

REVIEW

Signaling pathways in breast cancer metastasis - novel insights from functional genomics

Mario Andres Blanco¹ and Yibin Kang^{*1,2}

Abstract

The advent of genomic profiling technology has brought about revolutionary changes in our understanding of breast cancer metastasis. Gene expression analyses of primary tumors have been used to predict metastatic propensity with high accuracy. Animal models of metastasis additionally offer a platform to experimentally dissect components of the metastasis genetic program. Recent integrated studies have synergized clinical bioinformatic analyses with advanced experimental methodology and begun to uncover the identities and dynamics of signaling programs driving breast cancer metastasis. Such functional genomics studies hold great promise for understanding the genetic basis of metastasis and improving therapeutics for advanced diseases.

Introduction

Metastasis is the most poorly understood aspect of breast cancer, a disease that causes roughly half a million deaths each year worldwide and is the most common malignancy in women in the United States [1]. The field of metastasis research is at least a century old [2], and classical views hold that the metastatic phenotype is possessed by clonal variants within a tumor that happen to acquire the requisite mutations [1,3]. Progress in metastasis research, however, has stagnated because of a lack of effective tools to comprehensively understand the complex network of signaling pathways that drives the multistep process of the metastatic cascade [4,5]. The advent of genomic profiling technology has led to paradigm-shifting advances in the conceptual and mechanistic understanding of the metastatic process over the past decade. The early waves of clinical microarray studies found that gene expression profiles in

primary tumors could discriminate breast cancer patients with good prognosis from those with poor prognosis [6]. These works suggested that metastatic propensity may be selected for in the entire tumor and can be accurately assessed using bioinformatic approaches. Thus, an ensuing debate centered on whether there are any metastasis-specific genes, and, if so, how they could be identified [7,8]. Genomic profiling of clinical tumor samples alone, however, is fundamentally limited in providing functional insights, as it offers no method for testing hypotheses mechanistically. Though prognostically effective, such studies on their own have been unable to provide a satisfactory, functional understanding of the genetic and epigenetic underpinnings of metastatic progression.

In contrast, advances in animal models of metastasis have been applied to directly test the hypotheses generated by classical as well as modern genomic approaches to studying disease progression. Such studies have utilized the ability to create or isolate variants of breast cancer cell lines and quantitatively monitor their metastatic abilities in mice using various models of metastatic progression. These studies have provided critical insights into the mechanistic basis of metastatic progression and have suggested updated conceptual frameworks that have helped reconcile the differences between prior models of metastasis [4,9]. Considered alone, however, animal models of breast cancer progression will always have questionable applicability to human disease.

The combination of advances in bioinformatics approaches, animal model technology, and clinical dataset assembly has laid the groundwork for integrated studies to rapidly expand our knowledge of the breast cancer metastasis genetic program. While a mature understanding of this program has not yet been cemented, insights into the roles and functionality of metastasis-specific genes and pathways have recently emerged. Many studies have used powerful methodologies to define a gene expression program - such as a signal transduction pathway or physiological response program - and test its ability to significantly affect metastatic progression in the experimental setting, as well as test whether it shows elevated activity in large,

*Correspondence: ykang@princeton.edu

¹Department of Molecular Biology, Washington Road, LTL 255, Princeton University, Princeton, NJ 08544, USA

Full list of author information is available at the end of the article

clinical datasets and can thus be used for effective prognostication. In this review, our aim is not to exhaustively cover the understanding of any one such gene expression program in disease progression. Rather, we will instead focus our discussion on exemplary integrative studies that use functional genomics approaches to study the roles of various classical and novel signaling pathways in breast cancer metastasis.

Breast cancer subtypes - early portraits

Breast cancer has long been recognized as a heterogeneous disease that can be classified using a variety of characteristics and markers, such as histological grade, estrogen receptor, progesterone receptor and *HER2/ERBB2* status, and p53 mutational status. Around the turn of the century, nascent cDNA microarray technology made possible the first investigations into genome-wide expression patterns observed in breast cancer patients. The first wave of such breast cancer profiling studies performed microarray analyses on primary breast cancer tumor samples from small to medium sized patient cohorts [10-13]. In these works, unsupervised hierarchical clustering methodology was used to group patients according to patterns of gene expression, and the differentiating clusters of genes were scrutinized for biological meaning. In Perou and colleagues' landmark study [11], breast cancer patients were found to cluster into four discernable groups that, given immunohistochemical analyses and the identities of the differentiating genes, were annotated as basal-like, luminal-like, ERBB2⁺, and normal breast-like. These classifications were later validated in an independent cohort and it was further shown that the basal-like group patients had significantly worse prognoses than patients from other subgroups [12]. Notably, it was also observed here and in later work [14] that rare cases of matched primary tumors and metastatic lesions from the same patient always clustered together.

These initial works offered valuable insights into tumor biology and demonstrated that intrinsic gene expression patterns could be used in conjunction with histopathological characteristics for a far more sophisticated tumor classification system. However, they offered little information pertinent to the key question of what cohesive genetic programs underlie metastatic progression. In particular, the finding that matched primary and metastatic tumor samples cluster together could be interpreted in two quite different ways. One interpretation is that the genetic programs of primary tumors are fully maintained in metastatic lesions. An alternative explanation is that primary and secondary tumors are only more similar to one another than to tissue from another individual, with significant expression differences between primary and metastatic tumors still being possible.

