



Analysis of Pan-omics Data in Human Interactome Network (APODHIN)

*Nupur Biswas, Krishna Kumar, Sarpita Bose, Raisa Bera and Saikat Chakrabarti**

Structural Biology and Bioinformatics Division, CSIR-Indian Institute of Chemical Biology, Kolkata, India

OPEN ACCESS

Edited by:

Amit Kumar Yadav,
Translational Health Science
and Technology Institute (THSTI),
India

Reviewed by:

Bhanwar Lal Puniya,
University of Nebraska–Lincoln,
United States
Marco Vanoni,
University of Milano-Bicocca, Italy

*Correspondence:

Saikat Chakrabarti
saikat@iicb.res.in

Specialty section:

This article was submitted to
Systems Biology,
a section of the journal
Frontiers in Genetics

Received: 30 July 2020

Accepted: 11 November 2020

Published: 08 December 2020

Citation:

Biswas N, Kumar K, Bose S,
Bera R and Chakrabarti S (2020)
*Analysis of Pan-omics Data in Human
Interactome Network (APODHIN).*
Front. Genet. 11:589231.
doi: 10.3389/fgene.2020.589231

Analysis of Pan-omics Data in Human Interactome Network (APODHIN) is a platform for integrative analysis of transcriptomics, proteomics, genomics, and metabolomics data for identification of key molecular players and their interconnections exemplified in cancer scenario. APODHIN works on a meta-interactome network consisting of human protein–protein interactions (PPIs), miRNA–target gene regulatory interactions, and transcription factor–target gene regulatory relationships. In its first module, APODHIN maps proteins/genes/miRNAs from different omics data in its meta-interactome network and extracts the network of biomolecules that are differentially altered in the given scenario. Using this context specific, filtered interaction network, APODHIN identifies topologically important nodes (TINs) implementing graph theory based network topology analysis and further justifies their role via pathway and disease marker mapping. These TINs could be used as prospective diagnostic and/or prognostic biomarkers and/or potential therapeutic targets. In its second module, APODHIN attempts to identify cross pathway regulatory and PPI links connecting signaling proteins, transcription factors (TFs), and miRNAs to metabolic enzymes via utilization of single-omics and/or pan-omics data and implementation of mathematical modeling. Interconnections between regulatory components such as signaling proteins/TFs/miRNAs and metabolic pathways need to be elucidated more elaborately in order to understand the role of oncogene and tumor suppressors in regulation of metabolic reprogramming during cancer. APODHIN platform contains a web server component where users can upload single/multi omics data to identify TINs and cross-pathway links. Tabular, graphical and 3D network representations of the identified TINs and cross-pathway links are provided for better appreciation. Additionally, this platform also provides few example data analysis of cancer specific, single and/or multi omics dataset for cervical, ovarian, and breast cancers where meta-interactome networks, TINs, and cross-pathway links are provided. APODHIN platform is freely available at <http://www.hpppi.iicb.res.in/APODHIN/home.html>.

Keywords: meta-interactome, network analysis, pan-omics, multi-omics analysis, pathway cross links

INTRODUCTION

Technological advances have made different types of omics data accessible in large scale. Different types of omics data are outcomes of profiling of different bio-entities, namely RNA (RNA transcriptomics), miRNA (miRNA transcriptomics), proteins (proteomics, phosphoproteomics), genes (genomics, epigenomics), metabolites (metabolomics), lipids (lipidomics), and pharmacogenomics. These bio-entities are functionally inter-related in a complex fashion. Extrapolation from single omics data of one type of bio-entity fails to provide the true biological status of various linked bio-entities (e.g., RNA, protein, metabolites). Hence, to inquire the causative phenomena underlying the genesis and progression of systemic/genetic diseases, an integrative analysis considering the profiles of above mentioned bio-entities appears as a requisite. Moreover, because of the heterogeneous nature of the diseases, even if patients having similar pathological features are treated similarly, the disease prognosis differs a lot. It shows the inadequacy of symptom-based diagnosis and demands patient-specific analysis of omics data. Collective analysis of these multi-dimensional omics data is referred to as “pan-omics” (Sandhu et al., 2018) which are also considered as “big” data in the context of biological data analysis. Pan-omics data enable us to predict novel functional interactions between molecular mediators at multiple levels. Also, these data have the potential to uncover crucial biological observations into hallmarks and pathways that would otherwise not be obvious through single-omics studies. Patient-specific pan-omics data analysis is going to disclose the genetic, epigenetic, and other functional profiles responsible for the disease of an individual which might eventually lead to the development of individualistic “precision medicine” and will provide right treatment to right patient at right time.

Cancer is a leading cause of death worldwide, being responsible for 9.6 million deaths in 2018 (Bray et al., 2018). Cancer is a heterogeneous disease caused by aberrations of genes and proteins. “Precision oncology” promises identification of disease subtypes, specific biomarkers and subsequently prediction and translation toward the development of treatment procedures. Pan-omics or multi-omics analysis in breast cancer has revealed significant differences in molecular subtype distribution (Kan et al., 2018). Genomics and transcriptomics analysis of breast cancer data of Korean and Caucasian cohorts showed underlying molecular differences, which are responsible for the occurrence of breast cancer at the younger age in the Asian population compared to the western population (Kan et al., 2018). Multi-omics analysis extended to different types of cancers confirms the existence of broadly two types of cancers, cancers caused by recurrent mutations and cancers caused by copy-number variations (Mcgrail et al., 2018). Computational methodologies like, artificial intelligence are being used widely to extract patient-specific information from these big data, discussed in a recent review (Biswas and Chakrabarti, 2020). Machine learning based pan-omics analysis of pan-cancer data shows the existence of clusters within different types of cancers (Ramazzotti et al., 2018), identifies cell-model selective anti-cancer drug targets for breast cancer (Gautam et al., 2019).

Multiple data portals like TCGA (TCGA, 2020) and ICGA (Zhang et al., 2011) have been developed to make multi-omics data conveniently accessible. LinkedOmics contains pan-omics data of several types of cancers (Vasaikar et al., 2018). Databases like, GliomaDB (Yang et al., 2019) and MOBCdb (Xie et al., 2018) are dedicated to integrate multi-omics data for specific type of cancers. Standalone software packages and web-servers are also being developed for the analysis pan-omics data. **Table 1** compares the analytical tools which are being used by researchers. R package mixOmics (Rohart et al., 2017), based on multi-variate analysis is available for the integration of multi-omics data. It finds subsets of important features but excludes network analysis. OmicsNet provides a web-based platform to create different types of interactive molecular interaction networks for single or multiple types of omics data (Zhou and Xia, 2018). Network-based integration of multi-omics data using iOmicsPASS, allows to predict subnetworks of molecular interactions within a single type or multiple types of omics data (Koh et al., 2019). R package Miodin (Ulfenborg, 2019) provides a software infrastructure for vertical and horizontal integration of multi-omics data but lacks a comprehensive network analysis and visualization. PaintOmics allows integrated visualization of multiple types of omics data in KEGG pathway diagrams (Hern et al., 2018). Software package, Multi-Omics Factor Analysis (MOFA) (Argelaguet et al., 2018) integrates omics data in an unsupervised approach implementing generalized principal component analysis (PCA). pathfindR (Ulgen et al., 2019) finds active sub networks for genes in omics data and perform pathway enrichment analysis. R package Mergeomics (Shu et al., 2016) provides a pipeline to identify important pathways and key drivers in biological systems. However, platforms required for systematic analysis of the landscape of genetic, epigenetic, and metabolomics alterations and biological and clinical relevance of multi-layer signature in cancers are still limited.

Different types of omics data carry information on different types of bio-entities, e.g., genes, proteins, miRNAs, metabolites, etc. Hence, integrative analysis of pan-omics data needs a meta-interactome consisting of a protein–protein interaction network (PPIN) as well as different regulatory networks. The web server for the Analysis of Pan-omics Data in Human Interactome Network (APODHIN) provides a unique platform where users can analyze different types of omics data using a human cellular meta-interactome network. Graph theory based network analysis has become an essential tool for analysis of PPIN for extracting proteins important in the construction and information flow of the network (Jeong et al., 2001; Barabási and Oltvai, 2004; Mistry et al., 2017; Ashtiani et al., 2018), APODHIN provides options to identify topologically important nodes (TINs) such as hubs, bottlenecks, and central nodes (CNs) and their subsequent modules via protein–protein interaction (PPI) and regulatory relationship network analyses and pathway enrichment analysis. TINs are also correlated as prospective diagnostic and/or prognostic biomarkers. APODHIN can also analyze and compare multiple omics data set for a single omics layer, such as transcriptomics, proteomics data collected from different patient cohorts and/or different stage/grade of the same cohort.

TABLE 1 | Comparison of APODHIN with other existing pan-omics data analysis tools.

