Mechanisms of malignant progression

Robert A. Weinberg*  
Department of Biology, Ludwig Center for Molecular Oncology, Whitehead Institute for Biomedical Research, MIT, Cambridge, MA 02142, USA  
*To whom correspondence should be addressed.  
Email: weinberg@wi.mit.edu

While a coherent picture has begun to emerge about the biological and molecular mechanisms that create primary tumors, the processes that lead subsequently to invasion and metastasis have, until recently, been relatively obscure. However, over the past 5 years, research of diverse sorts has begun to generate the conceptual outlines that explain how high-grade malignancies arise. These discussions invariably are motivated by a widely accepted depiction of how metastatic dissemination occurs—the sequence termed the ‘invasion–metastasis cascade’ (1). Thus, primary tumor cells invade locally, enter into the circulation (intravasation), are transported through the circulation, are lodged in microvessels in distant tissues, invade the parenchyma of such tissue (extravasation) and form micrometastatic deposits, some of which eventually grow into macroscopic metastases, the last process being termed colonization.

Early determination

A major issue concerns the timing of acquisition of the ability to invade and metastasize. Is this ability a function of the great majority of cells within a primary tumor, or does it reflect the biology of a small subpopulation of cells in this tumor that have acquired the ability to invade and metastasize? If the great majority of cells in a primary tumor have a comparable if not identical ability to invade and metastasize, this suggests that such acquisition occurs early. Conversely, if these powers are an attribute of a relatively small subpopulation, this trait must be acquired relatively late during the course of primary tumor formation.

This issue might be addressed by arguing from first principles. For example, if a subpopulation of cells that are especially favored to metastasize exists within a primary tumor, the clonal expansion of these cells would seem to have been favored during the successive mutational and selection cycles that generated the clonal successions that drive multistep tumor pathogenesis forward. This model, however, confronts a conceptual difficulty: why should the phenotype of metastasis be selectively advantageous within the confines of a primary tumor? If it is not, this suggests that, to the extent that metastasis occurs, it does not represent a phenotype that has been actively selected during primary tumor formation (2). Accordingly, this phenotype may be an inadvertent consequence of the acquisition of alleles that did indeed confer growth advantage within the evolving primary tumor mass, that is, these alleles may act pleiotropically, encoding both selected phenotypes (growth and survival in the primary tumor) and unselected phenotypes (invasiveness and metastatic powers).

Still, these arguments from first principles must give ground to empirical observation, since only the latter can definitively settle the early versus late debate. Here, one confronts the complexity of the invasion–metastasis cascade, which rivals that of the prior steps that led initially to the formation of a primary tumor. Accordingly, one can pose the question of whether the acquisition of metastatic traits requires the accumulation of a significant number of mutations that rival in number and variety the genetic lesions that lead initially to primary tumor formation.

In fact, as evidence accumulates, it becomes increasingly plausible that metastatic dissemination does not depend on the acquisition of additional genetic lesions beyond those that are present in many primary tumors. (These discussions must focus on the mutant alleles that serve as ‘drivers’, i.e. those that were selected because they conferred advantageous biological phenotypes, rather than the ‘passenger’ mutations, which reflect the general increased mutability of tumor cell genomes and thereby constitute random genetic background noise.) (3) The evidence arguing against a necessary role of metastasis-specific mutations (acquired late) comes from at least three distinct sources.

First, as has been documented on several occasions, one can, with some measure of success, determine the prognosis of primary tumors, especially their tendency to progress toward metastatic relapse, by examining the gene expression profile of these tumors. These gene expression profiles must reflect the differentiation program of the normal cell of origin as perturbed by the subsequently acquired somatic mutations (as well as promoter methylation events).