Predicting metastasis using expression profiles - prognosis signatures

Given the difficulty in predicting metastatic progression based on histopathological and clinical criteria, most breast cancer patients receive adjuvant therapy. However, had they been left untreated, most of these patients would not have suffered from metastatic disease, rendering the therapy a cause of unnecessary suffering and expense. Recognizing the power of microarray approaches to discriminate breast cancer patients into clinically informative groups, several studies aimed at using clustering approaches to tackle the prognostication problem (Table 1). While methodologies varied, conclusions were similar: gene expression signatures can very effectively predict which patients survive and which succumb to metastatic disease, ostensibly supporting the view that metastatic propensity is selected for early in tumor progression.

The first prognosis signature study [15] used a supervised clustering approach to determine which genes could most effectively discriminate patients between those with good or poor prognosis. Such analysis led to the identification of a poor-prognosis signature consisting of 70 genes, many of which coded for proteins involved in processes such as cell cycle progression, invasion, and angiogenesis. Ultimately, the signature was able to correctly classify more than 80% of the patients as having good or poor prognosis, thus achieving a marked improvement in prognostication compared to standard methodologies. While this study used a relatively modest number of patient samples ($n = 78$), the efficacy of the prognostic signature was validated in a larger ($n = 295$), partially overlapping set of clinical samples [16].

While this original 70-gene signature has had direct clinical impact (commercialized as the MammaPrint, made available to patients in the United States in 2007), it is by no means the only effective prognosis signature. Using different patient cohorts, array platforms, and statistical methodologies, an alternative 76-gene signature was also reported [17], which provided comparably accurate prognostication to the 70-gene signature. However, while the 70- and 76-gene signatures consisted of similar classes and functional groups of genes [18], they had strikingly little actual overlap, with only three genes in common. This suggests that the given identities of signature genes are not nearly as important as the biological process of which they are but one representative. Cementing this point, it was shown that, using the same dataset and similar but non-identical methodology, many different 70-gene poor prognosis signatures of equivalent accuracy can be derived out of the original data [19]. Furthermore, a different type of approach compared expression data from primary tumors of various tissues to those of metastatic adenocarcinoma

Table 1. Gene expression signature analysis of breast cancer

Study	Signature	Size (genes)	Type	Validation
van't Veer <i>et al.</i> [15]	Poor prognosis	70	Classifier	Clinical
Wang <i>et al.</i> [17]	Poor prognosis	76	Classifier	Clinical
Ramaswamy <i>et al.</i> [20]	General metastasis	128	Classifier	Clinical
Chang <i>et al.</i> [34]	Wound healing	512	Physiological response	Clinical
Finak <i>et al.</i> [37]	Stromal	26	Classifier	Clinical
Troester <i>et al.</i> [38]	Stromal	155	Classifier	Clinical
Farmer <i>et al.</i> [39]	Stromal	50	Classifier, metagene	Clinical
Chi <i>et al.</i> [42]	Epithelial hypoxia	168	Physiological response	<i>In vitro</i> , clinical
Winter <i>et al.</i> [44]	Hypoxia	99	Classifier, metagene	Clinical
Buffa <i>et al.</i> [41]	Hypoxia	51	Classifier, metagene	Clinical
Hu <i>et al.</i> [43]	VEGF	13	Classifier	Clinical
Kang <i>et al.</i> [23]	Bone metastasis	102	Tissue tropism	Animal
Minn <i>et al.</i> [24]	Lung metastasis	95	Tissue tropism	Animal, clinical
Bos <i>et al.</i> [22]	Brain metastasis	243	Tissue tropism	Animal, clinical
Bild <i>et al.</i> [50]	Myc, E2F3, Ras, Src, β -catenin	248, 298, 348, 73, 98	Signaling pathway	Clinical, <i>in vitro</i>
Nguyen <i>et al.</i> [52]	WNT in lung cancer	81	Signaling pathway	Animal, clinical
Padua <i>et al.</i> [55]	TGF- β [53]	153	Signaling pathway	Animal, clinical
Zhang <i>et al.</i> [54]	Src [50]	172	Signaling pathway	Animal, clinical

TGF, transforming growth factor; VEGF, vascular endothelial growth factor.

lesions and found a discriminating 128-gene metastasis signature. This signature was furthermore shown to be active in a subset of primary tumors - with this subset having a significantly poorer prognosis than the rest of the patients [20].

These studies revealed strong implications to the metastasis genetic program debate. Specifically, they argued in different ways that metastatic propensity must indeed be captured within the phenotypes under selection at the primary tumor stage, otherwise no such prognostication would be possible. However, this argument suffers from conceptual difficulties - why would a metastatic phenotype be under selection in cells of the primary tumor? - and also contradicts the classic work of Fidler and others. Furthermore, the paucity of matched primary tumor and metastatic lesion pairs ($n = 2$ in [11] and $n = 8$ in [14]) renders these clinical studies unable to truly address the question of metastasis-specific genetic events. While these difficulties could perhaps be considered academic, an issue of more immediate concern is that the functional interchangeability and lack of overlap between these signatures has resulted in the proposal of few, if any, protein products as potential therapeutic targets for blocking metastatic progression. Functional - rather than purely bioinformatic - studies are therefore required to give further understanding to the metastasis genetic program.

Animal models of metastatic progression

To test whether there could be genes and signaling pathways whose activation specifically affects metastatic progression, experimental animal models of breast cancer progression have been utilized. Compared to clinical profiling studies, animal models of metastasis have several critical advantages, which stem largely from the ability to isolate and characterize both primary tumors and distant metastatic lesions, and to manipulate the expression levels of one to several genes at a time to directly test their roles in disease progression. Such methodology has profoundly advanced the understanding of how, on a mechanistic level, the metastatic program is executed, and has also provided further insights into the complexity of metastatic disease.

