



Multipath2.0: Extending Multilayer Reproducible Pathway Models with Omics Data

Zaynab Hammoud^{a,b,*}, Mohammad Al Maaz^c, Alicia D'Angelo^a, Frank Kramer^a

^a IT-Infrastructure for Translational Medical Research, University of Augsburg, Augsburg, Germany

^b Robert Bosch Hospital GmbH, IT-Department, Stuttgart, Germany

^c University of Applied Sciences Bonn-Rhein-Sieg, Department of Natural Sciences, Bonn, Germany

ARTICLE INFO

Keywords:

Multilayer graphs
Omics data
Data integration
Biological pathways
Reproducibility
Visualization

ABSTRACT

Background: Biological systems are often perceived as independent and consequently analyzed individually. In the field of omics, multiple disciplines target the study of specific types of molecules, such as genomics. The support of more data sources in these analyses is becoming more crucial for understanding the interplay of biological systems. However, this requires integration of heterogeneous knowledge, which is considered highly challenging in bioinformatics and biomedicine. Therefore, the R package Multipath was developed to model biological pathways as multilayered graphs and integrate influencing knowledge including proteins and drugs. In its previous form, Multipath generated multilayer models of BioPAX-encoded pathways and included features to integrate drug and protein information from DrugBank and UniProtKB respectively. Although the model showed remarkable utility, including further data sources ensures enriching and expanding its capabilities.

Results: In this paper, a new version Multipath 2.0 is presented. The update additionally supports the two databases KEGG Genes and OMIM, which serve as the source for gene and disease entries and interactions. Information on the interactions between the previously and newly added nodes are extracted and integrated. The Multipath 2.0 offers features to update the original multilayer model and integrate the corresponding nodes and edges into two additional layers referring to KEGG Genes and OMIM. Furthermore, the embedded nodes are inter- and intra-connected using interactions from the original and newly supported data sources.

Conclusion: The R Package Multipath is presented with the main functions that are newly developed to support the integration of the databases KEGG Genes and OMIM. The model comprises multiple information relevant to the analysis of pathway data, and offers a reproducible and simplified view of complex, intertwined systems. Through the application of such highly integrated models the inference of new knowledge becomes easier and contributes to many fields such as drug repurposing and biomarker discovery.

Background

In this manuscript, we present a newer version of our tool Multipath, which integrates multi-omics knowledge related to BioPAX encoded pathways in a multilayer model. Beside DrugBank and UniProtKB, this version supports two additional databases, KEGG Genes and OMIM. The model which is generated from a Reactome pathway now accommodates information on proteins, genes, diseases, and drugs. This is a great step that would lead to innovative research in many fields, including network biology, drug repurposing and pathway analysis.

Introduction

Over the past decades, omics research has made significant progress, in particular through the development of high-throughput technologies. These technologies allow to generate large amounts of data which are referred to as omics data. Omics data include among others genomics, transcriptomics, proteomics, and metabolomics. Specifically, these may include whole-genome sequencing data or genome-wide expression profiling on a microarray basis. They provide detailed information about biological systems at various levels. Omics data analysis can provide new insights into the molecular basis of disease, drug development and personalized therapies [1].

* Corresponding author.

E-mail address: zaynabhassanhammoud@gmail.com (Z. Hammoud).

<https://doi.org/10.1016/j.cmpbup.2025.100189>

Available online 21 March 2025

2666-9900/©2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

For this, the development of analysis tools, new models and procedures is essential. Since omics data is available in large quantities, integrating and analyzing it is a conceptual challenge. The extremely complex omics data comprise large amounts of information that should be examined simultaneously to identify relevant biological patterns and relationships. This requires powerful and efficient tools and models. If several omics data sets are used at the same time, one speaks of multi-omics data. The analysis of these represents a further challenge. Here, different omics levels are connected and influence each other [2].

Databases like The Cancer Genome Atlas (TCGA), Clinical Proteomic Tumor Analysis Consortium (CPTAC) or International Cancer Genomics Consortium (ICGC) contain publicly available multi-omics data sets of patients. With the help of newly developed tools for the analysis of multi-omics data, certain biological questions can be addressed and eventually answered. This includes subtyping and classifying diseases, predicting biomarkers for various applications, including diagnostics and driver genes for diseases, and deriving insights into disease biology [3].

The analysis of various omics layers often requires data integration, which aims to extract biological knowledge from multiple datasets. All omics data sets are analyzed simultaneously, and a common model is created comprising all data sets at the same time. Different sizes, formats, and dimensionalities of the data in question pose a challenge. In addition, the data is very complex, can be noisy and has a high information content. Gligorijević and Pržulj summarize the strategies for data integration into the categories of early, late and intermediate data integration. In the early data integration, the data model is built on a combination of the different input data sets. In contrast, with late data integration, models are first created for each individual data set, which are then combined into a unified model. Intermediate data integration infers a common model and then combines the data [4].

