Li et al. BMC Bioinformatics 2010, 11(Suppl 4):P29
http://www.biomedcentral.com/1471-2105/11/S4/P29

BMC Bioinformatics

POSTER PRESENTATION Open Access

Enabling proteomics-based identification of human cancer variations

Jing Li[1], Zeqiang Ma[1], Robbert JC Slebos[2], David L Tabb[1,3], Daniel C Liebler[1,2,3], Bing Zhang[1*]

From UT-ORNL-KBRIN Bioinformatics Summit 2010
Cadiz, KY, USA. 19-21 March 2010

Background

Shotgun proteomics is a powerful technology for protein identification in complex samples with remarkable applications in elucidating cellular and subcellular proteomes [1,2], and discovering disease biomarkers [3,4]. Shotgun proteomics data analysis usually relies on database search. Commonly used protein sequence databases in shotgun proteomics data analysis do not contain mutation information. This becomes a problem in cancer studies in which the detection of disease-related mutated peptides/proteins is crucial for understanding cancer biology [5]. Including protein mutation information into sequence databases can help alleviate this problem.

Results

Based on the human Cancer Proteome Variation Database developed by us recently [6], which comprises 41,541 nonsynonymous SNPs in 30,322 proteins from the dbSNP database and around 9000 cancer-related variations in 2,921 proteins, we created a variation-containing protein sequence database and a data analysis workflow for mutant protein identification in shotgun proteomics (Figure 1). Applying this workflow on colorectal cancer cell lines identified many peptides that contain either non-cancer-specific or very important cancer-related variations, such as a known somatic mutation in K-Ras in HCT116 cell line. Our workflow for mutant peptide identification has been tested for compatibility with various popular database search engines including Sequest, Mascot, X!Tandom as well as MyriMatch.

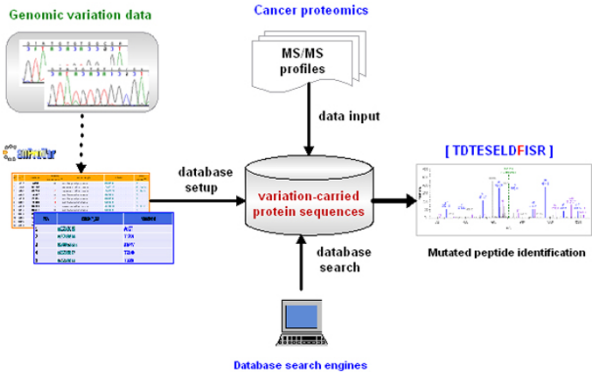

Figure 1 Architecture for identifying mutant peptides from cancer shotgun proteome data

Conclusion

Owing to its protein-centric nature, the approach we proposed can serve as a bridge between genomic variation data and proteomics studies in human cancer.

Acknowledgments

This work was supported by the National Institutes of Health (NIH)/ National Cancer Institute (NCI) through grant R01 CA126218 and the NIH/National Institute of General Medical Sciences through grant R01 GM88822.

Author details

[1]Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA. [2]Jim Ayers Institute for Precancer Detection and Diagnosis, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN 37232, USA. [3]Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN 37232, USA.

Published: 23 July 2010

References

1. Foster LJ, de Hoog CL, Zhang Y, Zhang Y, Xie X, Mootha VK, Mann M: **A mammalian organelle map by protein correlation profiling.** *Cell* 2006, **125**:187-199.

* Correspondence: bing.zhang@vanderbilt.edu
[1]Department of Biomedical Informatics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA

2. Kislinger T, Cox B, Kannan A, Chung C, Hu P, Ignatchenko A, Scott MS, Gramolini AO, Morris Q, Hallet MT, Rossant J, Hughes TR, Frey B, Emili A: **Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling.** *Cell* 2006, **125**:173-186.
3. Decramer S, Wittke S, Mischak H, Zürbig P, Walden M, Bouissou F, Bascands JL, Schanstra P: **Predicting the clinical outcome of congenital unilateral ureteropelvic junction obstruction in newborn by urinary proteome analysis.** *Nat Med* 2006, **12**:398-400.
4. Whiteaker JR, Zhang H, Zhao L, Wang P, Kelly-Spratt KS, Ivey RG, Piening BD, Feng LC, Kasarda E, Gurley KE, Eng JK, Chodosh LA, Kemp CJ, McIntosh MW, Paulovich AG: **Integrated pipeline for mass spectrometry-based discovery and confirmation of biomarkers demonstrated in a mouse model of breast cancer.** *J Proteome Res* 2007, **6**:3962-3975.
5. Wang Z, Moult J: **SNPs, protein structure, and disease.** *Hum mutat.* 2001, **17**:263-270.
6. Li J, Duncan DT, Zhang B: **CanProVar: a human cancer proteome variation database.** *Hum Mutat* 2010, **31(3)**:219-28.

doi:10.1186/1471-2105-11-S4-P29
Cite this article as: Li *et al.*: Enabling proteomics-based identification of human cancer variations . *BMC Bioinformatics* 2010 **11**(Suppl 4):P29.