Advancing Fidler's classic work [1,21], several studies have used *in vivo* selection approaches to ultimately determine which genes drive metastasis to which organs, with the breast-to-bone, -lung, and -brain tropisms each having been investigated to date (Table 1) [22-24]. Such investigations have involved experimental metastasis assay xenografts of weakly metastatic cells followed by isolation of secondary lesions in the tissue of interest. Microarray-based comparisons of the parental lines to aggressive, organ-tropic sublines have yielded the signatures of genes under selection during the late stage metastatic program of interest. Juxtaposing the breast-to-bone

and -lung studies, several findings are particularly informative to the conceptual framework of metastasis. Firstly, the bone and lung metastasis programs are distinct. While the bone and lung metastasis signatures (BMS and LMS) contain 102 and 95 genes, respectively, only six genes are common to both. Secondly, bone metastasis genes appear to be particular to bone microenvironment functionality, whereas lung metastasis genes have less obvious roles in the lung microenvironment and appear instead to facilitate general aggressive growth and invasiveness. BMS genes such as *CXCR4*, *CTGF*, and *IL-11* have been shown to play key roles in the 'vicious cycle' [25] of cancer cell-driven osteolysis [26-28], while LMS genes such as *IDI1*, *MMP1* and 2, and *SPARC* have been shown to promote the general phenotypes of growth, invasion, and adhesion, respectively [29-31]. Unsurprisingly, then, the bone metastasis gene expression program has little overlap with the 70-gene poor prognosis signature, while the LMS has significant overlap with multiple poor prognosis signatures and can indeed be used for effective prognostication.

The organ-specific metastasis studies have also laid the groundwork for advanced, mechanistic studies to further dissect the processes of breast cancer progression. Studies have assessed, for example, the physiological role of the metalloproteinases *MMP1* and *ADAMTS1* in breast cancer bone metastasis and uncovered a role for epidermal growth factor receptor inhibitors in targeting the reactive stroma in osteolytic metastasis [32]. In lung metastasis, the combinatorial effects of *COX2*, *EREG*, and *MMP1* and 2 were shown to promote primary tumor angiogenesis and extravasation of metastatic cells from the lung capillaries [33]. Here it was also found that pharmacological inhibition of these genes with targeting small molecule inhibitors ablated these phenotypes in aggressive lung metastasis breast cancer models.

Taken together, the results from these landmark clinical and experimental animal studies indicate that both sides of the metastasis genes debate are partially correct. On one hand, some degree of metastatic propensity is under selection at the primary tumor stage, as the prognostication studies would have failed were this not the case. On the other hand, some other components of the metastatic program must arise later, otherwise the animal studies would not have succeeded in finding such striking differences between primary and secondary, organ-specific lesions. Thus, it appears that, while there is indeed an early (primary tumor stage) metastatic program under selection, it should appropriately be considered necessary but not sufficient for distant metastasis to occur. More importantly, the tumors of metastatic disease have an at least moderately different genetic makeup to those of the primary tumor, and effective treatments will likely need to target the factors critical to

microenvironment-specific tumor survival. In short, the functional power of experimental models must be synergized with the relevance of clinical datasets to appropriately explore the genes and pathways that define and undergird breast cancer metastatic progression.

Integrated studies to understand breast cancer metastasis signaling

Central to the advances in understanding of metastasis-specific gene expression changes was the aforementioned recognition that, while the individual genes of various prognosis signatures may be interchangeable, the signaling pathways they represent are consistent. Pathway-level analyses therefore have several advantages over both single gene and gene expression profile studies. Compared to single gene studies, they can take advantage of the statistical power of gene sets, in which the activity readout is not dependent on the expression of any single gene but, rather, is determined by the concerted enrichment of the group overall. And in comparison to profiles, they test the activity of genes involved in a biologically defined (and therefore experimentally testable) phenotypic process.

Several studies have looked at signal transduction pathways or sets of genes of similar function as the unit of analysis to study metastatic progression and prognostication (Table 1). One approach started with the long standing observation that the physiology of the tumor-stroma interface appears to have much in common with that of a wound that is in the healing process, given the potent proliferative, invasive, and angiogenic stimulations in both contexts. Using a 512-gene 'core serum response' (CSR) signature as representative of the wound healing gene expression program, Chang *et al.* [34] found that CSR-active patients had significantly worse prognoses than CSR-inactive patients and were largely characterized as belonging to the 'basal-like' breast cancer subtype. Furthermore, several CSR genes and proteins involved in cell-cell communication (*ESDN* and *SDR1*) and extracellular matrix remodeling (*LOXL2*, *PLOD2*, and *PLAUR*) were shown to be upregulated in invasive ductal carcinoma samples by tissue microarray analysis. Thus, the CSR can be firstly considered a distinct prognosis classifier with similar power to previous signatures. However, in being defined by a specific physiological process (rather than general good versus poor prognosis), it is far more biologically coherent than previous signatures. In particular, such work has given rise to the concept of the 'reactive stroma' [35] as a crucial component in metastatic progression. As metastasis is defined by invasion into foreign tissue, stromal components such as cancer-associated fibroblasts have been shown to undergo inflammatory-like responses that help mediate tumor progression [36]. Furthermore, gene expression profiles

characteristic of tumor-associated stromal tissue can successively classify and prognosticate patients into appropriate subgroups and outcomes [37,38]. Interestingly, stromal signatures can not only distinguish good from poor prognosis, but also have been shown to predict response of breast cancer patients to chemotherapeutic treatment [39].

