35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Differential network biology reveals a positive correlation between a novel protein-protein interaction and cancer cells migration

Chia-Hung Liu, Tzu-Chi Chen, Chun-Houh Chen, Cheng-Yan Kao and Chi-Ying F. Huang

Abstract—**This paper introduces a differential network biology for discovering tumor migration. We applied statistical methods to prioritize PPI candidates and an in situ proximity ligation assay to verify 67 endogenous PPIs among 21 interlinked pathways in two hepatocellular carcinoma (HCC) cells, Huh7 (minimally migratory cells) and Mahlavu (highly migratory cells). Differential network biology analysis was applied to determine the novel interaction, CRKL-FLT1, has a high centrality ranking, and the expression of this interaction is strongly correlated with the migratory ability of HCC and other cancer cell lines. Knockdown of CRKL and FLT1 in HCC cells leads to a decrease in cell migration. This study demonstrated that functional exploration of a disease network with differential network in interlinked pathways via PPIs can be used to discover tumor migration.**

I. Introduction

Metastasis is one of the main causes of mortality from solid tumors, and metastasis is a poor prognostic factor for hepatocellular carcinoma (HCC). Understanding protein-protein interactions (PPIs) may uncover the generic organization of functional networks in cancer cells, when both the spatial and temporal aspects of the interactions are considered [1]. Recently, several studies applied protein network-based approach and differential network-based approach to identify markers to predict patient prognosis[2]. These computational approaches demonstrated great potential and could be further enhanced if more thorough PPI and pathway information is available especially at the cellular level and analyzed with a more sophisticated method.

Cancer can be considered as perturbations of highly interlinked cellular networks. Our hypothesis is that uncovering new PPI links within or between, referred to as interlinked PPIs (cross-talk PPIs), different signaling pathways could recapitulate the relationship between the genotype and phenotype in HCC. Multiple signaling cascades are interlinked in cancer cells via a variety of cross-talk connections with other pathways leading to several of the hallmarks of cancer (e.g., proliferative signaling, angiogenesis, invasion, and survival) [3]. Therefore, targeting these interlinked pathways could provide an opportunity for therapeutic application [3]. Here, we present a systems approach that computationally infers the interlinked pathways from numerous PPIs in HCC up-regulated genes and empirically detects endogenous PPIs using an *in situ* proximity ligation assay (PLA), which allows quantitative and localized detection of endogenous PPIs in cells [4]. Empirically, we validated 67 endogenous PPIs within or between signaling pathways in HCC. To the best of our knowledge, applying *in situ* PLA to this scale in cancer cells is unprecedented.

We demonstrate its effectiveness with the identification of a prioritized interaction, CRKL-FLT1, which links the c-Met, IGF1, PDGFR-alpha and VEGFR1/ VEGFR2 pathways together. CRKL-FLT1 was identified as a hub in the PPI network in HCC, and crucial for migration in HCC cells. Our analysis result shows the expression of CRKL-FLT1 is strongly correlated with the migratory ability of cancer cell lines.

In summary, this study provides broad insight into tumor migration by building an interlinked pathway map via PPIs in HCC.

II. Methods

A. Identification of interlinked PPIs in cross-talk pathways in human HCC

We systematically collected ~3000 differentially expressed HCC-related signatures [5, 6] and examined ~2,100 pathways from an integrated pathway database (ConsensusPathDB) [7]. Then, we prioritized 60 HCC-related pathways according to the HCC-related signatures and pathway datasets using a hypergeometric test with false discovery rate (FDR) adjustment. The proteins belonging to the HCC-related pathways (gray nodes on the left) were detected to determine whether collected PPIs (POINeT) [8] or predicted PPIs (PIPS) [9] with overexpressed patterns can link each pathway and enable potential interlinking. That is to say, each pathway member (single gene) among different HCC-related pathways was used to map each single protein for each PPI pair in the PPIs dataset, and thus it was possible to identify the cross-talk relationship. Because many genes/proteins are involved in multiple pathways, most interlinked PPIs map to several different pathways. These PPIs might link the cross-talk pathways together. In this step, we identified 97 proteins that participate in 375 PPIs occurring in overexpression patterns from HCC gene expression profiles, among the 60 HCC-related pathways.

