Applications of an integrated data warehouse system to investigate complex biological systems

Yi-An Chen[1]
chenyian@nibio.go.jp

Lokesh P. Tripathi[1]
lokesh@nibio.go.jp

Kenji Mizuguchi[1]
kenji@nibio.go.jp

[1] National Institute of Biomedical Innovation, 7-6-8 Saito-Asagi, Ibaraki-City, Osaka, 567-0085, Japan.

Keywords: Data integration, Target prioritization, Data warehouse, Drug discovery

We have developed TargetMine, an integrated data warehouse for the purpose of early-stage drug discovery [1]. TargetMine is based on InterMine, a flexible and effective framework for combining biological data from various repositories, which enables navigating through diverse biological data within a single user-friendly interface. TargetMine incorporates a wide range of biological associations including pathways, Gene Ontology annotations, protein structures, chemical compounds and different types of biomolecular interactions. Furthermore, TargetMine includes multiple application programming interfaces, thereby permitting complex pipeline analysis. We have proposed several protocols and also implemented accessory tools to enable streamlined data analysis using TargetMine. These approaches have been effectively employed for target selection in protein-protein interaction (PPI) network-based analyses of Hepatitis C virus (HCV) pathogenesis [2,3,5] and lung tumorigenesis [4]. The academic version of TargetMine is freely available at http://targetmine.nibio.go.jp .

[1] Chen YA, Tripathi LP, Mizuguchi K (2011) TargetMine, an Integrated Data Warehouse for Candidate Gene Prioritisation and Target Discovery. PLoS ONE 6(3): e17844.
[2] Tripathi LP, *et al*. (2010) Network based analysis of hepatitis C virus Core and NS4B protein interactions. MolBiosyst 6: 2539-2553.
[3] Tripathi LP, *et al*. (2012) Proteomic analysis of hepatitis C virus (HCV) core protein transfection and host regulator PA28γ knockout in HCV pathogenesis: A network-nased study. J Proteome Res 11: 3664-3679.
[4] Ihara S, *et al*. (2012) Inhibitory roles of signal transducer and activator of transcription 3 in anti-tumor immunity during carcinogen induced lung tumorigenesis. Cancer Res.72, 2990-2999.
[5] Tripathi LP, *et al*. (2013) Understanding the Biological Context of NS5A-Host Interactions in HCV Infection: A Network-Based Approach. J Proteome Res.