Signatures of breast cancer metastasis at a glance
George S. Karagiannis¹, Sumanta Goswami¹, Joan G. Jones¹,²,³,⁴, Maja H. Oktay¹,²,⁵ and John S. Condeelis¹,³,⁵,*

ABSTRACT
Gene expression profiling has yielded expression signatures from which prognostic tests can be derived to facilitate clinical decision making in breast cancer patients. Some of these signatures are based on profiling of whole tumor tissue (tissue signatures), which includes all tumor and stromal cells. Prognostic markers have also been derived from the profiling of metastasizing tumor cells, including circulating tumor cells (CTCs) and migratory–disseminating tumor cells within the primary tumor. The metastasis signatures based on CTCs and migratory–disseminating tumor cells have greater potential for unraveling cell biology insights and mechanistic underpinnings of tumor cell dissemination and metastasis. Of clinical interest is the promise that stratification of patients into high or low metastatic risk, as well as assessing the need for cytotoxic therapy, might be improved if prognostics derived from these two types of signatures are used in a combined way. The aim of this Cell Science at a Glance article and accompanying poster is to navigate through both types of signatures and their derived prognostics, as well as to highlight biological insights and clinical applications that could be derived from them, especially when they are used in combination.

KEY WORDS: Circulating tumor cell, CTC, Chemotaxis, Invadopodia, Human invasion signature, HIS, Intravasation, Mena, Tumor microenvironment of metastasis, TMEM

Introduction
High-throughput gene-expression profiling studies have yielded prognostic tests, such as MammaPrint-DX, Oncotype-DX and Prosigna (see below), that facilitate clinical decision making in breast cancer patients (see poster). However, these tests do not capture all of the patients who could benefit from adjuvant therapy or identify patients who do not need additional therapy [Arpino et al., 2013; Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Working Group, 2009; Sorlie et al., 2001; Sotiriou and Pusztai, 2009; Wirapati et al., 2008]. One of the reasons for the

¹Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, NY 10461, USA. ²Department of Pathology, Albert Einstein College of Medicine, Bronx, NY 10461, USA. ³Integrated Imaging Program, Albert Einstein College of Medicine, Bronx, NY 10461, USA. ⁴Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY 10461, USA. ⁵Gruss Lipper Biophotonics Center, Albert Einstein College of Medicine, Bronx, NY 10461, USA.

*Author for correspondence (john.condeelis@einstein.yu.edu)
hampered prognostic and predictive potential of these assays might be the fact that they are based on signatures that have been constructed from whole tumor tissue, which includes all tumor and stromal cells, instead of just the subpopulations of metastasizing tumor cells. To reveal the mechanisms responsible for the metastatic phenotype, gene expression has been performed on disseminating tumor cells, either circulating tumor cells (CTCs) (De Mattos-Arruda et al., 2013), or migratory–disseminating tumor cells isolated from the primary tumor (Goswami et al., 2004; Hernandez et al., 2009; Patsialou et al., 2013, 2012; Philippar et al., 2008; Roussos et al., 2011a,b,c; Wang et al., 2004, 2005, 2006, 2003; Wyckoff et al., 2011, 2000) (see poster). These signatures provide additional insights into the biological processes that are employed during metastasis, such as epithelial-to-mesenchymal transition (EMT), invasion and migration, resistance to apoptosis and anoikis, and intravasation. In this Cell Science at a Glance article and the accompanying poster, we discuss the discovery platforms and the clinical applications of clinically validated prognostic tests. We compare several prognostic tests derived from whole tissue profiling (i.e. MammaPrint DX, Oncotype DX, Prosigna) with prognostic tests derived from gene expression profiling of the metastasizing cancer cell subpopulation (e.g. migratory–disseminating tumor cells) and argue that their combined use has the potential to add additional information for treatment decisions for breast cancer patients.