B. Detection of protein-protein interaction by in situ proximity ligation assay

*Resrach supported by National Science Council (NSC99-3112-B-010-005-CC1) and National Health Research Institutes (NHRI--EX101-10029BI) to CYH. NSC/Interchange Association of Japan (2009 Summer Program) to CHL

C. H. L and C.Y.K. are with Graduate Institute of Biomedical Electronic and Bioinformatics, National Taiwan University, Taipei, Taiwan.

TCC and CYFH are with Institute of Clinical Medicine, National Yang-Ming University, Taipei, Taiwan (corresponding author to provide phone: 886-22826-7000; fax: 886-2 2874-5074; e-mail: cyhuang5@ym.edu.tw).

CHC is with Institute of Statistical Science, Academia Sinica, Taipei, Taiwan.

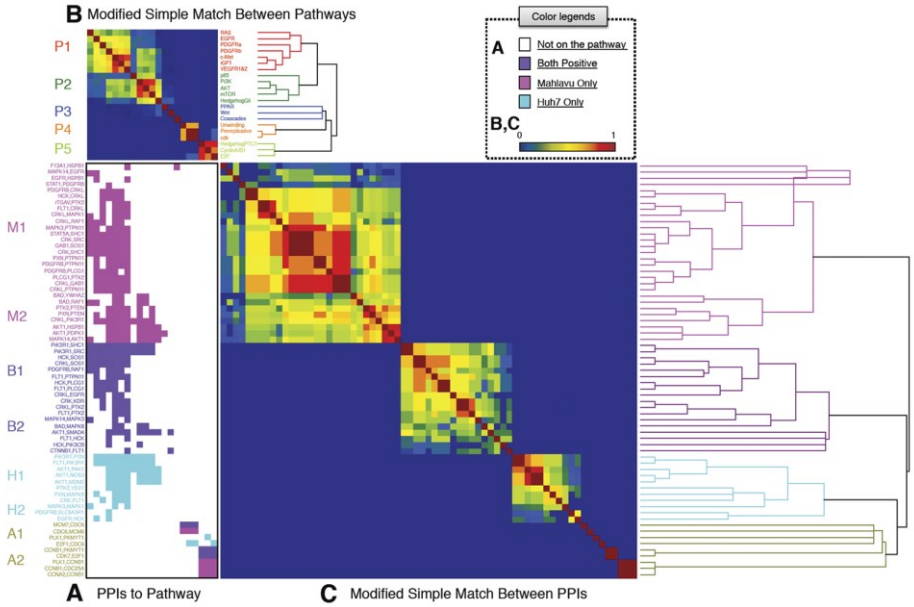

Figure 1. (A-C) The tripartite-association of PPIs, pathways and cell lines and analysis of the differential interaction hubs in the PPI networks

Recently, the PLA was developed to detect and visualize endogenous PPIs and post-translational modifications of proteins with a high sensitivity and specificity (11, 23). To detect PPIs, the dual targets of primary antibody pairs are added. If an antibody pair is in close proximity, secondary antibodies with oligonucleotides will be close enough to serve as templates for the ligation of two additional linear oligonucleotides into a DNA circle. The DNA circle can be amplified with the oligonucleotide in one of the secondary antibodies using rolling circle amplification (RCA). RCA can then be hybridized with fluorescent-labeled oligonucleotides to reveal dot-signal representing both their subcellular locations and the frequency of the PPI occurrences (11, 23). The detection of a dot signal with signal ratio was defined as $SR = (S_{pos} / C_{pos}) / (S_{neg} / C_{neg})$ The PPI was determined to be a positive PPI only if and (S_{pos} / C_{pos}) >10, S_{pos} was the signal and C_{pos} was the cell number for dual-recognition with one rabbit polyclonal antibody and one mouse monoclonal antibody; S_{neg} was the signal and C_{neg} was the cell number for the negative control, in which only one rabbit polyclonal antibody was added.