Feature	APODHIN	OmicsNet [14]	mixOmics [13]	iOmicsPASS [15]	Miodin [16]
Platform	Web	Web	Standalone	Standalone	Standalone
Programming language	Python, R, perl	R	R	C++	R
Types of omics data as input					
mRNA transcriptomics	Yes	Yes	Yes	Yes	Yes
miRNA transcriptomics	Yes	Yes	Yes	No	Yes
Proteomics	Yes	Yes	Yes	Yes	Yes
Phospho-proteomics	Yes	No	Yes	No	No
Genomics	Yes	Yes	Yes	Yes	Yes
Epi-genomics	Yes	No	No	No	Yes
Metabolomics	Yes	Yes	Yes	No	No
Multiple lists of same type of omics data	Yes	No	Yes	Yes	Yes
Finding deregulated proteins/genes/miRNAs	Yes	No	No	No	No
Map in meta-interactome	Yes	No	No	No	No
3D interactive network	Yes	Yes	No	No	No
Network topology analysis	Yes	No	No	No	No
Prognostic status of proteins/genes (in cancer)	Yes	No	No	No	No
Pathway enrichment analysis	Yes	Yes	Yes	Yes	No
Analysis for regulatory network protein links	Yes	No	No	No	No

Additionally, utilizing multi-omics data APODHIN calculates cross-pathway regulatory and PPI links connecting signaling proteins or transcription factors (TFs) or miRNAs to metabolic enzymes and their metabolites using network analysis and mathematical modeling. These cross-pathway links were shown to play important roles in metabolic reprogramming in cancer scenarios such as glioblastoma multiforme in a previous work (Bag et al., 2019).

In addition to the server part, APODHIN shares analysis of multi-omics data from various cancer cell lines where TINs and cross-pathway links were identified using publicly available omics datasets collected for various gynecological cancers. APODHIN platform is freely available at <http://www.hppi.iicb.res.in/APODHIN/home.html>.

MATERIALS AND METHODS

Server Description

Analysis of Pan-omics Data in Human Interactome Network web server is dedicated for the integration and subsequent analysis using single or multiple types of omics data. For single type of omics data, APODHIN can analyze multiple datasets (up to 3) which may correspond to either different stages of a disease from a single cohort or from dataset collected from multiple patient cohorts and/or cell lines.

For multiple types of omics data, APODHIN allows single input data file for each type of omics data. Following sections briefly describe the various analytical part of the APODHIN server.

Data Collection

Analysis of Pan-omics Data in Human Interactome Network web server is preloaded with a human cellular meta-interactome

network. This meta-interactome consists of human protein-protein interaction network (HPPIN), network of human miRNAs and their target genes and network of human TFs and their target genes. The PPI data was collected from STRING (Szklarczyk et al., 2019) database (version 11). Interactions having a medium threshold of experimental score ≥ 700 were considered (Ferretti and Cortelezzi, 2011) for construction of the PPIN. Target gene information of miRNAs was collected from the TarBase (Vergoulis et al., 2012) and miRTarBase (Chou et al., 2016) databases. From the TarBase database (version 6) we have taken reliable interactions supported only by low-throughput experiments (e.g., reporter gene assay, western blot, qPCR, etc.) whereas miRNA target interactions with strong confidence (i.e., validated by either of report assay, western blot, qPCR experiments) from miRTarBase (version 6) were considered for APODHIN meta-interactome network. We trusted on the more reliable low-throughput experimental data to build the parent miRNA-target mRNA interactome network. We found 2492 target genes for 544 miRNAs creating 6917 interactions. TFs and their target genes were downloaded from Human Transcriptional Regulation Interactions database (HTRIdb) (Bovolenta et al., 2012). We found 11887 target genes for 284 TFs creating 18153 interactions. These three networks were merged together to form the APODHIN meta-interactome consisting of two types of biomolecular nodes i.e., proteins/genes and miRNAs along with three types of interactions, i.e., protein-protein, miRNA-target gene, and TF-target gene, respectively.

Additionally, we have also included a network of metabolites as substrate and product with their corresponding metabolic enzymes in the APODHIN server. For constructing this network, we downloaded metabolic reactions from MetaNetX database (Moretti et al., 2016) and extracted the metabolites along with the corresponding metabolic enzymes and further filtered those

enzymes and metabolites which have been listed in the Human Metabolome Database (HMDB) database (Wishart et al., 2018).

Pan-omics Data Integration and Meta-Interaction Network Extraction

In APODHIN web server, user can upload single or multiple types of omics data. The server accepts RNA transcriptomics, miRNA transcriptomics, proteomics, phosphoproteomics, genomics, epigenomics, and metabolomics data. The current version of the server accept only processed format of the omics data where differential expression/abundance of corresponding biomolecules are provided with *logFC* for defining up and down regulation of genes/miRNAs/proteins and threshold probability or *p*-value. For RNA transcriptomics, miRNA transcriptomics and proteomics data user should select threshold values of *logFC* for defining up and down regulation of genes/miRNAs/proteins and corresponding adjusted *p*-value. Uploaded files should contain list of genes/miRNAs/proteins along with *logFC* and *p*-values. Sample file formats for different omics data are provided in the APODHIN help page. For genomics, epigenomics, and phosphoproteomics data, genes that are mutated and/or methylated and proteins, which are phosphorylated are considered, respectively. APODHIN help page also provides guidelines to process GEO (Barrett et al., 2013) transcriptomics data for using in APODHIN. Packages and tools for GEO series data are also enlisted in the APODHIN “Help” page. For other types, of omics data like, proteomics, genomics, metabolomics, useful links for data processing is provided in the APODHIN help page and it will be made more enriched gradually depending on the requirements from users.

Analysis of Pan-omics Data in Human Interactome Network web server extracts the interactome networks from the parent meta-interactome for the genes, mRNAs, miRNAs, proteins, and metabolites that are either deregulated or altered according to the user supplied single or multiple omics data. It creates a filtered meta-interactome network comprising of deregulated or altered nodes and their 1st or 2nd level (as chosen by user) interactors and/or regulators. For metabolomics data, the web server finds out the proteins linked with metabolites and constructs network. These single or multi omics data specific meta-interactome networks are subsequently displayed in an interactive three-dimensional (3D) network viewer within the APODHIN server. For creating omics data mapped network, and subsequently network analysis, APODHIN does not provide any special weight or scores to any type of omics data.

For the module “pathway connectivity analysis,” RNA transcriptomics, miRNA transcriptomics, and proteomics data were considered as primary data and submission of at least one of them is mandatory to define deregulated miRNAs and/or genes/proteins. In case of “pathway connectivity analysis,” the *logFC* values for each of the uploaded omics data is normalized in the scale of -1 to $+1$ following Eq. 1,

$$\log FC_{\text{normalized}} = \frac{\log FC}{|\log FC|_{\max}} \quad (1)$$

where positive and negative values indicate up and down regulated entities, respectively. If more than one primary omics data, for example, transcriptomics and proteomics are provided, APODHIN web server sums up the normalized *logFC* values from the different omics data for the same node (RNA/protein) and if the sum is non-zero, gene/protein is considered deregulated. Primary omics data determines whether the gene is deregulated or not. Also, if a gene is found not altered in supplied primary omics data, APODHIN does not consider this gene for further analysis, irrespective of its status in the supplied secondary omics data. Details of the utilization of the normalized omics values in mathematical modeling based pathway connectivity link identification are provided later. In this module, the information on metabolites for any enzyme can be obtained in the associated table on selection of enzyme.

Network Analysis and Identification of TINs

Once the context specific meta-interactome network is formed via utilization of user supplied single or multiple omics data, APODHIN web server primarily finds three types of TINs, namely, hubs, CNs (Bhattacharyya and Chakrabarti, 2015) and bottlenecks (BNs) (Yu et al., 2007). To find the important nodes, network and node indices like degree, betweenness, closeness and clustering coefficients are calculated from the extracted meta-interactome network. These node parameters were calculated using previously reported methods and protocols (Bhattacharyya and Chakrabarti, 2015). For transcriptomics and proteomics data, TINs are identified from the expressed nodes only. For phosphoproteomics, genomics, epigenomics and metabolomics data, TINs are identified from phosphorylated, mutated, methylated proteins/genes and metabolic enzymes, respectively.

Hubs are nodes that have high degrees. Degree distribution is normalized following Eq. 2,

$$x_{i,\text{normalized}} = \frac{x_i}{x_{\text{maximum}}} \quad (2)$$

where x_i is degree value of a node i and x_{maximum} is the maximum degree of the network. APODHIN web server converts normalized degree distribution to corresponding z-score distribution. The plot of probability distribution function (PDF) of z-scores for all nodes in network is sent to the user by email. This email shares intermediate results only. From the plot of PDF, users are asked to provide the threshold value for hub identification. After receiving the threshold value, APODHIN initiates hub identification program. Nodes having degree greater than the threshold value are considered as hub. It is also mentioned in the help page. Scores concerning individual centrality parameters like, betweenness, closeness and clustering coefficients are calculated and the cumulative centrality scores (CCS) are estimated by summing over the combined scores for first layer interactors (Bhattacharyya and Chakrabarti, 2015). CCSs are normalized following Eq. 2 where x is equal to CCS. Normalized CCS are converted into z-scores. The PDF of z-scores for all nodes of network are sent to the user by email and CNs are chosen based on the user provided threshold value of z-score following similar procedure as mentioned while identifying hubs.