Other physiological responses have been used as the basis for hypothesis-driven investigations into pathways that could be promoting metastatic progression. The hypoxia response is one such physiological program that is thought to enable metastatic invasion into the circulatory system. Under conditions of low oxygen (hypoxia), which are common in large tumors, the hypoxia-inducible factor-1 α transcription factor subunit is stabilized and activates a pro-angiogenic gene expression program that results in enhancement of local vascularization. The angiogenic response is thought to play a dual role in tumor progression. While first functioning to supply the growing (and starving) tumor with oxygen and other essential nutrients, angiogenesis also aids in tumor metastasis by providing entryways for primary tumor cells into the circulatory system [40]. Analyses in both breast and head and neck cancer [41-44] have described hypoxic responses and used them to prognosticate patient groups across a variety of cancer types. Response signature derivation methods varied considerably, as did signature size and gene identity (Table 1). While one approach used *in vitro* hypoxia-induced genes for signature building [42], others started with small sets of known hypoxia-response genes and built metagene networks off of them for prognostication efforts [41,44]. Despite methodology differences, all approaches had high rates of success in patient prognostication in various cancers. Interestingly, a vascular endothelial growth factor-based signature was found to be especially active in distant metastases compared to primary tumors or their local metastases, supporting the hypothesis that primary tumors and distant metastases do harbor significant gene expression differences despite overall clustering-based similarities [43]. Experimental analyses have further investigated the role of the hypoxic response in animal models of organ-specific breast cancer metastasis, finding that while bone and lung metastases utilize different hypoxic gene response programs and have different dependence on angiogenic response, both pathological conditions are highly responsive to hypoxia inhibitor treatments [45].

Although pathway-based analyses highlight the functional effects of concerted gene expression changes, they typically shed little light on one of the key questions in metastatic progression, which is how to find the underlying genetic mutations that drive these large-scale expression program changes. However, by treating

functional or pathway-based expression profiles as a phenotype that can be used for linkage analyses, methodology has been developed to find driver mutations in metastatic progression. Focusing on the CSR signature, a genomic method termed 'SLAMS' (stepwise linkage analysis of microarray signatures) was designed to find candidate master regulators within cytogenetic abnormalities linked to CSR activity [46]. A large region of genomic amplification on chromosome 8q was found to be most strongly linked to activation of the CSR profile, and mechanistic work indicated that overexpression of resident 8q genes *Myc* and *COP9* was sufficient to activate the CSR signature.

Such approaches have begun to bridge the gap between prognosis signatures and the underlying, driver mutations that activate them. While the *Myc* oncogene has long been known to be crucial for tumor progression [47], its role in promoting metastatic progression has remained unclear. The SLAMS approach highlighted *Myc* in a novel context as the potential activator of a metastasis signaling program, but the actual functional contribution of *Myc* transcriptional activity to tumor progression was not investigated. However, later work has directly tested the role of *Myc* signaling in metastatic progression using a variety of model systems. Building off of the SLAMS approach, Wolfer *et al.* [48] searched for potential regulators of multiple poor prognosis signatures using the MCF7 breast cancer cell line as a testing platform. Through a variety of informatics approaches, *Myc* activity was predicted and then validated to activate many (10 to 40%) of the genes in all of the poor prognosis signatures that were tested. Crucially, this cell line-based work was validated *in vivo* by demonstrating that stable knockdown of *Myc* in late stage MDA-MB-231 cells led to a dramatic reduction in lung metastasis burden without significantly affecting the growth of the primary tumor.

While work based on cell lines and xenograft mouse metastasis models has advantages in terms of tractability and efficiency, transgenic mouse models are often considered more biologically relevant. To study *Myc*-based profile induction and breast cancer progression, tumor subtypes were investigated in a mouse mammary tumor virus (MMTV)-*Myc* model of tumorigenesis [49]. Here it was observed that the MMTV-*Myc* transgene induced a striking variety of histological subtypes, with the 'epithelial to mesenchymal transition/squamous' type predicted to have poor prognosis by an independently derived metastasis signature. Accordingly, mice with tumors of this subtype indeed had far greater incidences of lung metastases than those of other subtypes. Furthermore, the epithelial to mesenchymal transition/squamous signature was found to be elevated in 'triple negative' (estrogen receptor, progesterone receptor, and Her2 negative) poor prognosis patients in a clinical

analysis, thus providing more evidence for *Myc* oncogene-based signaling in promoting metastatic progression.

The importance of the *Myc* pathway in metastatic progression underscores the concept that some degree of the metastatic signaling program could be driven by classic oncogenes or other well-known signaling cascades that can adapt to promote metastasis-specific gene expression changes. To aid in pathway-based analyses, key signatures have been derived for assessing the pathway activity of *Src*, *H-Ras*, *E2F3*, *Myc*, β -catenin [50], *TCF/Wnt* [51,52], and transforming growth factor (*TGF*)- β [53] by activating the pathway chemically or genetically and performing microarray profiling experiments. A key method for utilizing the power of these signatures has involved interrogating the activity of such pathways in clinical breast cancer datasets stratified by a phenotype of interest and then testing the effects of pathway activation on the relevant metastatic phenotype in the appropriate *in vivo* breast cancer progression model. In this fashion, powerful studies found that *Src* signaling mediates long-term survival (latency) and eventual outgrowth of clinical and experimental bone metastasis [54], whereas *TGF*- β activity promotes metastatic dissemination to the lung, rather than bone tissue [55]. Furthermore, the intersection of the *TGF*- β and lung metastasis signatures was effectively used to narrow the list of candidate dissemination mediators and effectively identify *ANGPTL4* as a novel, *TGF*- β -responsive lung metastasis gene. Pathway activity studies have been undertaken in other cancer types, with implications for breast cancer resulting from included analyses. For example, efforts to uncover signaling activity governing metastasis from lung carcinomas found that a lung *WNT* signaling program was functional in promoting metastasis from lung cancer lines and prognostic of lung cancer patients in clinical databases [52]. Notably, bioinformatic analyses indicated that the lung *WNT* signaling program was not successful in prognosticating breast cancer. By extension, then, *WNT* signaling may be considered of lesser importance in breast cancer progression, thus narrowing the focus of breast cancer metastasis to aforementioned candidate pathways such as *Src* and *TGF*- β .

