Lipid network and moiety analyses reveal enzymatic dysregulation and altered mechanisms from lipidomics

Tim D. Rose[†]
LipiTUM
Chair of Experimental Bioinformatics
TUM School of Life Sciences
Technical University of Munich
85354 Freising, Germany

Nikolai Köhler[†]
LipiTUM
Chair of Experimental Bioinformatics
TUM School of Life Sciences
Technical University of Munich
85354 Freising, Germany

Lisa Falk
LipiTUM
Chair of Experimental Bioinformatics
TUM School of Life Sciences
Technical University of Munich
85354 Freising, Germany

Lucie Klischat
LipiTUM
Chair of Experimental Bioinformatics
TUM School of Life Sciences
Technical University of Munich
85354 Freising, Germany

Olga E. Lazareva
Chair of Experimental Bioinformatics
TUM School of Life Sciences
Technical University of Munich
85354 Freising, Germany

Josch K. Pauling[*]
LipiTUM
Chair of Experimental Bioinformatics
TUM School of Life Sciences
Technical University of Munich
85354 Freising, Germany

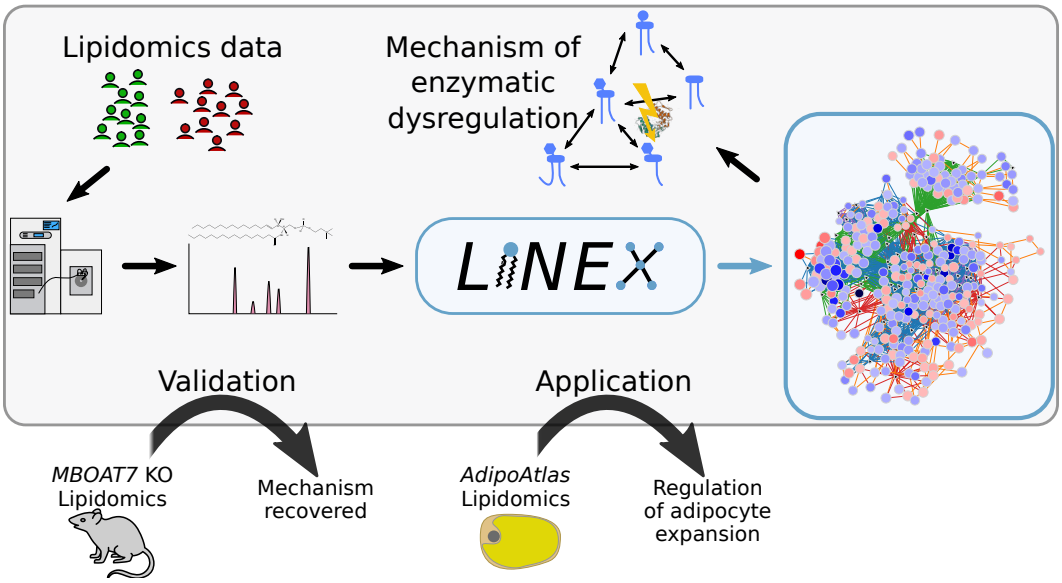
Abstract

Lipidomics is of growing importance for clinical and biomedical research due to an increasing number of discovered associations between lipid metabolism and diseases. However, sophisticated computational methods are required for biological interpretation including an understanding of metabolic processes, and their underlying (patho)mechanisms from lipidomics data. This can be achieved by using metabolic networks in combination with graph algorithms. Here, we present a lipid network analysis framework (Lipid Network Explorer, short LINEX) that allows biological interpretation of changes in lipidome composition. We developed an algorithm to generate data-specific lipid species networks from reaction database information. Using these networks, we developed a network enrichment algorithm that infers changes in enzymatic activity from lipidomics data, by leveraging multispecificity of lipid enzymes. Our inference method successfully recovered the MBOAT7 enzyme from knock-out lipidomics data . Additionally, we predict a PLA2 member to be at the center of dysregulation in adipose tissue in obesity. In our work, we showed the potential of lipidomics data to unravel mechanisms of metabolic regulation. Thereby our presented method can make lipidomics more clinically relevant by elucidating potential disease mechanisms.

***K*eywords** Network Enrichment · Lipid metabolic networks · Lipidomics · Disease mechanisms

[†] These authors contributed equally
[*] Correspondence to josch.pauling@tum.de

Synopsis



The Lipid Network Explorer (LINEX) is a framework to visualize and analyze quantitative lipidomics data. The included algorithms offer new perspectives on the lipidome and can propose potential mechanisms of dysregulation.