Bottleneck nodes are characterized based on their betweenness values. Normalized betweenness values were obtained from Eq. 2 where x is betweenness and subsequently, converted into z-scores. Similar to hubs, bottleneck nodes are also chosen based on the user provided threshold z-score, chosen from the PDF plot of z-score for all nodes.

Further, sub-network consisting of TINs and their first or second layer interactors are constructed and displayed in an interactive three-dimensional (3D) network viewer.

The overlap of TINs, as well as all nodes of the network, as prognostic cancer marker is checked after extraction of prognostic marker information from the Human Protein Atlas database (version 19) (Uhlen et al., 2017). The prognostic data was obtained from Kaplan-Meier survival analysis. The cancer type, for which prognostic status have minimum p -value, is shown in the “Node information” table in the page of “network view of identified important nodes.” On mouse hover on the cancer type, more detail information for other cancer types, is available.

Pathway Mapping and Network of Mapped Pathways

For each identified TIN, particularly for genes and proteins, APODHIN maps the corresponding pathways listed in the KEGG database (Kanehisa et al., 2017). APODHIN performs a hypergeometric Fishers Exact test and selects enriched pathways satisfying p -value (p_{HGD}) ≤ 0.05 using the following contingency table and formula.

$$\begin{bmatrix} a & b - a \\ c & d - c \end{bmatrix}$$

Where,

- a = Number of genes in the pathway.
- b = Number of genes in the gene list.
- c = Total number of genes in the pathway.
- d = Total number of genes in all pathways in KEGG.

$$p_{HGD} = \frac{\binom{b}{a} \binom{d}{c}}{\binom{b+d}{a+c}} \quad (3)$$

Further, a network representation of important nodes along with their enriched mapped pathways is displayed in an interactive three-dimensional (3D) network viewer. **Figure 1A** shows the flow chart of “pan-omics data mapping and network analysis” module of APODHIN.

Pathway Connectivity Analysis and Cross-Pathway Links

This module of the APODHIN web server aims to construct regulatory interaction networks and subsequently identifies cross-pathway interaction links connecting different cellular pathway proteins [e.g., signaling proteins (S)], regulatory proteins [e.g., transcription factor (TF)] or miRNAs with metabolic pathway proteins (M).

For this purpose, APODHIN web server was preloaded with cross-pathway links or paths where protein–protein interactors (P) connect X nodes (X can be S or target gene of TF or target genes of miRNAs) with M (metabolic) proteins. We have limited the number (n) of protein–protein interactors (P) to a maximum value of three between X and M proteins. This limit provides four types of paths, XM ($n = 0$), XPM ($n = 1$), XPPM ($n = 2$), XPPPM ($n = 3$). These cross-pathway linking paths are filtered and selected based on expression and/or abundance status of the biomolecules supplied by user uploaded pan-omics data for a given disease or context. The filtering criteria for any given path is set when the terminal nodes are found to be deregulated and the remaining nodes are at least expressed within the user provided single or multi-omics datasets.

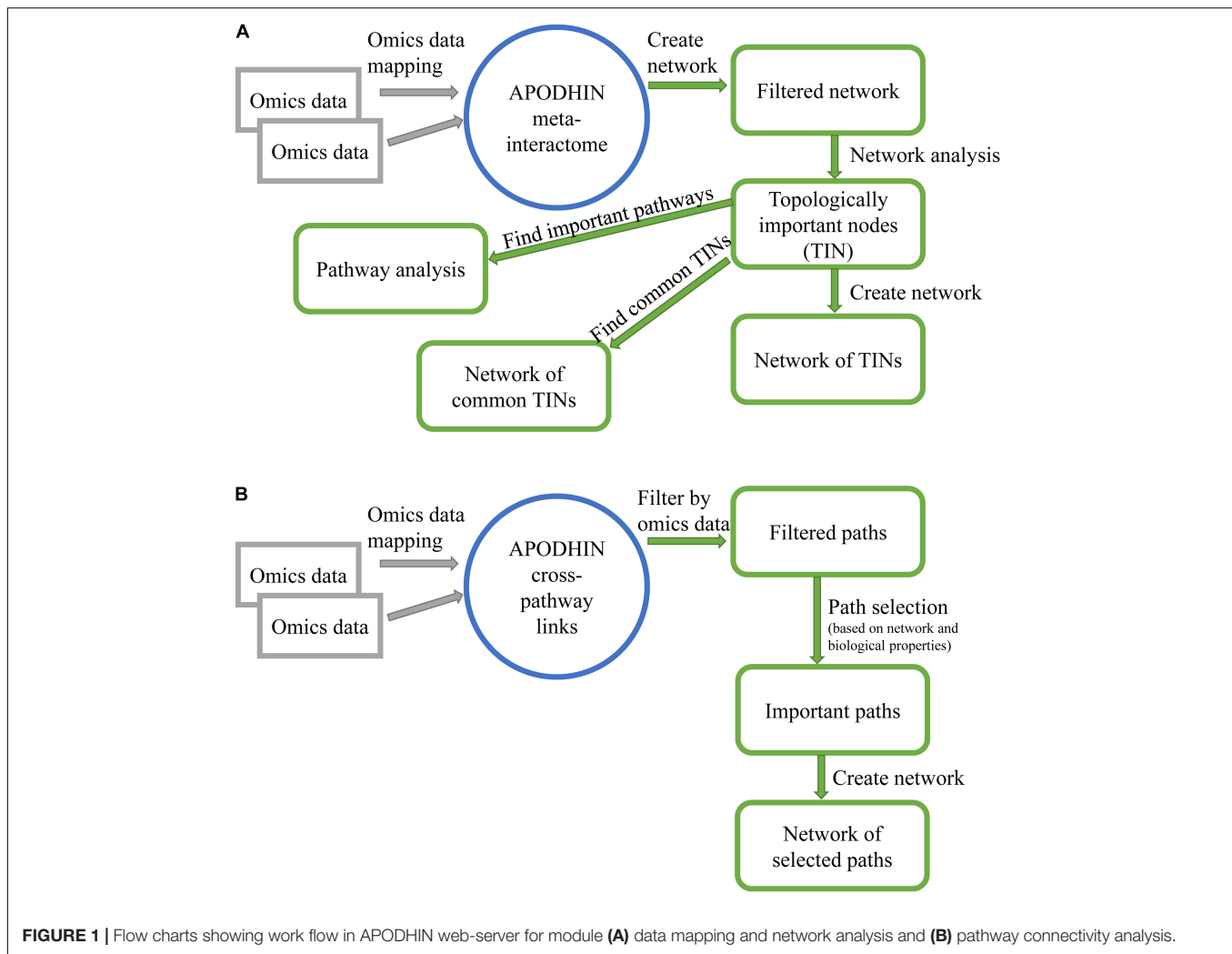
We implemented an established probabilistic approach based on the Hidden Markov Model (HMM) (Tuncbag et al., 2013; Vinayagam et al., 2014; Bag et al., 2019) utilizing the information of experimentally established PPIs and gene regulatory information to extract novel paths and interconnections between regulatory nodes such as signaling proteins, TFs and miRNAs and metabolic pathway proteins (M). Within these important X-M pairs, important cross-pathway connecting paths are again scored by considering all filtered paths between X-M pairs. To find important X-M pairs, weights are assigned on nodes and edges depending on network and biological properties. Edge weight is assigned in terms of normalized interaction probability which is proportional to the product of their expression scores.

Two types of node weights, network entropy, and effect-on-nodes are considered. Network entropy includes local entropy of the node. Another node weight parameter, “effect-on-node” considers the impact of interactors of a particular gene in the cross-connected network. The “effect-on-node” considers both biological and network properties of the node. Biological properties include deregulated gene, signaling crosstalk gene and rate limiting enzyme. Network properties include hubs, CNs, and bottlenecks.

Analysis of Pan-omics Data in Human Interactome Network web server allows the user to choose maximum four weight options out of the six weights. If a node satisfies any of the selected weight options, weight value 1 is assigned for each satisfied option. To identify important cross-connecting X-M pairs we have evaluated “path score” (PS) based on a HMM implemented within the core mathematical model that calculated the significant cross-pathway linking paths. “Path scores” are converted to z-scores and paths having z-score ≥ 1 are considered as important cross-connecting paths. A detailed description of the mathematical models and path calculation is available in our previous publication (Bag et al., 2019). **Figure 1B** shows the flow chart of “pathway connectivity analysis” module of APODHIN.

APODHIN Architecture

Analysis of Pan-omics Data in Human Interactome Network web server is created using HTML, PHP, PYTHON, and JAVA scripts. Client/user side scripts are written in HTML, PHP and JAVA scripts. User uploaded data is analyzed using PYTHON scripts. For network analysis, PYTHON package networkX (version



1.8.1) is used. For visualization of 3D presentation of networks JAVA scripts based open source technologies (*three.js* and *3d-force-graph.js*) were utilized.