For example, by using gene expression arrays to examine the transcription patterns of a large number of primary breast carcinomas, investigators have been able to predict, with some accuracy, which tumors are likely and which are unlikely to progress to metastatic relapse (4,5). This holds important implications for the mechanisms that govern metastasis, since they suggest that many cells, probably the great majority of neoplastic cells within a primary tumor, express a spectrum of genes that influence metastatic behavior. Conversely, this expression pattern cannot be that of a small minority subpopulation within the primary tumor, whose gene expression pattern would be fully obscured by the RNAs expressed by the majority cell population. Hence, it seems probably that eventual metastatic spread is determined relatively early in tumor progression, that is, early enough to be implanted in the great majority of the cells within a primary tumor (2). This in turn suggests that the genetic determinants that were selected during the initial multistep formation of the tumor also happen to be those that favor metastatic spread.

A second indication of the determinants of eventual metastatic spread comes from research in my own laboratory, in which a variety of normal human cell types have been transformed to a tumorigenic state through the introduction of a defined set of genes, specifically those encoded by the SV40 virus early region (which specifies the large T and small t oncoproteins), the hTERT gene (which encodes the catalytic subunit of the telomerase holoenzyme), and a ras oncogene. When early passage human mammary epithelial cells were propagated in two alternative culture media, they yielded populations that exhibited distinct gene expression patterns and, upon transformation by these introduced genes, yielded tumorigenic cells that grew into two histopathologically distinct tumors—one a squamous cell carcinoma and the other an invasive ductal adenocarcinoma of the breast (6). These outcomes conform with a widely held preconception—that the differentiation program of the normal cell of origin is a strong determinant of the eventual behavior of tumor cells arising following transformation. Indeed, in this case, the expression patterns of the two transformed cell types closely resembled that of their respective normal precursors and showed relatively minimal resemblance to one another.

More important for the present discussion, however, is the fact that one type of tumor—the invasive ductal carcinoma—metastasized to the lungs, whereas the other—the squamous cell carcinoma—did not metastasize to the lungs, whereas the other—the squamous cell carcinoma—did not (6). Since the two types of tumors had acquired the identical set of experimentally introduced genes, which formally mimic somatic mutations, this demonstrated that the differentiation program of the normal cell of origin represented a strong determinant of eventual metastatic spread. Hence, the nature of the normal cell of origin, which serves as the progenitor of all the neoplastic cells within a tumor, sets the stage for whether its descendants, years or decades later, will or will not show metastatic tendencies.

Abbreviations: EMT, epithelial–mesenchymal transition; TF, transcription factor.
An even more dramatic outcome was seen when normal human melanocytes were transformed with the same set of oncogenes. The resulting transformants, which represent a model of spontaneous melanomas, generated primary tumors from which hundreds of metastases arose in various organ sites throughout the body, vastly overshadowing the metastases that were formed by the above-described mammary adenocarcinomas (7). Here, one has an even more dramatic example of the fact that the differentiation program of the normal cell of origin exerts strong influence on whether or not metastasis will eventually occur.

A third indication of the lack of metastasis-specific genes comes from a variety of studies of the genomes of primary tumor cells with those of their derived metastases, including extensive sequencing of their respective genomes. It is generally the case that the same set of genetic lesions that are present in the genomes of primary tumor cells are found as well in the genomes of their derived metastases (8). Once again, there is every reason to believe that the driving force for metastasis has not been specific genetic lesions that were acquired late during the multistep formation of a primary tumor. Together, these diverse lines of evidence suggest, but hardly rigorously prove, that the dissemination of cancer cells from a primary tumor occurs as an almost inadvertent side-effect of primary tumor formation rather than a trait that is actively selected during this multistep process.

**Determinants of invasiveness**

The description of how metastases arise must also address the complexities of the invasion–metastasis cascade, as laid out above. If there are not a large suite of additional mutations required to drive invasion and metastasis beyond those needed for primary tumor formation, how then do carcinoma cells—the topic of most of the discussion that follows—acquire the ability to complete all these steps? The answer here seems increasingly to depend, as it does in many other types of tumor-generating processes, on the fact that cancer cells appropriate complex biological programs that play roles in normal cell and organismic physiology.

