Validation between gene expression measurement platforms

Serena Lee, Eveline Chen, Kevin Tran, Chunxiao Fu, Lili Du, William F. Symmans

ITERT Program

THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center

Making Cancer History®

Background

Treatments for breast cancer vary from endocrine therapy, chemotherapy, immunotherapy, to surgery. Dr. Symmans's lab works to determine the sensitivity to endocrine therapy using metastatic breast cancer biomarkers. The lab has developed an index to determine the sensitivity to endocrine therapy ($SET_{ER/PR}$)[1], which evaluates the gene expression of breast cancer related hormone receptors. $SET_{ER/PR}$ targets 18 genes correlating with estrogen and progesterone receptors, ESR1 and PGR. The lab also included the $PI3K_{ges}$, which measures the expression of the PI3K pathway activation due to mutations in PIK3CA.[2] This information can be used in clinical settings to determine if endocrine therapy will be effective in treating breast cancer patients. Different technology methods have been developed to measure gene expressions. The lab used the Quantigene Plex (QGP) platform[3] (quantitative gene expression through hybridization) to develop a 31 and 40 target gene panel. QGP31 has 31 target genes and can detect the $SET_{ER/PR}$ index, while QGP40 has 9 additional target genes and is able to detect both $SET_{ER/PR}$ and $PI3K_{ges}$. The Illumina MiSeq platform (RNA sequencing) was used to develop SET4, which includes both $SET_{ER/PR}$ and $PI3K_{ges}$.

The aim of this study is to validate the calibration of each platform's measured index. By doing so, labs can provide more consistent conclusions regarding patient's sensitivity regardless of the platform being used.

Methods

Formalin-fixed and paraffin-embedded (FFPE) breast cancer tissue blocks (n=20) were collected from individual breast cancer patients. The FFPE blocks were cut into 5 or 10 micron sections and extracted for RNA using Norgen FFPE RNA Purification kit. The gene expression of each sample was measured using 3 different platforms: The Affymetrix QuantiGene 31-Plex (QGP31), 40-Plex (QGP40), and SET4. QGP31 and QGP40 were measured using Luminex MAGPIX, and SET4 was sequenced and measured on the Illumina Miseq platform. The gene expression measurements were input into an established script written in the programming language R to calculate the $SET_{ER/PR}$ from all platforms and $PI3K_{ges}$ values from QGP40 and SET4. The equations for SETER/PR index and $PI3K_{ges}$ index are

$$SET_{ER/PR} = \frac{\sum_{i=1}^{18} T_j}{18} - \frac{\sum_{i=1}^{10} R_j}{10} + 2$$