Analysis of Pan-omics Data in Human Interactome Network has two separate parts A. APODHN server and B. APODHN example data analysis.

APODHN Server

Analysis of Pan-omics Data in Human Interactome Network web server is preloaded with human interactome network containing PPIN, target gene network of miRNAs and target gene network of TFs. Proteins participating in signaling and metabolic pathways are also marked separately. Metabolites along with their target enzymes are also included within APODHN. This meta-interactome network is used as framework of cellular interactions and is further used to map user supplied single or multiple types of “omics” data to perform the following analyses.

- Omics data mapping and network analysis: This module has two sub-modules. On clicking first submit button, this web server provides meta-interactome network filtered

by uploaded omics data where deregulated and/or altered nodes along with their interactors are included. Users can further proceed for finding important interacting nodes from the “pan-omics” data mapped interaction network by clicking second submit button. Tabular, graphical and 3D network representations of the identified TINs are provided for better appreciation. Overlap of the TINs is shown both in tabular and interactive 3D network visualization. Additionally, TINs and their enriched pathways are also shown in tabular and interactive 3D network visualization manner.

Sample input files for each omics data type and example analysis output are provided for the ease of use and apprehension.

- Pathway connectivity analysis: As mentioned before, this sub-module highlights significant PPI and regulatory paths connecting signaling proteins/TF/miRNAs to metabolic proteins. These cross-pathway links are thought to be supra-molecular regulatory links/signatures connected with metabolic rearrangement or reprogramming events that are observed during cancer. In APODHN, these

cross-pathway regulatory links can be constructed from three types of interaction networks.

1. Integrated network where signaling (S) and metabolic (M) pathway proteins are connected through protein–protein interactors (P).
2. Integrated network where target genes of TFs and metabolic (M) pathway proteins are connected through protein–protein interactors (P).
3. Integrated network where miRNA target genes and metabolic (M) pathway proteins are connected through protein–protein interactors (P).

Cross-pathway linking paths are filtered and selected based on expression and/or abundance status of the biomolecules supplied by user uploaded single or pan-omics data for a given disease or context. These paths are shown both in tabular and interactive 3D network visualization.

APODHIN Example Data Analysis

Analysis of Pan-omics Data in Human Interactome Network example data analysis page showcase few example analysis of multi-omics data for different cancer cell lines. We have used the APODHIN web server to construct individual cancer and dataset centric meta-interactome network using cell line specific single and/or multi-omics data collected from various resources such as GEO (Barrett et al., 2013), PRIDE (Perez-Riverol et al., 2019), publication reports and data sources for cervical, ovarian, and breast cancers, respectively. Further, these cancer and dataset specific meta-interactome networks were analyzed and important interacting nodes and cross-pathway links were identified and provided within the APODHIN example data analysis module. We have used cancer cell line derived omics data freely available from different public resources. Options are provided for the users to select single and/or multi-omics data to construct the meta-interactome networks and further analyze them to identify and important interacting nodes and cross-pathway links specific for the selected dataset.

RESULTS

Input Options

Analysis of Pan-omics Data in Human Interactome Network server provides two different but linked analysis options for the users who would like to utilize single or multiple types of omics data for a given context. APODHIN web server provides options to upload seven types of “omics” data comprising of mRNA transcriptomics, miRNA transcriptomics, proteomics, phosphoproteomics, genomics, epigenomics, and metabolomics. The file formats for each data type is specified in the “Help” page and sample input files are also available in the server input page. Information on preparing input files for using in APODHIN is also shared in the “Help” page. For transcriptomics and proteomics data, maximum and minimum threshold values for the differential expression/abundance ($\log_{2}FC$) and statistical significance of that (p -values) need to be provided. As the calculations are computation intensive, results are sent via email.

Similarly, for cross-pathway connectivity analysis users need to upload single or multiple types of “omics” data for a given context. At least one “primary” type (see Methods) of omics data need to be uploaded. Now, in this case, users also need to specify the type of connectivity they would like to explore, for example, signaling to metabolic proteins, TFs to metabolic proteins, or miRNAs to metabolic proteins. Only one type of pathway connectivity can be explored at a time for a given set of “omics” data. Additionally, users also need to select the kind of weights (see section “Materials and Methods”) that would be applied while calculating the scores of the selected cross-pathway regulatory and PPI paths. E-mail address needs to be supplied for APODHIN server to send the result link of the identified cross-pathway connections.

Output Options

Output option for the “Data mapping and network analysis” module has two stages. At first stage (Figure 2A), the context specific meta-interactome network (“filtered network”) can be visualized via a user interactive 3D network viewer where information regarding each node and edge are provided in graphical as well as tabular view (Figure 2B). Status of the “omics” data mapping is shown in various color codes for the nodes whereas different relationship like PPI, miRNA-target gene interaction, and TF and target in connections are shown varied color codes. Additional details about the protein nodes can be obtained via GeneCards (Stelzer et al., 2016) link while miRNA details can be found via miRTarBase (Chou et al., 2016) link. List of metabolites mapped onto the protein nodes are also provided both in the network viewer as well as in the adjacent tabular format. If network analysis is opted, along with filtered network, APODHIN provides the PDFs for the opted TINs (Figure 2C). Filtered nodes (genes/proteins/miRNAs) that satisfied the selected threshold criteria are characterized as TINs and further utilized for meta-interactome network construction. If multiple files of single type of omics data is uploaded, users can see the number of TINs (as hub, bottlenecks, and CNs) and their mutual overlap using interactive Venn diagram by clicking the “link for analysis” option for single or combination of “omics” data (Figure 2D). Combined analysis of multiple types of omics data files is shown if multiple types of omics data files are provided. Here also, the resultant page (Figure 2E) provides three output options. First, the regulatory and PPI connectivity specific to the hubs, bottleneck and CNs can be seen via corresponding link where networks of deregulated hubs, bottleneck, and CNs can be seen separately and saved accordingly (Figure 2F). Association to various kinds of cancers for the identified TINs as favorable/unfavorable prognostic markers are also provided here after mapping the TINs (see Methods) to the data provided in Human Protein Atlas (Uhlen et al., 2017). Another option provides the network of common TINs (Figure 2G) whereas a separate link provides network of enriched pathways with the identified TINs (Figure 2H). Enriched pathway networks of deregulated hubs, bottleneck, and CNs can be seen separately and saved accordingly. In all these three network output options, data can be downloaded in text format for further analysis.