In the present case, the specific normal biological process involves the epithelial–mesenchymal transition (EMT), which plays key roles in many steps of normal morphogenesis (9). In particular, a number of distinct morphogenetic steps involve the local movement of epithelial cells or their translocation to distant sites in the developing embryo. In general, true epithelial cells are incapable of such movements: they may move laterally in the plane of the epithelium while retaining adhesion to the underlying basement membrane/basal lamina. However, active movement in other directions appears to be forbidden to them.

Such departures from the plane of an epithelium depend on the shedding by epithelial cells of some of their native characteristics and the acquisition, instead, of mesenchymal cell traits; the latter cells are indeed capable of locomotion and invading the extracellular matrix that may impede their forward motion. In fact, EMTs play central roles in a number of distinct steps of embryogenesis, including gastrulation and the emigration of cells from the primitive neural crest to various destinations throughout the embryo (10). As has been demonstrated in recent years, EMTs can be programmed by a variety of transcription factors (TFs) that are activated transiently in various stages of embryogenesis and in specific locations within an embryo (11–16).

The transient expression of such TFs during embryogenesis indicates that, by necessity, the shutdown of their expression may lead to the loss of their EMT-inducing effects and thus a reversing of the EMT. Therefore, once a cell has passed through an EMT induced by such a TF, in the subsequent absence of this TF, the cell may revert, via a mesenchymal–epithelial transition, to the epithelial state in which its ancestors existed. This suggests a still unproven notion: that the epithelial state represents the ‘ground state’, and that cells that have been induced to enter into a mesenchymal state will revert to the epithelial state unless their mesenchymal phenotype is actively supported.

The fact that EMTs occur only in specific locations within an embryo suggests another important conclusion: they normally occur in cells in response to specific contextual signals that they receive from their surroundings, specifically the cells around them that create the local microenvironment. This suggests, by extension, that expression of EMT-inducing TFs may be determined in a neoplastic cell by the contextual signals that this cell receives from its neighbors.

These depictions, on their own, do not indicate the relevance of the EMT to cancer pathogenesis. However, two key aspects of carcinoma cells point to the relevance of these embryonic programs and TFs to tumor progression. First, many of the phenotypes of embryonic cells are recapitated by aggressive carcinoma cells. Second, many of the embryonic TFs that are known to play critical roles in orchestrating EMTs during embryogenesis are also found to be expressed in a variety of human tumor cells; indeed, their expression is often correlated with aggressive tumor cell-associated traits.

These TFs include Slug, Snail, Twist, Goosecoid, SIP-1, FOXC2 and ZEB1. Most of these TFs were discovered through the study of developmental genetics, often in organisms far removed from mammals, including *Xenopus* and *Drosophila*. Their presence, often in highly conserved form, in the genomes of distantly related animals testifies to their initial development early in metazoan evolution and their critical roles in the embryogenesis of these diverse organisms. Taken together, these various lines of evidence suggest that the various TFs are appropriated in order to enable carcinoma cells to acquire the traits of high-grade malignancy. Included among these is an ability to invade, to resist apoptosis and to secrete the proteases that are required to break down extracellular matrix. In truth, the EMT and the expression of most of these TFs are not limited to early embryonic development: during wound healing in the adult, various types of EMT are activated transiently in order to enable this process to reach completion (17).

**Induction of expression of EMT-inducing TFs**

As indicated above, during embryogenesis the expression of various EMT-inducing TFs appears to occur in response to certain contextual signals that are released by nearby cells. It seems likely that the same type of heterotypic signals impinge on various carcinoma cells during the process of carcinoma progression. Included among these are Wnts, Hedgehogs, members of the transforming growth factor-beta family, as well as ligands of tyrosine kinase receptors. It also seems likely that, in general, no one of these ligands is capable, on its own, of triggering an EMT; instead, in many circumstances they seem to act combinatorially to provoke the EMT in nearby carcinoma cells. The rules that define these interactions are still unexplored.

