The Impact of Gene Expression Patterns in Breast Cancer

Therese Sørlie*


In our 2001 article highlighted here, my colleagues and I demonstrated for the first time the clinical impact of the main 5 “intrinsic” molecular subtypes of breast carcinomas. Shortly before, we had published the identification of unique “molecular portraits” of human breast tumors based solely on variation in gene expression patterns (1). The pervasive order and interpretation of these patterns indicated that the tumors could be classified into distinct biological groups based on overall molecular commonalities. The basal-like, ERBB2 (erb-b2 receptor tyrosine kinase 2)-positive, and normal-like subtypes all showed low or no expression of the estrogen receptor (ER) and other genes characteristic of luminal epithelial cells and had additional features specific for each of the 3 subtypes. For example, basal-like tumors showed high expression of the basal cytokeratins 5 and 17 and epidermal growth factor receptor and represented the predominant breast cancer subtype among BRCA1 (breast cancer 1) carriers (2). Luminal subtypes A and B, on the other hand, expressed ER and luminal-related genes but were further discriminated from each other by expression of a cluster of genes whose coordinated function was unknown.

Whether these groups were clinically meaningful was yet unexplored. Many studies had provided data on molecular prognostic markers. Those studies, however, were based on single genes or proteins. Nevertheless, most prognostic factors for breast cancer were of clinicopathologic nature and hence did not capture the molecular heterogeneity of tumors. At that time, we had the opportunity to analyze breast tumor biopsies from patients enrolled in a prospective neoadjuvant study evaluating predictive factors for response to doxorubicin monotherapy (3). The study design allowed us to investigate the prognostic value of the gene expression–based molecular subtypes in a homogenously treated patient cohort. Interestingly, univariate survival analysis showed a highly significant difference in overall survival and relapse-free survival between the subtypes. The basal-like and ERBB2+ tumors were associated with the shortest survival times. Accordingly, diversity in gene expression across breast tumors reflected biological diversity, which influenced their clinical behavior. This was the first study that provided evidence for molecular heterogeneity within ER+ tumors and suggested at least 2 luminal subtypes of breast cancer. The estimated survival times for patients with luminal B tumors were significantly shorter than for those with luminal A tumors. This was not strictly correlated to expression of the ER protein, as determined by ligand binding assay. It has since become clear that luminal B tumors are more proliferative, a trait they share with basal-like tumors, and that the survival differences between luminal subtypes might be due to variation in response to endocrine therapy (4).

At that time, one of the unresolved questions was whether specific molecular properties of a tumor could predict its likelihood to respond to a given therapy. The introduction of DNA microarrays provided an unprecedented opportunity to systematically explore expression patterns and functions of thousands of genes simultaneously. There were great expectations that a comprehensive survey of gene expression patterns in repeated samples from the same tumor, separated by several weeks of chemotherapy treatment, would not only shed light on the molecular mechanisms for response and resistance but also provide a useful clinical predictive marker. Instead, by using a list of genes that robustly reflected the intrinsic properties of the tumors, we uncovered the uniqueness and homogeneity of the molecular portraits. The selection of the intrinsic genes was a key prerequisite for arriving at a robust taxonomy, since their distinct expression profiles would reflect which phenotypic relationships would protrude in the clustering pattern. The intrinsic gene set consisted of genes whose expression levels varied significantly from tumor to tumor but varied relatively little between successive samples of the same patient’s tumor. Using variable selection tools, a minimized gene set has since been built into a classifier termed PAM50 and optimized for reverse-transcriptase quanti-
tative PCR and the Nanostring digital technology platform (Prosigna assay) (5).

The intrinsic subtypes provide prognostic information and were implemented in international guidelines for treatment of early breast cancer such as the recommendations from the St. Gallen Consensus. Their predictive power is less profound, but this classification provides a basis for further refinement to identify predictive biomarkers for subgroups of tumors. The search for such markers for response to cancer treatment is as timely as ever using today’s deep sequencing technologies. The fundamental biological nature of the intrinsic molecular subtypes of breast cancer has since been repeatedly validated and supported by integrated molecular analyses across several molecular layers, including somatic mutation patterns (6). In retrospect, the intrinsic subtypes have become an established terminology, and they have provided an increased understanding of the biology of breast cancer and the signaling pathways involved in their progression. An important implication is the stratification of study populations based on subtypes in clinical trials, a necessary adaptation to modern personalized medicine.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: No authors declared any potential conflicts of interest.

References