- Using the Reactome and Rhea databases, a comprehensive set of lipid class reactions is included and utilized to map the lipidome on custom data-specific networks.
- With a novel network enrichment method, enzymatic dysregulation can be recovered from lipidomics data.
- We validate its usability on data with a central lipid enzymatic deficiency.
- LINEX is the first tool capable of such analysis and includes complementary analyses options for structural lipid analysis. It is freely available as a web service (`https://exbio.wzw.tum.de/linex`).

1 Introduction

Lipids play a fundamental role in cells. They are not only crucial for long term storage of energy, but also can influence the activity and occurrence of membrane proteins [4] and inflammatory processes [56, 11]. Therefore, diseases are also influenced by lipids. This is known not only for liver and metabolic diseases [7, 33], but also for e.g. various cancers [53, 59, 29, 45]. Despite their essential role in many biological processes, excessive accumulation of lipids, especially in non-adipose tissues can lead to lipotoxicity [54, 65]. Hence, to fully understand diseases on the molecular level, changes of the lipidome have to be characterized and their regulation understood.

Nowadays, an increasing part of the lipidome can be identified and quantified using mass spectrometry (MS). The field, also known as lipidomics, is becoming more relevant for clinical applications and biomarker research [66]. While MS based lipidomics is not yet used for diagnoses, potential biomarkers have been discussed [38, 67, 49] and disease stratifications based on lipidomics proposed [63, 58]. To gain more insights into disease mechanisms, it is necessary to go beyond classification and prediction by proposing functional interpretations of lipid changes and links to other omics layers. Due to the complexity of both acquired lipidomics data as well as the regulatory mechanisms behind lipid metabolism, dedicated computational tools are of great importance for unraveling these associations.

Such interactions can be studied through biological networks. On the metabolic level, these networks describe reactions between metabolites that are catalyzed by enzymes. Dynamics of metabolic

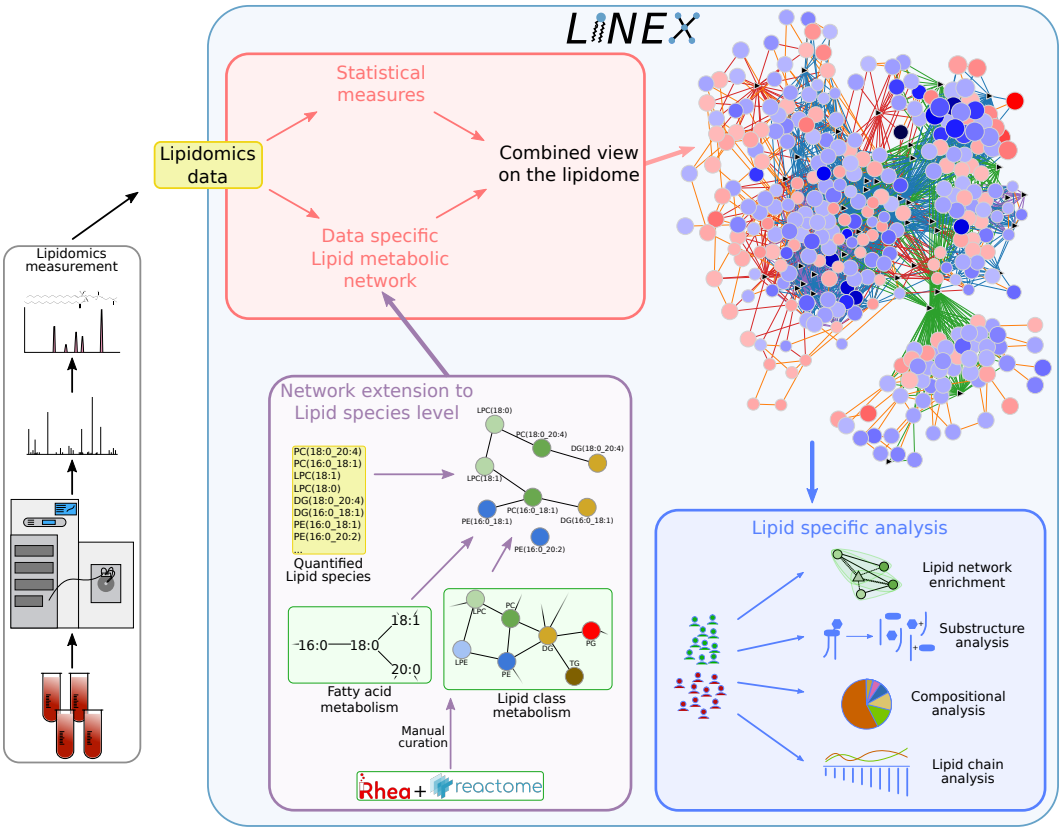


Figure 1: Lipidomics data is used as an input to LINEX. The lipids are then utilized to perform network extension that converters lipid class and fatty acid metabolic networks to lipid species, which are then visualized together with statistical measures such as t-tests or correlations. The network is also used as a basis for lipid substructure, compositional and lipid chain analysis. A lipid network enrichment algorithm, that takes enzymatic multi specificity into account, can be used to generate hypotheses for enzymatic dysregulation.