$$PI3K_{ges} = \frac{\sum_{i=1}^{10} P_i}{10} - \frac{\sum_{i=1}^{10} R_j}{10}$$

where T_i denotes the log2 expression of the eighteen SET biomarkers, R_j the log2 expression of the ten reference genes, and P_i the log2 expression of the ten $PI3K_{ges}$ biomarkers. These values were graphed against the same measurements of other platforms. A line of best fit with the equation, R^2 value, and concordance correlation coefficient (CCC) was generated.
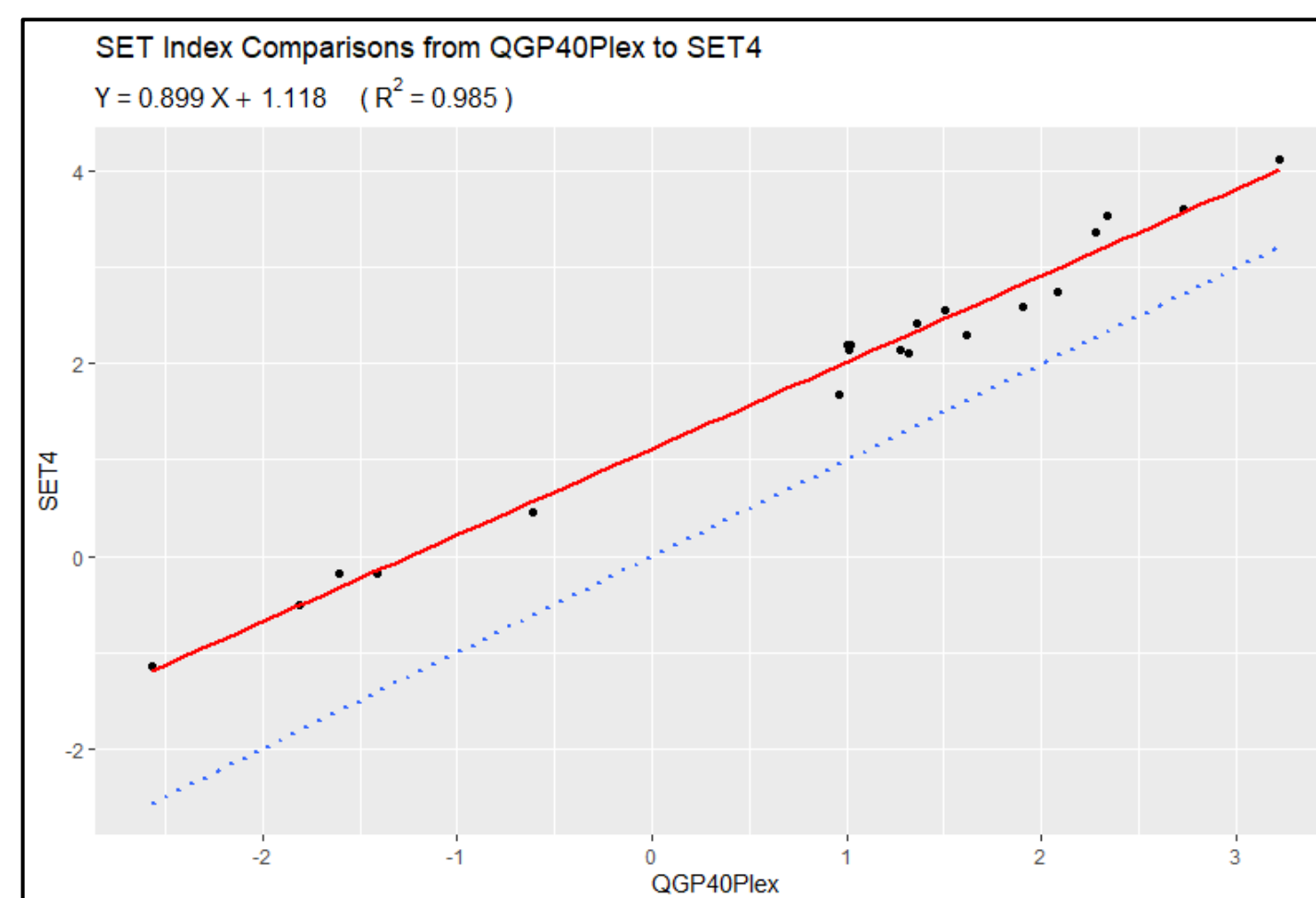
Results

