Introduction

Structure and function of extracellular matrix (ECM) differ spatially in tissues, normal and cancer, adult and developing ones. Accordingly, the roles of ECM are diverse (Figure 1) and depend on ECM’s composition, structure and location. It has also been acknowledged that ECM makes diverse contributions at different stages of development together with its respective cellular receptors (1). In development, the key role of the interaction between ECM components and their cellular receptors has been confirmed in shaping the organism, but its role in cell migration and determination of cell fate may not be strictly determined (1). Thus, diverse migratory events may be integrin-dependent as well as integrin-independent and function of ECM-cell receptor interaction in controlling cell fate may differ from well-known signaling pathways affecting cell fate decisions, such as Wnt or Notch (1). To what extent are ECM and ECM-cell receptor interactions directly involved in carcinogenesis and tumor biology is not clearly determined. On one hand ECM-mediated adhesion is believed to affect outcome of the signaling (6,7), on the other hand a number of similarities between early embryo development and tumorigenesis (8) suggests possible indirect contribution of ECM to signaling pathways in tumors. To clarify this point, reliable methods that take into account complexity of intricate interactions within tissues but also impact of spatial distributions of tissue components are necessary.

ECM and gene expression profiling

Gene expression profiling is used to study various aspects of cancer. Attempts to apply it to characterize tumor stroma and even ECM or to stratify tumors based on ECM-related genes have also been published. For interpretation of results of such studies, methodological aspects of the approach need to be realized. Thus, while tumor ECM represents acellular tissue compartment, the information provided by gene expression profiling concerns genes containing entities, i.e., cells. The properties of ECM can thus only be deduced from expression of genes coding for ECM proteins and ECM-related molecules synthesized by cells, and adhesion molecules and receptors of ECM components expressed on cellular surfaces. Even if gene expressions correlate with expressions of their respective products, which is not always the case (9-12), various phenomena occurring in vivo may not be easily tracked by gene expression studies. Among those, activation of ECM-related enzymes, often requiring specific spatial distribution of mutually interacting compounds (13), processing of ECM components or their receptors, that are known to be subjected to numerous post-translational modifications (14,15), effects of factors...
Numerous genes were found to be differentially expressed among individual cell types isolated from tumor tissues (21). Whether genome-wide expression analysis of several genes within distinct regions of invasive tumors [e.g., (19,24)] therefore can hardly be considered.

Gene expression signatures identified in various tumor types in association with poor clinical outcome contain ECM-related genes (25-30) (Table 1). Genes coding for structural ECM components, however, differ among studies. Many of them encode proteins that are different from ECM predictors of outcome identified (immuno) histochemically. Comparison of signatures of poor survival after adjuvant chemotherapy from three different serous ovarian cancer datasets yielded a signature of 10 common genes containing genes coding for six structural ECM components: collagen type XI α1, collagen type V α1, collagen type VI α2, peristin (POSTN), thrombospondin 2 (THBS2) and versican (VCAN) as well as for LOX, TIMP3 and other two genes (30). Only some of these genes were identified as signature genes by others analyzing the same cancer, depending on the used classification model (29) (Table 1).

ECM gene cluster within a gene signature associated with resistance to first-line tamoxifen therapy of patients with metastatic breast cancer contained genes for collagen type I α1, fibronectin 1 (FN1), SPARC, tenascin-C (TNC) as well as for LOX and TIMP3. Out of them only expressions of genes for fibronectin 1, LOX and SPARC were associated with shorter patient survival. Expression of gene for collagen type I α1 was associated neither with prognosis nor with metastasis-free survival after adjuvant tamoxifen therapy (27).

There seems to be no complete consistency with regard to clinical significance of ECM constituents judged from gene vs. protein expression studies. Genes for low abundance and minor ECM protein chains often appear among signature genes, those coding for ECM proteins, believed to have a clinical impact based on (immuno)histochemical methods (collagen type I, tenascin, laminin γ2 chain), appear only rarely. According to mRNA in situ expression studies, positive signals for frequently appearing ECM-related signature genes (e.g., COL11A1, COL3A2, SPARC) appear in stromal cells rather than in epithelial/tumor cells appearing in tumors as a result of special properties of tumors, such as e.g., hyperpermeable vasculature (16,17) or effects of ECM protein cleavage products contribute to the final assembly of ECM. Accordingly, it may not be surprising that no significant correlation was achieved when ECM gene expression profiles were compared with histological classification of primary breast carcinomas with dense fibrous stroma, loose connective tissue stroma or mixed stroma types (18). Expression profiles of 278 ECM-related genes were examined by unsupervised hierarchal clustering in these samples, which resulted in a separation of four main branches (ECM1-4). Only one subgroup (ECM3) was mainly enriched for genes coding for structural ECM proteins. In spite of upregulation of genes coding for proteins involved in maintenance of connective tissue, including collagen type I in this subgroup, the tumors showed low or no expression of collagen type I by immunohistochemistry (18).

