Re-assessment of the high-grade serous ovarian cancer phosphoproteome – Identification of kinase candidates in TCGA tumors for future clinical intervention

Frank Rolfs, Sander R. Piersma, Connie R. Jimenez

Cancer is a serious health problem and continues to be a major cause of disease-related deaths worldwide. Using latest technologies, researchers and clinicians are accumulating more and more knowledge about the mechanisms driving tumor development. Next to genomics, MS based proteomics is increasingly used in recent years to further uncover tumor biology. Constant improvement of instrumentation now allows deep quantitative measurements of protein expression and post-translational protein modification in short time and phosphoproteomics is the tool of choice to study how tumor cells hijack signaling pathways for their growth. The study of phosphorylation sites allows the identification of deregulated kinases of which some are known to be a possible cause and thus a potential Achilles’ heel of oncogenesis.

Ovarian cancer is a common cancer type in women and therapy options require improvement. To provide a landscape of molecular events and mechanisms associated with ovarian cancer development, the US Cancer Proteomics Tumor Analysis Consortium (CPTAC) extended the genomic and transcriptomic analysis of TCGA high-grade serous carcinomas (HGSC) using extensive proteomic (174 tumors) and phosphoproteomic (69 tumors) characterization [6]. Proteogenomic integration of the data gave insight in, for example, how changes of the genome or transcriptome correlate with the proteome. Moreover, protein abundance was used to subtype HGSC which resulted in the classification into five subgroups. In addition, proteogenomic data were correlated to TCGA patient survival information. Transcriptomic and proteomic data measured were mapped to known cellular pathways, associated with short patient survival and ordered. The RhoA regulatory, PDGFRα and integrin-like kinase pathways ranked highest, showing that tumor characterization is further extended via phosphoproteomic analysis. Perhaps due to the focus of CPTAC on combining genomic, transcriptomic and proteomic data streams, the phospho data were not analyzed to their full potential.

Now, in EBioMedicine, Tong et al. [4] revisited the phosphoproteomic data of the ovarian CPTAC study and considerably extended the analysis to bioinformatically derive kinase candidates connected with poor survival as potential starting points for future intervention in the clinic. Starting from CPTAC’s supplemental phosphopeptide table, data were further processed and filtered. Applying public information on kinase-substrate associations, relative kinase activities were calculated and predicted per patient. Then, these activities and phosphosite data were used to define phosphoproteomic subtypes of HGSC which resulted in five groups showing association with patient survival. Interestingly, phosphoproteome subtyping seemed to be better correlated to overall survival than grouping via the proteome. Next, the identification of kinase targets for potential intervention followed which yielded 29 candidates involved in the PI3K/AKT/mTOR pathway, cell cycle and MAP kinase signaling. Subsequent ranking of these candidates by integrating the top 3 kinases for each patient sample measured, resulted in a ranked list for the selection of most likely kinase targets for clinical intervention. Finally, by taking into account the pathways the activated kinases were mapped to, a possible drug priority list was proposed. Data on kinase inhibitor off-targets effects [1] were also taken into account. In the future, the identified kinases and drug combinations need to be validated and further shortlisted.

Since data analysis strategies and ideas constantly evolve we welcome the re-mining of cancer proteomic data to further extract biological information not yet uncovered in their primary studies. Moreover, we expect from future re-assessment investigations to start from raw (binary) data files rather than using pre-processed or paper supplemental data to account for evolution in the software used for the proteomics analysis workflow. It is thus of importance to deposit annotation-rich, unprocessed and edited data as well as bioinformatics tools in repositories like ProteomeXchange, GitHub or interactively online, to provide a common basis for knowledge gain.

Kinase activity prediction and cancer subtyping using phosphoproteomic data, in liaison with protein abundance data, is promising for fueling new therapeutic strategies. However, kinase prediction using kinase-substrate relations from repositories and prior information is only as powerful as the measurements and databases used are complete. For example, more than 95% of the currently known phosphosites are not yet mapped to their kinases or have annotated function [2] which affects kinase activity prediction analyses at the moment and likely overemphasizes well-studied kinases.
In the clinical setting, therapy decisions will be increasingly made on a personalized basis and single patient characteristics, such as mutational status [5]. Thus, in practical terms, the retrospective identification of possible kinases for therapeutic intervention based on survival data and group comparisons seems problematic. In the future, we envisage a shift to rather patient-tailored approaches identifying aberrant kinase activities from single patient samples. A soon available dedicated analysis pipeline for prioritizing active kinases in individual specimens is Integrative Inferred Kinase Activity analysis (INKA) [3].

Phosphoproteomics has great potential for use in precision oncology. However, given to complexity, data analysis is challenging. Multidisciplinary team work is thus of high importance, not only for high quality sample collection, state-of-the-art analysis and candidate identification, but also for validation and future therapies.

**Disclosure**

The authors declared no conflicts of interest.

**References**

[1] Klaeger S, Heinzlmeir S, Wilhelm M, Polzer H, Vick B, Koenig P-A, et al. The target landscape of clinical kinase drugs. Science 2017;358:eaan4368. https://doi.org/10.1126/science.aan4368.

[2] Needham EJ, Parker BL, Burykin T, James DE, Humphrey SJ. Illuminating the dark phosphoproteome. Sci Signal 2019;12:eaau8645. https://doi.org/10.1126/scisignal.aau8645.

[3] Pham TV, Beekhof R, van Alphen C, Knol JC, Henneman AA, Rolfs F, et al. INKA, an integrative data analysis pipeline for phosphoproteomic inference of active phosphokinases bioRxiv 259192;2018. https://doi.org/10.1101/259192.

[4] Tong M, Yu C, Zhan D, Zhang M, Zhen B, Zhu W, et al. Molecular subtyping of cancer and nomination of kinase candidates for inhibition with phosphoproteomics: reanalysis of CPTAC ovarian cancer. EBioMedicine 2018. https://doi.org/10.1016/j.ebiom.2018.12.039.

[5] Tung NM, Garber JE. BRCA 1/2 testing: therapeutic implications for breast cancer management. Br J Cancer 2018;119:141–52. https://doi.org/10.1038/s41416-018-0127-5.

[6] Zhang H, Liu T, Zhang Z, Payne SH, Zhang B, McDermott JE, et al. Integrated proteogenomic characterization of human high-grade serous ovarian cancer. Cell 2016;166:755–65. https://doi.org/10.1016/j.cell.2016.05.069.