Single Cell Transcriptome Meta-network Analysis Pipeline to Discover Regulators of Metastasis in Breast Cancer

Next Generation Sequencing (NGS) have provided us with a global transcriptome view of multiple biological contexts. Specifically, the systems biology approach using gene network analysis paved way for critical biological insights shifting our view from conceptualizing a system comprised of linear pathways to one with intertwined functional connections between molecules. This approach brought new therapeutic targets and made promising discoveries especially in the field of cancer and precision medicine. With the introduction of single cell RNA sequencing technology (scRNA-seq) we are now able to sequence the transcriptome of individual cells, and thus the prior network analysis is possible at a cell-type specific resolution. In this investigation we suggest a bioinformatic pipeline to construct a high confidence, cell-type specific co-expression network (scNET) using data imputation and Bayesian statistics. We constructed scNETs on various cancer cell lines with either primary or metastatic properties. By calculating the average rank difference (ARD) between primary scNETs and metastatic scNETS, we prioritized genes that are associated with breast cancer metastasis. The constructed scNET and our ARD ranking method was evaluated using 664 curated breast cancer metastastic genes, which showed higher enrichment fold compared DEG analysis and baseline expectation. Manual curation of these gene sets also confirms that of the 20 top gene through scNET and ARD, about 70 percent were experimentally validated to be functionally important in breast cancer metastasis. Finally, we show that the gene lists derived based on network topology have very low overlap with DEGs, implying that gene list prioritization through network topology and differential gene expression may be mutually informative.

Keywords: single cell RNA sequencing, breast cancer metastasis, single cell network biology, systems biology

* Correspondence: insuklee@yonsei.ac.kr

Junha Cha[1], Junhan Kim[1], and Insuk Lee[1,2*]

[1] Department of Biotechnology, College of Life Science & Biotechnology, Yonsei University, Seoul 03722, Korea

[2] Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul 03722, Korea