Novel pathways in metastasis

Clearly, well-known signaling pathways, such as the *Myc*, *TGF*- β , and *Src* pathways, are driving some components of breast cancer metastasis progression. However, given the complexity of the metastatic program, it would not be surprising to find novel master regulators or key mediators of metastatic progression. One study used a hypothesis-driven approach to investigate the role of *SATB1*, a so-called 'genome organizer' that localizes to

heterochromatin and recruits chromatin-remodeling enzymes and transcription factors to induce large scale transcriptional changes [56]. Cell line and large tissue array analyses found *SATB1* to be strongly correlated to poor prognosis. *In vivo* analyses showed that *SATB1* was both necessary and sufficient to promote both lung metastasis and primary tumor progression. Microarray analysis of *SATB1* signaling indicated remarkably penetrant gene expression changes, with significant regulation of multiple pertinent signatures, such as the 70-gene poor prognosis signature, and both the *BMS* and *LMS*. Curiously, despite the striking results, *SATB1* signaling has not been linked to a known signal transduction pathway, and has also been shown to not promote the initially reported phenotypes [57]. Conceptually, its role as a general 'genome organizer' is difficult to reconcile with the induction of such phenotypically specific and potent gene expression changes. Thus, *SATB1* presents a challenge to understanding metastasis signaling and suggests that large scale epigenetic regulators may play an important, yet underappreciated, role in tumor progression. Future studies will be required to explain the mystery.

Another recent analysis, much like the *SLAMS* approach, started with the motivating concept that mutations driving metastatic progression should be identifiable by their residence in areas of conserved cytogenetic abnormalities in poor prognosis tumor specimens. Using a computational approach termed 'ACE' (analysis of copy number abnormalities by expression data), Hu *et al.* [58] bypassed direct assessments of cytogenetic abnormalities and inferred them via interrogating clinical microarray expression data of genes according to chromosomal location. Combining the clinical expression datasets from three previous studies [15-17], the ACE approach identified a conserved window of amplification on chromosome 8q22 in poor prognosis breast cancer patients. A combination of *in vitro* and *in vivo* analyses led to the hypothesis that the gene *Metadherin* (*MTDH*) was the functional target of this amplification, and *in vivo* xenograft experiments strongly supported this view. Interestingly, further informatics analyses using the *NCI-60* database indicated that *MTDH* was also strongly associated with chemoresistance. This second phenotype was experimentally validated, highlighting *MTDH* as an example of a rare class of dual-functional genes that are active in two aspects of cancer progression. While *MTDH* was shown to affect the expression of many genes of relevance to the metastatic and chemoresistance phenotypes, the key signaling pathways upstream and downstream of *MTDH* remained elusive. Several other studies have recently explored *MTDH* signaling, with the *NF- κ B*, phosphoinositide 3-kinase-*AKT*, *Ha-Ras*, *FOXO3a*, and *Myc* pathways [59-62] all having been suggested as

activating, mediating, or augmenting MTDH functionality. Thus, MTDH represents another novel mediator of malignant breast cancer progression with exciting, yet inconclusive, effects on breast cancer signaling.

Conclusion

Clinical profiling studies, experimental models of disease progression, and especially the combination of both have greatly advanced the understanding of breast cancer metastasis since the turn of the century. However, despite the highlighted advances, metastasis remains a poorly understood biological process. Multiple genes and signaling pathways have been shown to have the ability to influence metastatic progression, but few universal signaling events have been established as truly essential to the metastatic program. Confounding issues - clinical, experimental, and technical - continue to pose problems for a firm understanding of the underlying biology. Clinical datasets, for example, very rarely contain expression data from metastatic lesions that can be matched to their corresponding primary tumors. And experimental studies, of course, always come with extensive assumptions that can never truly be shown to be valid. Examples include the assumptions that the effect of the immune system (for xenograft studies), the effect of genetic diversity in the host, and the differences between mice and humans are small if not negligible. Furthermore, technical challenges range from trivial to dramatic. With microarray platforms continuing to evolve, signatures from older studies are becoming more difficult to interpret in light of newer studies with more probes and different chemistries. Additionally, tumor specimens (unless obtained via laser capture microdissection) are typically in fact a mixture of tumor cells and stromal cells, making it difficult to determine whether a gene of interest is being expressed by the tumor, stroma, or both.

These challenges notwithstanding, the metastasis field is progressing rapidly and will continue to do so if it can take advantage of new methodologies, technologies, and conceptual ingenuities. Notably, the unabated 'omics' revolution is now offering avenues for several new approaches in metastasis prognostication and mechanistic hypothesis building. For example, several groups are utilizing next generation sequencing technology for whole-genome sequencing of primary tumors and matched metastases [63,64]. Such analyses will surely advance the ability to identify metastasis-specific driver mutations so long as the 'data overload' problem does not cripple the analyses. Additionally, proteomics-based approaches are advancing at a rapid pace as mass spectrometry technologies continue to evolve. While current sensitivity levels may make whole-cell proteomics approaches cumbersome, subcellular fractions are now being sequenced at the protein level with success [65,66].

This article is part of a review series on *New pathways of metastasis*, edited by Lewis Chodosh. Other articles in the series can be found online at http://breast-cancer-research.com/series/metastasis_pathway.