**Gene expression signatures for prediction of breast cancer metastasis**

**Signatures from whole tumor tissue**

Molecular profiling of breast cancers has confirmed the existence of four major subtypes traditionally identified using immunohistochemistry: the basal-like subtype [estrogen receptor negative (ER−), progesterone receptor negative (PR−) and human epidermal growth factor receptor factor 2 (also known as ERBB2) negative (HER2−)], the HER2+ subtype, and two different ER+ subtypes, the luminal A (ERhigh, HER2−, low Ki67) and the luminal B (ERlow, HER2+/−, high Ki67) subtypes (Perou et al., 2000). These subtypes show diverse clinical outcome and response to therapy. For example, luminal-A breast cancers respond to hormonal therapy better than any other subtype, and also have favorable prognosis, whereas the basal-like and HER2+ subtypes are more sensitive to chemotherapy (Rouzier et al., 2005). There is new evidence that breast cancers can be subdivided in up to ten integrative clusters when copy number variation and gene polymorphism data are co-analyzed with genomic profiles (Curtis, 2015). However, such signatures have not yet reached clinical validation and application.

Several prognostic assays that are based on gene expression profiling of whole tumor tissue complement the molecular subtypes and address certain clinical needs, such as the identification of patient subgroups with low risk of developing metastatic disease or need for additional systemic therapy. MammaPrint DX is a 70-gene signature (also known as NKI-70) that has been developed through inkjet-synthesized oligonucleotide microarray profiling of tumors less than 5 cm in size obtained from breast cancer patients younger than 55 with lymph-node-negative disease (i.e. the disease had not spread into the axillary lymph nodes); it distinguishes patients who are likely to develop metastases from those who are likely to remain metastasis-free within five years after diagnosis (van ’t Veer et al., 2002). MammaPrint has been validated in multiple independent validation studies (Bueno-de-Mesquita et al., 2009; Buyse et al., 2006; Mook et al., 2009, 2010; van de Vijver et al., 2002; Wittner et al., 2008), and is currently approved for stratification of patients of all ages with either hormone-receptor-positive or -negative disease and negative lymph nodes into ‘low-risk’ or ‘high-risk’ groups for developing distant metastases at 10 years after diagnosis without adjuvant chemotherapy (Curtis, 2015). The other two molecular tests, Oncotype DX and Prosigna, are primarily used for hormone-positive disease (see the summary Table on the poster for a more complete list of validated prognostic markers).

The initial breast cancer biomarker studies that eventually led to Oncotype DX started with the identification of a 250-gene signature that is associated with risk of breast cancer recurrence (Cobleigh et al., 2005; Cronin et al., 2004; Paik et al., 2004). Subsequent statistical analysis and modeling narrowed this list down to 21 candidate genes, of which 16 were cancer-related and primarily focused on the components of the estrogen receptor pathway and proliferation, and five reference genes, which were used for normalization of the gene-expression of the cancer-related genes (Paik et al., 2004). The Oncotype DX has been analytically and clinically validated in multiple studies in the past and diverse patient cohorts (Cronin et al., 2007; Esteva et al., 2005; Habel et al., 2006; Paik et al., 2006). The test result is presented as a recurrence score (RS), ranging from 0 to 100, which predicts recurrence in early-stage, node-negative ER+ disease and benefit from chemotherapy after breast surgery (Arpino et al., 2013; Brenton et al., 2005; Cobleigh et al., 2005; Daly et al., 2005; Simon et al., 2003; Stoeckli et al., 2005). In particular, the Tailorx study demonstrated that patients with RS of 10 and below had a 98% chance of being disease free when treated with endocrine therapy only (Sparano et al., 2015). The benefit of chemotherapy in the patients with the intermediate RS (67% of patients) is still being evaluated, and in general, RS values in the intermediate range have not been useful to stratify patients as to metastatic risk.

Another prognostic test is IHC4; it is based on four widely performed immunohistochemistry measurements, those for estrogen receptor, progesterone receptor, HER2 and the proliferation marker Ki-67. IHC4 has been shown to contain the same amount of prognostic information as that of Oncotype DX RS and it correlates well with RS, and thus can be used interchangeably with RS (Cuzick et al., 2011). This suggests that these prognostics based on proliferation signatures are measuring the same biology in the whole tumor.

