INTRODUCTION

Tumors frequently contain cancer stem cells (CSCs) or tumor-initiating subpopulations, with an ability to self-renew and regenerate all cell types within the tumor. Basal-like breast cancers exhibit features of CSCs, including expression of surface markers, even though these cells are rare. Given the role of CSCs in the recurrence and spread of cancer, there is an urgent need to develop new therapeutic agents that target CSCs. Development of CSC-targeted drugs will be greatly facilitated by determination of drug response. Defining the CSCs in tumors is gene expression and alternative splicing patterns in CSCs that are not observed in fibroblasts may provide valuable new CSC-specific markers.

METHODS

Gene expression and alternative splicing whole transcriptome microarray strategies were used to identify differentially expressed genes and exons between 10 Basal human breast cancer lines and a combination of 12 Luminal and 9 fibroblast cell lines. Genes signatures for Tumor Initiating (TI) and Epithelial–Mesenchymal Transition (EMT) were applied to resolve the patterns of expression.

Q-PCR analysis was conducted to validate candidate CSC genes from the above discovery method. Differential expression between Basal, Luminal, and Fibroblast cell lines, tumorsphere culture and human tumor xenograft experiments were evaluated.

RNA microarrays from n=178 TNBC patients (GEO GSE25066) were evaluated with Time to Recurrence (0-7 years) models from the Mesenchymal-High, Epithelial-Low, Fibroblast-Low genes in this study. The accuracy of the model was assessed by ROC curves.

In a separate discovery effort, alternative splicing discriminators were assembled according to the FlowChart:

- Assess differential gene expression data
- Refine alternative splicing prediction
- Classify samples into Groups
- Validate results based on analysis criteria

RESULTS

Expression levels of 11 genes were higher and 24 genes were lower in the Basal cell lines versus Luminal or fibroblastic cell lines. Comparison of Basal cell lines to the Luminal/Fibroblast cell lines identified 36 cassette exons that were included, and 26 that were excluded in Basal cell lines. Also, 19 genes were upregulated in Basal cell lines compared to the other groups as detected by Q-PCR.

Interestingly, the 19-gene model defined the Triple Negative Breast Cancer patients that were Likely to Recur under standard chemotherapy with a p = 1.90e-03 and AUC 0.723.

SUMMARY

- Differentially expressed genes and alternatively spliced exons were discovered as markers of cancer stem cells in breast cancer cell lines, distinct from fibroblasts.

- Gene and exon marker sets that distinguish CSCs versus fibroblasts and may be instructive in identifying patients that recur early in Triple Negative Breast Cancer.

- The CSC-associated RNA signatures identified here will be further refined to develop new CSC-specific diagnostic biomarkers to stratify breast cancer patients and monitor response to novel CSC-targeted therapies.