Appropriate utilization of omics-level DNA-, RNA-, and protein-based approaches can only be expected to synergize in unraveling the mystery of the breast cancer metastasis genetic program.

Abbreviations

ACE, analysis of copy number abnormalities by expression data; BMS, bone metastasis signature; CSR, core serum response; LMS, lung metastasis signature; MMTV, mouse mammary tumor virus; NF, nuclear factor; SLAMS, stepwise linkage analysis of microarray signatures; TGF, transforming growth factor.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We apologize to the authors of many excellent studies whose works we could not cite due to space limitations. Funding support of this work includes an NRSA pre-doctoral fellowship from the National Institutes of Health (to MAB) and grants from Champalimaud Foundation, Brewster Foundation, DOD (BC051647), and NIH (5R01CA134519 and 1R01CA141062) (to YK).

Author details

¹Department of Molecular Biology, Washington Road, LTL 255, Princeton University, Princeton, NJ 08544, USA. ²Breast Cancer Program, Cancer Institute of New Jersey, New Brunswick, NJ 08903, USA.

Published: 14 March 2011

References

1. Fidler IJ, Kripke ML: **Metastasis results from preexisting variant cells within a malignant tumor.** *Science* 1977, **197**:893-895.
2. Paget S: **The distribution of secondary growths in cancer of the breast.** *Lancet* 1889, **133**:571-573.
3. Fearon ER, Vogelstein B: **A genetic model for colorectal tumorigenesis.** *Cell* 1990, **61**:759-767.
4. Gupta GP, Massague J: **Cancer metastasis: building a framework.** *Cell* 2006, **127**:679-695.
5. Kang Y: **Functional genomic analysis of cancer metastasis: biologic insights and clinical implications.** *Expert Rev Mol Diagn* 2005, **5**:385-395.
6. Weigelt B, Peterse JL, van't Veer LJ: **Breast cancer metastasis: markers and models.** *Nat Rev Cancer* 2005, **5**:591-602.
7. Bernards R, Weinberg RA: **A progression puzzle.** *Nature* 2002, **418**:823.
8. Hynes RO: **Metastatic potential: generic predisposition of the primary tumor or rare, metastatic variants - or both?** *Cell* 2003, **113**:821-823.
9. Kang Y: **New tricks against an old foe: molecular dissection of metastasis tissue tropism in breast cancer.** *Breast Dis* 2006, **26**:129-138.
10. Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC, Lashkari D, Shalon D, Brown PO, Botstein D: **Distinctive gene expression patterns in human mammary epithelial cells and breast cancers.** *Proc Natl Acad Sci U S A* 1999, **96**:9212-9217.
11. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale AL, Brown PO, Botstein D: **Molecular portraits of human breast tumours.** *Nature* 2000, **406**:747-752.
12. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matise JC, Brown PO, Botstein D, Eystein Lønning P, Børresen-Dale AL: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci U S A* 2001, **98**:10869-10874.
13. Sørlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H,