A prognostic that can assist in predicting response to therapy is Prosigna, which was derived from the PAM50 signature of whole tissue (Parker et al., 2009). PAM50 is a quantitative real-time PCR (qRT-PCR)-based assay of 50 genes; it evolved from studies that have reproducibly demonstrated the prognostic significance and predictive ability of the four intrinsic subtypes of breast cancer (Bastien et al., 2012; Chia et al., 2012; Harvell et al., 2008; Kelly et al., 2012; Martin et al., 2013; Nielsen et al., 2010; Prat et al., 2014; Sorlie et al., 2001, 2003). Prosigna stratifies breast cancer patients who have been treated with surgery and endocrine therapy into risk groups regarding disease outcome for years 5–10 post surgery and endocrine therapy, which helps to reduce overtreatment with chemotherapy in patients with good prognosis (Parker et al., 2009).

Several factors partially limit the effectiveness of these whole-tissue-derived signatures in the clinical setting. First, these signatures are mostly useful for the ER+/HER2− subset of patients. Second, although the molecular signatures address a similar clinical question, that is they identify low-risk patients, which do not need systemic adjuvant treatment, there is only minimal overlap between their gene lists, thus raising questions about their biological interpretation. Third, a thorough examination
of the individual genes incorporated in MammaPrint-DX, Prosigna and IHC4 and Oncotype-DX tests reveals that they are dominated by those of the estrogen receptor, progesterone receptor and proliferation pathways, with only a few exceptions for genes that might be involved in metastasis, such as matrix metalloproteinase-11 (MMP11) and cathepsin-L2 (CTSL2, also known as CTSV). Finally, these signatures mostly refer to genes that are involved in tumorigenic pathways of cancer cells with little or no regard to the tumor microenvironment, which is known to dictate tumor cell dissemination, metastasis, survival and drug resistance (Hanahan and Weinberg, 2011).

**Molecular profiling of circulating tumor cells**

An alternative approach to address the metastatic potential of tumors is expression profiling of circulating tumor cells (CTCs). CTCs might be more representative of disease progression and resistance to treatment than random cells taken from the primary tumor (Bidard et al., 2014; Lianidou et al., 2012; Pantel and Alix-Panabieres, 2013). Early attempts to molecularly characterize CTCs in breast cancer patients yielded poor results, primarily because of the small number of CTCs collected and the inevitable contamination by circulating leukocytes. As such, those studies mostly pursued only small numbers of gene candidates that were already known to be involved in breast cancer progression, metastasis and therapeutic resistance, rather than global transcriptional profiles (Aktas et al., 2011; Bolke et al., 2009; Gervasoni et al., 2008; Gradlione et al., 2011; Theodoropoulos et al., 2010). Recent technological advancements allow for immunomagnetic-bead or microfluidic-based isolation of CTCs (Autebert et al., 2015; Magbanua and Park, 2013). When coupled to high-throughput microarray or RNA-sequencing technologies and unbiased bioinformatics analyses, the aforementioned obstacles are significantly surmounted (Curtis, 2015; Lianidou and Markou, 2012; Magbanua and Park, 2014). Overall, these new-generation studies have allowed single-cell and bulk-cell analyses of CTCs, which efficiently capture the transcriptional heterogeneity and identify the molecular portraits of metastasizing cells (Ascolani et al., 2015; Cassatella et al., 2012; De Mattos-Arruda et al., 2013; Fina et al., 2015; Lang et al., 2015; Lianidou et al., 2012; Polzer et al., 2014; Powell et al., 2012). For example, one study has demonstrated that CTCs express various biomarkers and mediators of EMT (Yu et al., 2013), a phenotypic change crucial for the completion of the metastatic cascade (Bill and Christofori, 2015; Hanahan and Weinberg, 2011; Kalluri, 2009; Kalluri and Weinberg, 2009; Radadelli et al., 2009). However, only certain fractions of CTC's represent clones that are capable of colonizing distant metastatic sites, and in addition, there are reported discordances between the profile of the primary tumor and CTCs, e.g. with regard to their epidermal growth factor receptor 2 (EGFR2) status. Strikingly, CTC incidence might be observed in the peripheral blood of breast cancer patients, even after removal of the primary tumor, posing additional layers of complexity with regard to the specific origin of these cells and their relationship to dissemination from the primary tumor (Baccelli et al., 2013; Bidard et al., 2014; Curtis, 2015; Lang et al., 2015; Lianidou et al., 2012; Magbanua and Park, 2014; Pakazhendhi and Glück, 2014). As such, molecular signatures derived from CTCs should always be interpreted with caution.