Figure 1a. Linear regression (red, solid) plot of $SET_{ER/PR}$ for QGP40 and SET4 in comparison to the line of perfect agreement (blue, dotted).
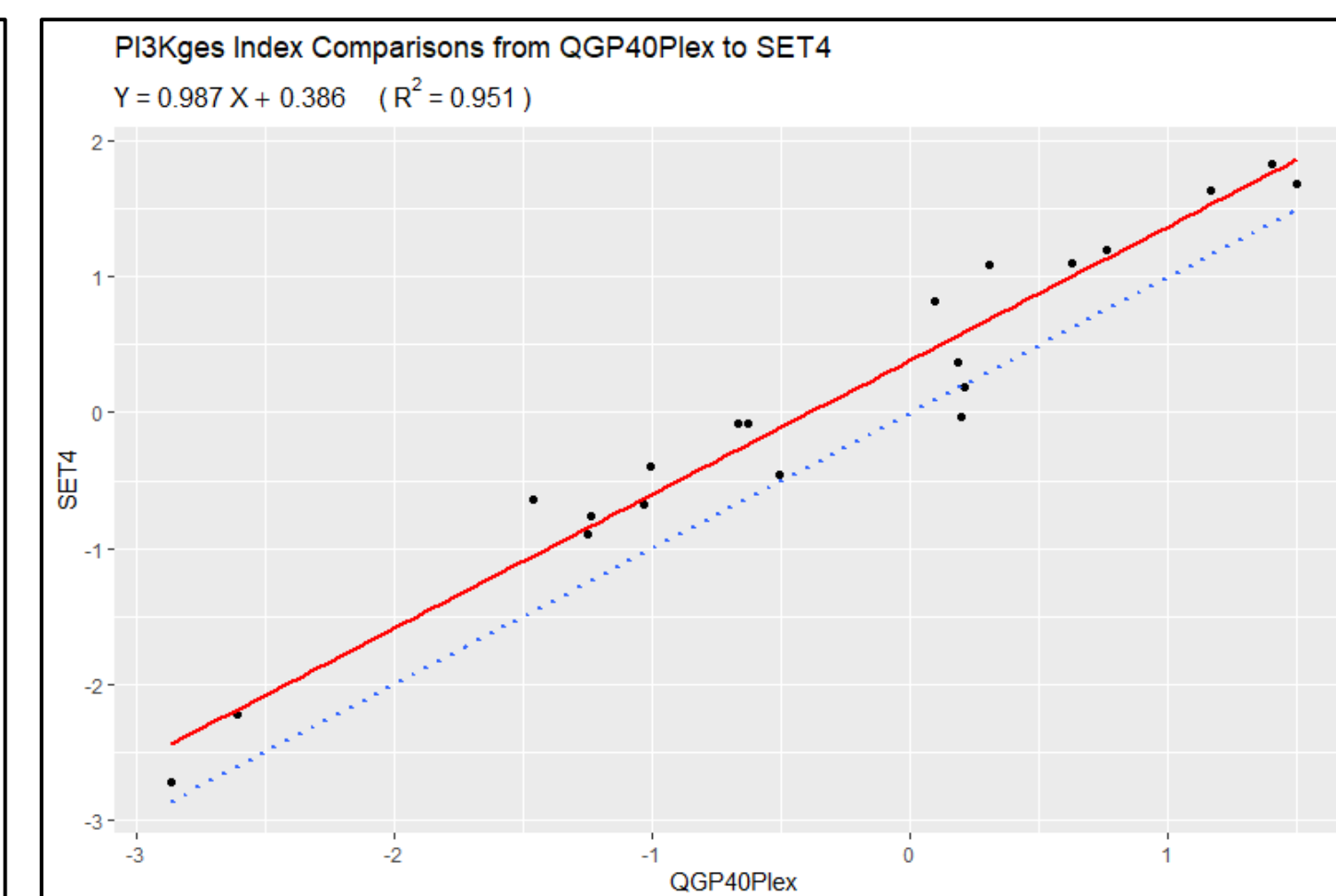