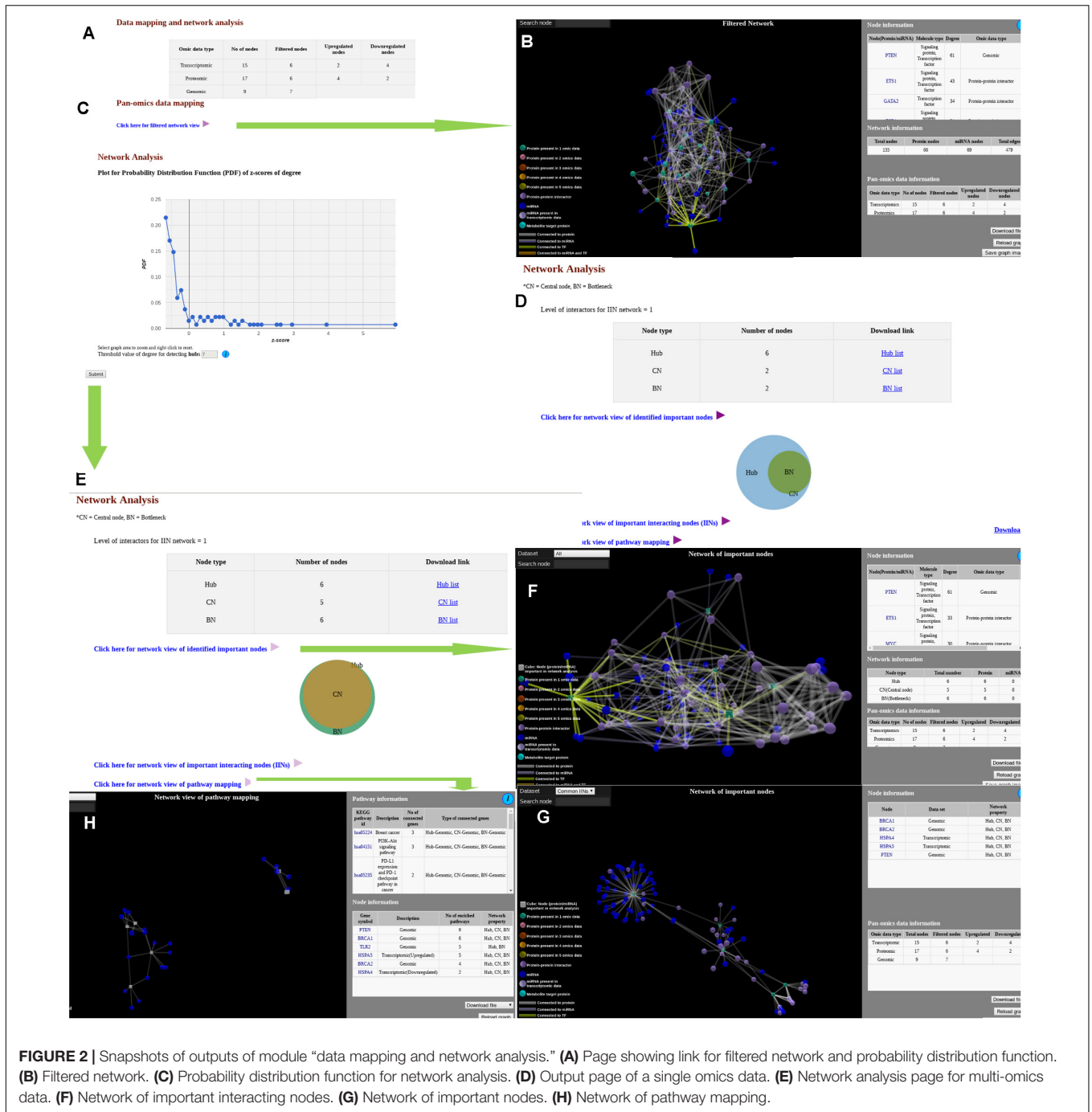


FIGURE 2 | Snapshots of outputs of module “data mapping and network analysis.” (A) Page showing link for filtered network and probability distribution function. (B) Filtered network. (C) Probability distribution function for network analysis. (D) Output page of a single omics data. (E) Network analysis page for multi-omics data. (F) Network of important interacting nodes. (G) Network of important nodes. (H) Network of pathway mapping.

Similar to “Data mapping and network analysis,” “Pathway connectivity analysis” module also provides a tabular result with a summary of the user uploaded data (Figure 3A). Users can see the cross-pathway links for multiple types of omics data (Figure 3B). For multiple types of files with single type of omics data (Figure 3C), the comparison (Figure 3D) is shown in Venn diagram as well as in network visualization. In the 3D network visualization window, significant PPI and regulatory paths connecting signaling proteins/TFs/miRNAs to metabolic proteins are shown in color coded fashion. As described before,

these cross-pathway links or paths connect X nodes (X can be S or target gene of TF or target genes of miRNAs) with metabolic (M) proteins. These linking paths are filtered and selected based on expression and/or abundance status of the biomolecules supplied by the users where for any given path the terminal nodes are found to be deregulated and the remaining nodes are at least expressed. The corresponding pathways and biological functions of the proteins are also provided in tabular format adjacent to the network viewer. Additionally, the metabolites connected to the metabolic proteins that are

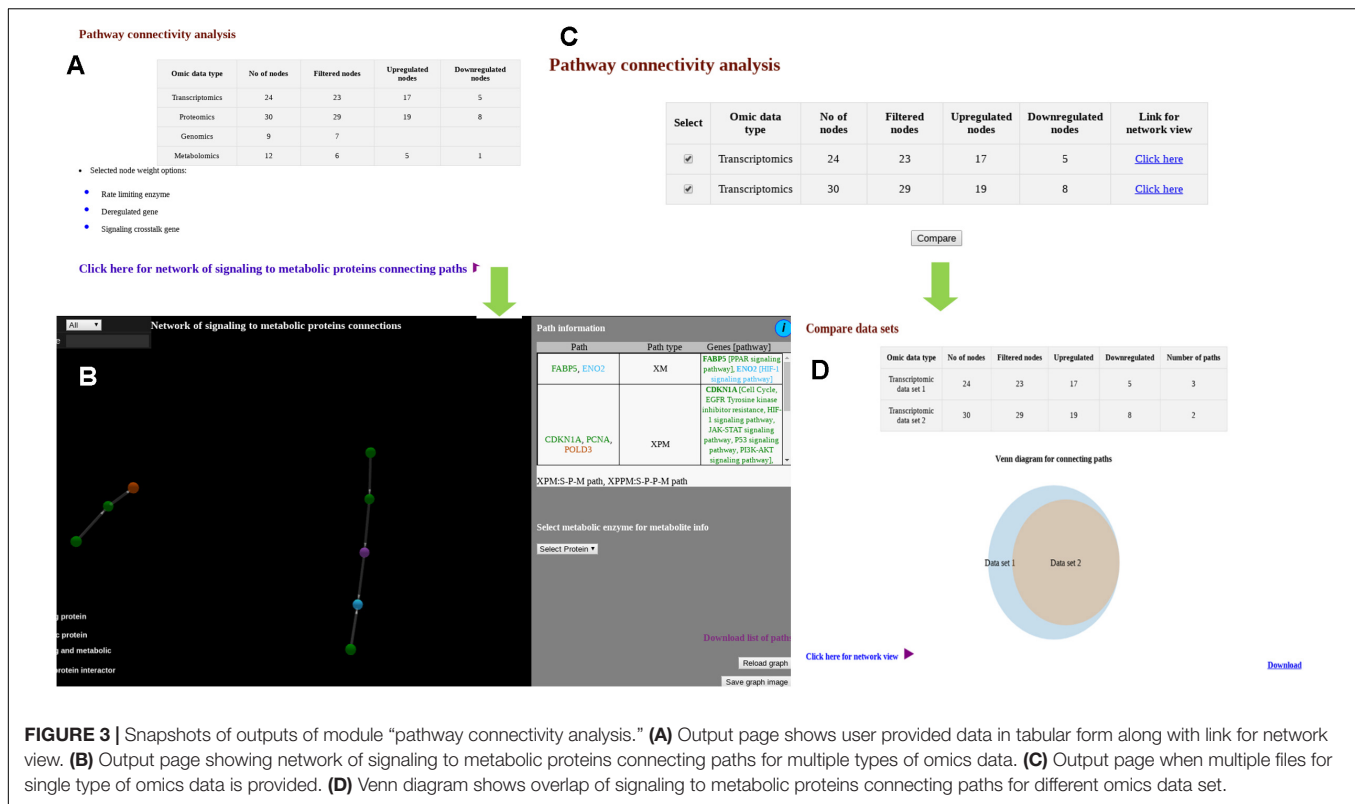


FIGURE 3 | Snapshots of outputs of module “pathway connectivity analysis.” **(A)** Output page shows user provided data in tabular form along with link for network view. **(B)** Output page showing network of signaling to metabolic proteins connecting paths for multiple types of omics data. **(C)** Output page when multiple files for single type of omics data is provided. **(D)** Venn diagram shows overlap of signaling to metabolic proteins connecting paths for different omics data set.

part of the selected cross-pathway links are also provided in the same page.

Example Data Analysis Option

Analysis of Pan-omics Data in Human Interactome Network example data analysis page contains important nodes (genes/proteins/miRNAs), pathways, and their networks with interacting partners specific for cancers affecting women such as cervical, ovarian, and breast cancer. This section also contains important paths linking signaling proteins/TFs/miRNAs to metabolic enzymes, which could perhaps be responsible for metabolic reprogramming in cancer. The example content is produced by APODHIN web server using publicly available cervical, ovarian, and breast cancer specific cell line based omics data. **Figure 4** briefs the statistics derived from APODHIN example analyses for mRNA transcriptomics data of different cell lines of cervical, ovarian, and breast cancer. **Figure 4A** shows the overlap of deregulated genes. It reveals lesser overlap among deregulated genes across cell lines for all cancers. However, there is almost complete overlap of pathways mapped by deregulated genes (**Figure 4B**). Nodes satisfying any two types of TINs are considered as important interacting nodes (IINs). **Figure 4C** shows overlap for common IINs between cell lines across cancer types are observed. Similarly, **Figure 4D** shows much higher overlap of common pathways mapped by IINs. This demonstrates that IINs and their pathways represent the common core genes and processes related to a cancer type in a better way than that achieved by the initial deregulated genes obtained from the omics data. We also checked whether the

mapped pathways are related to cancer pathways enlisted in KEGG database (Kanehisa et al., 2017). **Figure 5A** shows that pathways mapped by IINs are more cancer specific compared to the pathways mapped by deregulated genes for all cell lines. **Figures 4E,F** show the number and overlap of deregulated genes and IINs as prognostic markers of respective cancer type. **Figure 5B** shows that compared to the deregulated genes, IINs possess higher fractions of prognostic markers for all cancer cell lines, except MDAMB231. This advocates the usefulness of the IINs over deregulated genes. Moreover, as the number of IINs is much smaller than that of deregulated genes the false discovery rate is also expected to be lower.

Figure 6 shows overlap of cross-pathway links or paths connecting signaling (S) proteins, TFs, and miRNAs to metabolic (M) proteins identified using omics data derived from the cell lines of three types of cancers. For signaling to metabolic connection, four common paths for three cervical cell lines were observed. However, no such overlap was found for breast and ovarian cancer cell lines.