**Molecular profiling of metastatic breast cancer cell subpopulations**

A recent study has demonstrated an in-depth expression profile of breast cancer stem cells using a mass-spectrometry-based proteomic approach in an *in vitro* model of forced EMT induction (Lu et al., 2014). The authors identified and validated new markers of breast cancer stem cells, including the membrane glycoprotein Thy1 (also known as CD90), which was associated with recruitment of tumor-associated macrophages; specifically Thy1 contributed to the construction of a cancer stem cell niche and enhanced tumor induction and progression in patient-derived xenografts (Lu et al., 2014). In another study, by using fluorescence-activated cell sorting-based assays and gene-expression profiling at the single-cell level, Lawson et al. demonstrated that metastatic cells retained a distinct gene expression signature that is composed of stem cell-, EMT-, pro-survival- and dormancy-associated genes (Lawson et al., 2015). Overall, many biologically relevant stem-like expression signatures have been reported (Borcherdig et al., 2015; Feng et al., 2014), but, to our knowledge, no definite clinical validation or meta-analysis studies have been performed to provide robust molecular profiles of breast cancer stem cells. Thus, the extent to which breast cancer stem cell incidence and/or function holds important prognostic and therapeutic information remains to be elucidated (Ali et al., 2011; Rakha et al., 2009; Shipitsin et al., 2007).

An alternative approach of capturing metastasizing tumor cells in living tumors has been previously established in studies using an *in vivo* invasion assay. By this approach, migrating cells are collected from the primary tumor site using chemottractants [e.g. epidermal growth factor (EGF), hepatocyte growth factor (HGF), insulin-like growth factor-1 (IGF1) or colony-stimulating factor-1 (CSF-1)], which are involved in the migration and dissemination steps of metastasis (Hernandez et al., 2009; Roussos et al., 2011b; Wang et al., 2005, 2003; Wyckoff et al., 2011, 2000). Using this methodology, migratory–disseminating cancer cells were collected from mouse mammary cancer models and patient-derived xenografts, and subjected to comparative expression profiling with whole tumor tissue. The resulting profiles were specific to breast cancer cell invasion and dissemination and defined rodent and human invasion signatures (Goswami et al., 2009, 2004; Patsialou et al., 2013; Patsialou et al., 2012; Roussos et al., 2011b; Wang et al., 2007, 2004). These invasion signatures are enriched in genes that regulate embryonic and tissue development, EMT, cellular movement, DNA replication and repair, apoptosis evasion and resistance to treatment (Goswami et al., 2004; Patsialou et al., 2013, 2012; Wang et al., 2003; Wyckoff et al., 2011, 2000). Notably, the disseminating cancer cell subpopulation that gave rise to this signature was more resistant to conventional chemotherapies, such as doxorubicin, etoposide and cisplatin, when compared to the bulk tumor population (Goswami et al., 2004). In the discussion below, we focus on the clinically relevant, human invasion signature (HIS) (Patsialou et al., 2013, 2012) and the HIS-derived clinically validated prognostic markers TMEM and Mena-cale.