Not only the omics data sets are heterogeneous, but also the results in the multi-omics area are difficult to reproduce due to the use of different data analysis methods, tools and statistical processing. The research community is therefore striving to meet the standards of discoverability, accessibility, interoperability, and reusability and to take advantage of advances in data sharing and environmental replication [5].

Integration of drug information to biological pathways has contributed to different disciplines such as drug repurposing [6]. Extending such models with additional knowledge, specifically disease and gene data, can show huge potentials in inferring new interactions, enriching current perspectives of biological pathways, as well as improvement of the comprehension of the dynamics of the different components, individually and collectively.

Implementation

According to Hammoud and Kramer [7], Multipath is an R package that creates reproducible pathway models by transforming BioPAX [8] encoded pathways into multilayered graphs using the R packages mully [9] and rBiopaxParser [10]. The nodes representing the elements of pathways are classified into groups based on their biological types, and each group is embedded into one layer of the graph. The edges connecting the nodes are the interactions extracted from the BioPAX file. The generated mully graph is modifiable and retraceable, and modifications can be tracked using the track and undo feature. A mully object contains nodes, layers and edges, and their corresponding attributes. It also contains metadata on the content such as the name of the graph, edges directions (direct or indirect) and layers' information. Additionally, Multipath offers data integration functions to extract additional information from DrugBank [11] and UniProt [12]. A separate layer is created and added to the graph to accommodate the retrieved information on drugs and their interactions.

To generate the integrated pathway model using Multipath, the user starts by choosing a pathway from the Reactome database [13], where pathways can be downloaded as BioPAX encoded files (level 2 and 3).

The Reactome ID of the pathway along the BioPAX level should be provided to download the pathway in question, parse the corresponding file, and transform it into a multilayer mully graph with the various layers. Relevant data from UniProt and DrugBank can then be added to the model, and modifications can be made and retraced using provided functions.

In this version, the workflow was updated to support two additional knowledge databases, in order to elevate the pathway multilayer model with the diseases and genes' layers. This adjustment was done by implementing the two main functions `addGenesLayer()` and `addDiseasesLayer()`. This enables the user to automate genes and diseases' mapping from the databases KEGG Genes [14] and OMIM [15] respectively. With the help of Multipath 2.0 and the newly implemented functions, additional nodes are added to the model, and the relevant interactions are extracted from various databases and integrated as edges in the mully graph.

The contribution of several R Packages allows the release of this Multipath adjustment. The packages were utilized for implementing the main wrapper functions `addGenesLayer()` and `addDiseaseLayer()`, and further required ones. In the initial release, the package `rBiopaxParser` was used to parse BioPAX encoded files, and the packages `queryup` [16] and `dbparser` [17] to acquire information respectively on proteins from UniProtKB and drugs from DrugBank. Further additional packages are added to the list of dependencies in this version. The R packages `KEGGRest` [18] and `romim` [19] were used to query the databases KEGG Genes and OMIM respectively. Furthermore, `KEGGGraph` [20] was used to map gene IDs to KEGG IDs, and XML [21], XML2R [22], and `xml2` [23] to parse the returned result of the query to the OMIM database.

For the scope of this article, the Wnt Signaling pathway is used to demonstrate the abilities of the Multipath R Package. The Wnt signaling pathway is an ancient and evolutionarily conserved pathway that regulates crucial aspects of cell fate determination, cell migration, cell polarity, neural patterning, and organogenesis during embryonic development [24]. The multilayered model of the Signaling by Wnt pathway (BioPAX level 3) [25] from the Reactome Database was again used in this version of the package to demonstrate its usage.

Results

KEGG genes

KEGG is a database that combines genomic knowledge with higher-order functional information. It involves linking genes in the genome to a network of interacting molecules in the cell. KEGG comprises a set of databases with several focuses, one of which is called KEGG Genes. KEGG Genes is a database that provides information on the functions and pathways of genes from various organisms, which can be useful for understanding biological processes and diseases [14]. For the scope of this article, KEGG Genes database is used as the main source of information on genes.

The genes that are translated to the proteins in the original pathway model are identified after creating the mully graph. To do so, a series of functions must be executed (Fig. 1). The external IDs of the proteins on the protein layer are extracted using the function `getExternalIDs()` which was integrated in the main function `getRelatedGenes()` for this specific action. The latter also calls the function `getUPKBInfo()` to find the KEGG genes that are present as a cross reference for the protein IDs in the UniProtKB database. Afterwards, we incorporated the retrieved genes into the function `getKEGGtoDatabase()` to verify if the genes in the KEGG Genes database have references to UniProtKB. Then, we compared the two obtained dataframes and merged them together using the new function `getKEGGUPKBRelations()`.

