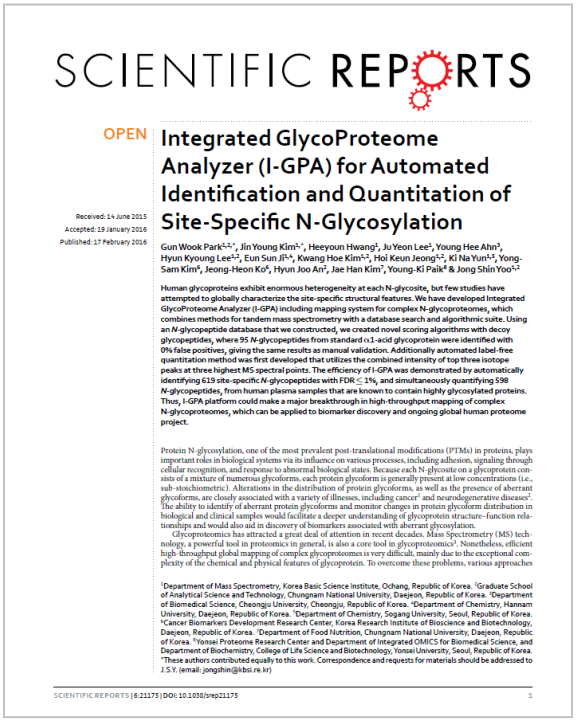
신속 정확하게 암, 질병 진단 마커 발견 및 진단에 활용하는 빅데이터 기반 당단백질 분석 신기술 개발

• 교신저자 : 유종신(바이오융합)

• Scientific Reports / 2016. 2.

연구내용

바이오 분석 분야의 난제로 남았던 인간 혈액내 당단백질을 정확하고 빠르게 분석할 수 있는 질량분석 빅데이터 기반 당단백질 분석 신기술(GlycoProteome Analyzer : GPA)을 개발함. 당단백질은 구조적 복잡성 때문에 분석이 어려웠지만, 2만여개의 단백질과 400여개 당성분의 조합에서 얻어지는 대량의 질량분석 데이터를 예측한 후 실험값과 비교하여 찾아내는 생물정보학 기반 소프트웨어(GPA)를 개발하여 당단백질을 빠르고 정확하게 확인할 수 있음. 이 기술을 바탕으로 암 특이적인 당단백질 바이오마커를 포함하여, 세계 최초로 혈액 내 600여 개 이상의 다양한 형태를 가지는 당단백질들을 동시에 확인하여 정상인과 암 환자의 차이를 비교하였음.



SCIENTIFIC REPORTS

OPEN

Integrated GlycoProteome Analyzer (I-GPA) for Automated Identification and Quantitation of Site-Specific N-Glycosylation

Received: 14 June 2015
Accepted: 19 January 2016
Published: 17 February 2016

Gun Wook Park[1,2,*], Jin Young Kim[3,*], Heeyoun Hwang[1], Ju Yeon Lee[1], Young Hee Ahn[3], Hyun Kyoung Lee[1,2], Eun Sun Ji[3,4], Kwang Hoe Kim[1,2], Hoi Keun Jeong[1,2], Ki Na Yun[1,2], Yong-Sam Kim[4], Jeong-Heon Ko[4], Hyun Joo An[5], Jae Han Kim[3], Young-Ki Paik[6] & Jong Shin Yoo[1,2]

Human glycoproteins exhibit enormous heterogeneity at each N-glycosite, but few studies have attempted to globally characterize the site-specific structural features. We have developed Integrated GlycoProteome Analyzer (I-GPA) including mapping system for complex N-glycoproteomes, which combines methods for tandem mass spectrometry with a database search and algorithmic suite. Using an *N*-glycopeptide database that we constructed, we created novel scoring algorithms with decoy glycopeptides, where 95 *N*-glycopeptides from standard α1-acid glycoprotein were identified with 0% false positives, giving the same results as manual validation. Additionally automated label-free quantitation method was first developed that utilizes the combined intensity of top three isotope peaks at three highest MS spectral points. The efficiency of I-GPA was demonstrated by automatically identifying 619 site-specific *N*-glycopeptides with FDR ≤ 1%, and simultaneously quantifying 598 *N*-glycopeptides, from human plasma samples that are known to contain highly glycosylated proteins. Thus, I-GPA platform could make a major breakthrough in high-throughput mapping of complex N-glycoproteomes, which can be applied to biomarker discovery and ongoing global human proteome project.

