



Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb



Integrated enrichment analysis and pathway-centered visualization of metabolomics, proteomics, transcriptomics, and genomics data by using the InCroMAP software

Johannes Eichner^{a,1}, Lars Rosenbaum^{a,*,1}, Clemens Wrzodek^a, Hans-Ulrich Häring^{b,c,e}, Andreas Zell^a, Rainer Lehmann^{b,d,e,**}

^a Center for Bioinformatics, University of Tübingen, Tübingen, Germany

^b Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Centre Munich at the University of Tübingen, Tübingen, Germany

^c Division of Endocrinology, Diabetology, Vascular Medicine, Nephrology and Clinical Chemistry, Department of Internal Medicine IV, University Hospital Tübingen, Tübingen, Germany

^d Division of Clinical Chemistry and Pathobiochemistry, Department of Internal Medicine IV, University Hospital Tübingen, Tübingen, Germany

^e German Center for Diabetes Research (DZD), Germany

ARTICLE INFO

Article history:

Received 30 September 2013

Accepted 15 April 2014

Available online xxx

Keywords:

Metabolomics

Transcriptomics

Proteomics

Genomics

Enrichment analysis

Pathway visualization

ABSTRACT

In systems biology, the combination of multiple types of omics data, such as metabolomics, proteomics, transcriptomics, and genomics, yields more information on a biological process than the analysis of a single type of data. Thus, data from different omics platforms is usually combined in one experimental setup to obtain insight into a biological process or a disease state. Particularly high accuracy metabolomics data from modern mass spectrometry instruments is currently more and more integrated into biological studies. Reflecting this trend, we extended InCroMAP, a data integration, analysis and visualization tool for genomics, transcriptomics, and proteomics data. Now, the tool is able to perform an integrated enrichment analysis and pathway-based visualization of multi-omics data and thus, it is suitable for the evaluation of comprehensive systems biology studies.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Today, high-throughput methods for the analysis of biological systems, such as microarrays, next generation sequencing, as well as mass spectrometric and NMR approaches generate a wealth of omics-scale data. To develop hypotheses about a biological process or a disease state, a variety of omics platforms for measuring different genomic, proteomic, and metabolomic features are combined in one experimental setup. Probably the most frequently

analyzed genomic feature is messenger RNA (mRNA), which can be quantified on a genome-wide level by gene expression chips (microarrays) or next generation sequencing techniques. Other important (epi)genomic features include microRNA (miRNA), single-nucleotide polymorphisms (SNPs), and epigenetic information, such as the promoter methylation status (DNAm). Proteomic features can be derived from abundance profiling of particular protein patterns and posttranslational modifications, such as acetylation and phosphorylation. Furthermore, mass spectrometric and NMR-based metabolomics platforms allow to investigate complex metabolite patterns with high accuracy. In recent years, increasing attention has been drawn to the integration of metabolomics with molecular profiling on the transcriptional and protein level [1,2]. The sensible and integrated visualization of omics data at different levels of abstraction is crucial to obtain biological insight without being overwhelmed by the intrinsic complexity of the data [3]. A key concept for detecting alterations in cell

* Corresponding author.

** Corresponding author at: Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Centre Munich at the University of Tübingen, Tübingen, Germany. Tel.: +49 7071 29 8 31 93.

E-mail addresses: johannes.eichner@uni-tuebingen.de (J. Eichner),

rainer.lehmann@med.uni-tuebingen.de (R. Lehmann).

¹ These authors contributed equally to this work.

signalling or metabolism in a biological system are pathway-based visualizations.

A plethora of tools were developed for the inspection of data from individual platforms (see [3] for examples). Furthermore, solutions for the combined visualization of transcriptomics (mRNA) and metabolomics data exist [4,5]. Both Paintomics [4] and MassTRIX [5] are web-service based tools that are capable of visualizing mRNA microarray and identified metabolomics data in KEGG pathways [6]. MassTRIX is able to handle unidentified metabolic data by comparing each mass against theoretical adducts stored in metabolomics databases. An example of a tool that can handle several heterogeneous types of omics data is the commercial Ingenuity Pathway Analysis software (www.ingenuity.com). However, Ingenuity does not provide an integrated visualization of multi-level omics data. In summary, the comprehensive analysis of multi-omics data is presently limited because current high-level analysis tools are not able to perform an integrated analysis of data from multiple types of omics platforms, are focused on certain specialized platforms, or not freely available. Thus, novel analysis tools and appropriate visualizations are required, since complex interactions between multiple layers of a biological system can only be inferred by the integration of omics data across multiple platforms.