C. Clustering and Visualization

First, we elucidated the tripartite-interactions of positive PPIs, pathways, and cell lines. In all, 67 of the 194 PPIs tested via the *in situ* PLA method were considered positive PPIs in one or both of the cell lines (Mahlavu or Huh7). Thus, the positive PPIs can be grouped into three types: (1) Mahlavu Only, (2) Huh7 Only, and (3) Both Positive. A matrix M (Fig. 2A) was prepared to store information regarding the 67 PPIs (rows), 21 pathways (columns) and cell lines:

$$M_{ij} = \begin{cases} 0, \text{the } i\text{th PPI does not engage the } j\text{th pathway (displayed in white in Fig. 2A)} \\ 1, \text{the } i\text{th PPI engages the } j\text{th pathway only in Huh7 (in magenta)} \\ 2, \text{the } i\text{th PPI engage the } j\text{th pathway only in Mahlavu (in cyan)} \\ 3, \text{the } i\text{th PPI engages the } j\text{th pathway in both cell lines (in purple)} \end{cases}$$

, $i = 1, \ldots, 67;\ j = 1, \ldots, 21.$

In this study, we adopted a modified version of SMC, SMC_0, which excludes SMC counts for attributes with both objects belonging to the non-existent state ($M_{ij} = 0$ below) in the calculation for both denominator and numerator. SMC_0 is more robust than SMC for sparse data (data with many non-existent states), as in matrix M (Fig. 1A). SMC_0 can also be termed the nominal version of Jaccard coefficient.

To have a more structural and visual representation, we used SMC_0 for representing the between rows (PPIs) and between columns (pathways) association structure and to identify PPI clusters with pathway groups. Here we sort the rows (PPI and columns (pathways) as follows: Two proximity matrices C (Fig. 1B) and R (Fig. 1C) were calculated using the modified simple matching coefficient representing the between rows (PPIs) and between columns (pathways)

$$\begin{cases} C_{ij} = (c_{ij}^m - c_{ij}^0)/(21 - c_{ij}^0), i = 1, \ldots, 67;\ j = 1, \ldots, 67; \\ \text{where} c_{ij}^m = \underset{k=1,\ldots,21}{count}(M_{ik} = M_{jk}), c_{ij}^0 = \underset{k=1,\ldots,21}{count}(M_{ik} = M_{jk} = 0) \end{cases} \quad \text{(Fig. 2B)}$$

and

$$\begin{cases} R_{ij} = (r_{ij}^m - r_{ij}^0)/(67 - r_{ij}^0), i = 1, \ldots, 21;\ j = 1, \ldots, 21; \\ \text{where} r_{ij}^m = \underset{k=1,\ldots,67}{count}(M_{ki} = M_{kj}), r_{ij}^0 = \underset{k=1,\ldots,67}{count}(M_{ki} = M_{kj} = 0) \end{cases} \quad \text{(Fig. 2C)}$$

D. Migration Assay

For Mahlavu stable clones (Vehicle, shCRKL and shFLT1), 1 x 10^4 Mahlavu cells were suspended in 200 µl of DMEM