networks are commonly studied with dynamic modeling or flux balance analysis. Parameterization of such models requires huge amounts of data [52] Especially metabolite concentrations and fluxes are necessary, which are usually not available for clinical data. Another constraint is the inherent complexity of the lipidome and its chemical reactions. In contrast to the metabolome, where enzymes often catalyze only one specific reactions, multispecific enzymes are central in the lipid metabolism [18], which can catalyze conversions for a variety of lipid species. A plethora of lipid species can be synthesized because of various fatty acids (FAs) that can be attached to lipids of a certain class. Therefore, the creation of lipid-metabolic networks already poses a challenge in itself.

Another way to analyze biological networks is through network enrichment, which combines molecular networks with experimental data. By comparing two experimental conditions, the goal is to find highly connected molecular subnetworks that are enriched with significant genes, proteins or metabolites. The rationale behind this approach is to find mechanisms for dysregulation by assuming that e.g. a set of genes with close regulatory proximity marks the center of biological alteration. While this assumption is not always valid, it can provide hypotheses about potentially disturbed mechanisms for further validation experiments. Many algorithms have been developed over the years [3, 35, 14, 41, 20], mainly with a focus on protein-protein interaction (PPI) or gene-regulatory networks. A dedicated method for metabolomics data is included in the MetExplore analysis and visualization software [12]. Their MetaboRank [16] algorithm is an enrichment and network-based fingerprint recommendation method. In contrast to de-novo enrichment on large scale biological networks, pathway enrichment identifies significantly altered categorized pathways. This can be performed with the KEGG [30] or Reactome database [28]. As already mentioned, these methods cannot be directly applied to lipid species. However, dedicated pathway analysis methods have been

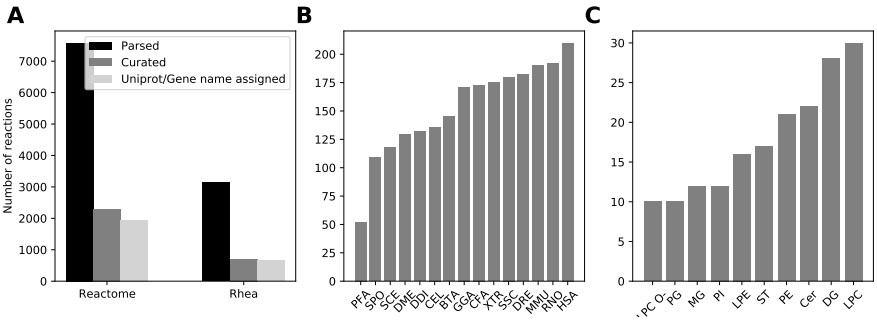


Figure 2: **A** Number of reactions parsed from Reactome and Rhea databases, after curation for available lipid classes and number of curated reactions, for which Uniprot or gene name annotations were available. **B** Number curated reactions for each organism from the Reactome database. Rhea does not list details about organisms. **C** Ten lipid classes with the most curated unique reactions.

developed. The Lipid Ontology web service (LION/web) performs an ontology based enrichment incorporating biological and chemical properties about lipids [43]. A dedicated tool to create lipid networks and find dysregulated linear reaction chains is BioPAN [19]. While being the first method to enable lipid de-novo pathway enrichment analysis for lipidomics, it only considers modifications of headgroups or fatty acids separately. We previously developed the Lipid Network Explorer (LINEX) [31]. A novel method to generate and visualize lipid-metabolic networks with statistical properties of the data that can be mapped onto networks to reveal global alterations of the lipidome. However, this method was limited in several aspects. Lipid class reactions, catalyzed by multispecific enzymes have to be entered by users. This increases customizability, but requires users to have detailed knowledge about lipid metabolism, reactions beyond the default were required. LINEX only offers a visualization of the lipidome and does not provide computational methods for automated pathway analysis. To overcome these shortcomings, we redesigned and extended the LINEX framework. First, we curated reactions from the Rhea [39] and Reactome [28] databases to construct complex data specific lipid networks, which gives a new view on the lipidome. Second, we developed an algorithm to derive hypotheses for mechanisms of enzymatic dysregulation in lipid metabolism. This is done with a novel network enrichment algorithm for lipid networks that takes the multispecificity of enzymes into account. This might help to translate lipidomics into clinical application [40, 69] and improve our understanding of the role of lipid metabolism in disease mechanisms. We successfully applied our method to knock-out lipidomics data and reveal potential dysregulations of the lipid metabolism in the adipose tissue of obese humans.