In this contribution, we present an extended version of the tool InCroMAP (Integrated analysis of Cross-platform MicroArray and Pathway data) [7,8]. InCroMAP is a stand-alone Java software originally developed for the enrichment analysis and pathway-based visualizations of genomic and proteomic data, where multiple biological layers were monitored in the same set of samples. The application was specifically designed to provide a high ease of use for investigators of all kinds of analytical disciplines, i.e. detailed experiences in bioinformatics are not necessary. Consequently, all information required, for example, for the mapping between different metabolite identifiers or the annotation of miRNAs with mRNA targets, is either directly included in the tool or dynamically downloaded in the background. Previous versions of InCroMAP focused on genomics and transcriptomics data and were designed for the combined analysis of a wide variety of omics platform types, where each feature can be linked to a gene or genomic interval. Up to the present, metabolomics data was not supported. In the extended version of InCroMAP presented here, we support the enrichment analysis and pathway-based visualization of annotated metabolomics data and thereby complemented the tool for comprehensive systems biology data evaluation. For convenience, InCroMAP supports several commonly used metabolite identifiers, such as specific database identifiers (e.g., HMDB [9]), common synonyms, and InChIKeys.

InCroMAP is freely available under the LGPL3 license at <http://www.cogsys.cs.uni-tuebingen.de/software/InCroMAP>, including a comprehensive users guide and several example data files to test the capabilities of the tool.

2. Methods

Before systems biology data from heterogeneous platforms can be integrated and visualized with high-level data analysis tools like InCroMAP, the data has to be preprocessed in a platform-dependent manner (see Fig. 1). A typical workflow for metabolomics data from nontargeted mass spectrometric measurements includes the following steps. First, metabolite features, which are characterized by mass and relative intensity, are extracted from the raw data files by applying software tools, like our recently developed FeatureFinder software [10] or commercial tools implemented in the software of the instrument. Ideally, the feature induced by a metabolite should contain all signals that were generated by the metabolite

(including isotopic peaks) [10]. Second, the extracted features have to be aligned and linked between the different samples. The aforementioned preprocessing steps can be performed with sophisticated open-source software tools like OpenMS [11]. The features are then subjected to quality controls and low-level statistical data analysis, which includes normalization and calculation of measures for differential abundance between conditions (e.g., *p*-values, fold changes, or log ratios). Mostly, these tasks are performed with commercial statistics software, open-source omics analysis applications like Mayday [12], or directly with the statistical programming language R (www.r-project.org). In a final preprocessing step, the metabolite features are annotated with candidate identifiers using information provided by metabolomics databases, such as PubChem [13], HMDB [9], or LIPID MAPS [14]. Similar preprocessing workflows exist for other platforms, like microarrays, next generation sequencing, and proteomics. The processed data can then be imported and analyzed with InCroMAP.

In a typical use case of InCroMAP (see Fig. 1), the user first imports his preprocessed multi-level omics data, given in tabular format. Then, the differing metabolites, genes and proteins are determined for each platform, based on appropriate cutoffs. Next, relevant pathways related to the biological background of the experiments are inferred. For the detection of relevant pathways, InCroMAP employs a special pathway enrichment algorithm, which integrates differing metabolites, genes, and proteins across multiple platforms. The resulting pathways can then be selected for further visual inspection from a table in which each pathway is associated with a significance value. Alternatively, the metabolic overview function of InCroMAP can be used to generate an interactive global map of cellular metabolism, in which each subordinate metabolic pathway is colored according to the significance of its enrichment. The results of the enrichment analysis can be exported in tabular format and the pathway-based visualizations can be easily stored as JPEG images.

2.1. Data import

InCroMAP software uses an intuitive, largely uniform input format for the import of heterogeneous types of processed omics data. The software accepts tabular input files (e.g., CSV files exported from MS Excel) which consist of platform-dependent meta-data columns containing appropriate identifiers (e.g., Affymetrix, EntrezGene, HMDB IDs, etc.) and data columns corresponding to the fold-changes or *p*-values calculated for a particular sample group.