**HIS describes migration and dissemination pathways critical for breast cancer metastasis**

The HIS contains 445 transcripts that are significantly altered in the migratory tumor cells, of which 185 are annotated genes with known protein products (Patsialou et al., 2013, 2012). Gene transcripts altered in the HIS are ranked according to functional categories, which provides valuable clues for the molecular events that precede tumor cell intravasation and hematogenous dissemination, and allows insight into a detailed biological context that supports tumor cell invasion and dissemination. Interestingly, in contrast to MammaPrint-DX, Oncotype-DX and Prosigna that do not provide such insight, selected genes from HIS are functionally required for *in vivo* invasion and hematogenous dissemination. Moreover, prominent dissemination/metastasis-associated pathways, including transforming-growth factor β
(TGF-β) and HGF were found to serve as molecular hubs in the identified overrepresented pathways (Patsialou et al., 2013, 2012), further confirming that this approach provides mechanistic insights underlying breast cancer dissemination and metastasis.

One of the prominently upregulated dissemination pathways in HIS is the mammalian-enabled (Mena, also known as ENAH)–cofilin pathway that regulates actin polymerization during chemotaxis and invasion of tumor cells (Bravo-Cordero et al., 2013; Philippar et al., 2008; Roussos et al., 2011b). Computational systems analysis of multiphoton microscope imaging datasets taken from large tumor volumes has been used to categorize the migration phenotypes of tumor cells in mammary tumors in vivo (Gligorijevic et al., 2014). High-resolution analysis of these categories indicates that the locomotion category, characterized by the streaming migration phenotype, and the invasion category, characterized by invadopodium-containing tumor cells near blood vessels and their associated intravasation, together contribute to tumor cell dissemination (Gligorijevic et al., 2014; Harney et al., 2015). Of significance, the Mena–cofilin pathway of the HIS contains the actin-binding proteins and their regulatory molecules (blue and lilac highlighted pathways in the HIS motility pathway poster panel) that define the locomotory (blue) and invasive (lilac) phenotypes of tumor cells in mammary tumors in vivo.

Mena is an actin-binding protein that is involved in the regulation of cofilin-stimulated actin polymerization, the key cofilin activity that determines chemotactic direction and invasion (Gertler and Condeelis, 2011; Philippar et al., 2008; Roussos et al., 2011a,b,c) (see poster). Subsequent investigation of the HIS gene expression pathways has revealed that differentially spliced Mena isoforms, including MenaINV, are upregulated, whereas the invasion-suppressing Mena11a isoform is downregulated in the migratory–disseminating tumor cell subpopulation (Goswami et al., 2009; Patsialou et al., 2013, 2012; Roussos et al., 2011a,b,c; Wang et al., 2004). This isoform-splicing pattern of Mena (MenaINV-high and Mena11alow) is functionally crucial for tumor cell invasion and migration during the metastatic cascade. In particular, it is required for directional cell migration in vivo, called streaming, towards chemotactic factors such as EGF and HGF, matrix degradation and transendothelial migration, as well as development of distant metastases (Patsialou et al., 2013, 2012; Patsialou and Condeelis, 2014; Philippar et al., 2008; Roussos et al., 2011a,c). In addition, MenaINV dramatically increases the sensitivity of receptor tyrosine kinases to their ligands EGF, IGF1 and HGF to increase cell protrusion and locomotion (Hughes et al., 2015b; Roussos et al., 2011a).

Identification and thorough examination of the HIS has led to development of clinically validated prognostic markers of metastasis, called MenaCalc and ‘tumor microenvironment of metastasis’ (TMEM; microanatomical structures that regulate transendothelial migration of breast tumor cells in primary tumors and possibly in secondary sites). MenaCalc is a measure of the isoform expression pattern of MenaINV-high and Mena11alow, and is commonly measured in formalin-fixed paraffin-embedded tissue as the subtraction of the immunofluorescence intensity of anti-metastatic-Mena11a antibody staining from total anti-Mena antibody staining (i.e. MenaCalc = PanMena – Mena11a), thus indirectly reflecting the expression of the invasive Mena isoforms. MenaCalc is predictive of metastatic relapse and survival in breast cancer patients (Agarwal et al., 2012; Forse et al., 2015). The relative expression of MenaINV to Mena11a is directly related to TMEM assembly because tumor cells that both assemble and use TMEM for dissemination require the expression pattern MenaINV-high and Mena11alow (Pignatelli et al., 2014) (see poster).