Analysis of pan-omics data considering transcriptomics, genomics, epigenomics, metabolomics data in different combinations are available for different cell lines in the example data analysis section of APODHIN.

DISCUSSION

Large-scale genomics, transcriptomics and proteomics approaches have made it possible to characterize different

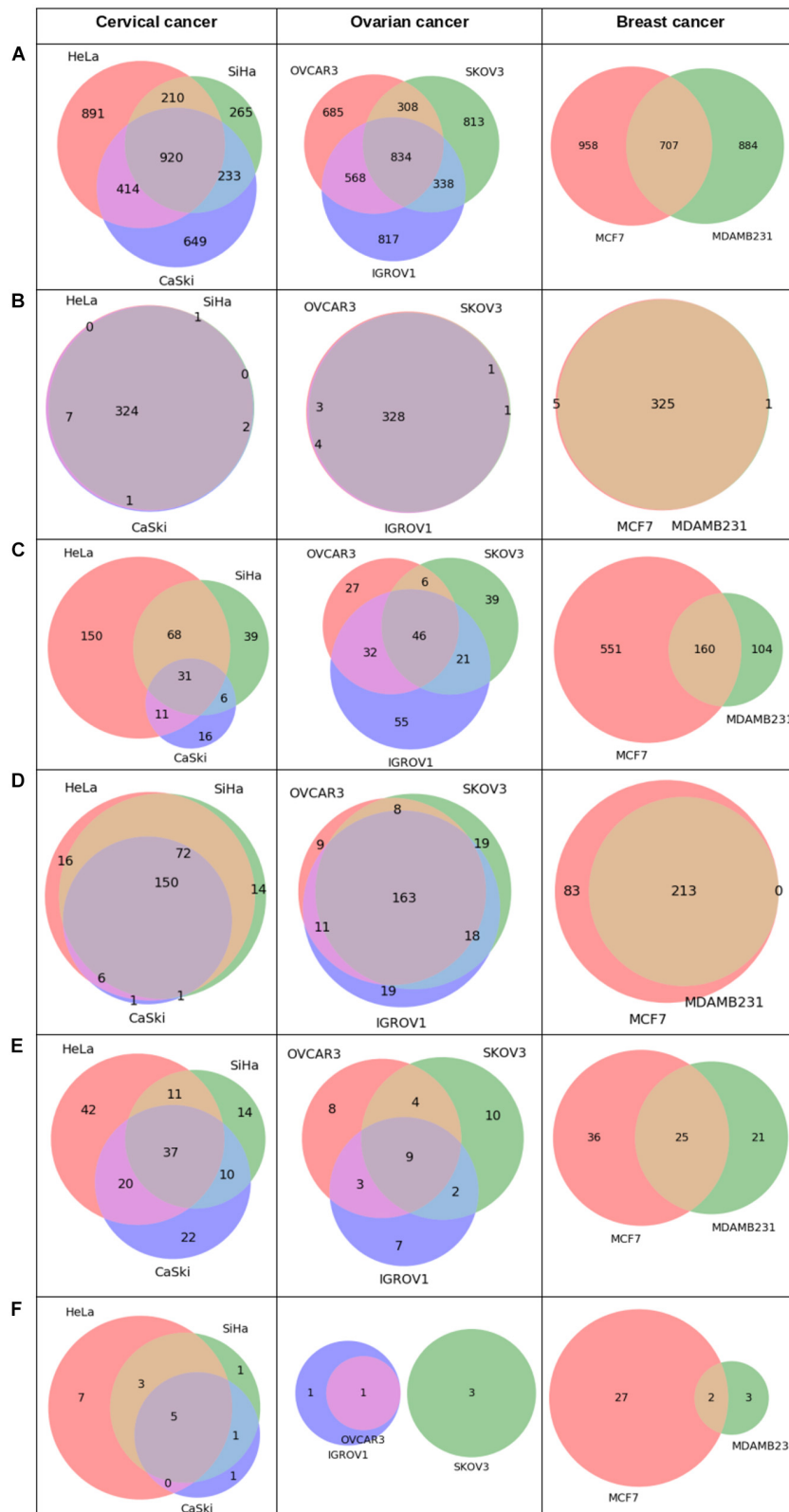


FIGURE 4 | Statistics derived from APODHDIN database for mRNA transcriptomics data derived from different cell lines of cervical (HeLa, SiHa, and CaSki), ovarian (IGROV1, SKOV3, OVCAR3), and breast cancer (MCF7 and MDAMB231). Transcriptomics data was derived from the GEO datasets GSE9750, GSE19352, and GSE71363, respectively. **(A)** Deregulated genes, **(B)** Overlap of pathways mapped by deregulated genes, **(C)** Overlap of IINs, **(D)** Overlap of pathways mapped by IINs, **(E)** Deregulated genes as prognostic marker, and **(F)** IINs as prognostic marker.

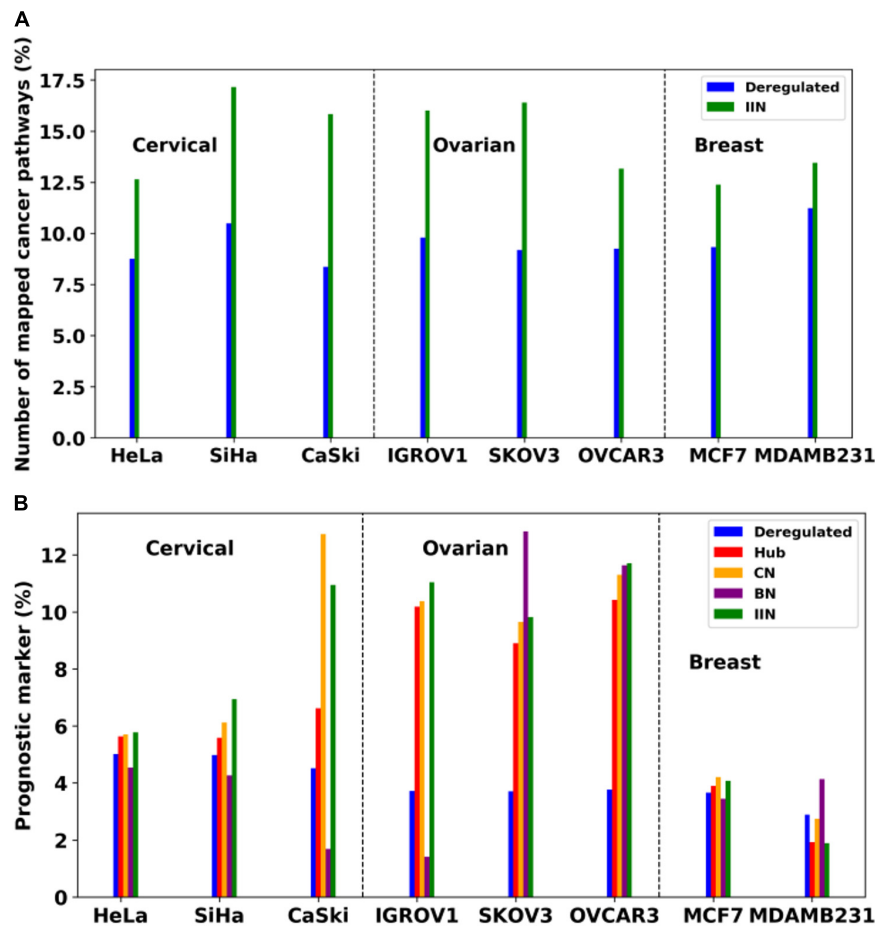
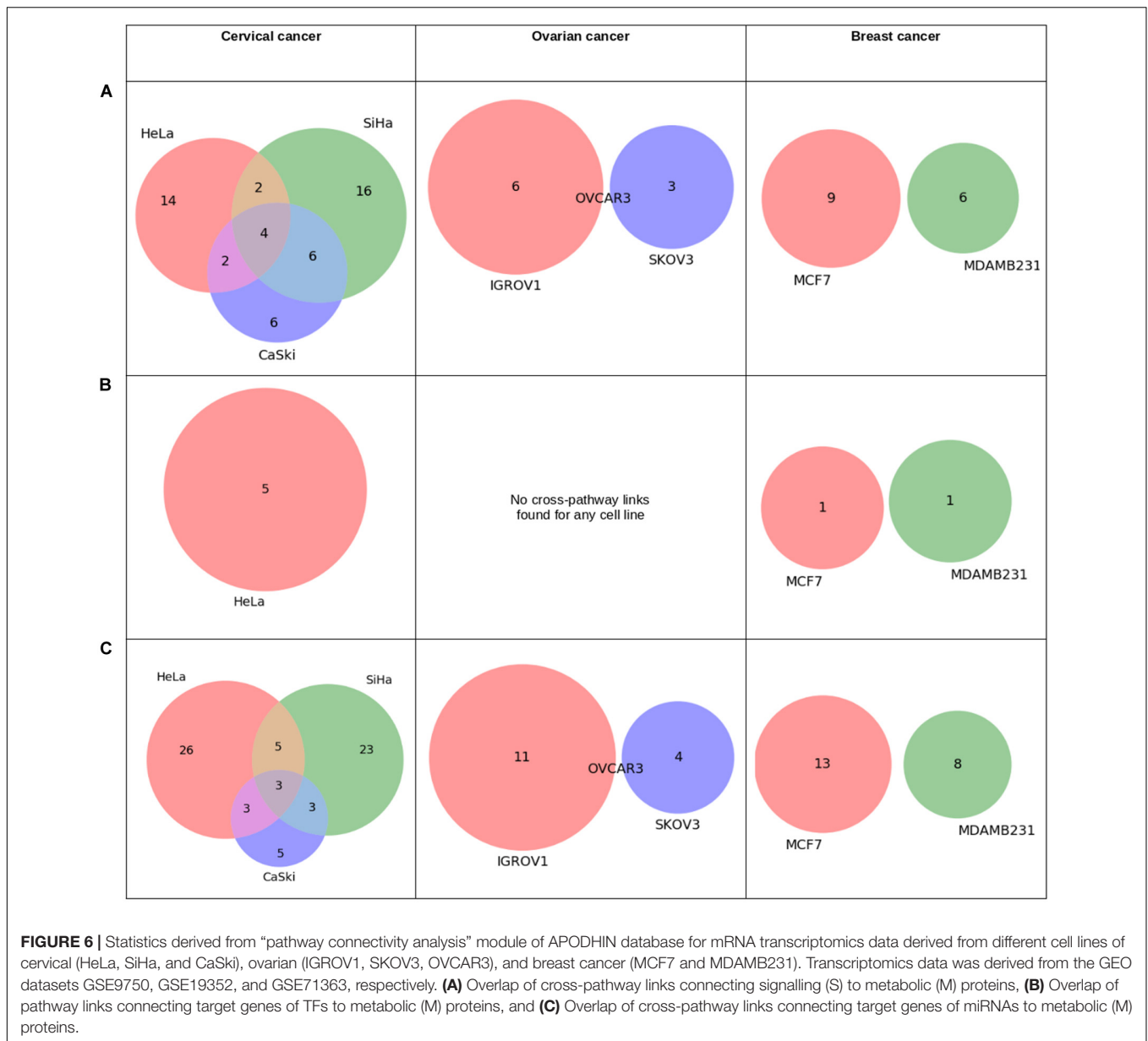