In majority of cases bulk tumors are homogenized to extract RNA. Spatial distributions of individual factors, including those related to ECM, are thus lost. Expression levels of genes are averaged throughout the tissue sample, although in situ hybridization studies documented expression of several genes within distinct regions of invasive tumors only (19). Whether genome-wide expression analysis of distinct regions of tumors is able to reliably map spatial distributions of genes is not completely clear. Distinct gene expression alterations were identified in epithelial and stromal compartments after their separation (20) and among individual cell types isolated from tumor tissues (21). Numerous genes were found to be differentially expressed between invasion front-containing tissues of moderate to well differentiated colorectal adenocarcinomas with high nuclear β-catenin expression and corresponding central tumor region (22). In contrast, genome-wide expression patterns of tumor invasion front did not reveal significant differences compared to inner tumor mass when unselected colorectal tumors were analyzed (23). It is also clear that without supplementary studies precise cellular origin of “characteristic” transcripts cannot be determined (19).

Biological consequences of expression of the same factor in distinct cell types and within distinct regions of invasive tumors [e.g., (19,24)] therefore can hardly be considered.

Table 1
### Table 1

<table>
<thead>
<tr>
<th>Type of malignancy</th>
<th>ECM-related signature genes identified by gene expression analysis</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epithelial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td>Poor survival after adjuvant chemotherapy</td>
<td>(30)</td>
</tr>
<tr>
<td></td>
<td>Higher hazard of death</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Resistance to first-line tamoxifen therapy</td>
<td>(27)</td>
</tr>
<tr>
<td></td>
<td>Worse survival probability, only grade III</td>
<td>(31)</td>
</tr>
<tr>
<td></td>
<td>Tumours, separation by hierarchical clustering of ECM-related genes</td>
<td>(18)</td>
</tr>
<tr>
<td>Breast cancer ECM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ECM3</td>
<td>Bad outcome</td>
<td>(26)</td>
</tr>
<tr>
<td>Hodgkin’s disease</td>
<td>Paediatric osteosarcoma Chemotherapy resistance</td>
<td>(25)</td>
</tr>
<tr>
<td>Paediatric osteosarcoma</td>
<td>Chemotherapy resistance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thrombosis</td>
<td>(21)</td>
</tr>
<tr>
<td>Non-epithelial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hodgkin’s disease</td>
<td>Bad outcome</td>
<td>(26)</td>
</tr>
<tr>
<td>Paediatric osteosarcoma</td>
<td>Chemotherapy resistance</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**BGN**, biglycan; **DCN**, decorin; **FBN1**, fibrillin-1; **FN1**, fibronectin 1; **HSPG2**, heparan sulfate proteoglycan 2 (perlecan); **POSTN**, periostin; **SPARC**, secreted protein acidic and rich in cysteine; **THBS**, thrombospondin; **VCAN**, versican; **TNC**, tenascin C.

(30,32,33). The products of these genes are often detected in stromal cells with no evidence of extracellular localization (30,31) or in interstitial tumor stroma, as documented for collagen type V in breast carcinomas (34). Such gene products can therefore hardly directly affect tumor cell behavior. Surprisingly, genes coding for receptors of ECM components, such as integrins, appear only rarely in the gene signatures (usually upregulated genes). Their lack does not seem to be due to methodology. Thus, one of the breast cancer subtypes discussed above (ECM1), was found to be characterized by upregulation of various genes coding for integrins and cell surface receptors related to immune infiltration. This group of tumors was associated with abundant lymphoid infiltration and upregulation of adhesion molecules by immunohistochemistry (18), indicating that appearance of novel cell adhesions, typical for diseased as compared to healthy tissue, can indeed be reflected by gene expression profiles. In contrast, in the subgroup characterized by gene expression signature enriched mainly in genes for structural ECM components (ECM3), no transcripts for integrin receptors appeared among representative genes (18). In extended study (ECM-related gene list consisted of 738 instead of 278 genes), 58 genes emerged to be the most relevant in determining the ECM3 subtype (31). Genes for **THBS2**, **SPARC**, **FN1**, collagen VI α1, collagen V α2, collagen V α1, TIMP3 and collagen I α2 were included in the most influential ECM3 genes. Only two of the 58 genes coded for integrin receptor chains—**ITGB1** and **ITGBL1** (31). The respective products of these genes do not belong to main adhesion receptors of collagens or laminins, but mainly of vitronectin, at least for integrin β5 chain.

Experimental evidence suggests role of ECM in response to therapeutic effects, however, the absence of genes for receptors of ECM proteins among signature genes separating tumors with different clinical outcomes raises a question about significance of direct ECM-tumor cell interactions for treatment outcome. Interactions of tumor cells with newly expressed ECM products can therefore be disclosed from gene expression studies. ECM is highly related to protein changes; therefore, gene expression data corresponding to mRNA levels might not truly represent the properties of ECM. Certainly, proteomic
studies would be more appropriate to characterize ECM and its changes caused by diseases. In spite of challenges related to proteomic analysis of ECM (35,36), proteomic studies identifying cancer related ECM signatures start to appear in the literature (37).

Acknowledgements

Technical assistance of Ing. E. Čunderlíková is highly acknowledged.

Footnote

Conflicts of Interest: The author has no conflicts of interest to declare.

References