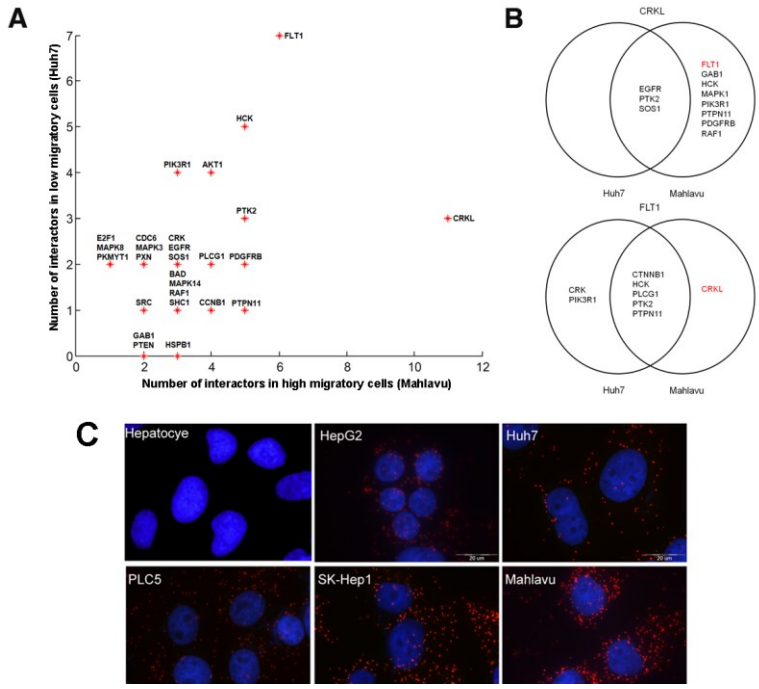

Figure 2. (A-B)The differential network biology in HCC cell lines (C) The images of in situ PLA signal for CRKL-FLT1 interaction in hepatocyte and five HCC cell lines were shown.

without serum and were seeded into the upper chamber, while 750 µl of DMEM containing 10% FBS was added to the outer side of the chamber. For measuring migratory ability in different HCC cell lines (HepG2, PLC5, Huh7, SK-Hep1, and Mahlavu), 1 x 10^5 cells were seeded into the upper chamber with 200µl of DMEM without serum. After being cultured in a 37°C, 5% CO^2/95% air environment and allowed to adhere for 12-16 hours and then incubated, cells on the upper surface of the membrane were removed by a cotton tip applicator and migratory cells on the lower membrane surface were fixed by methanol and stained with Giemsa (Sigma-Aldrich). Cell migration values were determined by counting from three independent membranes and then normalized using vehicle cells to give a relative ratio.

III. Results

An *in situ* proximity ligation assay (PLA) with the available paired antibodies was used to detect, validate and quantify the endogenous presence of 194 PPIs in two HCC cell lines, Huh7 (minimally migratory cells) and Mahlavu (highly migratory cells). We observed 67 PPIs (49 proteins) among 21 pathways in either Huh7 or Mahlavu cells. There are two key features of these datasets. First, according to a literature survey, 17 PPIs (including 11 PPIs from PIPS) of the 67 PPIs are novel PPIs identified in this study. Second, there are contrastingly different distributions of the *in situ* PLA signal for the 67 validated PPIs between Huh7 and Mahlavu, providing the opportunity to apply differential network biology to characterize tumor migration.

In addition to revealing individual PPIs, our purpose here is to provide a global view of the 67 analyzed PPIs and 21 pathways in two HCC cell lines (with different migratory abilities). Briefly, pair-wise modified version of simple matching coefficient (SMC_0) was employed to construct both the between-pathway proximity matrix $\boldsymbol{C}$ (with C_2^{21} = 210 pairs in Fig. 1B) and the between-PPI proximity $\boldsymbol{R}$ (with C_2^{67} = 2211 pairs in Fig. 1C) in the first step. In the second step, we applied two dendrograms (hierarchical clustering trees) to sort the 21 pathways in $\boldsymbol{C}$ into five clusters of pathways with 7, 5, 3, 3, or 3 pathways each and 67 PPIs in $\boldsymbol{R}$ as eight groups of PPIs. Heatmaps was applied to elucidate the tripartite-interactions (Fig. 1A-1C). Then we created three PPIs that are specific in Huh7, Mahlavu, either cell lines or none are color coded as cyan, magenta, purple, or white.