Figure 1b. Linear regression (red, solid) plot of $PI3K_{ges}$ for QGP40 and SET4 in comparison to the line of perfect agreement (blue, dotted).
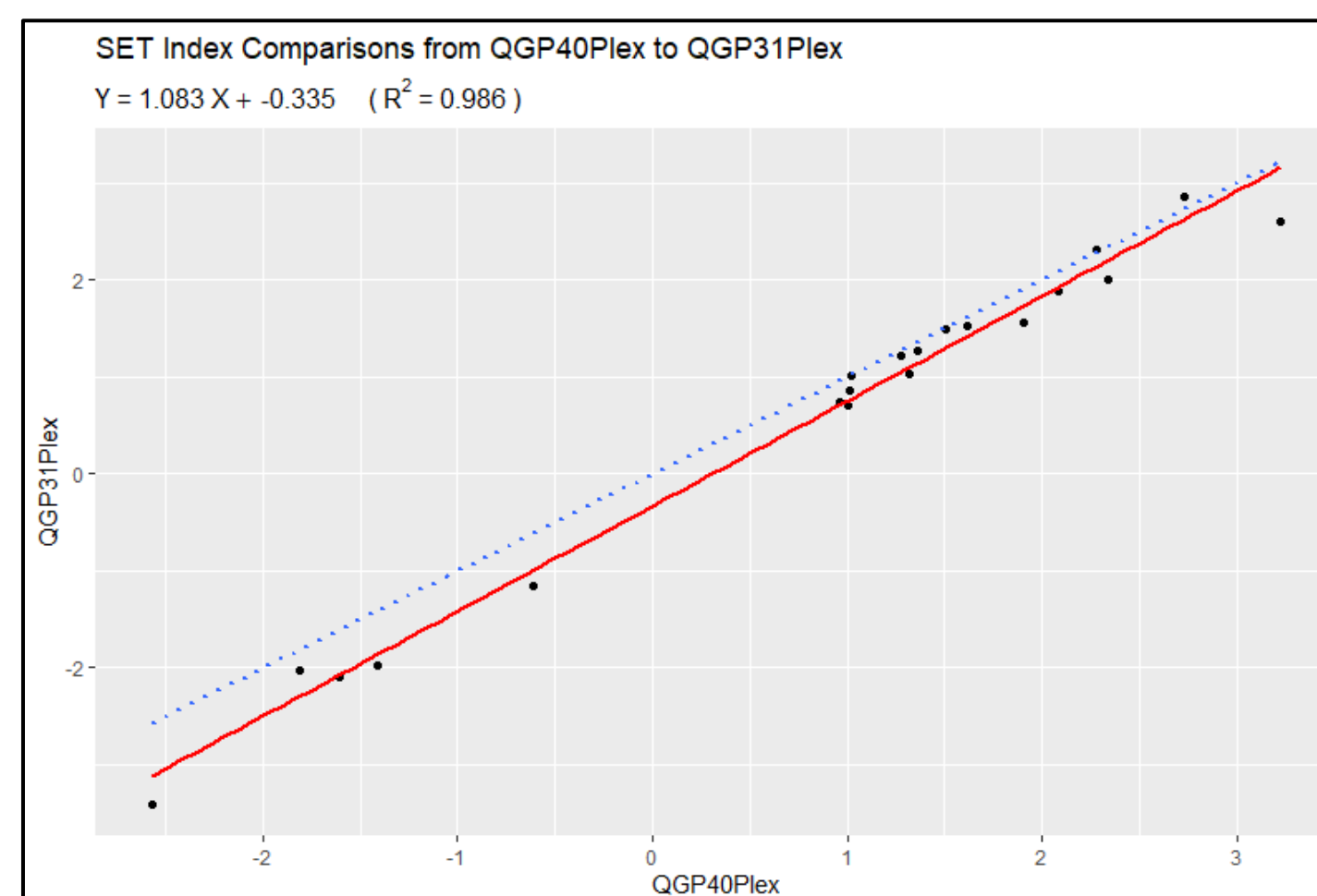

Figure 1c. Linear regression (red, solid) plot of $SET_{ER/PR}$ for QGP40 and QGP31 in comparison to the line of perfect agreement (blue, dotted).
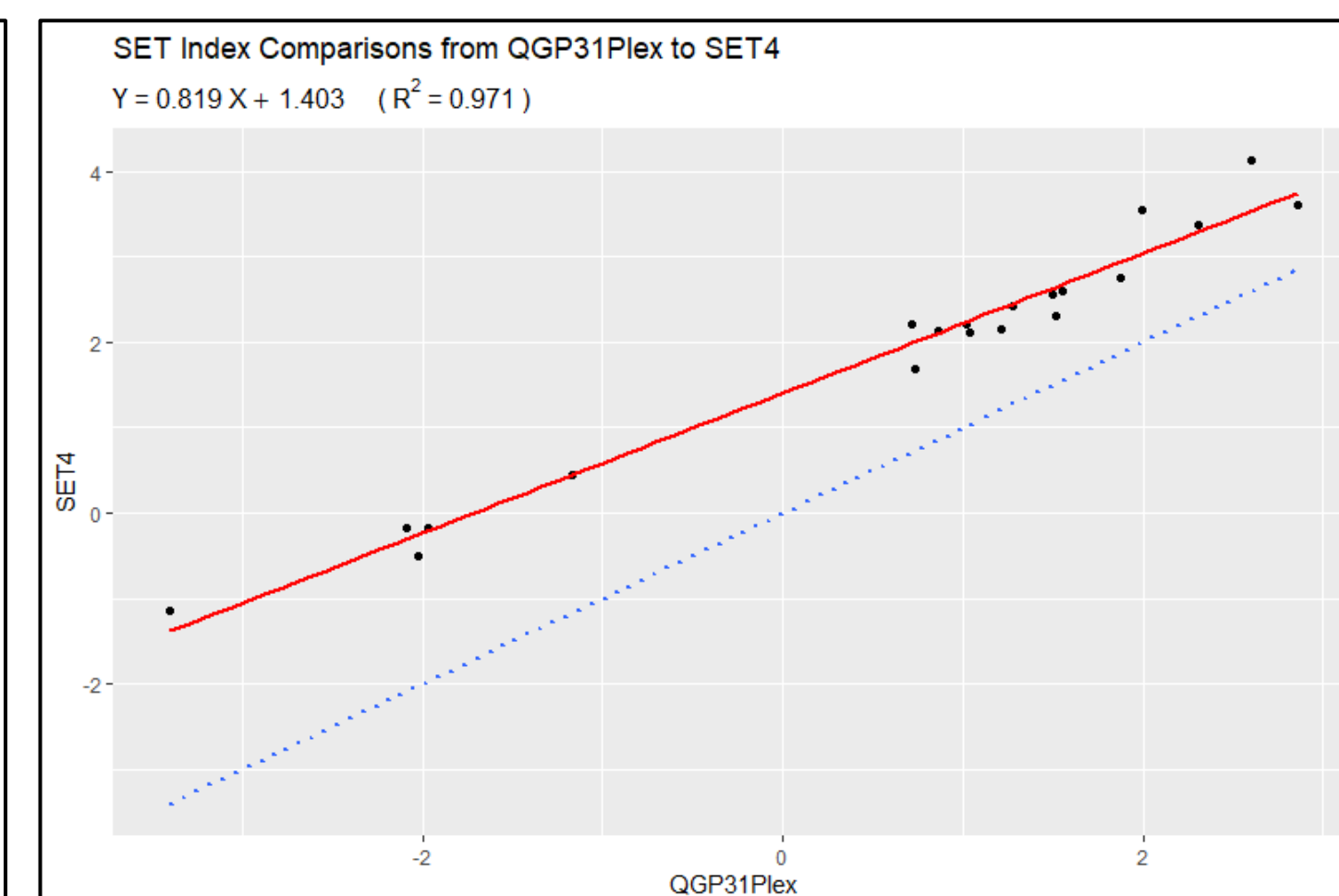