InCroMAP is able to handle various types of identifiers for genomics, transcriptomics, proteomics, and metabolomics data. It supports the recognition of identifiers used by the most common oligonucleotide microarray manufacturers (e.g., Affymetrix, Agilent, etc.). Additionally, generic formats facilitate the import of processed data, provided that each measurement can be either associated with a certain gene, genomic region, protein, or metabolite. For miRNA data the systematic name (i.e., miRBase ID) has to be provided. The miRNA transcripts can then be automatically connected to canonical pathways based on confirmed or predicted interactions with potential target mRNAs. Protein data is simply linked to the corresponding genes. Since diverse proteins corresponding to the same gene may have been measured, a gene identifier as well as an arbitrary identifier for the protein modification has to be provided. Metabolomics data may be annotated by a database identifier (e.g., KEGG, HMDB, PubChem, LIPID MAPS), an InChIKey, or a common synonym. Internally, all metabolites are mapped to InChIKeys. Pathways of interest can either be automatically downloaded from KEGG or imported from other sources (e.g., Reactome, BioCarta, etc.) in BioPAX format [15].

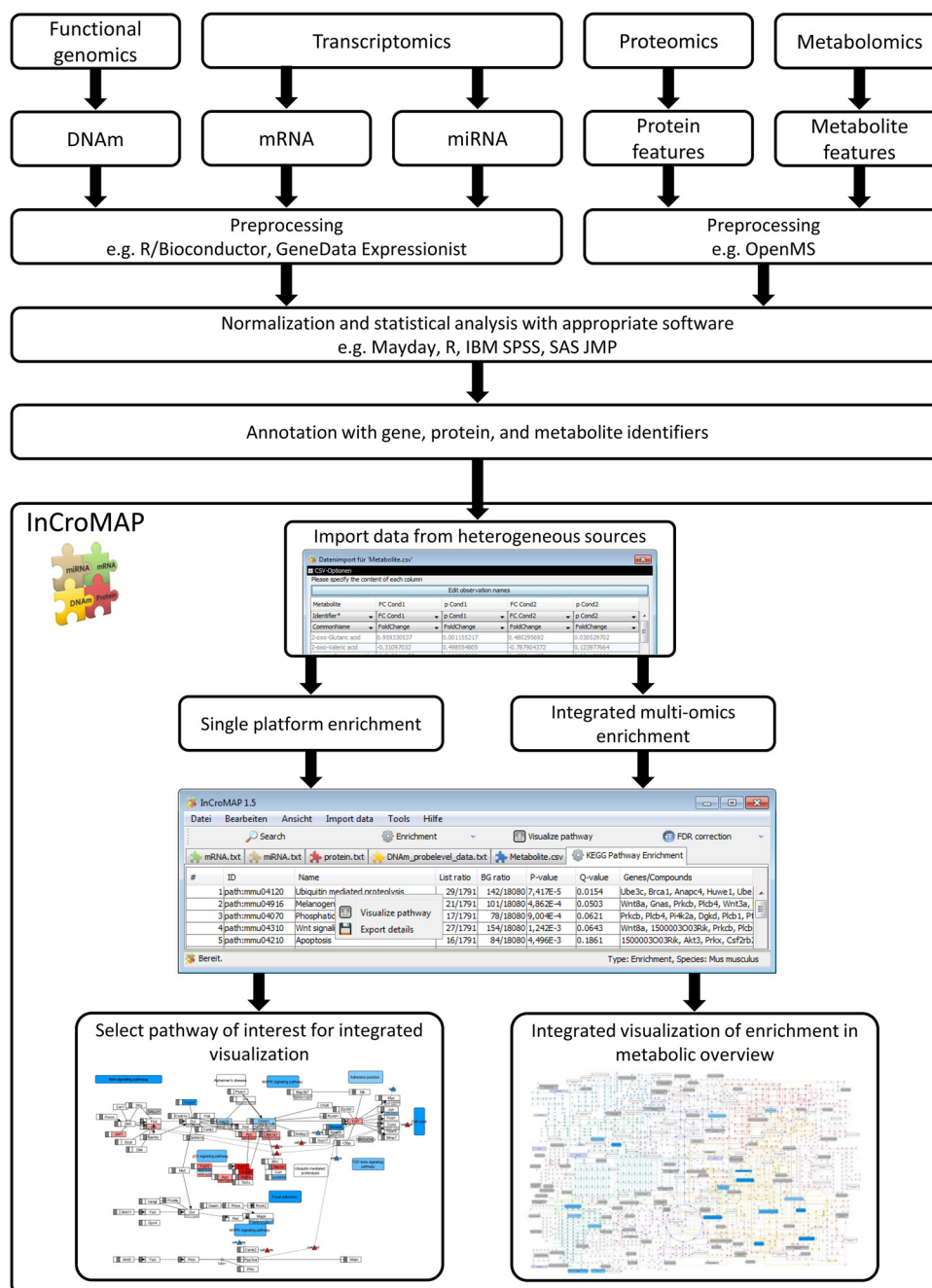


Fig. 1. The flow chart depicts the general workflow of an analysis with InCroMAP. The raw data from each platform is preprocessed using appropriate, platform-specific algorithms for data preparation (e.g., feature extraction) and the alleviation of experimental artifacts (e.g., normalization techniques). Then, the data is analyzed with a statistics software, which yields *p*-values or fold changes for each of the features. Finally, the gene, protein, and metabolite features are annotated with candidate identifiers. After these preprocessing steps the data can be easily imported to InCroMAP and an enrichment analysis can be performed. On the basis of the enrichment, the user may select a pathway of interest, which can be overlaid with the corresponding omics data. Alternatively, a visualization showing a global metabolic overview of the alterations in pathways can be generated.

We decided to use InChIKeys as internal identifier for metabolites because they are a compact, almost certainly unique identifier, which does not change as frequently as most database identifiers. We merged the dumps of four different compound databases (HMDB, LIPIDMAPS, KEGG, and PubChem) based on the InChIKeys with self-made python scripts and kept only entries with information for at least two different databases. Entries without InChIKey information, e.g. information on compound classes, were discarded. In principal, the information can be mapped on actual compounds using taxonomies, but this procedure might introduce errors in the generated mapping database.

2.2. Enrichment analysis