In the context of carcinoma pathogenesis, it is likely that these heterotypic signals are released by mesenchymal cells that form the tumor-associated stroma. Such mesenchymal cells have a quite different origin than the carcinoma cells that have become mesenchymal via passage through an EMT. These stromal cells are recruited either from the stroma of the tissue in which the tumor arises or, alternatively, from the bone marrow, which appears to generate a number of distinct types of mesenchymal progenitor cells that are released into the circulation and become available for local recruitment by carcinoma cells (18). Such cells appear to enter into the tumor-associated stroma and thereafter differentiate into a variety of mesenchymal cell types, including myofibroblasts and endothelial cells. In fact, the stroma of most carcinomas is assembled from a variety of mesenchymal cell types whose precise origins are still quite unclear.

While still poorly supported by direct experimental observations, there is already considerable circumstantial evidence to indicate that the EMT-inducing heterotypic signals are not released by the stroma of early stage tumors. Instead, as tumor progression proceeds, the stroma becomes increasingly ‘activated’, ‘reactive’ and desmoplastic, indeed taking on the attributes of tissues that are in the midst of active wound healing or are chronically inflamed. It is these inflamed stroma that are the likely sources of the heterotypic signals that evoke EMTs in nearby carcinoma cells.
This depiction of the activated stroma and its role in evoking an EMT within primary tumor cells holds an important implication for the behavior of carcinoma cells once they have left a primary tumor: while the stroma within the primary tumor site may have induced them to undergo an EMT, the stromata that these cancer cells encounter in distant sites of dissemination are likely to be quite different. In these secondary sites, the stromata have not been perturbed by long-term stimulation by nearby-carcinoma cells and therefore will not have these activated attributes. Lacking these, such stromata are unlikely to release EMT-inducing heterotypic signals. In the absence of these signals, it seems likely that the disseminated carcinoma cells will lose expression of EMT-inducing TFs and revert, via a mesenchymal-epithelial transition, to the epithelial phenotype of their ancestors in the primary tumor. The absence of mesenchymal phenotypes in metastases has been cited by some as a disproof of the notion that cancer cells must pass through an EMT in order to disseminate. However, the known reversibility of the EMT renders such arguments moot.

The fact that heterotypic signals may induce an EMT in carcinoma cells reveals another important aspect of malignant progression: these cells do not need to undergo additional genetic changes in order to acquire the cellular phenotypes associated with high-grade malignancy. Instead, when confronted with an appropriate mix of contextual signals, primary carcinoma cells will develop such phenotypes without suffering additional mutations.

This scheme raises yet another question: what variable factors determine whether or not the cancer cells within a primary tumor will or will not undergo an EMT? As proposed above, the appropriate mix of heterotypic signals appears to be essential to this change. But in addition, the carcinoma cells must be responsive to these signals. It seems highly likely that the differentiation program of the normal cell of origin represents one critical determinant of this responsiveness, echoing the assertion made earlier that this differentiation program can set the stage for eventual malignant progression occurring years or decades after tumor progression has been initiated. In addition, the suite of somatic alterations that are accumulated during primary tumor progression, including mutations and promoter methylation events, are also likely to play a key role in determining such responsiveness. However, at present, we are still far away from being able to assess how the combinatorial actions of initial differentiation programs and subsequent somatic changes determine responsiveness to EMT-inducing signals.

The suggested critical role of responsiveness to stromal signals also holds another implication for the evolution of malignant traits: if these traits are only manifested in response to heterotypic signals from an activated stroma, then the traits themselves cannot have been the objects of selection during the multistep evolution that led previously to the formation of a primary tumor. Instead, it is the somatically generated alleles that confer ‘responsiveness’ to these signals, rather than alleles that are selected during primary tumor formation directly specifying malignant traits that are selected during primary tumor formation. This adds to the weight of arguments, enumerated above, that the expression of highly malignant traits occurs as an inadvertent (‘unintended’) consequence of the actions of alleles that were initially selected because they confer traits that are, on their own, unrelated to the phenotypes of invasion and metastasis.

