Development of a Prognostic Test for Breast Cancer Recurrence Using a Highly Efficient Molecular Profiling Platform

Paul Capewell[1], Elizabeth Mallon[2], Joseph Loane[2], Jeff Robson[1], Harper VanSteenhouse[1], Joanne Edwards[3], Ditte Andersen[1] 1. BioClavis Ltd 2. NHS Greater Glasgow and Clyde 3. University of Glasgow

Abstract Although the clinical utility of gene expression signatures for breast cancer has been established across multiple prospective trials for several commercially available testing platforms, broad access to these tools is limited due to international sample delivery, prohibitive costs and limited laboratory resources. The purpose of our study was to identify a set of transcripts and clinical factors that associated with the likelihood of recurrence in HR+ breast cancer patients that could provide a more cost-effective testing service. TempO-Seq™ is a highly cost effective targeted sequencing platform capable of a wide array of analyses of genomic panels(1), ranging from the identification of somatic mutations to whole transcriptome profiling. The methodology accommodates various sample types, including more challenging targets, such as cultured cells, fresh, frozen or fixed tissue lysates, and has been successfully used for the transcriptomic analysis of decades-old archived FFPE tissue and <1mm^2 FFPE tissue input(2). The technology avoids the typical requirements of extraction, reverse transcription or pre-amplification. Conversely, standard methods to analyse gene expression using FFPE samples are time-consuming, expensive and require many sections of tissue to obtain RNA yields appropriate for downstream analysis (even with pre-amplification). TempO-Seq bypasses these limitations to provide higher precision measurements of gene expression from crude lysates.

Methods:

We used the TempO-Seq targeted expression technology platform(1) to perform whole transcriptome profiling of crude FFPE lysates made from hormone receptor-positive (HR+) breast cancer samples (n=292). The collection was provided by the Greater Glasgow and Clyde biorepository (UK) and contained patients with 5–10+ years follow-up that had received heterogeneous adjuvant treatment (Table 1). Patient outcomes and both clinical and pathological variables were also provided. Machine learning-based expression analysis of a training and discovery set containing a subset of 235 samples revealed 16 transcripts with recurrence-associated expression. We subsequently trained a regression-based risk model for 10-year recurrence using this sample set and assessed its prognostic potential with an independent panel of samples not used for training or factor identification (n=57). A regression-based risk model for 10-year recurrence was also trained and tested that included node status as a binary factor.

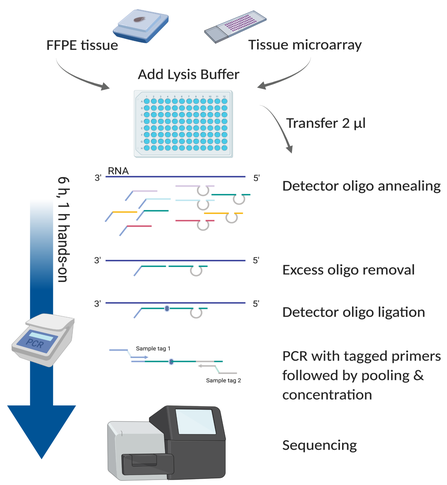

Figure 1. Overview of the TempO-Seq Technology Pipeline for FFPE and tissue microarrays.

Results

Our TempO-Seq-based analysis identified 16 transcripts with recurrence-associated expression in the HR+ breast cancer discovery cohort. Many of these transcripts have previously been shown to be prognostic for breast cancer outcome and are implicated in cell cycle control, the oestrogen pathway, and tumour suppression (Figure 2). When used to construct a regression-based risk model for 10-year recurrence, the expression of these transcripts provided prognostic information in node-negative (LRχ2 =39.66, p<0.001; Hazard Ratio=3.27) and node positive patients (LRχ2=29.24 p<0.001; Hazard Ratio=1.82; Figure 3A). This was also shown in the independent test group in node negative (LRχ2=26.14, p<0.001; Hazard Ratio=2.32) and node positive patients (LRχ2=41.50, p<0.001; Hazard Ratio=2.16). Individual recurrence risk estimates allowed patients to be grouped into low and high risk cohorts, with significantly different (p<0.05) Kaplan-Meier curves for recurrence-free survival and cancer-associated mortality (Figure 2A). Including node status as a binary factor (N0 or N+) allowed for the construction of a unified model for recurrence in HR+ patients that could differentiate low and high risk groups for recurrence that were also prognostic for cancer-related mortality (p<0.05; Figure 3B). The prognostic information provided by the model was found to be independent of the UK-PREDICT V2 tool (Figure 4), suggesting complementary applications for the technology. Ongoing refinement of the TempO-Seq-based assay includes validation with additional patient samples, comparison to other prognostic scores to aid interpretation, and commercialisation in an ISO 15189-accredited service laboratory.