The InCroMAP software allows for the calculation of single-platform and cross-platform enrichments, based on omics experiments from individual and multiple platforms, respectively. In the former case, a hypergeometric test is applied to assess the significance of the overrepresentation of predefined gene sets (e.g., KEGG pathways) within a list of differing genes derived from a certain experiment based on fold-change and/or *p*-value cutoffs. The same statistical approach is used to determine pathways enriched with differing metabolites. In both cases, the union

of all genes/metabolites in KEGG is considered as the universe. For the calculation of metabolite set enrichments a variety of other tools (e.g., MSEA, MBRole, MPEA, etc.) have been developed [16–18], some of which also include implementations of more sophisticated algorithms accounting for metabolite concentrations. However, these tools do not offer integrated enrichment analysis across multiple platforms. For this purpose, InCroMAP uses a straightforward extension of the single-platform method. Specifically, a hypergeometric test is used, where the sample corresponds to the union of the molecules which were changed in terms of abundance on any of the platforms and the universe is given by the union of the platform-specific background sets. To our knowledge, the only two alternative methods for calculating an integrated enrichment based on both transcriptomics and metabolomics data are implemented in the tools IMPaLA [19] and Paintomics [4]. Paintomics also calculates the integrated enrichment p -values based on a hypergeometric test, but the details remain unclear from the publication. In contrast to our software, IMPaLA calculates the enrichment p -values independently for each platform using either a hypergeometric test or Wilcoxon enrichment analysis. As the experiments are considered independent, the joint p -value is then computed based on the product of the p -values calculated for the individual platforms. Like IMPaLA, InCroMAP supports a wide variety of source databases, such as KEGG pathways (including proteins/compounds), gene sets from the molecular signatures database (MSigDB), and gene ontology (GO) terms. The result list of an exemplary KEGG pathway enrichment is depicted in Fig. 1, where each pathway is assigned a p -value and an FDR corrected p -value (q -value).

During the enrichment analysis, InCroMAP discards unidentified features that could not be mapped to a gene or metabolite. While this procedure results in a loss of information, there is currently no better way to handle the problem of unidentified features. Generally, InCroMAP is not limited to targeted metabolomics data because dedicated tools, such as OpenMS or XCMS, are able to assign multiple candidate identifiers to metabolic features. However, the annotation with multiple candidate identifiers usually yields many-to-many mappings because a metabolic feature can be mapped to several candidate identifiers and vice versa. These many-to-many mappings are problematic for the overrepresentation analysis. A user should be careful when interpreting enrichment p -values of untargeted metabolomics data. Solutions to this problem in the setting of gene set enrichment have been proposed [18] and will be evaluated for implementation in a future version of InCroMAP.

