Integrative Meta-Analysis of Gene-Expression Profiles in Breast Cancer: Toward a Unified Understanding of Breast Cancer Sub-typing and Prognosis Signatures

Christos Sotiriou1,2, Pratyaksha Wirapati1, Susanne Kunke1, Pierre Farmer2, Sylvain Prognavand1, Benjamin Haibe-Kains1,3, Christine Desmedt1, Thierry Sengstag1, Frédéric Schütz1, Darlene R. Goldstein1, Mauro Delorenzi1, Martine Piccart1

Jules Bordet Institute, Centre Institute of Bioinformatics, Université Libre de Bruxelles

Abstract

Background: Breast cancer sub-typing and prognosis have been extensively studied by gene expression profiling, revealing disparate signatures with little overlap in their constituent genes. The biological roles of individual genes in a signature, the equivalence of several signatures and their relation to conventional prognostic factors are still unclear.

Methods: Here we undertook a comprehensive meta-analysis of publicly available gene-expression and clinical data to (1) identify stages of breast cancer (2) to define expression modules (especially NFRM 504) that were highly correlated with survival, and (3) to evaluate the impact of these modules on survival. The concept of co-expression modules (comprehensive lists of genes with highly correlated expression) was used extensively to reveal the common thread connecting molecular sub-typing and several prognostic signatures, as well as to reveal potential predictive markers.

Results: Breast tumors were consistently grouped into three main subsets corresponding roughly to ER+/ERBB2- (luminal), ERBB2+ and ER- (basal) tumors. ERBB2+ tumors showed an intermediate estrogen receptor mediator score which is not obvious from the traditional ER and ERBB2 marker status combination. Both, ER+/ERBB2- and ERBB2+ subtypes were characterized by high proliferation, whereas the ER- subtype appeared to be more heterogeneous. Using our meta-analytical approach we were able to identify 504 genes which were significantly associated with survival. Of the 244 prognostic genes, 184 were strongly co-expressed with proliferation, 14% with ER, 0.6% with ERBB2, 2.7% with tumor invasion, 1.5% with immune response and 16% with none of our co-expression modules.

All previously reported prognostic signatures examined in this meta-analysis (N=9) despite the disparity in their gene lists, shared similar information with regard to proliferation, with proliferation genes being the common driving force. They were all very useful for determining the risk of recurrence in the ER+ subgroup and much less informative for ER- and ERBB2+ disease. Combining the risk did not improve their performance.

Finally, in multivariate analysis nodal status and tumor size still remained independent prognostic information.

Conclusions: This meta-analysis unifies various results of previous gene-expression studies in breast cancer. It reveals connections between traditional prognostic factors, expression-based sub-typing and prognostic signatures, highlighting the important role of proliferation in breast cancer prognosis.

Key Messages

• All signatures show similar performance
• Proliferation is the common denominator (better quantification...!) or informative only in ER+ tumoral
• Still need stage information.

Acknowledgments

FNR, MEDIC Foundation, Breast Cancer Research Foundation (BCRF, Evelyn Lauder)