A dataframe is an R data structure format used to store data of any type in a tabular two-dimensional manner. After establishing the dataframe containing all proteins and genes of interest, we proceeded by adding a new gene layer called "KEGGGENES" containing the gene

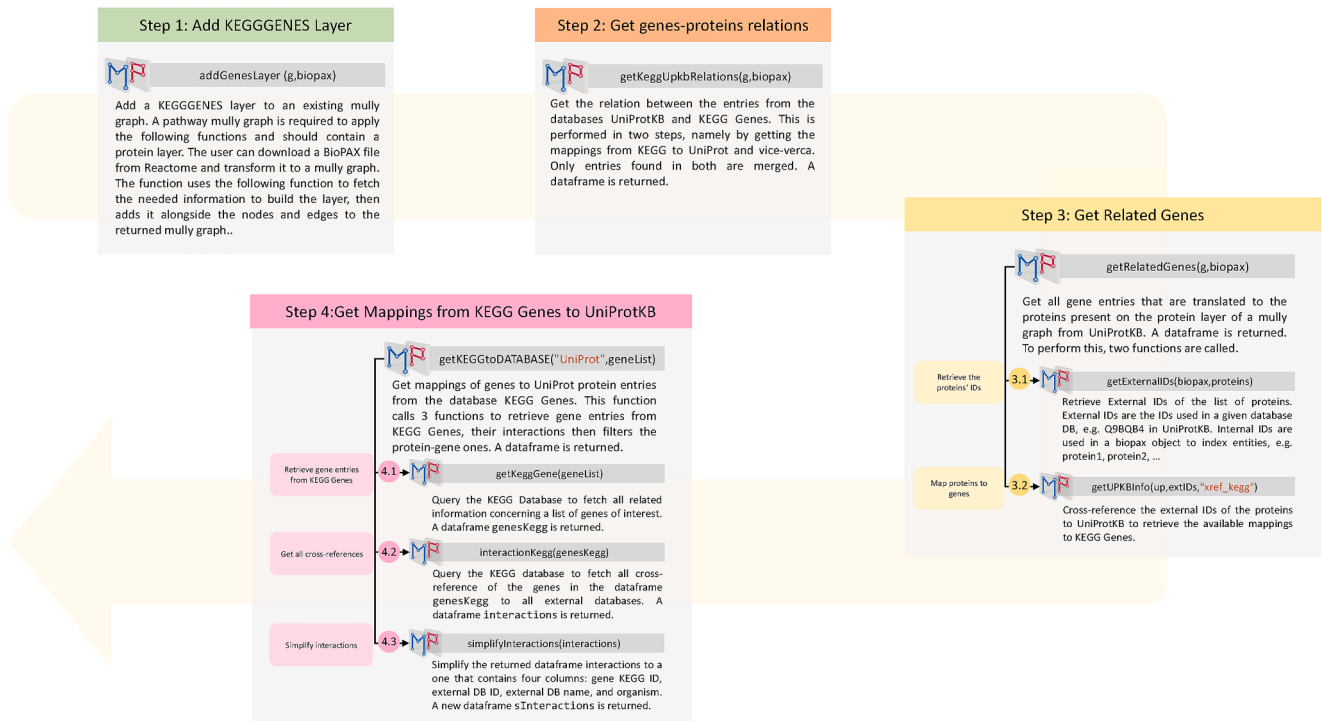


Fig. 1. KEGG Genes Workflow Diagram. The Diagram shows the steps to fetch information on genes from KEGG Genes using Multipath and the R Package KEGGREST with the help of the R Package queryup. The diagram also shows how each wrapper function works and what the needed arguments are.

information retrieved, which we were able to identify through the protein IDs present in the protein layer. Afterwards, we added each gene ID as a node on the KEGGGENES layer. Finally, the edges which connect the gene with its respective protein were then added one at a time (Fig. 1).

To perform this and benefit from all the aforementioned functions, it is essential to provide the function `addGenesLayer()` with a heretofore parsed BioPAX object and its corresponding graph `g`, containing a protein layer.

OMIM

The N Mendelian Inheritance in Man (OMIM) database is a comprehensive and authoritative compendium of human genes and genetic disorders. OMIM provides detailed information on genes, genetic disorders, and their phenotypic manifestations, as well as links to relevant literature and other resources [15]. The OMIM database is used as the main source for diseases that could be potentially related to genes and proteins.