Table 1. HR+ Cohort Sample Details

Characteristic	Value	Node Negative (n=172)		Node Positive (n=120)	
		Train	Test	Train	Test
Adjuvant Treatment	Chemotherapy	18 (12%)	8 (35%)	48 (56%)	22 (65%)
	Hormone Therapy	119 (80%)	16 (70%)	73 (85%)	29 (85%)
	No Adjuvant Therapy	2 (1%)	0 (0%)	0 (0%)	0 (0%)
Grade	G1	40 (27%)	7 (30%)	14 (16%)	5 (15%)
	G2	77 (52%)	10 (43%)	47 (55%)	21 (62%)
	G3	32 (21%)	6 (26%)	25 (29%)	8 (24%)
Tumour Size	≤ 2cm	84 (56%)	11 (48%)	26 (30%)	10 (29%)
	> 2 cm	65 (44%)	12 (52%)	60 (70%)	24 (71%)
HER2	Positive	9 (6%)	2 (9%)	11 (13%)	5 (15%)
Outcomes	Early Recurrence (0–5 y)	12 (8%)	2 (9%)	22 (26%)	6 (18%)
	Late Recurrence (5–10 y)	7 (5%)	1 (4%)	6 (7%)	2 (6%)
	Cancer-related Death	21 (14%)	2 (9%)	28 (33%)	6 (18%)

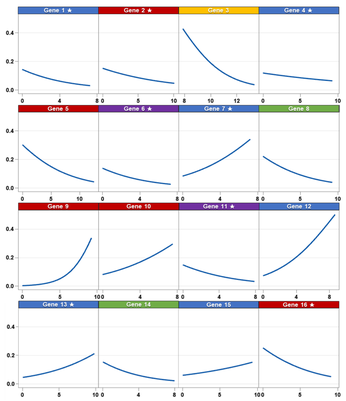

Figure 2. Main effects plots for the transcripts identified in the study, demonstrating the impact of expression on the likelihood of recurrence in a patient. The x-axis for each plot shows normalized expression and the y-axis shows the probability of recurrence. Panel headings are coloured to indicate evidence in the literature for involvement in the cell cycle (blue), oestrogen pathway (red), tumour suppression (yellow), cell stress responses (green) or transport (Purple). ★ indicates where a gene has previously been shown to be prognostic for outcome in breast cancer.