3. Results and discussion

3.1. Metabolic overview

The results of a pathway enrichment analysis are typically presented to the user as a table or barplot of pathways, sorted according to the associated p -values, which score the significance of the enrichment with differing genes. As these traditional presentations of an enrichment result do not highlight any functional relations between pathways we devised an alternative method, which provides the user with a more structured view of the metabolic pathways altered by a certain condition. Using the metabolic overview feature, the user can display enrichment results in the context of KEGG's global 'Metabolic Pathways' map (KEGG: map01100), which contains references to all subordinate metabolic KEGG pathways. In this illustration, pathways, which are significantly enriched with up- or down-regulated genes or metabolites, are highlighted in different shades of blue (Fig. 2).

Since KEGG does not offer a meta-pathway for cellular signal transduction, this feature is limited to metabolic pathways.

In a recent study, we employed InCroMAP to detect changes in the intermediary metabolism of mouse liver tumors caused by mutations in the oncoproteins CTNNB1 and Ras which play a central role in the Wnt and MAPK signaling pathway, respectively [20]. In order to exemplify the utility of the metabolic overview feature demonstrated in Fig. 2A, we applied InCroMAP to multi-level omics data, which are available from GEO under accession GSE51358. Besides metabolite levels detected by NMR this dataset comprises mRNA, miRNA and protein expression as well as DNA methylation profiles.

3.2. Pathway-based visualization

Metabolic or signaling pathways of interest can either be automatically downloaded from KEGG or imported from other sources in BioPAX format. Being rendered in an interactive graph viewer, the pathway nodes representing genes/proteins can be overlaid with expression data from mRNAs and multiple protein products (Fig. 2B and C). Additionally, miRNAs can be connected to a given pathway based on experimentally confirmed or predicted interactions to their target mRNAs. If desired, the tool also visualizes differential methylation of proximal gene promoters, which is by default computed based on the largest peak observed in the upstream region of the transcription start site. In the extended version of InCroMAP we also support the visualization of metabolomics data. For each pathway node corresponding to a gene, protein, or metabolite, the individual measurements derived from multiple platforms and/or probesets as well as the DNA methylation profile of the proximal promoter can be inspected in a separate details panel. The integrated visualization of these data in a pathway-centered manner allows for drawing hypotheses and gaining novel biological insights. To illustrate the analytical power of InCroMAP we selected a complex multi-omics dataset including metabolite levels detected by NMR, expression of mRNA, miRNA and proteins as well as DNA methylation profiles which ideally exemplifies the performance of the software tool. These data were achieved by the investigation of a hepatic tumor characterized by a mutation in the oncoprotein CTNNB1. Exemplarily, tumor based effects on glucose catabolic and anabolic pathways (Fig. 2B) as well as of taurine and hypotaurine metabolism (Fig. 2C) are presented. The InCroMAP evaluation clearly shows tumor associated alterations in DNA methylation, mRNA expression, microRNA as well as metabolite levels. For instance, the two key enzymes G6pc and Pck1 catalyzing rate-limiting steps in glycogenolysis and gluconeogenesis, respectively, were both downregulated in CTNNB1-mutated tumors (G6pc, Fig. 2B). Concurrently, a slight downregulation of glucose levels was observed. Given the integrated view of the pathway one may hypothesize if G6pc expression is silenced by potentially interacting miRNAs or if the downregulation of Pck1 is caused by the hypermethylation present in the proximal promoter of the gene. Since this figure only serves for illustrative purposes, the reader is referred to Unterberger et al. for a more detailed description of the metabolic alterations found in CTNNB1-mutated mouse liver tumors [20].