Overall, the clustering analysis shows that most of the PPIs are involved in P1 pathway group, suggesting pathways in P1 group might play an important role in hepatocarcinogenesis. PPIs in M1 can only be observed in Mahlavu cell (highly migratory cells) and belong to the P1 pathway group. Therefore, this allows for intuitive selection of candidate PPIs (e.g., CRKL-FLT1) from M1 for further functional characterization.

A recent study suggests that a differential network biology approach, such as differential interaction hubs, is a promising approach to dynamic network discovery under disease state or speciation [10, 11]. These observations led us to measure the number of interactions in order to identify the differential interaction hubs for ranking essential proteins/pairs, which might be involved in migration in HCC from this PPI network (Fig. 2A). For each protein in the differentially expressed network, CRKL is the hub with the highest degree of centrality in Mahlavu cells (Fig. 2B). This topological analysis suggests that CRKL may play a crucial role in the malignant network, especially in the Mahlavu cells. FLT1,

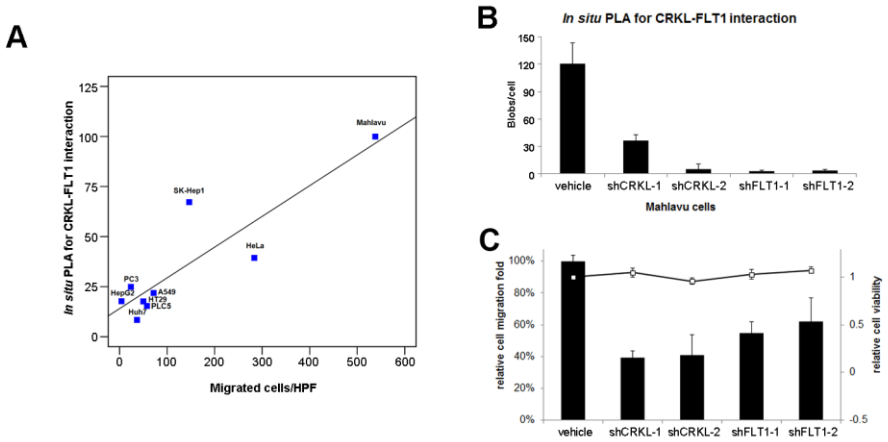

Figure3. The characterization of the CRKL-FLT1 interaction in HCC cell lines and other cancer cell lines.

another hub with seven partners in Huh7 cells and six partners in Mahlavu cells, ranks second in degree centrality (Fig. 2A). These two hubs, CRKL and FLT1, interact with each other in the network of Mahlavu cells. Interaction between CRKL and FLT1, thus, has a higher connectivity than other interactions, suggesting that the CRKL-FLT1 interaction might be important. More importantly, FLT1 has seven interaction partners but only CRKL occurred in highly migratory cells (Fig. 2B), implying an unusual relationship between these two proteins.

The role of CRKL, FLT1, and the prioritized CRKL-FLT1 interaction, in HCC remains unclear. Thus, we measured CRKL-FLT1 interaction in hepatocyte and five different HCC cell lines (HepG2, Huh7, PLC5, SK-Hep1, and Mahlavu) (Fig. 2C). Comparing with migratory ability of five HCC cell lines, the CRKL-FLT1 interaction correlates with the migratory ability of the cells analyzed, which is consistent with our hypothesis.