TMEM sites are the sites of intravasation in polyoma middle T oncprotein (PyMT)-induced mouse mammary tumors and patient-derived xenograft (PDX) mammary tumors, and are therefore necessary for the dissemination step of the metastatic process (Harney et al., 2015). The number of TMEM sites (TMEM score) in ER+ breast tumor tissue is predictive of an increased risk of distant metastasis. TMEM score is more strongly associated with risk of distant breast cancer metastasis than the IHC4 score and, by extension, Oncotype DX recurrence score (Rohan et al., 2014). In addition, a TMEM composite score is defined by combining the TMEM score with standard clinical variables including tumor grade, lymph node status, age, size and diagnosis year (Rohan et al., 2014). The ‘AUC’ curves [i.e. the area under the receiver operating characteristic (ROC) curve, which is the true positive rate versus false positive rate] for the TMEM composite score yield a value of 0.78 compared to those for MammaPrint (0.68), Oncotype DX
(0.69) and IHC4 (0.64). These AUC values, where 0.8 is good, 0.7 is fair and 0.6 is poor discrimination (Metz, 1978), indicate that TMEM composite is better at predicting metastasis than these other prognostics (Rohan et al., 2014).

TMEM sites are composed of one Tie2\textsuperscript{high}, VEGF\textsuperscript{high} and MRC1\textsuperscript{+} perivascular macrophage, which is in physical contact with a Mena-overexpressing cancer cell, as well as an underlying endothelial cell (see poster) (Harney et al., 2015; Oktay and Jones, 2015; Pignatelli et al., 2014; Robinson et al., 2009; Rohan et al., 2014; Roussos et al., 2011c; Wyckoff et al., 2007). Previous reports on cancer angiogenesis and vasculogenesis have collectively demonstrated that newly constructed endothelia are composed of myeloid precursors expressing the receptor tyrosine kinase Tie2 (also known as TEK) (Riabov et al., 2014; Squadrito and De Palma, 2011), and, more recently, the mannose-receptor c-type 1 (MRC1) (Hughes et al., 2015a). Despite the fact that Tie2 expression has been primarily associated with endothelial precursors, it has also been linked to specific bone-marrow-derived monocyte lineages; these transdifferentiate into non-inflammatory, blood-vessel-resident macrophages that are committed to mediate proangiogenic programs and tissue remodeling, which are both frequently phenocopied by tumors (Forget et al., 2014; Pucci et al., 2009; Venneri et al., 2007).

TMEM sites enforce transient and localized blood vessel permeability through the disruption of homotypic cell–cell adhesions in the underlying endothelium (Harney et al., 2015). The detailed mechanism of TMEM function and the involvement of Mena isoforms in intravasation at TMEM sites have been described (Harney et al., 2015; Pignatelli et al., 2014; Roh-Johnson et al., 2014) and it has been shown that activation of the Tie2 signaling pathway in the perivascular macrophage stimulates the release of macrophage vascular endothelial growth factor (VEGF). VEGF in turn, promotes reversible local loss of endothelial adherens junctions at TMEM sites (Harney et al., 2015). The TMEM macrophage also initiates Mena-dependent invadopodium assembly in the TMEM tumor cell, with which it is in direct contact, thereby completing a required invasive step for transendothelial migration at TMEM sites (Pignatelli et al., 2014; Roh-Johnson et al., 2014). The resulting reversible, local permeability of the endothelium is visualized in real time by multiphoton microscopy as a dramatic release of luminal blood vessel components into the subluminal space (termed ‘bursting’). During bursting, lines of migrating cancer cells (called streaming cells; Roussos et al., 2011b) with the Mena\textsuperscript{INV-high} and Mena11a\textsuperscript{low} splicing pattern intravasate (Harney et al., 2015; Pignatelli et al., 2014; Roussos et al., 2011b). This finding directly links migration to dissemination (see poster). Indeed, conditional ablation of macrophages, disruption of Tie2 or VEGF-signaling, or inhibition of Mena\textsuperscript{INV}-dependent invadopodium assembly dramatically inhibits TMEM assembly and function, and, consequently, tumor cell intravasation (Harney et al., 2015; Pignatelli et al., 2014).

