

論文名 : **Designing and implementation of Proteomics and peptidomics approaches to improve the knowledge of complex biological systems** (要約)

新潟大学大学院自然科学研究科

氏名 AMR EL GUOSHY ABD ELDAYEM SHALABY

(以下要約を記入する)

Genomics, transcriptomics, proteomics, and peptidomics are different technologies aiming at understanding the biological systems at the molecular level. Proteomics and peptidomics technologies still remain one step behind genomics and transcriptomics. This lag makes the proteome of the *Homo sapiens* still incomplete, which encouraged many research laboratories worldwide to take initiatives toward the completion of the human proteome project (HPP). The rationale behind the human proteome completion is to integrate proteomics data into a genomic framework that will lead to improving the knowledge of complex biological systems and lay a foundation for preventive and therapeutic medical applications. In the first part of this thesis (1st and 2nd chapter), I proposed and implemented different proteomics approaches, which helped the proteomics community to reduce the gap between the human genome and human proteome.

In details, in the first chapter of this thesis, I introduced a comprehensive bioinformatics workflow to analyze the “missing part” of human proteome and then proposing different approaches, which contribute in the reduction of the “the missing part” of the human proteome. One of the most important proteomic approaches introduced in this chapter was using customized mature protein database to identify missing proteins by targeting their unique N-terminal tryptic peptide from mature protein database and or C-terminus tryptic peptide from signal peptidome database. The implementation of this approach resulted in the identification of one missing

protein “putative adrenomedullin-5- like protein (C9JUS6) (PE2)” in human brain tissues. In the second chapter, I proposed a pipeline to analyze, identify and validate human missing and uncertain proteins in open-access transcriptomics and proteomics databases. And this approach resulted in the identification and validation of 41 missing proteins using at least 2 unique peptides.

In the second part of this thesis (3rd and 4th chapters), I focused on peptidomics. Peptidomics is a part of modern proteomics, by which all native peptides present in a biological sample of interest can be identified and characterized. Understanding of peptidome data is a big challenge more than proteome data, where the *in vivo* cleavage, which resulted in native peptide generation, is unknown.

Thus, I proposed peptidomics workflow in order to reach a comprehensive understanding of peptidome data through simulating the molecular processes occurred *in vivo*, and which resulted in native peptide generation. Then, I implemented this peptidomics approach on 2 different species; in chapter 3, I implemented the approach on *Homo sapiens*. One of the main findings is that Matrix metalloproteinases (MMP) are the highest presented peptidases in urinary precursor peptide generation with the high-predicted contribution of MMP-9 and MMP12 peptidases. Whereas, in the last chapter, I implemented the proposed approach on *Oryza sativa* subsp. *Japonica*. And the most important finding is that the activity of the Putative cysteine proteinase (Q6ZJ18) was reduced due to the upregulation of cysteine proteinase inhibitor in rice under heat stress conditions.