The fact that carcinoma cells that undergo an EMT adopt mesenchymal phenotypes and invade into the tumor stroma and then into adjacent normal tissues creates an experimental difficulty, since these invading neoplastic cells are, at least superficially, indistinguishable from the true mesenchymal cells that surround them in the tumor stroma and, later on during the course of invasion, in the stroma of normal tissues lying outside the initial margins of tumors. This complication has caused some pathologists to dismiss the EMT as a laboratory artifact.

It seems likely, however, that this controversy will be settled, sooner or later, because of two factors. First, it is probably the case that most carcinoma cells undergoing an EMT do so incompletely, i.e. by partially shutting down epithelial markers (such as E-cadherin and cytokeratins) while acquiring mesenchymal markers (such as N-cadherin, vimentin and fibronectin). Accordingly, future attempts at finding cells coexpressing both epithelial and mesenchymal markers are likely to reveal the invading neoplastic cells hiding among the true mesenchymal cells in the stroma.

A second fact is likely to help reveal otherwise occult cancer cells that have undergone an EMT: carcinoma cells that have passed through an EMT express certain markers that are appear not to be expressed by true mesenchymal cells (19). These two factors should reveal the elusive wolves hiding in sheep’s clothing—the invasive carcinoma cells present in small nests and large aggregates in the otherwise-normal tissues of carcinoma patient.

The EMT and the invasion–metastasis cascade

One key question addresses how the actions of EMT-inducing TFs empower cells to complete the various successive steps of the invasion–metastasis cascade. Given the multiple distinct cell biological traits that these pleiotropically acting TFs can elicit, how effective are they in enabling cancer cells to complete these complex steps? To enumerate the traits once again, these TFs confer increased resistance to apoptosis, cell motility, release of degradative enzymes and invasiveness.

Given current knowledge, it is plausible that the expression of one of these TFs should enable a primary tumor cell to invade locally, intravasate, survive in the circulation, extravasate and survive for a limited period of time in the parenchyma of a foreign tissue in which it has landed. The subsequent fate of such a disseminated cell is less clear, however. Thus, a breast cancer cell landing in the brain, the bone marrow, or the liver must confront an array of extracellular matrix components, signaling molecules and stromal cell types to which it is, at least initially, poorly adapted. This lack of instantaneous compatibility between newly arrived cancer cells and their newfound homes is likely to explain the very low success rate of the last step of the invasion–metastasis cascade—the growth of a micrometastasis into a macroscopic metastasis that is, as mentioned, termed colonization. It is apparent that only a small number of micrometastases out of the thousands that are initially seeded ever succeed in growing into a macroscopic metastasis.

It also seems apparent that colonization is not a problem that is readily addressed by the multiple traits programmed by an EMT-inducing TF. To be sure, the increased resistance to apoptosis associated with an EMT program should increase the survival of the cells within a micrometastasis. This acquired trait does not, however, deal with the fact that these cells are otherwise maladapted to the foreign microenvironment of the tissue in which they have landed.

Still, if this scenario is eventually justified by direct experimentation, this would indicate that the expression of one of these TFs may choreograph almost all the steps of the invasion–metastasis cascade save the final one of colonization. On the one hand, this would represent an enormous conceptual simplification of this process, since it would rationalize the many complex steps in terms of the actions of single or small groups of centrally acting controllers. On the other, it would demonstrate how very dangerous these TFs are and the necessity of keeping them under tight control in adult tissues, lest they wreak havoc on normal tissue homeostasis.