Two different approaches exist that allow the identification of

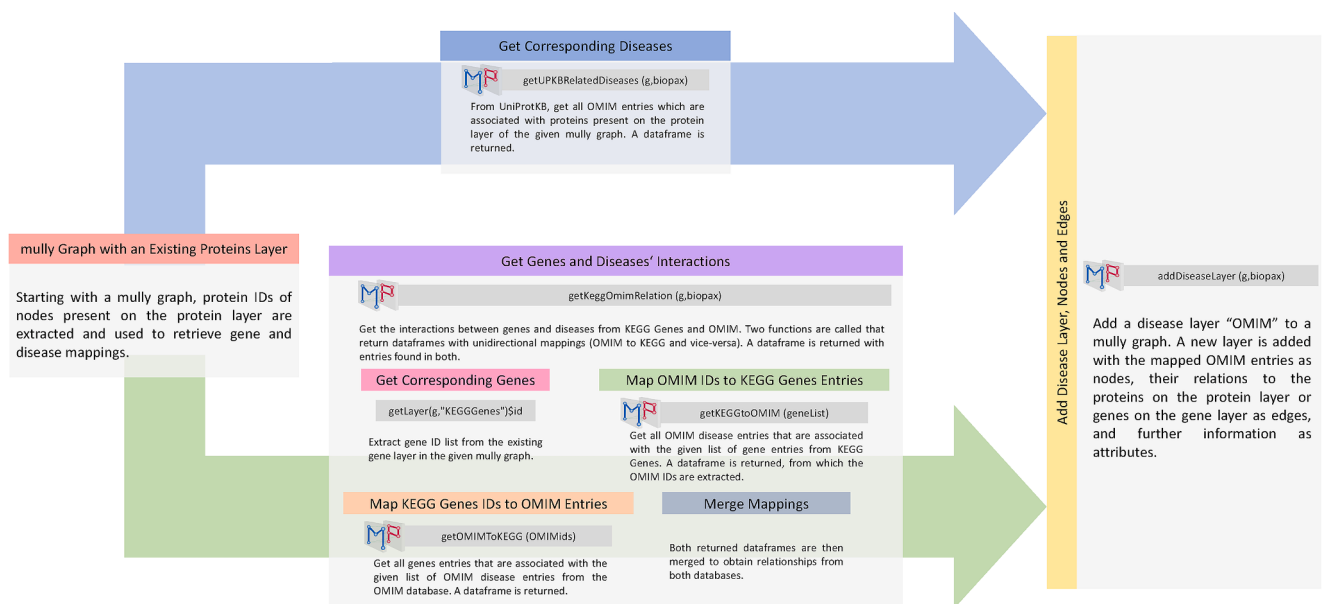


Fig. 2. OMIM Workflow Diagram. The Diagram shows the steps to fetch information on diseases from OMIM using Multipath and the R Package romim with the help of other packages including queryup, KEGGRest and KEGGgraph. The diagram shows the two possible approaches, approach 1 following the blue arrows, and approach 2 following the green arrows. The green arrow follows the function `getKeggOmimRelation()` with its underlying sub-functions.

diseases associated with a pathway of interest, namely through the proteins or the genes that are involved in this pathway. The first method is to directly find the diseases that are associated with the proteins, whereas the second method would be to find the diseases that are directly related to genes (Fig. 2).

First approach

To identify the diseases that are known to be associated with proteins, all proteins were extracted from the mully graph. This was done by calling the function `getUPKBRelatedDiseases()` which extracts the external IDs of the proteins from the protein layer and feeds the entries as a list into the function `getUPKBInfo()`, returning a dataframe containing OMIM IDs that are related to proteins. These OMIM IDs represent diseases that are available in the OMIM database. Then, a new disease layer called “OMIM” is added to the existing mully graph, to which each OMIM ID is added as a node.

Second approach

Moreover, finding diseases could also be accomplished when a gene layer exists on the mully graph using the function `getKeggOmimRelation()`. Firstly, gene IDs are extracted from the pathway mully graph as a list by calling the function `getLayer(g,"KEGGGenes")$id`. The gene list is passed to the function `getKEGGtoOMIM()` which queries the KEGG Genes database to identify genes that have a cross reference to OMIM and returns a dataframe. The returned OMIM IDs are then passed to the function `getOMIMtoKEGG()` which returns a dataframe containing OMIM and KEGGGENE IDs. Both returned dataframe are then joined and merged into one. Finally, we add the diseases' layer called “OMIM” and all the OMIM IDs as nodes into this layer (Fig. 2).

Discussion

The new Wnt signaling pathway model

In the previous release of Multipath, we used the Wnt signaling pathway to demonstrate the usage of the package. We decided to reuse the same pathway also in this paper to showcase the differences and updates to the model. To create the model, the same code from our last paper was used and updated to add the two additional layers to the graph using the newly implemented functions.

The following script (Fig. 3) was used to create the new Signaling by Wnt multilayer model. The script from the original release can also be used to add a drug layer.

The function `pathway2Mully()` transforms the pathway parsed in the BioPAX object into mully graphs. The functions `addGenesLayer()` and

```
downloadPathway("R-HSA-195721",biopax=3)
wntBiopax=readBiopax("R-HSA-195721.owl")
pathwayID=listPathways(wntBiopax)$id[1]
wntmully=pathway2Mully(wntBiopax,pathwayID)
romim::set_key("KEY")
wntmully=addGenesLayer(wntmully,wntBiopax)
wntmully=addDiseaseLayer(wntmully,wntBiopax)
plot3d(wntmully)
```

Fig. 3. Model creation script. The script to generate and plot the Signaling by Wnt Pathway model with the genes and diseases layer.

`addDiseaseLayer()` find the respective genes and diseases for the Wnt Pathway and create the new layers. To plot the graph, we used the function `plot3d()` from the mully package.

