Universality of a mesenchymal transition signature in invasive solid cancers

Dimitris Anastassiou

Center for Computational Biology and Bioinformatics and Department of Electrical Engineering, Columbia University, New York, NY, USA, email: anastas@ee.columbia.edu

Abstract

In this brief communication, additional computational validation is provided consistent with the unifying hypothesis that a shared biological mechanism of mesenchymal transition, reflected by a precise gene expression signature, may be present in all types of solid cancers when they reach a particular stage of invasiveness.

Introduction

We recently identified a signature consisting of a set of coordinately expressed genes, many of which are epithelial-mesenchymal transition (EMT) markers including the EMT-inducing transcription factor Slug (SNAI2) [1]. The signature was derived using computational methods of systems biology after observing that the genes of the signature become significantly overexpressed only when cancer progresses to a particular invasive stage specific to each cancer type. For example, the gene signature is significantly triggered roughly when colon cancer progresses to stage II, ductal carcinoma in situ progresses to invasive ductal carcinoma of stage I, and ovarian cancer progresses to stage IIIc.

We also found [2] that many of the genes of the signature, including α-SMA, and Slug being the only present EMT-inducing transcription factor, are expressed by the cancer cells themselves *in vivo*, and not by the peritumoral stoma, at least in one xenograft model of neuroblastoma that we tried, confirming that cancer cells have undergone a mesenchymal transition (the term EMT may not be accurate in this case, as the same signature is also present in nonepithelial cancers, such as neuroblastoma and Ewing's sarcoma). We refer to this multi-cancer signature as the "cancer mesenchymal transition signature." It is characterized by a prominent presence of co-expressed genes COL11A1, THBS2, INHBA. A set of the top 64 genes comprising the signature was presented in Table 1 of [2] (the number 64 is arbitrarily chosen as corresponding to the distinct genes from the top 100 probe sets originally presented in [1]).

Results

We used the 64 genes listed in Table 1 of [2] as input for Gene Set Enrichment Analysis (GSEA) provided by the Broad Institute against the Molecular Signatures Database (MSigDB) (www.broadinstitute.org/gsea/msigdb). The results included many "hits" with *P* value exactly equal to "zero," providing biological insights into the nature of the underlying mechanisms. Among those (with "$P = 0e^{0}$"), there were many occurrences of data sets of genes expressed in higher-stage samples from many cancer types, such as nasopharyngeal, head and neck, urothelial, lymphomas, etc. Such cancer types had not participated in any way whatsoever in the derivation of the signature. This remarkable validation of the signature by pointing to all kinds of cancer types in MSigDB suggests that the signature may reflect a universal biological mechanism present in the invasive stage of all solid cancers. The following table provides a sample of such GSEA results:

SENGUPTA_NASOPHARYNGEAL_CARCINOMA_UP [286]	**Genes up-regulated in nasopharyngeal carcinoma relative to the normal tissue.**
GRUETZMANN_PANCREATIC_CANCER_UP [346]	**Genes up-regulated in pancreatic ductal adenocarcinoma (PDAC) identified in a meta analysis across four independent studies.**
LINDGREN_BLADDER_CANCER_CLUSTER_2B [389]	**Genes specifically up-regulated in Cluster IIb of urothelial cell carcinoma (UCC) tumors.**
PICCALUGA_ANGIOIMMUNOBLASTIC_LYMPHOMA_MA_UP [207]	**Up-regulated genes in angioimmunoblastic lymphoma (AILT) compared to normal T lymphocytes.**
SCHUETZ_BREAST_CANCER_DUCTAL_INVASIVE_VE_UP [355]	**Genes up-regulated in invasive ductal carcinoma (IDC) relative to ductal carcinoma in situ (DCIS, non-invasive).**
VECCHI_GASTRIC_CANCER_ADVANCED_VS_EARLARLY_UP [167]	**Up-regulated genes distinguishing between two subtypes of gastric cancer: advanced (AGC) and early (EGC).**
CROMER_TUMORIGENESIS_UP [44]	**Tumorigenesis markers of head and neck squamous cell carcinoma (HNSCC): up-regulated in the 'early' tumors vs normal samples.**

Discussion

The epithelial-mesenchymal transition, when induced by transcription factors Snail or Twist, is known to generate cells with properties of stem cells [3]. As we show below in two examples, our computational results suggest that the same may also be true for the Slug-based mesenchymal transition reflected by this signature.

First, we found that the same set of 64 genes (in the form of a corresponding "metagene") is associated with time to recurrence in glioblastoma [4]. Specifically, all glioblastoma patients with exceptionally long time to recurrence following treatment had exceptionally low levels of the signature. This is consistent with the hypothesized reduced stemness in the malignant cells of those patients.

Second, we have also identified a "nonfibroblastic" version of the Slug-based EMT that only contains a subset of the 64 co-expressed genes [5], mainly SNAI2, DCN, LUM, COL1A1, COL1A2, COL3A1, COL6A3. This version is not necessarily associated with cancer; in fact the corresponding signature is differentially expressed even in normal tissues in a tissue specific manner [5], which may also induce cells to acquire stem cell properties: At one extreme, brain samples do not express the signature at all. At the other extreme, reproductive system samples do. This is consistent with the notion that stemness is most prominent in the cells of the reproductive system and least prominent in the highly differentiated cells of the brain. This lack of stemness in normal brain cells is also consistent with the above-mentioned association with prolonged time to recurrence in glioblastoma.

