridization for DNA microarray analysis of the cell lines have been described previously (12). For the primary breast tumor data, tissues from primary breast cancers were obtained from therapeutic procedures performed as part of routine clinical management. Samples were “snap-frozen” in liquid nitrogen and were stored at  $-80^\circ\text{C}$ . Each sample was examined histologically with H&E-stained cryostat sections. Regions were manually dissected from the frozen block to provide a consistent tumor cell content of more than 70% in tissues used for analysis. All studies were conducted under protocols approved by the Memorial Sloan-Kettering Cancer Center Institutional Review Board. RNA was extracted from frozen tissues by homogenization in TRIzol reagent (GIBCO-BRL; Invitrogen Corp.) and was evaluated for integrity. Complementary DNA was synthesized from total RNA using a T7 promoter–tagged dT primer. RNA target was synthesized by in vitro transcription and was labeled with biotinylated nucleotides (Enzo Biochem). Labeled target was assessed by hybridization to Test3 arrays (Affymetrix).

All gene expression analysis was carried out using the Affymetrix U133A chip. Analysis of the poor-prognosis signature was performed using GeneSpring 6.1 (Silicon Genetics) with a list of genes from the 70 genes comprising the poor-prognosis signature that are present on the U133A chip. For multidimensional scaling and hierarchical clustering, Affymetrix data were imported into BRBArray Tools 3.1 (developed by Richard Simon and Amy Peng Lam; <http://linus.nci.nih.gov/BRB-ArrayTools.html>). Hierarchical clustering was performed using either Euclidean distance or Pearson correlation. Cluster reproducibility was reported as an R index (18). To obtain a list of genes that are broadly differentially expressed among the SCPs, we applied a filter to the 22,238 genes; this filter eliminated genes in



which expression levels differed by at least either 1.5-fold or twofold from the mean expression level in less than half of the data sets. An additional filter was applied to eliminate genes with an absent detection call in all of the datasets. The final filtered list comprised 1,267 genes (1.5-fold filter) or 286 genes (twofold filter). This list was used in both multidimensional scaling and hierarchical clustering. Other filtering criteria were also tested and gave comparable results.

**CXCR4 staining for flow cytometry.** Subconfluent cells were trypsinized and were washed twice in cold PBS. Phycoerythrin-conjugated anti-human CXCR4 (BD Pharmingen) or control IgG was incubated in FACS buffer (0.1% sodium azide and 1% bovine serum albumin in PBS) for 1 hour at 4°C. Cells were subsequently washed twice in PBS and, finally, were resuspended in FACS buffer. Cells were analyzed by flow cytometry using a BD FACSCalibur unit, and subsequent data analysis was done using FlowJo software.

### Acknowledgments

We are indebted to Juri Gelovani for invaluable discussions; Weiping Shu for expert technical assistance; Agnes Viale and Julia Zhao of the Genomics Core Facility for microarray data analysis; and Katia Manova and the staff of the Molecular Cytol-

ogy Core Facility for assistance with immunohistochemistry. A.J. Minn is a recipient of the Leonard B. Holman Research Pathway fellowship, and Y. Kang is the recipient of a postdoctoral fellowship from the Irvington Institute for Immunological Research. G.P. Gupta is supported by the NIH Medical Scientist Training Program grant GM07739 and a fellowship from the Katherine Beineke Foundation. J. Massagué is an Investigator of the Howard Hughes Medical Institute. This research is also supported by NIH grant P01-CA94060 to J. Massagué and US Army Medical Research grant DAMD17-02-0484 to W.L. Gerald.

Received for publication June 2, 2004, and accepted in revised form November 2, 2004.

Address correspondence to: Joan Massagué, Cell Biology Program, Box 116, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA. Phone: (212) 639-8975 Fax: (212) 717-3298; E-mail: j-massague@ski.mskcc.org.

Dilip D. Giri's present address is: Department of Pathology and Laboratory Medicine, Brown University, Providence, Rhode Island, USA.

1. Fidler, I.J. 2003. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat. Rev. Cancer*. **3**:453-458.
2. Chambers, A.F., Groom, A.C., and MacDonald, I.C. 2002. Dissemination and growth of cancer cells in metastatic sites. *Nat. Rev. Cancer*. **2**:563-572.
3. Tarin, D., et al. 1984. Clinicopathological observations on metastasis in man studied in patients treated with peritoneovenous shunts. *Br. Med. J. (Clin. Res. Ed.)*. **288**:749-751.
4. Tarin, D., Vass, A.C., Kettlewell, M.G., and Price, J.E. 1984. Absence of metastatic sequelae during long-term treatment of malignant ascites by peritoneovenous shunting. A clinico-pathological report. *Invasion Metastasis*. **4**:1-12.
5. Fidler, I.J. 1970. Metastasis: quantitative analysis of distribution and fate of tumor emboli labeled with 125 I-5-iodo-2'-deoxyuridine. *J. Natl. Cancer Inst.* **45**:773-782.
6. Cameron, M.D., et al. 2000. Temporal progression of metastasis in lung: cell survival, dormancy, and location dependence of metastatic inefficiency. *Cancer Res.* **60**:2541-2546.
7. van 't Veer, L.J., et al. 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. **415**:530-536.
8. van de Vijver, M.J., et al. 2002. A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* **347**:1999-2009.
9. Ramaswamy, S., Ross, K.N., Lander, E.S., and Golub, T.R. 2003. A molecular signature of metastasis in primary solid tumors. *Nat. Genet.* **33**:49-54.
10. Bernards, R., and Weinberg, R.A. 2002. A progression puzzle. *Nature*. **418**:823.
11. Fidler, I.J., and Kripke, M.L. 2003. Genomic analysis of primary tumors does not address the prevalence of metastatic cells in the population [correspondence]. *Nat. Genet.* **34**:23; author reply, 25.
12. Kang, Y., et al. 2003. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell*. **3**:537-549.
13. Cailleau, R., Young, R., Olive, M., and Reeves, W.J., Jr. 1974. Breast tumor cell lines from pleural effusions. *J. Natl. Cancer Inst.* **53**:661-674.
14. Ponomarev, V., et al. 2004. A novel triple-modality reporter gene for whole-body fluorescent, bioluminescent, and nuclear noninvasive imaging. *Eur. J. Nucl. Med. Mol. Imaging*. **31**:740-751.
15. Weigelt, B., et al. 2003. Gene expression profiles of primary breast tumors maintained in distant metastases. *Proc. Natl. Acad. Sci. U. S. A.* **100**:15901-15905.
16. Paget, S. 1989. The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev.* **8**:98-101.
17. Schmidt-Kittler, O., et al. 2003. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. *Proc. Natl. Acad. Sci. U. S. A.* **100**:7737-7742.
18. McShane, L.M., et al. 2002. Methods of assessing reproducibility of clustering patterns observed in analyses of microarray data. *Bioinformatics*. **18**:1462-1469.