FIGURE 5 | (A) Comparison of number of cancer specific pathways mapped by deregulated genes and IINs. **(B)** Comparison of fraction of prognostic markers within the deregulated genes and network analysis derived important nodes, such as IIN and various TIN (e.g., Hubs, CN, and BN, respectively). Dashed lines are drawn to separate cell lines of different cancer types.

clinical spectra associated with cancers. Use of pan-omics platforms and approaches in the analysis of systemic disease like cancer will not only help to identify numerous useful biomarkers but also will expose areas for further improvement in therapeutic intervention. Here, we present APODHIN web server, which extracts cellular interactome networks from the parent meta-interactome for the genes, mRNAs, miRNA, proteins, and metabolites that are either deregulated or altered according to the user supplied single or multiple omics data. These single or multi-omics data specific meta-interactome networks are utilized to identify TINs and their sub-modules enriched with PPI and regulatory relationship via utilization of graph theory based network analyses and biological pathway enrichment analysis. Important interacting nodes (proteins and miRNAs), IINs are identified based on the overlap of key nodes such as hubs and bottlenecks. Using data from The Human Protein Atlas database, APODHIN provides the probable prognostic status of the IINs. Also, as observed in our earlier works (Bhattacharyya and Chakrabarti, 2015), IINs extracted from network topology, could correlate to be prospective diagnostic

and/or prognostic biomarkers or even turn out to be potential therapeutic targets.

Molecular mechanisms for cancer progression and development of potential therapeutics to inhibit these complex diseases are difficult from the independent knowledge of signaling, TFs, miRNAs, and metabolic pathways. Metabolic reprogramming is an essential hallmark of cancer (Hanahan and Weinberg, 2011). Understanding the coordination among various cellular pathways, such as gene-regulatory, signaling and metabolic pathways is crucial and may provide clues into the molecular mechanism of metabolic adaptation in cancer and associated cells. Therefore, there is an urgent need for systems biology model, which can coordinate among signaling-induced proliferation of tumor cells/growth, transcription factor/miRNA based gene regulation and metabolic processes. Hence, we emphasized to design a mathematical approach to identify significant proteins forming interconnections between signaling, regulatory and metabolic pathways. We have constructed an integrated network where signaling (S), regulatory (TFs and miRNAs), and metabolic (M) pathway entities are connected



through protein–protein and gene regulatory interactions. Interconnections between regulatory components such as signaling proteins/TFs/miRNAs and metabolic pathways need to be elucidated rigorously to understand the role of oncogene and tumor suppressors in regulation of metabolism alongside their normal functions. Analyses of such cross-connected network and linking paths will facilitate probable way(s) to inhibit cancer progression in a more specific manner.

Considering the growing demand of multi-omics data integration followed by systems biology based analytical interpretation of the large-scale “omics” data, implementation of a robust and user-friendly web-based platform is very much due. In order to make better sense out of the various “omics” data, it is imperative to utilize them in a way so that the global

scenario of the complex and multi-layer cellular interactome can be recapitulated. Several data portals have been coming up to make multi-omics data accessible, visible and more importantly, interpretable. Various programs and web portals are being made to interpret omics data in different perspectives. Each of these tools has its own merits and limitations also. **Table 1** provides a qualitative comparison of features and functionalities of APODHIN with respect to existing omics data analysis tools. Web servers like OmicsNet (Zhou and Xia, 2018) is a technically powerful web based platform specifically meant for better visualization of molecular networks. It mainly provides varied and efficient ways of network visualization including different components. However, it provides minimal emphasis on networks analysis and identification and interpretation of

important interacting nodes and cross-pathway links. Similarly, this server only accept differential omics data for genes/proteins and metabolites, it does not have the option to include the epigenetic modification, miRNA expression data, and phosphoproteomics data. mixOmics (Rohart et al., 2017) is a software package which is based on multi-variate analysis. It performs data reduction, and then identifies combination of biomarkers. It offers a network visualization but does not consider network topology. It does not consider any meta-interactome. Software package iOmicsPASS (Koh et al., 2019) considers a meta-interactome by including PPIN and TF regulatory network within omics data. But it excludes miRNA-mRNA regulatory network. It considers only three types of omics data, transcriptomics, proteomics, DNA copy number data, thus limiting its applicability. Another R package Miodin (Ulfenborg, 2019) provides opportunity of creating a workflow of data analysis. It considers different omics data, but not metabolomics data. It requires pre-installation of several R packages. Miodin provides Venn diagram of differentially expressed genes, overlapped within different datasets. However, Miodin does not consider any meta-interactome and does not construct any network. None of these tools perform network topology analysis and provide cross-pathway connectivity information of proteins. APODHIN is perhaps the only available web based platform that offers to (a) integrate multi-omics data onto an exhaustive multi layered cellular meta-interactome network, (b) extract and analyze the context specific networks and sub-networks to identify TINs that could serve as potential biomarkers and/or therapeutic targets (c) rationalize the role of the identified TINs to the given context via pathway enrichment and prognostic marker correlation, and (d) identify cross-pathway interconnections between regulatory components such as signaling proteins/TFs/miRNAs and metabolic pathways for better understanding the role of oncogenes and tumor suppressors in regulation of metabolic reprogramming during cancer. Additionally, being a web based tool, APODHIN requires no installation of software, good computing systems, and technical expertise. We believe these features make APODHIN useful as well as a user-friendly application.

However, there is still scope for improvement for the APODHIN server. The example data analysis part can be enriched to upgrade as a database. For example, in future we

would like to equip the server to accept and process raw “omics” data directly and further create the processed data for genetic or epigenetic alterations, differential expression and abundance, respectively. We would also like to add components for handling large number of datasets which will be able to analyze cohort data. Current version is mostly aimed to patient-specific personalized data. Similarly, the server and along with a database should be enriched in such a way that it could be utilized for deep learning and artificial intelligence based tools to predict the disease outcome, recurrence and drug resistance, respectively.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article.

AUTHOR CONTRIBUTIONS

NB and SC designed the web server. NB created the web server. KK, SB, and RB provided the data for meta-interactome network. NB and SC analyzed the data and drafted the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The authors acknowledge CSIR-Indian Institute of Chemical Biology for infrastructural support. SC acknowledges the Systems Medicine Cluster (SyMeC) grant (GAP357), Department of Biotechnology (DBT) for funding. NB acknowledges the Systems Medicine Cluster (SyMeC) grant (GAP357), Department of Biotechnology (DBT) for fellowship. KK, SB, and RB acknowledge Department of Biotechnology (DBT), Council of Scientific and Industrial Research (CSIR), respectively for their fellowships. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. This manuscript has been released as a pre-print at bioRxiv (Biswas et al., 2020).

