BITS :: Call for Abstracts 2022 - Oral communication

Type	Oral communication
Session	Biological Networks
Title	Detection of pan-cancer surface protein biomarkers via a network-based approach on transcriptomics data
All Authors	Mercatelli D(1), Cabrelle C(1), Guzzi PH(2), Giorgi FM(1).

Affiliation

(1) Department of Pharmacy and Biotechnology, University of Bologna, 40138 Bologna, Italy
(2) Department of Surgical and Medical Sciences, Magna Graecia University, 88100, Catanzaro, Italy

Motivation

Cell surface proteins have been used as diagnostic and prognostic markers in cancer research and as targets for the development of anticancer agents. Many of these proteins lie at the top of signaling cascades regulating cell responses and gene expression, therefore acting as “signaling hubs”. It has been previously demonstrated that the integrated network analysis on transcriptomic data is able to infer cell surface proteins activity in breast cancer. Such an approach has been implemented in our publicly available protocol called “SURFACER”. SURFACER implements a network-based analysis of transcriptomic data focusing on the overall activity of curated surface proteins, with the final aim to identify those proteins driving major phenotypic changes at network level, named Surface Signaling Hubs (SSHs). Here, we show the ability of SURFACER to discover relevant surface biomarker knowledge within and across cancer datasets.

Methods

SURFACER is a protocol to identify pan-cancer deregulated cell surface markers serving as SSHs, i.e. surface proteins analogous to a MR. It is organized into 3 major tasks, each divided into 3 simple steps (Fig. 1). Task 1 includes data acquisition from publicly available datasets: RNA-seq data from TCGA and the Genotype-Tissue Expression (GTEx) (step 1). Raw counts are then scaled by considering the total coverage per sample, obtaining a Recount-Scaled Experiment matrix, and expression matrices are then obtained by Variance Stabilizing Transformation (VST) (step 2). Differential expression analysis of the cancer group with respect to normal reference tissue was performed on RSE matrices (step 3), returning a list of differentially expressed genes for each tumor subtype. Task 2 includes all the network analyses. Healthy tissue (normal reference) cell surface proteins coexpression networks are obtained using tissue-specific GTEx samples (step 4). These networks are then analyzed to identify surface proteins whose activity is mostly deregulated in cancer, i.e. cancer SSHs, ranked and filtered on the basis of statistical significance (steps 5 and 6). Task 3 includes surface biomarkers prioritization by differential expression filtering of SSHs (step 7). The resulting SSHs are then evaluated for prognostic significance (step 8) and the druggable surfaceome explored (step 9). Twenty TCGA solid tumor subtypes were analyzed.