Figure 1d. Linear regression (red, solid) plot of $SET_{ER/PR}$ for QGP31 and SET4 in comparison to the line of perfect agreement (blue, dotted).

Comparisons	R2	CCC	95% LCI	95% HCI
SET QGP40 vs SET4	0.985	0.801	0.665	0.886
PI3Kges QGP40 vs SET4	0.951	0.925	0.843	0.965
SET QGP40 vs QGP31	0.986	0.977	0.951	0.989
SET QGP31 vs SET4	0.971	0.727	0.562	0.837

Discussion

$SET_{ER/PR}$ is a reliable and robust analytical system to accommodate the various platforms used in laboratory molecular testing. Through the $SET_{ER/PR}$ index comparison of 20 individual breast cancer FFPE derived RNA samples between QGP31, QGP40, and SET4, $SET_{ER/PR}$ was observed to be highly concordant (CCC 0.727-0.977) between all three platforms. This is also true with the PI3Kges index comparison between QGP40 and SET4 (CCC 0.925). Index pairs involving SET4 were less concordant than the QGP40-QGP30 pair. This is possibly due to the difference in assay technology. The high concordance and R^2 (0.951-0.986) indicate that the index values could be accurately calibrated between platforms. Equations from the linear regression indicate that the $SET_{ER/PR}$ for QGP31-SET4 and QGP40-SET4 require larger adjustments for standardization between platforms.

Conclusion

The high concordance for each index comparison show that the calculated index of the platforms can be comparable with calibration. Labs will be able to use any platform to provide accurate conclusions on a patient's sensitivity to endocrine therapy. Future direction of the study can be done with interlaboratory testing of a larger cohort size using the adjusted algorithm.

References

1. Sinn BV, Fu C, Lau R, et al. SETER/PR: a robust 18-gene predictor for sensitivity to endocrine therapy for metastatic breast cancer. *npj Breast Cancer*. 2019;5(16). https://doi.org/10.1038/s41523-019-0111-0
2. Fu C, Marczyk M, Samuels M, *et al.* Targeted RNAseq assay incorporating unique molecular identifiers for improved quantification of gene expression signatures and transcribed mutation fraction in fixed tumor samples. *BMC Cancer*.2021;21(114). https://doi.org/10.1186/s12885-021-07814-8
3. Lau R, Du L, Fu C, et al. Technical validity of a customized assay of sensitivity to endocrine therapy using fixed breast cancer tissue sections. *Clin Chem*.2020;66(7):934-945. doi:10.1093/clinchem/hvaa105.