- Pesich R, Geisler S, Demeter J, Perou CM, Lønning PE, Brown PO, Børresen-Dale AL, Botstein D: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc Natl Acad Sci U S A* 2003, **100**:8418-8423.
14. Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ: **Gene expression profiles of primary breast tumors maintained in distant metastases.** *Proc Natl Acad Sci U S A* 2003, **100**:15901-15905.
15. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH: **Gene expression profiling predicts clinical outcome of breast cancer.** *Nature* 2002, **415**:530-536.
16. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R: **A gene-expression signature as a predictor of survival in breast cancer.** *N Engl J Med* 2002, **347**:1999-2009.
17. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer van Gelder ME, Yu J, Jatkoe T, Berns EM, Atkins D, Foekens JA: **Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer.** *Lancet* 2005, **365**:671-679.
18. Solé X, Bonifaci N, López-Bigas N, Berenguer A, Hernández P, Reina O, Maxwell CA, Aguilar H, Urruticoechea A, de Sanjosé S, Comellas F, Capellá G, Moreno V, Pujana MA: **Biological convergence of cancer signatures.** *PLoS One* 2009, **4**:e4544.
19. Ein-Dor L, Kela I, Getz G, Givol D, Domany E: **Outcome signature genes in breast cancer: is there a unique set?** *Bioinformatics* 2005, **21**:171-178.
20. Ramaswamy S, Ross KN, Lander ES, Golub TR: **A molecular signature of metastasis in primary solid tumors.** *Nat Genet* 2003, **33**:49-54.
21. Fidler IJ, Nicolson GL: **Organ selectivity for implantation survival and growth of B16 melanoma variant tumor lines.** *J Natl Cancer Inst* 1976, **57**:1199-1202.
22. Bos PD, Zhang XH, Nadal C, Shu W, Gomis RR, Nguyen DX, Minn AJ, van de Vijver MJ, Gerald WL, Foekens JA, Massague J: **Genes that mediate breast cancer metastasis to the brain.** *Nature* 2009, **459**:1005-1009.
23. Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Cordon-Cardo C, Guise TA, Massague J: **A multitgenic program mediating breast cancer metastasis to bone.** *Cancer Cell* 2003, **3**:537-549.
24. Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, Viale A, Olshen AB, Gerald WL, Massague J: **Genes that mediate breast cancer metastasis to lung.** *Nature* 2005, **436**:518-524.
25. Mundy GR: **Metastasis to bone: causes, consequences and therapeutic opportunities.** *Nat Rev Cancer* 2002, **2**:584-593.
26. Richert MM, Vaidya KS, Mills CN, Wong D, Korz W, Hurst DR, Welch DR: **Inhibition of CXCR4 by CTCE-9908 inhibits breast cancer metastasis to lung and bone.** *Oncol Rep* 2009, **21**:761-767.
27. Shimo T, Kubota S, Yoshioka N, Ibaragi S, Isovva S, Eguchi T, Sasaki A, Takigawa M: **Pathogenic role of connective tissue growth factor (CTGF/CCN2) in osteolytic metastasis of breast cancer.** *J Bone Miner Res* 2006, **21**:1045-1059.
28. Singh B, Berry JA, Shohar A, Lucci A: **COX-2 induces IL-11 production in human breast cancer cells.** *J Surg Res* 2006, **131**:267-275.
29. Swarbrick A, Akerfeldt MC, Lee CS, Sergio CM, Caldon CE, Hunter LJ, Sutherland RL, Musgrove EA: **Regulation of cyclin expression and cell cycle progression in breast epithelial cells by the helix-loop-helix protein Id1.** *Oncogene* 2005, **24**:381-389.
30. Kawamata H, Kameyama S, Kawai K, Tanaka Y, Nan L, Barch DH, Stetler-Stevenson WG, Oyasu R: **Marked acceleration of the metastatic phenotype of a rat bladder carcinoma cell line by the expression of human gelatinase A.** *Int J Cancer* 1995, **63**:568-575.
31. Everitt EA, Sage EH: **Expression of SPARC is correlated with altered morphologies in transfected F9 embryonal carcinoma cells.** *Exp Cell Res* 1992, **199**:134-146.
32. Lu X, Wang Q, Hu G, Van Poznak C, Fleisher M, Reiss M, Massague J, Kang Y: **ADAMTS1 and MMP1 proteolytically engage EGF-like ligands in an osteolytic signaling cascade for bone metastasis.** *Genes Dev* 2009, **23**:1882-1894.
33. Gupta GP, Nguyen DX, Chiang AC, Bos PD, Kim JY, Nadal C, Gomis RR, Manova-Todorova K, Massague J: **Mediators of vascular remodelling co-opted for sequential steps in lung metastasis.** *Nature* 2007, **446**:765-770.
34. Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, Chi JT, van de Rijn M, Botstein D, Brown PO: **Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds.** *PLoS Biol* 2004, **2**:E7.
35. Angeli F, Koumakis G, Chen MC, Kumar S, Delinassios JG: **Role of stromal fibroblasts in cancer: promoting or impeding?** *Tumour Biol* 2009, **30**:109-120.
36. Erez N, Truitt M, Olson P, Arron ST, Hanahan D: **Cancer-associated fibroblasts are activated in incipient neoplasia to orchestrate tumor-promoting inflammation in an NF-kappaB-dependent manner.** *Cancer Cell* 2010, **17**:135-147.
37. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, Chen H, Omeroglu G, Meterissian S, Omeroglu A, Hallett M, Park M: **Stromal gene expression predicts clinical outcome in breast cancer.** *Nat Med* 2008, **14**:518-527.
38. Troester MA, Lee MH, Carter M, Fan C, Cowan DW, Perez ER, Pirone JR, Perou CM, Jerry DJ, Schneider SS: **Activation of host wound responses in breast cancer microenvironment.** *Clin Cancer Res* 2009, **15**:7020-7028.
39. Farmer P, Bonnefoi H, Anderle P, Cameron D, Wirapati P, Becette V, André S, Piccart M, Campone M, Brain E, Macgrogan G, Petit T, Jassem J, Bibeau F, Blot E, Bogaerts J, Aguet M, Bergh J, Iggo R, Delorenzi M: **A stroma-related gene signature predicts resistance to neoadjuvant chemotherapy in breast cancer.** *Nat Med* 2009, **15**:68-74.
40. Folkman J: **Fundamental concepts of the angiogenic process.** *Curr Mol Med* 2003, **3**:643-651.
41. Buffa FM, Harris AL, West CM, Miller CJ: **Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene.** *Br J Cancer* 2010, **102**:428-435.
42. Chi JT, Wang Z, Nuyten DS, Rodriguez EH, Schaner ME, Salim A, Wang Y, Kristensen GB, Helland A, Børresen-Dale AL, Giaccia A, Longaker MT, Hastie T, Yang GP, van de Vijver MJ, Brown PO: **Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers.** *PLoS Med* 2006, **3**:e47.
43. Hu Z, Fan C, Livasy C, He X, Oh DS, Ewend MG, Carey LA, Subramanian S, West R, Ikpatt F, Olopade OI, van de Rijn M, Perou CM: **A compact VEGF signature associated with distant metastases and poor outcomes.** *BMC Med* 2009, **7**:9.
44. 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: **Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers.** *Cancer Res* 2007, **67**:3441-3449.
45. Lu X, Yan CH, Yuan M, Wei Y, Hu G, Kang Y: **In vivo dynamics and distinct functions of hypoxia in primary tumor growth and organotropic metastasis of breast cancer.** *Cancer Res* 2010, **70**:3905-3914.
46. Adler AS, Lin M, Horlings H, Nuyten DS, van de Vijver MJ, Chang HY: **Genetic regulators of large-scale transcriptional signatures in cancer.** *Nat Genet* 2006, **38**:421-430.
47. Sinn E, Muller W, Pattengale P, Tepler I, Wallace R, Leder P: **Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo.** *Cell* 1987, **49**:465-475.
48. Wolfer A, Wittner BS, Irimia D, Flavin RJ, Lupien M, Gunawardane RN, Meyer CA, Lightcap ES, Tamayo P, Mesirov JP, Liu XS, Shioda T, Toner M, Loda M, Brown M, Brugge JS, Ramaswamy S: **MYC regulation of a "poor-prognosis" metastatic cancer cell state.** *Proc Natl Acad Sci U S A* 2010, **107**:3698-3703.
49. Andrechek ER, Cardiff RD, Chang JT, Gatzka ML, Acharya CR, Potti A, Nevins JR: **Genetic heterogeneity of Myc-induced mammary tumors reflecting diverse phenotypes including metastatic potential.** *Proc Natl Acad Sci U S A* 2009, **106**:16387-16392.
50. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, Joshi MB, Harpole D, Lancaster JM, Berchuck A, Olson JA Jr, Marks JR, Dressman HK, West M, Nevins JR: **Oncogenic pathway signatures in human cancers as a guide to targeted therapies.** *Nature* 2006, **439**:353-357.
51. van de Wetering M, Sancho E, Verweij C, de Lau W, Oving I, Hurlstone A, van der Horn K, Batlle E, Coudreuse D, Haramis AP, Tjon-Pon-Fong M, Moerer P, van den Born M, Soete G, Pals S, Eilers M, Medema R, Clevers H: **The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells.** *Cell* 2002, **111**:241-250.
52. Nguyen DX, Chiang AC, Zhang XH, Kim JY, Kris MG, Ladanyi M, Gerald WL, Massague J: **WNT/TCF signaling through LEF1 and HOBX9 mediates lung adenocarcinoma metastasis.** *Cell* 2009, **138**:51-62.
53. Kang Y, Chen CR, Massague J: **A self-enabling TGFbeta response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells.** *Mol Cell* 2003, **11**:915-926.
54. Zhang XH, Wang Q, Gerald W, Hudis CA, Norton L, Smid M, Foekens JA,

