

# LocPPI : A Protein-Protein Interaction System for Analysis of Cellular Proteome

**Pan-Gyu Kim<sup>1</sup>**  
pgkim@smallsoft.co.kr

**Hwajung Seo<sup>1</sup>**  
hjseo@smallsoft.co.kr

**Suk Hoon Jung<sup>2</sup>**  
jsh@icu.ac.kr

**Kiejung Park<sup>1</sup>**  
kjpark@smallsoft.co.kr

<sup>1</sup> Information and Technology Institute, SmallSoft Co., Ltd. Jang-Dong 59-5, Yusung-Gu, Daejeon, 305-811, South Korea

<sup>1</sup> School of Engineering, Information and Communications University, 119, Munjiro, Yuseong-Gu, Daejeon, 305-714, South Korea

## Abstract

We have developed, Loc-PPI, an integrated system for protein-protein interactions which has three major features, to display protein-protein interaction diagrams by graphical mode, to show information for protein localization, and to provide predictions of protein-protein interactions extracted from abstracts in PubMed.

**Keywords:** LocPPI, protein-protein interaction, localization

## 1 Introduction

Protein-protein interactions, which is a critical information to understand cellular processes. Protein-protein interactions occurring in cells are very intricate networks and are very difficult to analyze and comprehend because of their multi-cross interactions.

The sub-cellular localization of a protein can provide valuable clues as to its function as well as help to verify protein interaction information from public PPI databases.

We have developed Loc-PPI which provides protein-protein interaction information by merging major PPI databases and each protein's sub-cellular localization. Loc-PPI has a very convenient PPI viewer which helps to confirm PPI information and location of protein in a cell.

## 2 Methods

Loc-PPI is a stand-alone java application program to show the protein interaction graph with localization and functional information of the proteins in a cell. Loc-PPI consists of two main components: (1) integrated PPI database in a Linux server, and (2) a stand-alone Java application to display protein interaction networks. The integrated PPI database contains not only protein-protein interaction information from public PPI databases but also localization information for each protein from public localization databases. We are adding text mining tool which can verify the PPI information from public PPI databases by extracting additional PPI information from PubMed.

Public database	No. of proteins	No. of protein-protein interactions
Total	980,058	60,344,728
DIP	19,872	96,460
MINT	26,869	117,752
BIND	29,684	106,056
INTACT	36,632	127,317
STRING	959,928	60,232,332

Table 1 : Total numbers of constructed protein-protein interactions data obtained five major public protein database.

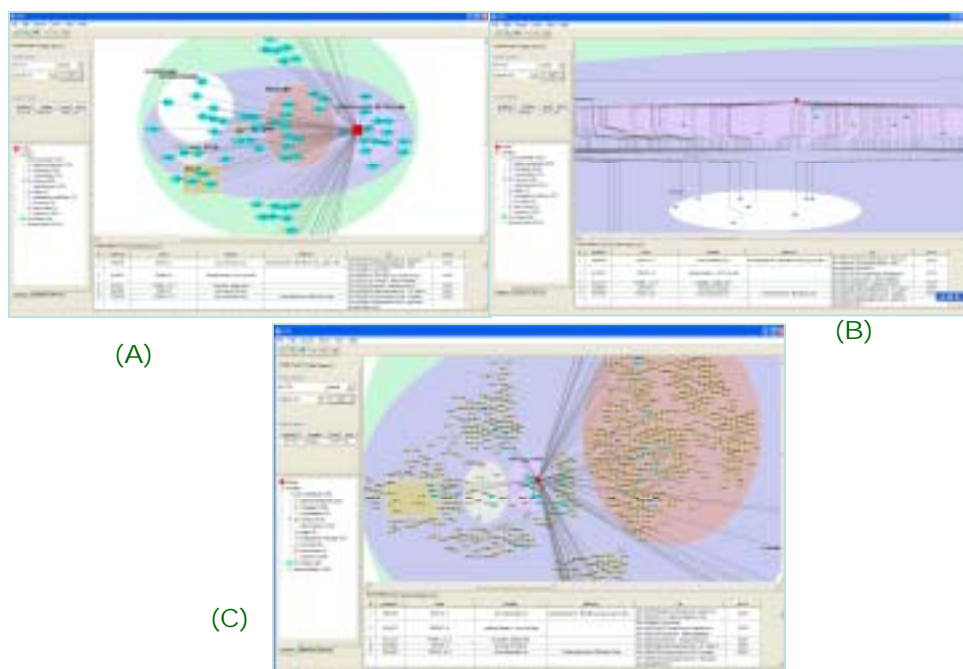


Fig. 1. The snapshots of Loc-PPI viewer: (A) organic layout, (B) hierarchical layout, (C) showing 2-depth proteins.

### 3 Results and Discussion

We have developed Loc-PPI for protein-protein interactions and localizations of proteins. Loc-PPI can display protein-protein interaction diagrams by graphical mode and show information for protein localization. The database part in server-side contains the data for modified proteins and protein-protein interactions obtained from public databases for DIP, BIND[1], INTACT, MINT, and STRING. Localization information database was configured by a number of resources including databases such as PSORTdb, DbSubLoc[2], eSLDB, WoLF\_PSORT[3] for bacteria and eukaryotes. In future, by extracting new protein-protein interaction information from abstracts of journals in PubMed, we can verify protein-protein interaction information from public protein-protein interaction database. The protein-protein interactions data are displayed with four graphical layout types, organic layout, hierarchical layout, pathway layout and circular layout. Protein localization information is assigned to each protein with sub-cellular organelles according to defined category. A lot of features were implemented for analysis and visualization, which include grouping proteins by metabolic functions, changing network node structures and attributes such as color and shape, and calculating shortest path is formed from between protein pairs.

### References

- [1] Bader GD, Betel D, Hogue CW. BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res.* 2003 Jan 1;31(1):248-50
- [2] Guo T, Hua S, Ji X, Sun Z. DBSubLoc: database of protein subcellular localization. *Nucleic Acids Res.* 2004 Jan 1;32(Database issue):D122-4.
- [3] Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 2007 Jul;35(Web Server issue):W585-7. Epub 2007 May 21.