REFERENCES

- Argelaguet, R., Velten, B., Arnol, D., Dietrich, S., Zenz, T., Marioni, J. C., et al. (2018). Multi-Omics factor analysis — a framework for unsupervised integration of multi-omics data sets. *Mol. Syst. Biol.* 14:e8124.
- Ashtiani, M., Salehzadeh-yazdi, A., Razaghi-moghadam, Z., Hennig, H., and Wolkenhauer, O. (2018). A systematic survey of centrality measures for protein-protein interaction networks. *BMC Syst. Biol.* 12:80. doi: 10.1186/s12918-018-0598-2
- Bag, A. K., Mandloi, S., Jarmalavicius, S., and Mondal, S. (2019). Connecting signaling and metabolic pathways in EGF receptor-mediated oncogenesis of glioblastoma. *PLoS Comput. Biol.* 15:e1007090. doi: 10.1371/journal.pcbi.1007090
- Barabási, A.-L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113. doi: 10.1038/nrg1272
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: Archive for functional genomics data sets - Update. *Nucleic Acids Res.* 41, 991–995.
- Bhattacharyya, M., and Chakrabarti, S. (2015). Identification of important interacting proteins (IIPs) in *Plasmodium falciparum* using large-scale interaction network analysis and in-silico knock-out studies. *Malar J.* 14, 1–17.
- Biswas, N., and Chakrabarti, S. (2020). Artificial intelligence (AI) based systems biology approaches in multi-omics data analysis of cancer. *Front. Oncol.* 10:588221. doi: 10.3389/fonc.2020.588221
- Biswas, N., Kumar, K., Bose, S., Bera, R., and Chakrabarti, S. (2020). Analysis of pan-omics data in human interactome network (APODHIN). *bioRxiv* [Preprint], doi: 10.1101/2020.04.18.048207
- Bovolenta, L. A., Acencio, M. L., and Lemke, N. (2012). HTRIdb: an open-access database for experimentally verified human transcriptional regulation interactions. *BMC Genom.* 13:405. doi: 10.1186/1471-2164-13-405

- Bray, F., Ferlay, J., and Soerjomataram, I. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Chou, C.-H., Chang, N.-W., Shrestha, S., Hsu, S.-D., Lin, Y.-L., Lee, W.-H., et al. (2016). miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res.* 44, 239–247.
- Ferretti, L., and Cortelezzi, M. (2011). Preferential attachment in growing spatial networks. *Phys. Rev. E* 84:016103.
- Gautam, P., Jaiswal, A., Aittokallio, T., Al-ali, H., and Wennerberg, K. (2019). Phenotypic screening combined with machine learning for efficient identification of breast cancer-selective therapeutic targets. *Cell Chem. Biol.* 26, 970–979. doi: 10.1016/j.chembiol.2019.03.011
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674. doi: 10.1016/j.cell.2011.02.013
- Hern, R., Tarazona, S., Mart, C., Balzano-nogueira, L., Furi, P., Pappas, G. J., et al. (2018). PaintOmics 3: a web resource for the pathway analysis and visualization of multi-omics data. *Nucleic Acids Res.* 46, W503–W509.
- Jeong, H., Mason, S. P., Barabási, A.-L., and Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *Nature* 411, 41–42. doi: 10.1038/35075138
- Kan, Z., Ding, Y., Kim, J., Jung, H. H., Chung, W., Lal, S., et al. (2018). Multi-omics profiling of younger Asian breast cancers reveals distinctive molecular signatures. *Nat. Commun.* 9:1725.
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361.
- Koh, H. W. L., Damian, F., Vogel, C., Choi, K. P., Ewing, R. M., and Choi, H. (2019). iOmicsPASS: network-based integration of multiomics data for predictive subnetwork discovery. *NPJ Syst. Biol. Appl.* 5:22.
- Mcgrail, D. J., Federico, L., Li, Y., Dai, H., Lu, Y., Mills, G. B., et al. (2018). Multi-omics analysis reveals neoantigen-independent immune cell infiltration in copy-number driven cancers. *Nat. Commun.* 9:1317.
- Mistry, D., Wise, R. P., and Dickerson, J. A. (2017). DiffSLC: a graph centrality method to detect essential proteins of a protein-protein interaction network. *PLoS One* 12:e0187091. doi: 10.1371/journal.pcbi.0187091
- Moretti, S., Martin, O., Van Du Tran, T., Bridge, A., Morgat, A., and Pagni, M. (2016). MetaNetX/MNXref - Reconciliation of metabolites and biochemical reactions to bring together genome-scale metabolic networks. *Nucleic Acids Res.* 44, D523–D526.
- Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., et al. (2019). The PRIDE database and related tools and resources in: Improving support for quantification data. *Nucleic Acids Res.* 47, D442–D450.
- Ramazzotti, D., Lal, A., Wang, B., and Batzoglou, K. (2018). Serafim, Sidow A. Multi-omic tumor data reveal diversity of molecular mechanisms that correlate with survival. *Nat. Commun.* 9:4453.
- Rohart, F., Gautier, B., Singh, A., and Cao, K. A. L. (2017). mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* 13:e1005752. doi: 10.1371/journal.pcbi.1005752
- Sandhu, C., Qureshi, A., and Emili, A. (2018). Panomics for precision medicine. *Trends Mol. Med.* 24, 85–101. doi: 10.1016/j.molmed.2017.11.001
- Shu, L., Zhao, Y., Kurt, Z., Byars, S. G., Tukiainen, T., Kettunen, J., et al. (2016). Mergeomics: multidimensional data integration to identify pathogenic perturbations to biological systems. *BMC Genom.* 17:874. doi: 10.1186/s12864-016-3198-9
- Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., et al. (2016). The GeneCards suite: from gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinform.* 54, 1.30.1–1.30.33.
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-cepas, J., et al. (2019). STRING v11: protein - protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, 607–613.
- TCGA (2020). Available at: <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga> (accessed 22 July 2020).
- Tuncbag, N., Braunstein, A., Pagnani, A., Huang, S. S. C., Chayes, J., Borgs, C., et al. (2013). Simultaneous reconstruction of multiple signaling pathways via the prize-collecting steiner forest problem. *J. Comput. Biol.* 20, 124–136. doi: 10.1089/cmb.2012.0092
- Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhor, G., et al. (2017). A pathology atlas of the human cancer transcriptome. *Science* 357:660.
- Ulfenborg, B. (2019). Vertical and horizontal integration of multi-omics data with miodin. *BMC Bioinform.* 20:649. doi: 10.1186/s12859-019-3224-4
- Ulgen, E., Ozisik, O., and Sezerman, O. U. (2019). pathfindR: an R package for comprehensive identification of enriched pathways in omics data through active subnetworks. *Front. Genet.* 10:858. doi: 10.3389/fgene.2019.00858
- Vasaikar, S. V., Straub, P., Wang, J., and Zhang, B. (2018). LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res.* 46, D956–D963.
- Vergoulis, T., Vlachos, I. S., Alexiou, P., Georgakilas, G., Maragkakis, M., Reczko, M., et al. (2012). TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. *Nucleic Acids Res.* 40, D222–D229.
- Vinayagam, A., Zirin, J., Roesel, C., Hu, Y., Yilmazel, B., Samsonova, A. A., et al. (2014). Integrating protein-protein interaction networks with phenotypes reveals signs of interactions. *Nat. Methods* 11, 94–99. doi: 10.1038/nmeth.2733
- Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Azquez-Fresno, R. V., et al. (2018). HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* 46, D608–D617.
- Xie, B., Yuan, Z., Yang, Y., Sun, Z., Zhou, S., and Fang, X. (2018). MOBcDb?: a comprehensive database integrating multi - omics data on breast cancer for precision medicine. *Breast Cancer Res. Treat.* 169, 625–632. doi: 10.1007/s10549-018-4708-z
- Yang, Y., Sui, Y., Xie, B., Qu, H., and Fang, X. (2019). GliomaDB: a web server for integrating glioma omics data and interactive analysis. *Genom. Proteom. Bioinform.* 17, 465–471. doi: 10.1016/j.gpb.2018.03.008
- Yu, H., Kim, P. M., Sprecher, E., Trifonov, V., and Gerstein, M. (2007). The importance of bottlenecks in protein networks: Correlation with gene essentiality and expression dynamics. *PLoS Comput. Biol.* 3:e59. doi: 10.1371/journal.pcbi.0030059
- Zhang, J., Baran, J., Cros, A., Guberman, J. M., Haider, S., Hsu, J., et al. (2011). International cancer genome consortium data portal-a one-stop shop for cancer genomics data. *Database* 2011:bar026. doi: 10.1093/database/bar026
- Zhou, G., and Xia, J. (2018). OmicsNet: a web-based tool for creation and visual analysis of biological networks in 3D space. *Nucleic Acids Res.* 46, W514–W522.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Biswas, Kumar, Bose, Bera and Chakrabarti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.