With the help of the new functions in Multipath 2.0, it becomes possible to add the new genes and diseases' layer and re-plot the mully graph of the Signaling by Wnt pathway. The original model is seen in Fig. 4(a), and the new in Fig. 4(b). Fig. 4(a) shows the initial model of the given pathway using the functionalities of Multipath 1.0. The graph contains 6 layers derived from the BioPAX level 3 file of the Signaling by Wnt pathway. As shown in the script (Fig. 3), we used the newly implemented functions `addGenesLayer()` and `addDiseaseLayer()` to add two additional layers, namely the disease (OMIM Layer in green) and genes' (KEGGGenes Layer in purple) layers. In contrast to the first model, the graph now contains 8 layers, 864 nodes and 1245 edges. The additional nodes constitute the diseases and genes' entries retrieved from OMIM and KEGG Genes respectively. The interactions are subsequently retrieved jointly from UniProtKB, OMIM, KEGG Genes and the parsed BioPAX information from Reactome. Hence, the model was updated with additional 706 edges between the OMIM, KEGGGENES and Protein layers, making a total of 1245 edges.

Multipath against existing tools

Pathway Commons is a web-based interface that offers a collection of publicly available pathway data of various organisms. It focuses on pathway data that is encoded in BioPAX and PSI-MI formats [26]. While Pathway Commons is one of the widely-used resource to access pathway information, it differs in its purpose from Multipath. Our package focuses on integrating multiple database resources that we believe to be very essential in analysing pathway models. Comparing the Multipath Signaling by wnt model to the one offered by pathway commons, the first can include further information influencing the modelled pathway, such as drug and disease information. This offers a wider coverage of biological resources, and significantly expands the number of nodes and edges included in the graph.

On the other hand, tools like Cytoscape [27] are also used by the research community to model biological and biomedical knowledge. However, Multipath offers multilayer and layer-based clustering features that are currently not available elsewhere. Additionally, data integration must be performed by the user before employing Cytoscape to create the intended visualizations.

Advantages and limitations

In addition to the main function `addGenesLayer()`, users can utilize several sub-functions that are combined together to create it. These functions include but are not limited to `getKEGGGene()` and `getKEGGtoOMIM()`, which return a dataframe containing the genes, their reference to another database (e.g. UniProt or OMIM), and primary information about the genes that are also available in the KEGG Genes database. More detailed information about the different sub-functions and their usage can be found in the Vignette and Reference Manual of the Multipath R Package as Additional File 1 and 2 respectively.

The two different approaches for mapping diseases to an existing pathway provide users with full flexibility to decide on the best approach that fits their needs. Hence, users can choose to map diseases from either genes or proteins present in a multi-layered graph.

To deal with a multi-layered graph that already contains a protein layer, users can call the function `addDiseaseLayer(g, BioPAX)`, where `g` is the multi-layered graph and `BioPAX` is a parsed BioPAX object, to add diseases from proteins. To add diseases from both proteins and genes simultaneously, we recommend calling the function `addDiseaseLayer(g, BioPAX, addGenesLayer=TRUE)`. This approach could yield faster results as it enables a one-time query from KEGG Genes and UniProtKB databases to map genes and diseases in one step.