Permanent EMT

While epithelial cancer cells have been depicted here as highly plastic, in that they can enter reversibly into a mesenchymal state, this plasticity may not be apparent in some human tumors, in which cancer cells seem to be locked irreversibly in a mesenchymal state. Such irreversible EMT might, in principle, be caused by any of a number of changes at the genetic and biological levels. However, only one has come into clear view in recent years—this change involving the cell-surface E-cadherin protein, which is involved in forging the side-by-side interactions between adjacent epithelial cells that are termed adherens junctions.
Given its prominent role in knitting together epithelial cell sheets, E-cadherin has, quite appropriately, been considered as one of the archetypal epithelial markers. Significantly, the promoter of its encoding gene contains binding sites for a number of EMT-inducing TFs that serve to repress transcription (20–22). The repression of E-cadherin expression seems to be one of the key targets of action of these TFs. At the same time, they shut down, through still obscure mechanisms, cytokeratin expression; this shutdown is co-ordinated with an induction of the various aforementioned mesenchymal genes.

This scenario depicts E-cadherin expression as the target of regulation by EMT-inducing TFs. However, in certain tumors, E-cadherin expression is compromised by alterations in the gene itself, including point mutations and deletions that cause the production of truncated or fully unstable proteins. Like the TF-programmed EMTs, these changes deprive carcinoma cells of the functions of this critical protein, but in addition, they evoke another change in the lives of these cells. In the absence of functional E-cadherin, the cytoplasmic proteins that usually serve to physically link it to the actin cytoskeleton are cast adrift. Some of them, lacking this tethering, may suffer rapid degradation. But one of them—β-catenin—may survive, especially if it escapes phosphorylation by glycogen synthase kinase-3 β and subsequent degradation in proteasomes. This state may exist, for example, in cells that have hyperactivated phosphatidylinositol-3 kinase and Akt kinases since the latter functions to inactivate glycogen synthase kinase-3 β.

The resulting liberated β-catenin may then localize to the nucleus, where it can associate with the T-cell factor group of TFs, allowing it, together with other afferent signals, to trigger expression of a wide variety of downstream target genes, many of them involved in triggering an EMT, including ‘Twist’ itself (23). Hence, Twist represses E-cadherin expression, and loss of E-cadherin results in induction of Twist expression, resulting in a self-reinforcing positive feedback (i.e. feed forward) loop that may help to maintain the mesenchymal state following induction of an EMT.

An emerging pattern

Mechanisms of the sort depicted here, involving components of the EMT program, may ultimately serve to explain how carcinoma cells are able to leave the primary tumor and ultimately arrive at distant anatomical sites. It is unclear at present whether aggressive tumors arising from different embryonic cell lineages, specifically hematopoietic, neuroectodermal and mesenchymal tumors, deploy the same mechanisms or whether their metastatic dissemination depends on an entirely different set of factors and molecular mechanisms. Indeed, at present, it remains possible that the motility and invasiveness of some carcinomas derives from mechanisms that have nothing whatsoever to do with the EMT, although we suspect otherwise.

We suggest that the mechanisms of physical dissemination of a variety of tumor cells will come into clear view over the next 5 years. This will leave the mechanisms underlying the last step of the invasion–metastasis cascade—colonization—unresolved. As proposed earlier, colonization appears to depend on the adaptation of disseminated cancer cells to unfamiliar tissue microenvironments. Solutions to such adaptation may happen to occur almost accidentally, even if they have nothing whatsoever to do with the EMT, although we suspect otherwise.

Once they arrive at such solutions, such disseminated cancer cells can begin to generate rapidly growing macroscopic metastases, which may in turn serve as sources for a secondary wave of metastatic dissemination, a “shower” of metastases that signals the final, aggressive step of malignant progression. Importantly, however, these adaptive solutions are likely to involve multiple concomitant changes in the cancer cells and on different combinations of such changes, depending on the tissue- and cell-of-origin of the primary tumor and the identities of the tissues in which they have landed. Hence, while physical dissemination may soon be understood in terms of a relatively small number of unifying principles, the mechanisms of colonization in various tissues may one day be described in terms of long catalogs of changes in cell phenotype.

Funding

Breast Cancer Research Foundation; U.S. National Cancer Institute

Acknowledgements

Conflict of Interest Statement: None declared.

References


Received April 25, 2008; revised April 25, 2008; accepted April 25, 2008