In summary, we presented an extended version of InCroMAP, which allows the integrated pathway-based visualization of data from multiple omics platforms, such as DNA methylation, mRNA, miRNA, proteins, protein modifications, and metabolomics data. The tool enables a user to interactively browse through and visualize different KEGG pathways. Furthermore, the metabolic overview function of InCroMAP can be used to generate an interactive global map of cellular metabolism, which provides a structured view of the metabolic changes present in a certain experimental condition. The tool provides a useful overview of multi-omics data in the

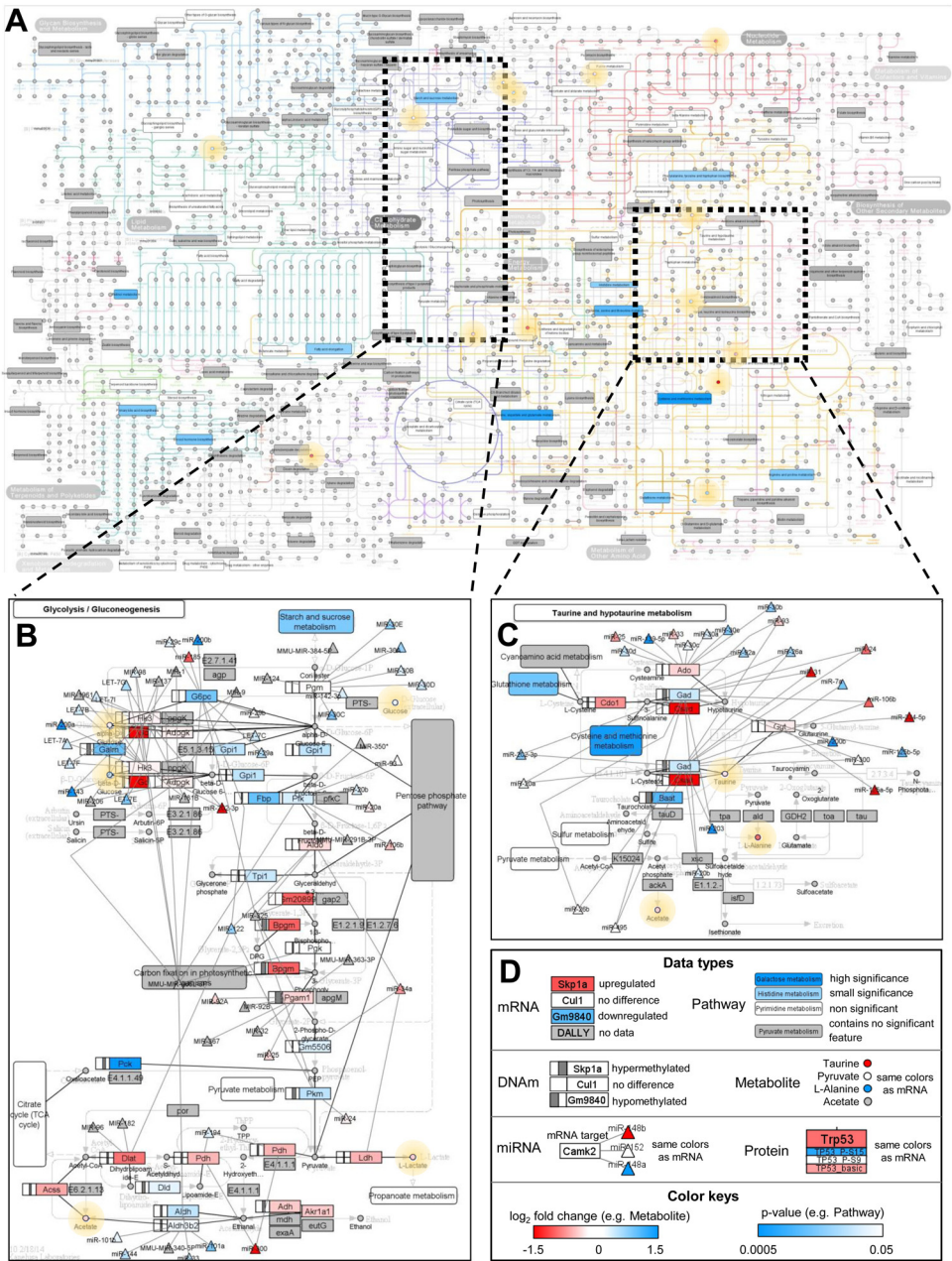


Fig. 2. The evaluation and visualization of complex multiomics data in the context of metabolic pathways affected in CTNNB1-mutated mouse liver tumors by the application of InCroMAP is exemplarily shown. (A) shows the metabolic overview in which subordinate metabolic pathways are colored according to their enrichment *p*-values. Additionally, changes in concentration are indicated for specific metabolites (highlighted in yellow) which could be identified based on NMR spectroscopy. (B) depicts an excerpt from the glycolysis/gluconeogenesis pathway overlaid with omics data obtained from diverse molecular levels (e.g., mRNA/miRNA expression, DNA methylation status of promoters, metabolite concentrations determined by NMR analysis). (C) Taurine and hypotaurine metabolism pathway overlaid with multi-level omics data; (D) shows a legend that explains the visualization of the data types and the coloring schemes.

