**Clinically relevant gene expression subsets deciphered from breast carcinomas classified by quantified expression levels of estrogen and progestin receptors**

James L. Wittliff1,2, William A. Stern1,2, Zohair R. Hamid1,2 and Michael W. Daniels3

1Department of Biochemistry & Molecular Genetics, 2Institute for Molecular Diversity & Drug Design, University of Louisville, Louisville, KY 40292 and 3Department of Biostatistics, University of Colorado at Aurora, CO

**Background:** Prior to standardization of IHC parameters by CAP/ASCO Guidelines in 2010 (Hammond et al., Arch Pathol Lab Med), ER/PR measurements were highly variable (e.g., differences in fixation/staining, observer scoring). While serving as the National Reference Laboratory for NCI supported cooperative clinical trials by the NSABP, SWOG, CALGB, NCCTG, ECOG and SECSG, both EIA and radio-ligand binding assays (LBA) were employed that quantified ER/PR content and activities. Using quantified ER/PR levels, we assessed relationships of qPCR expression of 85 candidate genes from 569 primary breast cancers (without selection bias) with clinical outcomes to decipher molecular signatures that predict risk of recurrence.

**Methods:** This retrospective study used a unique dataset of de-identified ER and PR biomarker results of 1220 primary cancers from patients with clinical follow-up, stored in an IRB-approved Database. ER/PR protein levels expressed as fmol/mg cytosol protein were quantified earlier from each carcinoma with FDA-approved kits using either enzyme immunoassay (n=751 EIA, Abbott Labs) or radio-labeled ligand binding assay (n=552 LBA, NEN/DuPont). Relationships between ER/PR content assessed by LBA compared to that from EIA were examined. Results obtained from 1988-1996 and de-identified clinical outcomes were examined using REMARK criteria in a CLIA licensed laboratory. Gene expression, ER/PR quantified by LBA or EIA, features of primary breast cancers and clinical outcomes were analyzed by univariable and multivariable Cox regressions, Kaplan Meier plots and LASSO with R software v3.6.0. Gene expression levels were validated by qPCR while molecular signatures were externally validated with SurvExpress. (Aguirre-Gamboa et al. PLoS One e742502013).

**Results:** Examination of patient survival by Kaplan-Meier analyses using cancers classified by quantified ER or PR (regardless of by EIA or by LBA) revealed RFS and OS were related to increased content. Relationships with Kd values were examined. Also, differences in ER/PR status of breast carcinomas were associated with RFS and OS of patients (i.e., ER+/PR+ was related to increased RFS and OS while ER-/PR- status was associated with poor RFS and OS). Univariable Cox regression of independent expression of the 85 genes identified 18 that predicted RFS while 22 genes predicted OS, confirmed by Kaplan-Meier Analyses. Multiple imputation by chained equations (‘*MICE’*, Zhang, Ann Transl Med. 2016) imputed missing qPCR values in order to utilize all 569 patients and 85 genes. With the ‘*SGL’* package in R, an optimal L2 penalty determined the best fit LASSO model (algorithm sets minimum size for each beta coefficient and its associated gene to be included in the final Cox survival models). The 16 gene subsets that were deciphered according to quantified ER, PR or ER/PR status predicted either RFS or OS or both reflecting the increased or decreased hazard for a one-unit change in a gene while controlling for other gene associations.

**Conclusion:** Expression of each gene in each of the 16 subsets exhibits a hazard ratio, when taken jointly, predicts either PFS or OS. Multidimensional portrayal of variation in the outcome PFS or OS is controlled by gene expression in the signature for an overall effect. Using expression results validated by qPCR, subsets of ER and/or PR related genes in primary lesions were identified that predict a patient's risk of recurrence and overall survival. Collectively, results suggest that gene expression profiles based on quantified ER/PR content of a biopsy serve as value-added biomarkers for assessing prognosis of breast carcinomas and identifying candidates for developing companion diagnostics.

Supported in part by grants to JLW from Phi Beta Psi Sorority Charity Trust, University of Louisville, Office of the EVPR, and a CTSP award, Commonwealth of Kentucky. WAS is the recipient of a fellowship from Bates College, ME.