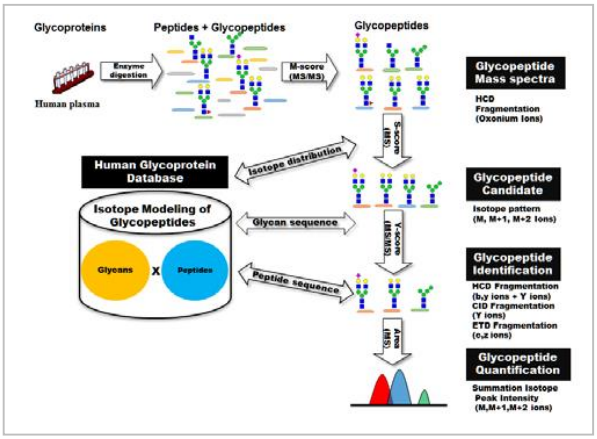
Protein N-glycosylation, one of the most prevalent post-translational modifications (PTMs) in proteins, plays important roles in biological systems via its influence on various processes, including adhesion, signaling through cellular recognition, and response to abnormal biological states. Because each N-glycosite on a glycoprotein consists of a mixture of numerous glycoforms, each protein glycoform is generally present at low concentrations (i.e., sub-stoichiometric). Alterations in the distribution of protein glycoforms, as well as the presence of aberrant glycoforms, are closely associated with a variety of illnesses, including cancer[1] and neurodegenerative diseases[2]. The ability to identify of aberrant protein glycoforms and monitor changes in protein glycoform distribution in biological and clinical samples would facilitate a deeper understanding of glycoprotein structure–function relationships and would also aid in discovery of biomarkers associated with aberrant glycosylation.

Glycoproteomics has attracted a great deal of attention in recent decades. Mass Spectrometry (MS) technology, a powerful tool in proteomics in general, is also a core tool in glycoproteomics[3]. Nonetheless, efficient high-throughput global mapping of complex glycoproteomes is very difficult, mainly due to the exceptional complexity of the chemical and physical features of glycoprotein. To overcome these problems, various approaches

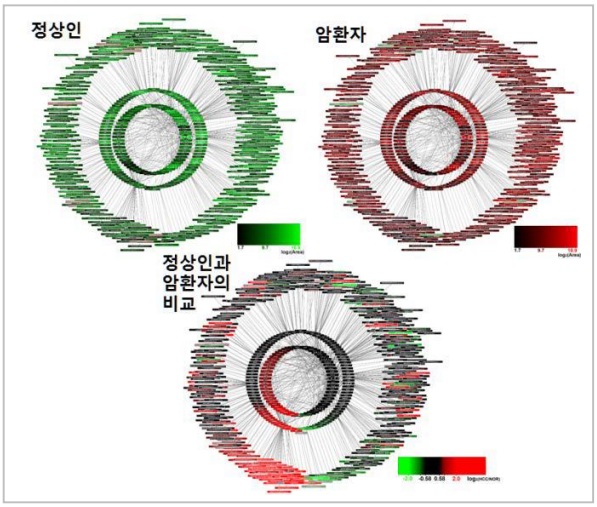
[1]Department of Mass Spectrometry, Korea Basic Science Institute, Ochang, Republic of Korea. [2]Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon, Republic of Korea. [3]Department of Biomedical Science, Cheongju University, Cheongju, Republic of Korea. [4]Department of Chemistry, Hannam University, Daejeon, Republic of Korea. [5]Department of Chemistry, Sogang University, Seoul, Republic of Korea. [6]Cancer Biomarkers Development Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon, Republic of Korea. [7]Department of Food Nutrition, Chungnam National University, Daejeon, Republic of Korea. [8]Yonsei Proteome Research Center and Department of Integrated OMICS for Biomedical Science, and Department of Biochemistry, College of Life Science and Biotechnology, Yonsei University, Seoul, Republic of Korea. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.S.Y. (email: jongshin@kbsi.re.kr)

기대효과

당단백질의 양적변화를 효율적이고 정확한 분석을 통해, 인체 시료로부터 암 또는 질병 진단 마커를 발견하여 예측, 진단할 수 있는 기술로 활용가능하며, 고분해능 질량 분석기와 GPA 기술을 이용하면 당단백질 신약개발 및 바이오시밀러를 포함한 신약평가에 활용될 수 있음.



[그림 1] 당단백질의 정성/정량 분석을 위한 플랫폼



[그림 2] 정상인과 암환자 혈액에서 유래한 당단백질의 정성/정량 분석 결과