Table 1
Comparison between chosen features of four different freely available tools for the analysis of multi-omics data. Supported features are marked by a bullet (•), whereas non-supported features are indicated by a circle (○).

	InCroMAP	Paintomics	MassTRIX	IMPALA
Mass annotation	○	○	•	○
Supporting probe-level data	•	•	•	○
Integrated enrichment analysis	•	•	•	•
Visualization of pathways	•	•	•	○
Gradual coloring	•	•	•	○
Interactive analysis	•	○	○	○
Metabolic overview	•	○	○	○
Analysis of more than two data types	•	○	○	○

context of metabolic and signalling pathways, which allows a user to gain insight in complex multi-omics data sets. Thus, it facilitates the interpretation of otherwise cumbersome data and enables the generation of initial biological hypotheses.

Table 1 compiles a comparison between InCroMAP and three others tools for the analysis of multi-level omics data: Paintomics, MassTRIX, and IMPaLA. Paintomics is most similar compared to InCroMAP. However, it only visualizes user specified pathways as image, whereas InCroMAP allows for interactively browsing through the pathways. MassTRIX was specifically designed for metabolomics data and can annotate metabolic features with candidate identifiers, which is not supported by any of the other tools. However, it does not feature an integrated enrichment analysis and does not perform a gradual coloring of the pathways nodes based on the abundance profiles. IMPaLA is an advanced enrichment analysis tool but it does not support probe-level data and does not provide any means of visualization. Furthermore, InCroMAP provides unique features for data analysis and visualization which are not offered by any published software. Most notably, a structured, global view of the changes in cellular metabolism can be generated by using the metabolic overview feature. Secondly, InCroMAP facilitates the integrated analysis of more than two types of data.

Currently, the pathway enrichment method of InCroMAP is focused on targeted metabolomics data because the overrepresentation analysis is not designed to handle many-to-many mappings, which usually occur when assigning multiple candidate identifiers to metabolic features. Future versions of InCroMAP must implement a more dedicated enrichment analysis to handle the problem of many-to-many mappings.

Further future improvements concern the integration of visualizations from additional metabolomics and lipidomics databases, such as LIPID MAPS [14]. Particularly the inclusion of visualizations based on LIPID MAPS pathways can overcome limitations of KEGG with respect to lipidomics data. Additionally, an automatic annotation with candidate identifiers for nontargeted metabolomics data based on the comparison against theoretical adducts of metabolites from metabolomic databases is planned for a future release of InCroMAP. The tool automatically notifies a user if an update is available on the project homepage (<http://www.cogsys.cs.uni-tuebingen.de/software/InCroMAP>).

To conclude, InCroMAP is a useful tool for the analysis and visualization of complex metabolomics, proteomics, transcriptomics, and genomics data. It is one of the first freely available tools for the interactive visualization of systems biology data, thereby supporting the identification of (pathobiological) alterations in complex multi-omics data sets

Acknowledgements

This project was supported by the Competence Network for Diabetes mellitus funded by the BMBF (FKZ 01GI1104A) and a grant from to the German Center for Diabetes Research (DZD eV; 01GI0925). Furthermore, the project received funding from the Innovative Medicine Initiative Joint Undertaking (IMI JU) [115001] (MARCAR project).