Although the HIS contains additional prognostic markers that are currently being explored (Weitsman et al., 2014), the few that have already been investigated have provided an in-depth understanding of the process of tumor cell intravasation and dissemination, and more importantly have resulted in the development of clinically useful predictive markers of metastasis called TMEM and Mena\textsuperscript{Calc}.

**Clinical applications and future perspectives**

Prognostic signatures derived from the study of whole tumor tissues, such as Oncotype DX, and its related test IHC4, are dominated by proliferation pathways (Paik et al., 2006), whereas signatures obtained from disseminating and metastatic tumor cell subpopulations (e.g. the HIS) are associated with cancer-progression-specific pathways, including invasion, EMT, metastasis (Patsialou et al., 2012), as well as TMEM assembly, intravasation and dissemination (Harney et al., 2015; Robinson et al., 2009; Rohan et al., 2014). The fact that most whole-tumor tissue gene-expression profiles are dominated by proliferation markers has been recently criticized, suggesting that such signatures might lose prognostic value regarding breast cancer outcome if tumors are statistically adjusted for the prevalence of proliferation-associated genes by the reference dataset (Venet et al., 2011). That is to say, any random gene set will contain a high proportion of proliferation-associated genes given that these genes are so prevalent, and have prognostic value similar to the whole-tumor tissue gene-expression profiles. However, proliferation is indeed a rate-limiting step of distant colonization (Falato et al., 2014; Nielsen et al., 2010; Wishart et al., 2014; Zurrida et al., 2013) and thus particularly important for assessment of breast cancer prognosis if combined with additional information. For these reasons, we speculate that prognostic score based on proliferation signature combined with a prognostic score derived from a signature of dissemination could provide complementary and more personalized prognostic and predictive information for breast cancer patients. An example of this is the possible combination of Oncotype Dx or IHC4, two correlated proliferation prognostics (Cuzick et al., 2011), with the TMEM prognostic marker, which provides independent prognostic information from IHC4 (Rohan et al., 2014). How a combined marker analysis of patients might be used to derive additional information about patient risk and treatment options is illustrated in the text box and accompanying chart (see Box 1).

In the future, patients whose disease shows a high risk of dissemination (i.e. a high TMEM score) would need to be monitored very closely after the removal of the primary tumor. Although they might not have clinically detectable metastases at the time of initial treatment, elevated markers of dissemination (either TMEM or Mena\textsuperscript{Calc}) would indicate that these patients already have cancer cells in distant organs which in time might lead to growth of metastatic tumors. These patients would be candidates for further treatment with drugs to suppress the growth of, and destroy dormant disseminated tumor cells.

In summary, future efforts should encompass further clinical validation of the use of combined signatures in larger patient cohorts. As noted in the Tailorx trial (Sparano et al., 2015), up to 67% of breast cancer patients have a disease with an intermediate Oncotype DX RS, which cannot be stratified as to the risk of recurrence by RS alone and would benefit from additional tests that measure different aspects of tumor biology, such as dissemination.

**Competing interests**

J.G.J. and S.G. have ownership interest (including patents) in MetaStat and J.G.J. is a consultant and advisory board member for the same. J.S.C. has ownership interest (including patents) in MetaStat and is a consultant and advisory board member for Deciphera Pharmaceuticals and MetaStat. No potential conflicts of interest are disclosed by the other authors.

**Funding**

This article was supported by a National Cancer Institute provocative question grant [grant number 1R01 CA170507-01]; the National Institutes of Health [grant numbers CA100324, CA163131, CA164468], the Integrated Imaging Program; and the Gruss Lipper Biophotons Center. Deposited in PMC for release after 12 months.

**Cell science at a glance**