2 Results

2.1 A framework for lipid network analyses

The idea of LINEX is to provide a comprehensive view on the lipidome that combines quantitative statistical changes with prior knowledge about biochemical reactions. With this, we previously showed [31] on several studies that new insights into lipidome-wide data can be generated and that often central alterations are metabolically highly related. However, it was missing the capability to evaluate complex reactions with more than two lipids involved, a comprehensive lipid reaction database, and dedicated analysis methods for the lipid networks. Therefore, we redesigned and extended LINEX, now referred to as LINEX2. Since LINEX2 from now on incorporates LINEX, the terms will be used synonymously in this manuscript. LINEX2 provides extensive lipid metabolic information, including links to other omics-disciplines, as well as robust algorithms to identify mechanistic insights for hypothesis generation and deliver specialized analysis methods for lipidomics researchers.

The workflow of a lipidomics experiment can be divided into five steps: sampling, sample preparation, data acquisition, data processing and data interpretation [70]. LINEX is dedicated to the biological interpretation of lipidomics data, requiring high quality quantified lipid data as an input (Figure 1). Additionally, labels for samples can be uploaded. This is required for many comparative statistics such as fold changes or t-tests. The core of the analysis relies on the network extension. This

means that metabolic reactions on the lipid class level and fatty acid reactions are combined and extended to the lipid species level (Figure 1). For instance, the lipid class reaction PC → DG, where the phosphocholine headgroup is cleaved off from the PC, is applied to the measured lipid species PC(16:0_18:1) → DG(16:0_18:1) (for detailed description see Materials & Methods). Also fatty acid reactions, such as elongation or desaturation are added to the network e.g. for LPC(18:0) → LPC(18:1). While such a reaction usually does not occur on the lipid directly, but rather for activated free fatty acids, it helps to visualize fatty acid specific effects on the network as we showed in previously [31]. The network computation is possible with molecular species (e.g. DG(16:0_18:1)) or sum species data (e.g. DG(34:1)). For sum species, they are internally converted to molecular species, to properly evaluate reactions incorporating modifications or additions/removals of fatty acids. This is achieved using sets of common possible fatty acids for lipid classes. For a detailed description, please refer to the Methods section. Additionally, studies sometimes report individually identified (but not quantified) molecular species [63]. These can be used as input for LINEX to refine the network extension.

The network is then the basis for further analysis methods. These include class and fatty acid compositional analysis, lipid chain analysis [42] and lipid substructure deconvolution, inspired by a method for glycans [5]. Additionally, we developed a novel lipid network enrichment that can reveal enzymatic dysregulation from lipid data. Before going into detail on the analysis methods, the basis for network extension needs to be discussed.

2.2 Comprehensive curation of lipid-metabolic reactions

To provide a comprehensive overview of lipid metabolism, we curated lipid class reactions from the Rhea [39] and Reactome [28] databases (Figure 2A). As a reference for lipid classes, we updated the lipid classes from the ALEX123 lipid database [46]. Since LINEX only covers lipid species transformations, all transport reactions, which leave lipids unmodified, are not included in the final model. Furthermore, we removed specialized modifications such as oxidations or fatty acid branching, which are not generalizable or cannot be annotated to standardized lipid classes. This generalizability is an important aspect of the network extension to generate data-specific lipid-metabolic networks. That resulted in over 3000 annotated reactions from both databases combined (Figure 2A). Reactome additionally includes organism specific reactions, which are also available through LINEX. The top three organisms including the most reactions are homo sapiens (HSA), rattus norvegicus (RNO) and mus musculus (MMU) (Figure 2B).

Many reactions are duplicates for class reactions, because they describe reactions catalyzed in different compartments or confirmed reactions between enzymes and certain lipid species. In LINEX such reactions are aggregated into single class reactions, resulting in LPC, being the lipid class participating in most reactions (Figure 2C), followed by DG. All reaction identifiers are individually linked, providing a reference to the original database entries in the network.