References

- [1] K. Gruden, M. Hren, A. Herman, A. Blejec, T. Albrecht, J. Selbig, C. Bauer, J. Schuchardt, M. Or-Guil, K. Zupancic, U. Svajger, B. Stabuc, A. Ihan, A.N. Kopitar, M. Ravnikar, M. Knežević, P. Rožman, M. Jeras, *PLoS ONE* 7 (1) (2012) e28761.
- [2] N. Amieur, S. Imbaud, G. Clément, N. Agier, M. Zivy, B. Valot, T. Balliau, P. Armengaud, I. Quilleré, R. Cañas, T. Tercet-Laforgue, B. Hirel, *J. Exp. Bot.* 63 (14) (2012) 5017–5033.
- [3] N. Gehlenborg, S.I. O'Donoghue, N.S. Baliga, A. Goesmann, M.A. Hibbs, H. Kitano, O. Kohlbacher, H. Neuweger, R. Schneider, D. Tenenbaum, A.-C. Gavin, *Nat. Methods* 7 (3s) (2010) S56–S68.
- [4] F. García-Alcalde, F. García-López, J. Dopazo, A. Conesa, *Bioinformatics* 27 (1) (2011) 137–139.
- [5] B. Wägele, M. Witting, P. Schmitt-Kopplin, K. Suhre, *PLoS One* 7 (7) (2012) e39860.
- [6] M. Kanehisa, S. Goto, Y. Sato, M. Furumichi, M. Tanabe, *Nucleic Acids Res.* 40 (Database issue) (2012) D109–D114.
- [7] C. Wrzodek, J. Eichner, A. Zell, *Bioinformatics* 28 (23) (2012) 3021–3026.
- [8] C. Wrzodek, J. Eichner, F. Büchel, A. Zell, *Bioinformatics* 29 (4) (2012) 506–508.
- [9] D.S. Wishart, C. Knox, A.C. Guo, R. Eisner, N. Young, B. Gautam, D.D. Hau, N. Psychogios, E. Dong, S. Bouatra, R. Mandal, I. Sinelnikov, J. Xia, L. Jia, J.A. Cruz, E. Lim, C.A. Sobsey, S. Shrivastava, P. Huang, P. Liu, L. Fang, J. Peng, R. Fradette, D. Cheng, D. Tzur, M. Clements, A. Lewis, A. De Souza, A. Zuniga, M. Dawe, Y. Xiong, D. Clive, R. Greiner, A. Nazzyrova, R. Shaykhtudinov, L. Li, H.J. Vogel, I. Forsythe, *Nucleic Acids Res.* 37 (Database issue) (2009) D603–D610.
- [10] E. Kenar, H. Franken, S. Forcisi, K. Wörmann, H.-U. Häring, R. Lehmann, P. Schmitt-Kopplin, A. Zell, O. Kohlbacher, *Mol. Cell. Proteomics* 13 (1) (2014) 348–359.
- [11] M. Sturm, A. Bertsch, C. Gröpl, A. Hildebrandt, R. Hussong, E. Lange, N. Pfeifer, O. Schulz-Trieglaff, A. Zerck, K. Reinert, O. Kohlbacher, *BMC Bioinform.* 9 (2008) 163.
- [12] F. Battke, S. Symons, K. Nieselt, *BMC Bioinform.* 11 (2010) 121.
- [13] Y. Wang, J. Xiao, T.O. Suzek, J. Zhang, J. Wang, S.H. Bryant, *Nucleic Acids Res.* 37 (Web Server issue) (2009) W623–W633.
- [14] M. Sud, E. Fahy, D. Cotter, A. Brown, E.A. Dennis, C.K. Glass, A.H. Merrill Jr., R.C. Murphy, C.R.H. Raetz, D.W. Russell, S. Subramaniam, *Nucleic Acids Res.* 35 (Database issue) (2007) D527–D532.
- [15] P. D'Eustachio, *Methods Mol. Biol.* 694 (2011) 49–61.
- [16] J. Xia, D.S. Wishart, *Nucleic Acids Res.* 38 (Web Server issue) (2010) W71–W77.
- [17] M. Chagoyen, F. Pazos, *Bioinformatics* 27 (5) (2011) 730–731.
- [18] M. Kankainen, P. Gopalacharyulu, L. Holm, M. Oresic, *Bioinformatics* 27 (13) (2011) 1878–1879.
- [19] A. Kamburov, R. Cavill, T.M.D. Ebbels, R. Herwig, H.C. Keun, *Bioinformatics* 27 (20) (2011) 2917–2918.
- [20] E.B. Unterberger, J. Eichner, C. Wrzodek, H. Lempiäinen, R. Luisier, R. Teranova, U. Metzger, S. Plummer, T. Knorpp, A. Braeuning, J. Moggs, M.F. Templin, V. Honndorf, M. Piotto, A. Zell, M. Schwarz, *Int. J. Cancer* (2014), <http://dx.doi.org/10.1002/ijc.28798>.