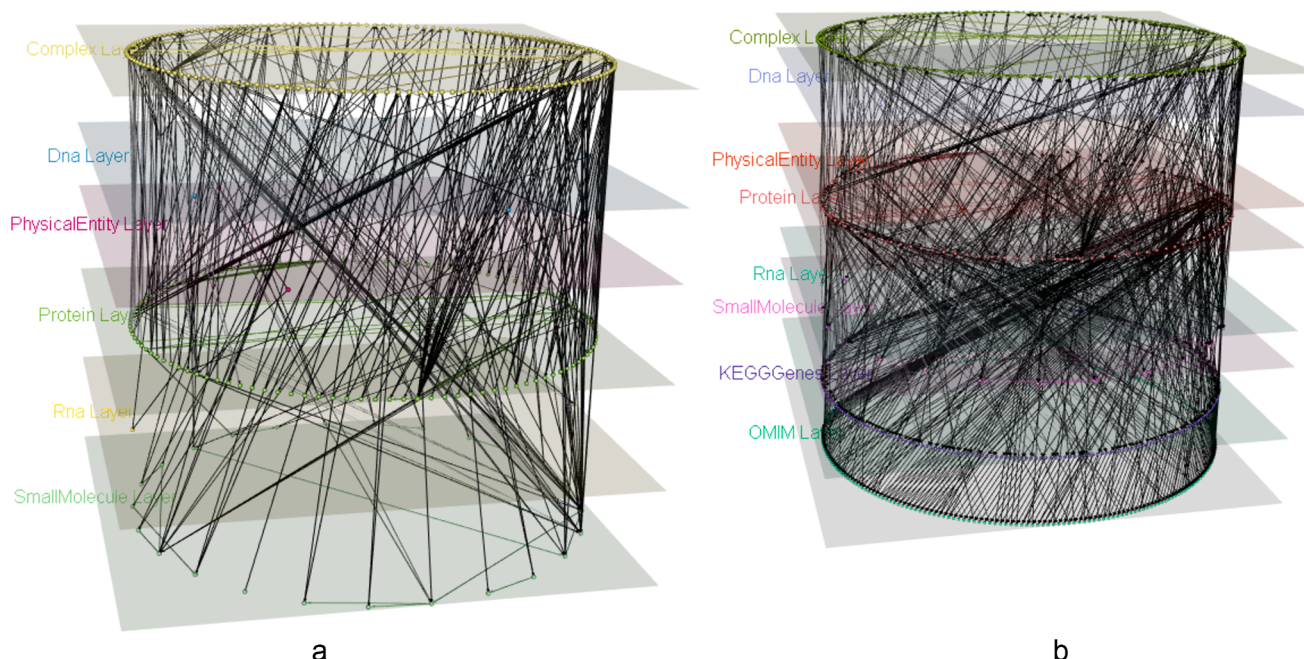


Fig. 4. Updated Signaling by Wnt mully model. The model was built using the BioPAX level 3 Signaling by Wnt from Reactome. (a) The model had initially 6 layers, 311 nodes and 539 edges. (b) After adding the genes and diseases layers, the model reached 8 layers, 864 nodes and 1245 edges.

External factors

In the first release of Multipath, protein information retrieval was accomplished through the UniProt.ws R Package [28]. However, slow responses from the UniProt server were encountered multiple times during the process, with no detection of response speed patterns (i.e. specific time of the day or the amount of information to be retrieved). Moreover, this process does not require any memory space. The web service also recommends querying by using a maximum number of 50 UniProt entries' IDs in a single query. To address limitations in UniProt.ws query speed and reliability, we substituted the usage of this package with the queryup, an R package that queries the UniProtKB database through its REST API. Additionally, we implemented parallel processing using the future.apply [29] and future [30] R packages to reduce runtime. UniProt.ws has proven so far slow response speeds, inconsistent or incomplete responses especially upon high number of user requests, and limited batch process with 50 IDs per query. Directly interacting with the web service of UniProt and querying its database majorly solves these issues. It is optimized for performance and can handle larger datasets efficiently, querying up to 10,000 IDs at a time. The retrieved set of results is also more reliable and consistent with the intended purpose, and offers more flexibility for data format and processing such as filtering.

Queries to the KEGG Genes database were handled via the R Package KEGGREST [18]. Compared to UniProt.ws, the KEGGREST package performed queries at a much faster pace, was however limited to 10 entries per query. To overcome KEGGREST's 10-ID query limit, we split input gene lists into chunks and processed them in parallel using future_lapply(). It is important to mention that using the parallel processing feature for smaller lists of genes is time-exhausting. However, this function insures a smoother interaction with the KEGG server and allows the user to specify specific chunk sizes. Another R Package called KEGGgraph was utilized to translate GeneID to KEGGID using the function getOmimToKEGG(). The OMIM database was queried using the R Package romim, which provided a fast and direct usage of the OMIM database API by returning an XML object for every query. To chunk and extract the necessary information from this XML object, we used the R packages XML, XML2R, and xml2. However, romim, similar to all other

packages dealing with the OMIM API, requires calling the function romim::set_key('key') before performing any disease mapping. This key is provided exclusively by OMIM via its official website.

Data processing and API related information

The Signaling by Wnt pathway file (4.19 MB) is parsed in 13.11 s and occupies ~5.01 MB of memory space. The transformation to a mully graph is conducted in 12.07 s, and the resulted graph's size is 118.9 KB. Mapping the genes and creating the gene layer required an approximate of 2 min and 28 s to finalize. Accordingly, the graph size increased from 121.744 KB to 180.1 KB. Whereas mapping the diseases and creating the OMIM layer required almost 1 min and 32 s to finalize, taking into consideration that the graph size increased from 121.744 KB to 179.5 KB. However, the speed might vary depending on several factors such as the processing speed of the user's CPU, the internet speed, and the speed at which the API is returning information.

Future optimization

In general, introducing an internal database or a caching system to store precomputed pathway graphs, or often request nodes' information would contribute to enhancing the performance of Multipath, and overcoming many challenges such as server request errors, limitations in the query size, and the response speed. We also plan to introduce more efficient approaches to query OMIM information, especially using batch processing.