Results

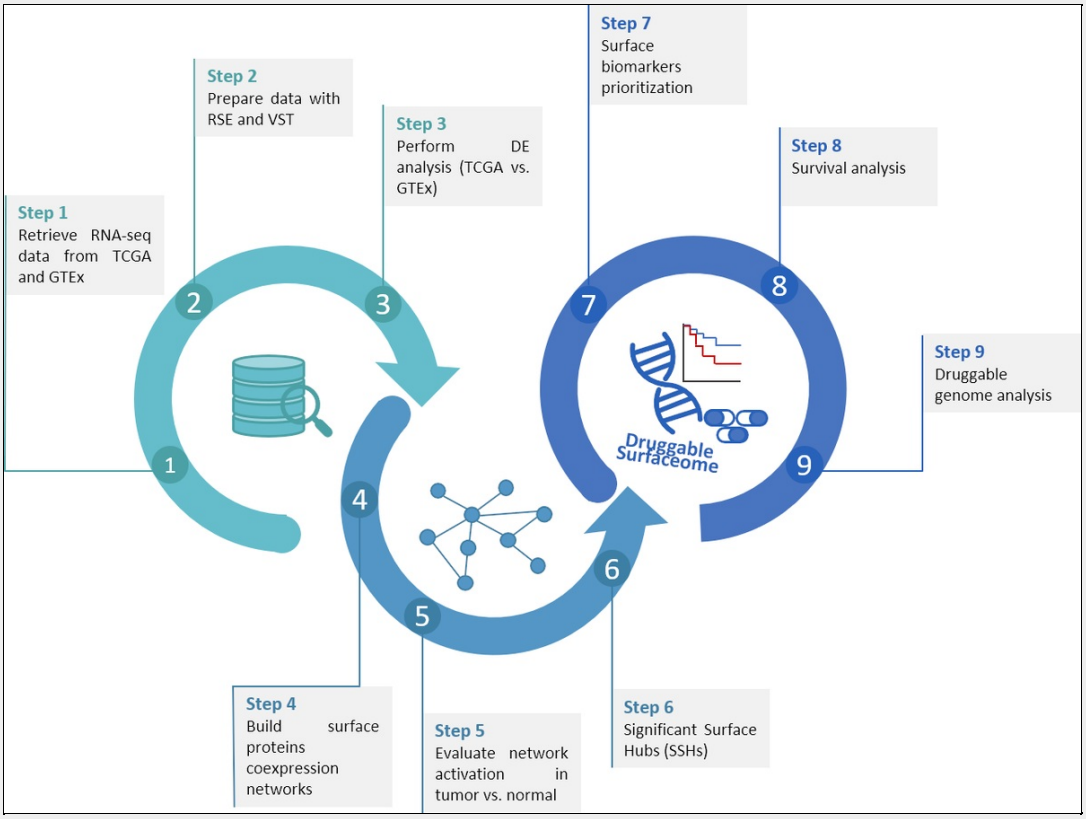
We identified a total of 134 SSHs. A tendency for primary tumors arising from the same tissue of origin to cluster together according to surface activity patterns was detected, e.g. neural tumors (glioblastoma multiforme, low-grade glioma), gynecological (ovarian and uterine cancers), or lung cancers, suggesting tissue-specific patterns of surface proteins alterations. Approximately 40 SSHs showed enhanced activity in most of the TCGA subtypes. These SSHs are of particular interest, considering that overexpressed markers are frequently used for diagnostic purposes and also that enhanced activity could be overcome by targeted therapy with selective inhibitors. We then performed a survival analysis to understand whether each SSH could be used as a survival predictor in each tumor subtype. We found that most of these SSHs were significantly ($p < 0.01$) associated with survival in at least one tumor subtype. While no pan-cancer surface prognostic marker can be detected, the activity of the inferred surface proteins may be useful to stratify patients according to surface patterns associated with survival. We defined surface activity patterns on a single-patient base, and performed a survival analysis taking the overall survival of TCGA patients in each tumor subtype as outcome variable. Among the 20 TCGA subtypes analyzed, 4 subtypes (20%) showed a significant prognostic value for SURFACER-defined patient subtypes. We found that patient stratification according to inferred surface activity models can also show a prognostic value in breast cancer, low-grade glioma, lung and pancreatic adenocarcinoma.

Info

Figure caption
Fig. 1 The SURFACER workflow. Our protocol for the detection of pan-cancer surface protein biomarkers includes 3 major tasks (plotted as cyan, light blue, and blue arrows) each divided into 3 sub-steps. Task 1 (steps 1-3) includes data retrieval, manipulation and differential expression analysis between tumor and normal reference RSE raw counts matrices. Task 2 (steps 4-6) include both building and evaluating context-specific surface proteins centered coexpression networks to identify surface proteins whose inferred activity is significantly different from normal tissue (i.e., SSHs). Task 3 (steps 7-9) includes the analysis of SSHs to prioritize differentially expressed targets, to identify clinically relevant prognostic biomarkers, and to explore their potential druggability.

filename	Fig_1.png

Figure



Availability	-

Corresponding Author

Name, Surname	Daniele, Mercatelli
Email	daniele.mercatelli2@unibo.it

Submitted on	26.04.2022