Many cancer types have been classified into subtypes using traditional bioinformatics methods, such as non-negative matrix factorization consensus clustering, from rich gene expression data. These techniques often lead to the identification of mesenchymal "subtypes." Our results suggest that these subtypes may simply be reflections of the fact that cancer cells have undergone a cancer mesenchymal transition reflected by the same core signature, though there are slight type-specific variations. In some cases there has been recognition that some cancer cells of "mesenchymal subtypes" may have undergone some transdifferentiation related to EMT. For example, the term "proneural-to-mesenchymal transition" [6] has been coined in glioblastoma. However, such transitions in solid cancers, including glioblastoma, may be amenable to being unified under the umbrella of the universal cancer mesenchymal transition signature. Interestingly, in glioblastoma the full signature is present but COL11A1 is not coexpressed as highly (perhaps because cancer cells do not encounter adipocytes [2]), while in other solid cancers COL11A1 is consistently the most reliable proxy of the signature.

We found most of the genes in the signature expressed by cancer cells, but not by stromal cells, at least in one xenograft model [2] (though some genes of the signature, such as MMP11, were not expressed by the cancer cells). This suggests that the signature is largely produced by the cancer cells undergoing mesenchymal transition as a result of contextual microenvironmental interactions. However, related versions of the cancer mesenchymal transition signature have often been labeled as "stromal," because the signature is fibroblastic and it is mostly found in the stroma following laser capture microdissection. Any presence of the signature in the tumor may then be interpreted as due to stromal infiltration. However, the truth may be exactly the opposite: the weak presence of the signature in the tumor may be genuine, as cancer cells start undergoing a mesenchymal transition. And the fully transdifferentiated, myofibroblast-like, cancer cells may well play the roles of cancer associated fibroblasts (CAFs) within the stroma [7]. Of course, the stroma contains fibroblasts from many other sources, some of which may express some of the genes in the signature. However, the remarkable co-expression of the genes in the particular mesenchymal transition signature may signify the presence of one of these sources as coming from the cancer cells themselves, resulting in a subpopulation of CAFs in the stroma that cannot

be easily distinguished from pure stromal cells. For example, the spindle-shaped cells expressing COL11A1 (the proxy of the mesenchymal transition signature) in Figure 2D of www.progenika.com/eu/images/stories/pdf/publicaciones/Barneo%20et%20al%20%20MIP%202006.pdf may conceivably be the transdifferentiated pancreatic cancer cells confused with desmoplastic stromal cells.

It is believed [7] that EMT in cancer reactivates early embryogenesis programs. For example Slug is known to be active in the neural crest and was also found to be required for the metastasis of transformed melanoma cells [8]. By analyzing publicly available datasets, we further found that the "fibroblastic" Slug-based cancer mesenchymal transition signature (including COL11A1, etc.) is also highly enriched in the set of genes expressed by invasive trophoblasts during implantation in the maternal tissue. This finding suggests that invasive cancer cells undergoing mesenchymal transition may reactivate pathways of invasive trophoblasts involved in placental formation. Expression of Slug during implantation has already been known [9]. These concepts are strikingly reminiscent of the "trophoblast theory of cancer" proposed as early as 1902 by John Beard [10].

The cancer mesenchymal transition signature is found in significant amounts only if cancer has exceeded a particular invasive stage. On the other hand, the signature may or may not be detected in samples that have reached or exceeded this invasive stage. The absence of the signature in a particular high-stage sample does not necessarily imply that the signature had not been present earlier in time or at some other neighboring location in the heterogeneous tumor. It is also unclear to what extent the underlying mechanism of mesenchymal transition is causal for invasion and metastasis. It is conceivable, however, that, at least in some cases, it plays a causal role, leading to the exciting possibility that its inhibition may lead to reduction of recurrence and metastasis applicable to multiple cancer types.

References

1. **Kim H, Watkinson J, Varadan V, Anastassiou D: Multi-cancer computational analysis reveals invasion-associated variant of desmoplastic reaction involving INHBA, THBS2 and COL11A1. *BMC Med Genomics* 2010, 3:51.**
2. **Anastassiou D, Rumjantseva V, Cheng W, Huang J, Canoll PD, Yamashiro DJ, Kandel JJ: Human cancer cells express Slug-based epithelial-mesenchymal transition gene expression signature obtained in vivo. *BMC Cancer* 2011, 11(1):529.**
3. **Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY, Brooks M, Reinhard F, Zhang CC, Shipitsin M *et al*: The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 2008, 133(4):704-715.**
4. **Cheng WY, Kandel J, Yamashiro DJ, Canoll P, Anastassiou D: Slug-based epithelial-mesenchymal transition gene signature is associated with**

prolonged time to recurrence in glioblastoma. Available from Nature Precedings <hdl.handle.net/10101/npre.2011.6544.1>. 2011.

5. **Cheng WY, Anastassiou D: A subset of co-expressed genes in Slug-based cancer mesenchymal transition signature remains coexpressed in normal samples in a tissue-specific manner. Available from Nature Precedings <hdl.handle.net/10101/npre.2012.6813.1>. 2012.**
6. **Skinner M: Cancer stem cells: TAZ takes centre stage. *Nat Rev Cancer* 2012, 12(2):82-83.**
7. **Hanahan D, Weinberg RA: Hallmarks of cancer: the next generation. *Cell* 2011, 144(5):646-674.**
8. **Gupta PB, Kuperwasser C, Brunet JP, Ramaswamy S, Kuo WL, Gray JW, Naber SP, Weinberg RA: The melanocyte differentiation program predisposes to metastasis after neoplastic transformation. *Nat Genet* 2005, 37(10):1047-1054.**
9. **Du F, Yang R, Ma HL, Wang QY, Wei SL: Expression of transcriptional repressor Slug gene in mouse endometrium and its effect during embryo implantation. *Appl Biochem Biotechnol* 2009, 157(2):346-355.**
10. **Gurchot C: The trophoblast theory of cancer (John Beard, 1857-1924) revisited. *Oncology* 1975, 31(5-6):310-333.**