Conclusions

The Multipath package, which utilizes BioPAX encoded pathway information, is capable of producing customizable and trackable multilayered models. Moreover, it can access various knowledge databases such as UniProt, DrugBank, KEGG Genes, and OMIM to incorporate relevant information into the multilayered graph. This research paper highlights the novel functionalities of Multipath, including its ability to construct multilayered models from BioPAX files and integrate diverse pathway knowledge. To illustrate its capabilities, we employed

the Signaling by Wnt pathway as an example, added genes and diseases layers that correspond to the protein entries in the pathway, and demonstrated a 3D visualization of the pathway model. Furthermore, the package has proven to be very effective in various fields of biomedicine, including drug repurposing, and revealing potential therapeutic targets for host-directed antileishmanial therapy [6]. Multipath is an open-source software that can be downloaded from our GitHub Repository [31]. The reference manual and the vignette of the package were added as Additional Files, and the link to the GitHub repository is provided in the Data Availability Statement.

Availability and requirements

Project Name: Multipath

Project home page: <https://github.com/frankkramer-lab/Multipath>

Operating system(s): Platform independent

Programming language: R

Other requirements: R ($\geq 4.2.2$), RStudio, and the R packages mully ($\geq 2.1.26$), queryup, dbparser, rBiopaxParser, svMisc, uuid, dplyr, crayon, igraph, RCurl, graph, romim, KEGGREST, KEGGgraph, XML, XML2R, xml2, future, future.apply and stringr

License: GNU GPL

Any restrictions to use by non-academics: No restrictions

Declarations

Ethics approval and consent to participate

Not Applicable.

Consent for publication

Not Applicable.

Availability of data and material

The R package Multipath is open-source and can be accessed on the corresponding Repository on GitHub under the following link: <https://github.com/frankkramer-lab/Multipath>.

Competing interests

The authors declare that they have no competing interests.

Funding

This work is a part of the Multipath and MoMoTuBo projects funded by the GERMAN MINISTRY OF EDUCATION AND RESEARCH (Bundesministerium für Bildung und Forschung), grant FKZ01ZX1508 and FKZ01ZZ2008 respectively.

Authors' contributions

MAM and ZH were responsible for the software. ZH, AD and MAM were responsible for the writing (preparation of the original draft). ZH was responsible for the methodology and validation. ZH and FK were responsible for the conceptualization, supervision and writing (review and editing). FK was responsible for funding acquisition and project administration. All authors have read and agreed to the published version of the manuscript.

Ethics statement

We hereby declare that our manuscript does not require an ethics statement. Our research does not use any experimentation with human subjects, and is not affected by any laws or institutional guidelines.

CRediT authorship contribution statement

Zaynab Hammoud: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Mohammad Al Maaz:** Writing – original draft, Visualization, Software. **Alicia D'Angelo:** Writing – original draft. **Frank Kramer:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to acknowledge the help of our colleagues at the chair of IT-Infrastructure for Translational Medical Research, and Prof. Dr. Martin Sieber at the University of Applied Sciences Bonn-Rhein-Sieg.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.cmpub.2025.100189](https://doi.org/10.1016/j.cmpub.2025.100189).

References

- [1] A.R. Joyce, B.Ø. Palsson, The model organism as a system: integrating “omics” data sets, *Nat Rev Mol Cell Biol* 7 (3) (2006 Mar) 198–210.
- [2] D. Gomez-Cabrero, I. Abugessaisa, D. Maier, A. Teschendorff, M. Merkenschlager, A. Gisel, et al., Data integration in the era of omics: current and future challenges, *BMC Syst Biol* 8 (2) (2014 Mar 13). II.
- [3] I. Subramanian, S. Verma, S. Kumar, A. Jere, K. Anamika, Multi-omics Data integration, interpretation, and its application, *Bioinform Biol Insights* 14 (2020 Jan 1) 1177932219899051.
- [4] V. Gligorijević, N. Pržulj, Methods for biological data integration: perspectives and challenges, *J R Soc Interface* 12 (112) (2015 Nov 6) 20150571.
- [5] M. Krassowski, V. Das, S.K. Sahu, B.B. Misra, State of the field in Multi-Omics research: from computational needs to data mining and sharing, *Front Genet* 11 (2020 Dec 10) 610798.
- [6] J.E. Martinez-Hernandez, Z. Hammoud, A.M. de Sousa, F. Kramer, Monte-Neto RL do, V. Maracaja-Coutinho, et al., Network-based approaches reveal potential therapeutic targets for host-directed antileishmanial therapy driving drug repurposing, *Microbiol Spectr* 9 (2) (2021 Oct 20) e01018–e01021.
- [7] Z. Hammoud, F. Kramer, Multipath: an R package to generate integrated reproducible pathway models, *Biology (Basel)* 9 (12) (2020 Dec 21) 483.
- [8] E. Demir, M.P. Cary, S. Paley, K. Fukuda, C. Lemer, I. Vastrik, et al., The BioPAX community standard for pathway data sharing, *Nat Biotechnol* 28 (9) (2010 Sep) 935–942.
- [9] Z. Hammoud, F. Kramer, mully: an R package to create, modify and visualize multilayered graphs, *Genes (Basel)* 9 (11) (2018 Oct 23) 519.
- [10] F. Kramer, M. Bayerlová, F. Klemm, A. Bleckmann, T. Beissbarth, rBiopaxParser—an R package to parse, modify and visualize BioPAX data, *Bioinformatics* 29 (4) (2013 Feb 15) 520–522.
- [11] D.S. Wishart, C. Knox, A.C. Guo, D. Cheng, S. Shrivastava, D. Tzur, et al., DrugBank: a knowledgebase for drugs, drug actions and drug targets, *Nucleic Acids Res* 36 (Database issue) (2008 Jan) D901–D906.
- [12] R. Apweiler, A. Bairoch, C.H. Wu, W.C. Barker, B. Boeckmann, S. Ferro, et al., UniProt: the Universal protein knowledgebase, *Nucleic Acids Res* 32 (Database issue) (2004 Jan 1) D115–D119.
- [13] M. Gillespie, B. Jassal, R. Stephan, M. Milacic, K. Rothfels, A. Senff-Ribeiro, et al., The reactome pathway knowledgebase 2022, *Nucleic Acids Res.* 50 (D1) (2022 Jan 7) D687–D692.
- [14] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, *Nucleic Acids Res* 28 (1) (2000 Jan 1) 27–30.
- [15] A. Hamosh, A.F. Scott, J.S. Amberger, C.A. Bocchini, V.A. McKusick, Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders, *Nucleic Acids Res.* 33 (suppl_1) (2005 Jan 1) D514–D517.
- [16] G. Voisinne, queryup: Query the UniProt REST API Using R. Centre d'Immunologie De Marseille-Luminy, Aix Marseille Université, INSERM, CNRS, Marseille, France, 2019. <https://github.com/VoisinneG/queryup/>.
- [17] Ali, M. et al. (2023). dbparser: drugs databases parser. Available online: <https://cran.r-project.org/web/packages/dbparser/index.html> (accessed on 11 September 2023).

