Glyco(proteo)mic workflows for cancer biomarker discovery

Moran, A.B.

English Summary

In prostate (PCa) and colorectal (CRC) cancer, there is a need to improve patient stratification techniques that aid diagnostic and prognostic decision-making. To fulfill this unmet clinical need, the measurement of disease-related biological parameters known as “biomarkers” from biofluids is an approach with the potential to develop non-invasive tests as well as achieve greater clinical accuracy and personalized medicine. Thus, the aim of this thesis was to develop a better understanding of biomarkers relevant to PCa and CRC as well as advancing analytical methodology and achieving methodological advancements for the purpose of biomarker discovery. In this regard, prostate-specific antigen (PSA) proteoform (PCa) and serum *N*-glycosylation (CRC) profiles were investigated using capillary electrophoresis-electrospray ionization-mass spectrometry (CE-ESI-MS) and reversed phase-liquid chromatography (RPLC)-MS, respectively.

In *part one* of this thesis (**Chapters 2** and **3**), the intact proteoform profile of both urinary and seminal prostate-specific antigen (PSA) was explored using CE-ESI-MS. **Chapter 2** focused on the development of the CE-ESI-MS method in order to explore urinary PSA proteoforms, in particular as these forms had previously not been well defined. In addition, previous studies often focused on further defining the glycosylation – rather than the proteoform – profile whereas both were equally assessed here. Furthermore, proteoforms had largely been described solely as “pI isoforms” whereas the assignment of cleaved proteoforms is provided in **Chapter 2**. Additionally, **Chapter 3** addressed some of the challenges associated with the data processing of intact proteoforms, namely semi-automatic proteoform annotation and quantification, thereby enhancing the throughput and accuracy of the workflow. Interestingly, in-depth bottom-up and middle-up approaches were integrated into the analysis in order to further support proteoform assignment in seminal and urinary PSA. Furthermore, it was also shown that extracted ion electropherogram and deconvolution approaches yielded similar quantification results with respect to the relative abundance of PSA proteoforms Overall, the method is now poised for the analysis of a larger number of patient samples and our results may inform future studies regarding this protein’s proteofoms as well as any related clinical implications thereof.

In *part two* (**Chapters 4** and **5**), the potential of serum *N*-glycosylation as a biomarker for CRC was investigated. In this sense, important developments to the methodology were made in **Chapter 4**. Here, fluorescent labeling of total plasma released *N*-glycans via procainamide was combined with sialic acid-specific derivatization via ethyl esterification and amidation. This allowed the retention of all *N*-glycan species by RPLC-MS and, in particular, α2,3- and α2,6-sialylated *N*-glycans were differentiated based upon retention time, precursor mass, and fragmentation spectra, and additional sialylated isomers were resolved. Other isomeric species such as antennary and core-fucosylated *N*-glycans were also separated. In addition, a new quantification approach was derived by combining the fluorescent detection signal with *N*-glycan ratios determined by MS. This improved feature coverage as well as the precision of the method in comparison with performing either quantification approach alone. Finally, the performance of the method was benchmarked against the gold-standard method for released glycan analysis, namely HILIC-MS.

Chapter 5 demonstrated the application of the developed method in order to analyze serum *N*-glycosylation from pre- and post-operative CRC. In this chapter, previous findings with regard to specific glycosylation patterns in CRC were corroborated. In addition, new results were obtained as it was demonstrated that specific *N*-glycan isomers, such as a trisialylated structure with antennary fucosylation, are implicated in the disease. It was also illustrated that pre-operative abundances of serum *N*-glycans may differentiate adenocarcinoma from other histological types and, in particular, differences between histological types were eradicated following surgery. Thus, these results promote the use of serum *N*-glycosylation signatures in order to monitor patient profiles in response to clinical interventions such as surgery.

In **Chapter 6**, the challenges and limitations of the presented work were discussed. In addition, future directions with regard to the biomarkers assessed in this thesis, namely PSA proteoforms and total serum *N*-glycosylation, were explored. Importantly, perspectives were provided on the translation of the results and methods described by this thesis to the clinics.