- Massague J: **Latent bone metastasis in breast cancer tied to Src-dependent survival signals.** *Cancer Cell* 2009, **16**:67-78.
55. Padua D, Zhang XH, Wang Q, Nadal C, Gerald WL, Gomis RR, Massague J: **TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4.** *Cell* 2008, **133**:66-77.
56. Han HJ, Russo J, Kohwi Y, Kohwi-Shigematsu T: **SATB1 reprogrammes gene expression to promote breast tumour growth and metastasis.** *Nature* 2008, **452**:187-193.
57. Iorns E, Hnatyszyn HJ, Seo P, Clarke J, Ward T, Lippman M: **The role of SATB1 in breast cancer pathogenesis.** *J Natl Cancer Inst* 2010, **102**:1284-1296.
58. Hu G, Chong RA, Yang Q, Wei Y, Blanco MA, Li F, Reiss M, Au JL, Haffty BG, Kang Y: **MTDH activation by 8q22 genomic gain promotes chemoresistance and metastasis of poor-prognosis breast cancer.** *Cancer Cell* 2009, **15**:9-20.
59. Emdad L, Sarkar D, Su ZZ, Randolph A, Boukerche H, Valerie K, Fisher PB: **Activation of the nuclear factor kappaB pathway by astrocyte elevated gene-1: implications for tumor progression and metastasis.** *Cancer Res* 2006, **66**:1509-1516.
60. Lee SG, Su ZZ, Emdad L, Sarkar D, Fisher PB: **Astrocyte elevated gene-1 (AEG-1) is a target gene of oncogenic Ha-ras requiring phosphatidylinositol 3-kinase and c-Myc.** *Proc Natl Acad Sci U S A* 2006, **103**:17390-17395.
61. Emdad L, Lee SG, Su ZZ, Jeon HY, Boukerche H, Sarkar D, Fisher PB: **Astrocyte elevated gene-1 (AEG-1) functions as an oncogene and regulates angiogenesis.** *Proc Natl Acad Sci U S A* 2009, **106**:21300-21305.
62. Lee SG, Su ZZ, Emdad L, Sarkar D, Franke TF, Fisher PB: **Astrocyte elevated gene-1 activates cell survival pathways through PI3K-Akt signaling.** *Oncogene* 2008, **27**:1114-1121.
63. Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, Burleigh A, Delaney A, Gelmon K, Guliiany R, Senz J, Steidl C, Holt RA, Jones S, Sun M, Leung G, Moore R, Severson T, Taylor GA, Teschendorff AE, Tse K, Turashvili G, Varhol R, Warren RL, Watson P, Zhao Y, Caldas C, Huntsman D, Hirst M, Marra MA, Aparicio S: **Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution.** *Nature* 2009, **461**:809-813.
64. Ding L, Ellis MJ, Li S, Larson DE, Chen K, Wallis JW, Harris CC, McLellan MD, Fulton RS, Fulton LL, Abbott RM, Hoog J, Dooling DJ, Koboldt DC, Schmidt H, Kalicki J, Zhang Q, Chen L, Lin L, Wendl MC, McMichael JF, Magrini VJ, Cook L, McGrath SD, Vickery TL, Appelbaum E, Deschryver K, Davies S, Guintoli T, Lin L, et al: **Genome remodelling in a basal-like breast cancer metastasis and xenograft.** *Nature* 2010, **464**:999-1005.
65. Goodison S, Urquidi V: **Breast tumor metastasis: analysis via proteomic profiling.** *Expert Rev Proteomics* 2008, **5**:457-467.
66. Roesli C, Borgia B, Schliemann C, Gunthert M, Wunderli-Allenspach H, Giavazzi R, Neri D: **Comparative analysis of the membrane proteome of closely related metastatic and nonmetastatic tumor cells.** *Cancer Res* 2009, **69**:5406-5414.

doi:10.1186/bcr2831

Cite this article as: Blanco MA, Kang Y: **Signaling pathways in breast cancer metastasis - novel insights from functional genomics.** *Breast Cancer Research* 2011, **13**:206.