- [18] Tenenbaum D., Volkening J., Maintainer B.P. KEGGREST: client-side REST access to the Kyoto Encyclopedia of Genes and Genomes (KEGG) [Internet]. Bioconductor version: Release (3.17); 2023 [cited 2023 Sep 11]. Available from: <https://bioconductor.org/packages/KEGGREST/>.
- [19] Tang, D. (2023) romim. Available online: <https://github.com/davetang/romim> (accessed on 11 September 2023).
- [20] J.D. Zhang, S. Wiemann, KEGGgraph: a graph approach to KEGG PATHWAY in R and bioconductor, *Bioinformatics* 25 (11) (2009 Jun 1) 1470–1471.
- [21] Lang D.T., Kalibera T. XML: tools for parsing and generating XML within R and S-plus [Internet]. 2023 [cited 2023 Sep 4]. Available from: <https://cran.r-project.org/web/packages/XML/index.html>.
- [22] Stevert C. XML2R: easieR XML data collection [Internet]. 2014 [cited 2023 Sep 11]. Available from: <https://cran.r-project.org/web/packages/XML2R/index.html>.
- [23] Wickham H., Hester J., Ooms J., RStudio, example) RF (copy of R project homepage cached as. xml2: parse XML [Internet]. 2023 [cited 2023 Sep 4]. Available from: <https://cran.r-project.org/web/packages/xml2/index.html>.
- [24] Y. Komiya, R. Habas, Wnt signal transduction pathways, *Organogenesis* 4 (2) (2008 Apr) 68–75.
- [25] Kimelman D. Signaling by Wnt. Reactome [Internet]. 2007 May 15 [cited 2023 Sep 11];21. Available from: <https://reactome.org/content/detail/R-HSA-195721.1>.
- [26] E.G. Cerami, B.E. Gross, E. Demir, I. Rodchenkov, O. Babur, N. Anwar, N. Schultz, G.D. Bader, C. Sander, Pathway Commons, a web resource for biological pathway data, *Nucleic Acids Res* 39 (Database issue) (2011 Jan) D685–D690, <https://doi.org/10.1093/nar/gkq1039>. Epub 2010 Nov 10. PMID: 21071392; PMCID: PMC3013659.
- [27] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: a software environment for integrated models of biomolecular interaction networks, *Genome Res* 13 (11) (2003 Nov) 2498–2504, <https://doi.org/10.1101/gr.1239303>. PMID: 14597658; PMCID: PMC403769.
- [28] Carlson M., Ramos M. (2024). UniProt.Ws: r interface to UniProt Web services. R package version 2.46.1, <https://github.com/Bioconductor/UniProt.ws>.
- [29] Bengtsson H. (2021). “A unifying framework for parallel and distributed processing in R using futures.” *R J*, 13(2), 208–227. <https://doi.org/10.32614/RJ-2021-048>.
- [30] Bengtsson H. (2021). “A unifying framework for parallel and distributed processing in R using futures.” *R J*, 13(2), 208–227. <https://doi.org/10.32614/RJ-2021-048>.
- [31] Multipath [Internet]. frankkramer-lab on GitHub; 2023 [cited 2023 Sep 11]. Available from: <https://github.com/frankkramer-lab/Multipath>.