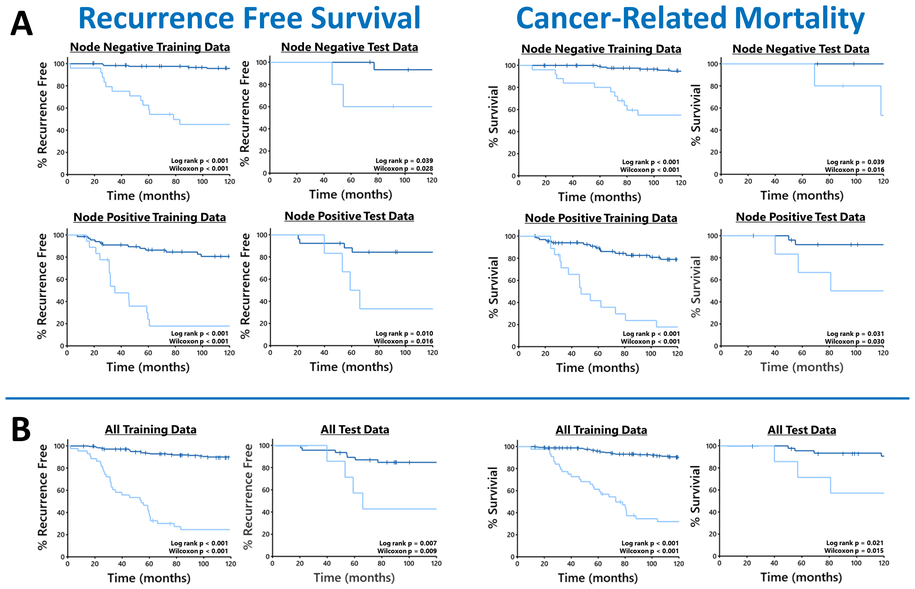

Figure 3. (A) Right-censored Kaplan-Meier curves for recurrence free survival (left) and cancer-related mortality (right) in node negative and node positive patients. Patients were assigned to low or high risk groups using a 10-year risk of recurrence estimated using a regression model constructed from TempO-Seq-derived expression of 16 transcripts and an optimal cut-off of 32% that was automatically determined from the training data. **(B)** Kaplan-Meier curves for high and low risk patients using an expanded model of recurrence risk that included node status as a binary factor (N0 or N+) in addition to the expression of the 16 transcripts. Survival curves were compared using Log-Rank and Wilcoxon tests.

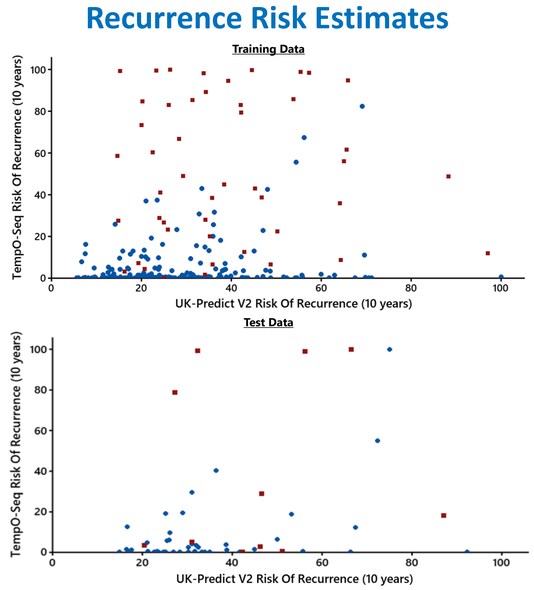

Figure 4. Scatter plots showing the likelihood of 10-year recurrence free survival estimated using UK-PREDICT V2 (x axis) and the TempO-Seq-derived recurrence risk model (y axis) for matched samples in the training data (upper) and test set (lower). Patients that experienced recurrence within 10 years are shown in red and recurrence-free patients are shown in blue.

Conclusion

Our novel prognostic signature and testing method based on TempO-Seq profiling was developed using whole transcriptome analysis of a heterogeneously treated cohort in an adjuvant setting. This approach overcomes many of the technical limitations inherent to developing genetic tests (e.g., does not require extraction) and cost barriers typically encountered with traditional molecular approaches such as qPCR, RNA-Seq or molecular barcoding. The high throughput test can rapidly generate results from large numbers of patient samples for efficient turnaround times. More widely, the study demonstrates that TempO-Seq technology can be used for accelerated diagnostic panel development and clinical research. The described approach to estimate the risk of breast cancer recurrence will be further refined and validated in prospectively designed clinical trials.

Disclosure Statement: PC has no conflicts of interest to declare.

References: (1) Yeakley, J. M., Shepard, P. J., Goyena, D. E., Vansteenhouse, H. C., Mccomb, J. D., and Seligmann, B. E. (2017) "A trichostatin A expression signature identified by TempO-Seq targeted whole transcriptome profiling." PLoS ONE 12:e0178302
(2) Trejo, C. L., Babić, M., Imler, E. Gonzalez, M., Bibikov, S. I., Shepard, P. J., VanSteenhouse, H. C., Yeakley, J. M., and Seligmann, B. E. (2019) "Extraction-free whole transcriptome gene expression analysis of FFPE sections and histology-directed subareas of tissue." PLoS One 14:e0212031

Final Publication Number: 54P