Moreover, we extend the measurement of CRKL-FLT1 interaction and migratory ability in other cancer cells. The intensity of *in situ* PLA for CRKL-FLT1 interaction and migrated cells/HPF (high-power field) showed a positive correlation with a correlation coefficient of 0.886 ($p < 0.001$), as estimated p value by t-test for Pearson product-moment correlation when we observed five different HCC cell lines (HepG2, Huh7, PLC5, SK-Hep1, and Mahlavu) and four other cancer cells , including cervical cancer (HeLa), lung adenocarcinoma (A549), colon cancer (HT29), prostate cancer (PC3) (Fig. 3A). The result indicated the expression of CRKL-FLT1 interaction is correlated with cell migration not only in HCC, but also in other cancer cells. Next, we established five different Mahlavu stable clones with knockdown of CRKL or FLT1 to investigate the biochemical function of CRKL and FLT1. The expression of CRKL-FLT1 interaction was decreased in Mahlavu stable clones with knockdown of CRKL or FLT1 compared with vehicle control cells (Fig. 3B). It shows the specificity of knockdown ability. We then tested the effect of reducing CRKL and FLT1 on migration in highly migratory Mahlavu cells. Depletion of CRKL and FLT1 decreased the migration of Mahlavu cells approximately 60% and 40%, respectively (Fig. 3C). Together, CRKL and FLT1 may play a crucial role in the regulation of metastasis in HCC.

IV. Conclusion

In summary, we anticipate that our integrated approach and analysis will improve the interpretation of interlinked PPIs and pathways and facilitate the development of novel prognosis markers and drug targets in HCC research.

REFERENCES

[1] A.-L. Barabasi, N. Gulbahce, and J. Loscalzo, "Network medicine: a network-based approach to human disease," *Nat Rev Genet,* vol. 12, pp. 56-68, 2011.

[2] H.-Y. Chuang, E. Lee, Y.-T. Liu, D. Lee, and T. Ideker, "Network-based classification of breast cancer metastasis," *Molecular systems biology,* vol. 3, 2007.

[3] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: the next generation," *Cell,* vol. 144, pp. 646-74, Mar 4 2011.

[4] O. Söderberg, M. Gullberg, M. Jarvius, K. Ridderstråle, K. J. Leuchowius, J. Jarvius, *et al.*, "Direct observation of individual endogenous protein complexes in situ by proximity ligation," *Nature methods,* vol. 3, pp. 995-1000, 2006.

[5] C. N. Hsu, J. M. Lai, C. H. Liu, H. H. Tseng, C. Y. Lin, K. T. Lin, *et al.*, "Detection of the inferred interaction network in hepatocellular carcinoma from EHCO (Encyclopedia of Hepatocellular Carcinoma genes Online)," *BMC bioinformatics,* vol. 8, p. 66, 2007.

[6] M. H. Chen, W. L. R. Yang, K. T. Lin, C. H. Liu, Y. W. Liu, K. W. Huang, *et al.*, "Gene Expression-Based Chemical Genomics Identifies Potential Therapeutic Drugs in Hepatocellular Carcinoma," *PLoS ONE,* vol. 6, p. e27186, 2011.

[7] A. Kamburov, U. Stelzl, H. Lehrach, and R. Herwig, "The ConsensusPathDB interaction database: 2013 update," *Nucleic Acids Research,* November 11 2012.

[8] S. A. Lee, C. H. Chan, T. C. Chen, C. Y. Yang, K. C. Huang, C. H. Tsai, *et al.*, "POINeT: protein interactome with sub-network analysis and hub prioritization," *BMC bioinformatics,* vol. 10, p. 114, 2009.

[9] M. D. McDowall, M. S. Scott, and G. J. Barton, "PIPs: human protein–protein interaction prediction database," *Nucleic acids research,* vol. 37, pp. D651-D656, 2009.

[10] T. Ideker and N. J. Krogan, "Differential network biology," *Molecular systems biology,* vol. 8, 2012.

[11] S. Bandyopadhyay, M. Mehta, D. Kuo, M. K. Sung, R. Chuang, E. J. Jaehnig, *et al.*, "Rewiring of genetic networks in response to DNA damage," *Science,* vol. 330, p. 1385, 2010.