Using microarray-based subtyping methods for breast cancer in the era of high-throughput RNA sequencing

**Using microarray-based subtyping methods for breast cancer in the era of high-throughput RNA sequencing**

Breast cancer is a highly heterogeneous disease that can be classified into multiple subtypes based on the tumor transcriptome. Most of the subtyping schemes used in clinics today are derived from analyses of microarray data from thousands of different tumors together with clinical data for the patients from which the tumors were isolated. However, RNA sequencing is gradually replacing microarrays as the preferred transcriptomics platform, and although transcript abundances measured by the two different technologies are largely compatible, subtyping methods developed for probe-based microarray data are incompatible with RNA sequencing as input data. Here we present an RNA sequencing data processing pipeline, which relies on the mapping of sequencing reads to the probe set target sequences instead of the human reference genome, thereby enabling probe-based subtyping of breast cancer tumor tissue using sequencing-based transcriptomics. By analyzing 66 breast cancer tumors for which gene expression was measured using both microarrays and RNA sequencing, we show that RNA sequencing data can be directly compared to microarray data using our pipeline. Additionally, we demonstrate that the established subtyping method CITBCMST (Guedj et al., 2012), which relies on a 375 probe set-signature to classify samples into the six subtypes basL, lumA, lumB, lumC, mApo, and normL, can be applied without further modifications. This pipeline enables a seamless transition to sequencing-based transcriptomics for future clinical purposes.

**General information**

**State:** Accepted/In press  
**Organisations:** Department of Biotechnology and Biomedicine, Cancer Genomics, Department of Bio and Health Informatics, Copenhagen University Hospital  
**Contributors:** Pedersen, C. B., Nielsen, F. C., Rossing, M., Olsen, L. R.  
**Number of pages:** 21  
**Publication date:** 2018  
**Peer-reviewed:** Yes

**Publication information**

**Journal:** Molecular Oncology  
**ISSN (Print):** 1574-7891  
**Ratings:**  
BFI (2018): BFI-level 1  
Web of Science (2018): Indexed yes  
BFI (2017): BFI-level 1  
Scopus rating (2017): CiteScore 5.17 SJR 2.526 SNIP 1.119  
Web of Science (2017): Impact factor 5.264  
Web of Science (2017): Indexed yes  
BFI (2016): BFI-level 1  
Scopus rating (2016): CiteScore 5.25 SJR 2.461 SNIP 1.164  
Web of Science (2016): Impact factor 5.314  
Web of Science (2016): Indexed yes  
BFI (2015): BFI-level 1  
Scopus rating (2015): CiteScore 5.65 SJR 2.989 SNIP 1.301  
Web of Science (2015): Impact factor 5.367  
Web of Science (2015): Indexed yes  
BFI (2014): BFI-level 1  
Scopus rating (2014): CiteScore 5.3 SJR 2.991 SNIP 1.338  
Web of Science (2014): Impact factor 5.331  
BFI (2013): BFI-level 1  
Scopus rating (2013): CiteScore 5.94 SJR 3.644 SNIP 1.415  
Web of Science (2013): Impact factor 5.935  
ISI indexed (2013): ISI indexed yes  
Scopus rating (2012): CiteScore 5.59 SJR 3.325 SNIP 1.498  
Web of Science (2012): Impact factor 6.701  
ISI indexed (2012): ISI indexed yes  
Scopus rating (2011): CiteScore 3.86 SJR 2.489 SNIP 1.25  
Web of Science (2011): Impact factor 5.082  
ISI indexed (2011): ISI indexed no