To keep the LINEX database up-to-date, user contributions can be made using an online form. This way users can submit new lipid classes and lipid-metabolic reactions with references. After validation of the references, these contributions will be curated and added to the LINEX database. This way LINEX can be updated in a community effort to provide better analysis possibilities for all researchers and offer support for less studied parts of the lipidome.

2.3 Visualizing global lipidome alterations

We used LINEX to visualize the lipidome of white adipose tissue from the AdipoAtlas comparing obese to lean individuals [32]. Out of 674 Lipids, 647 were mapped, showing that a whole lipidome atlas can be analyzed with the LINEX database integration. The majority of the lipidome is connected, confirming the high coverage (Figure 3).

As expected of adipocytes, roughly one fourth of measured lipids consists of TG species. This can also be seen in the network visualization of the lipidome (Supplementary Figure S5). The network shows a global trend of fat accumulation. The biggest network community consists of TG and DG species, where the majority increases in obese individuals. Especially the lower right part of this community shows higher fold changes. These species include many unsaturated fatty acyls with three or more double bonds, indicating fatty acid specific regulation of the TG metabolism. The neutral lipid community has high connectivity to a subnetwork consisting of PC and PE species, in contrast

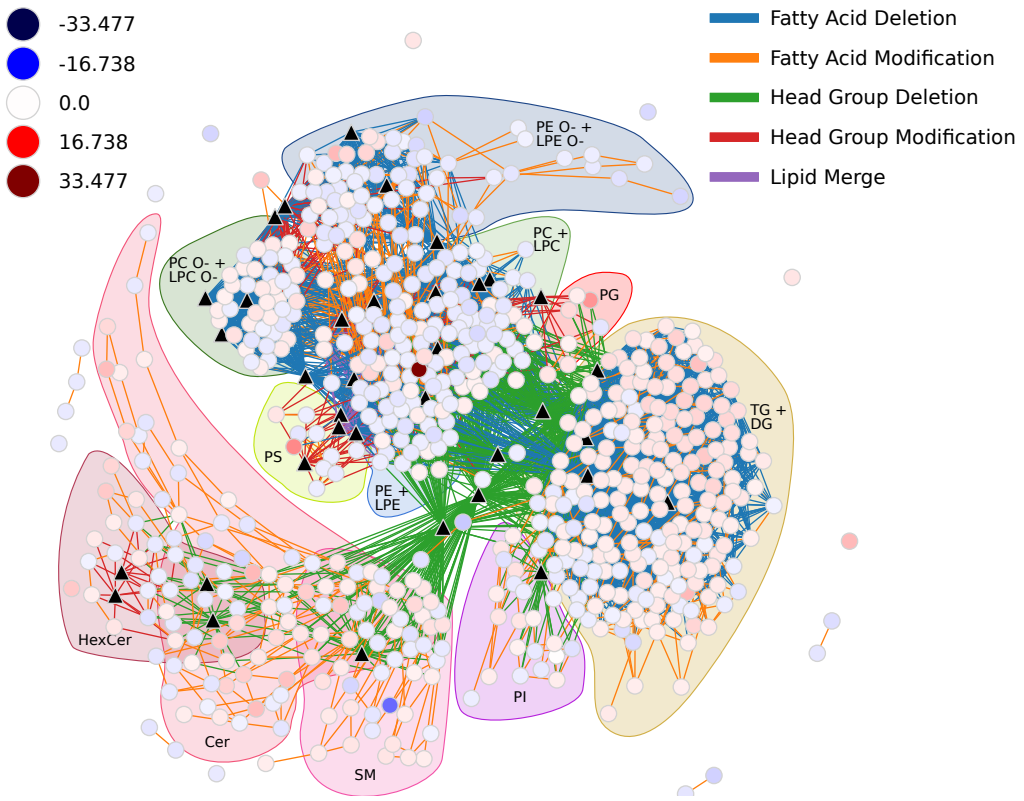


Figure 3: Lipidomics data from the AdipoAtlas visualized with LINEX. The network shows lipids as circular nodes and lipid class reactions as triangles. Red color of lipid nodes represents a positive fold change from lean to obese condition, and blue a negative fold change. Edge color indicates the type of reaction connecting two nodes. Abbreviations: DG - Diacylglycerol, TG - Triacylglycerol, PE (O-) - (Ether) Phosphatidylethanolamine, LPE (O-) - (Ether) Lysophosphatidylethanolamine, PC (O-) - (Ether) Phosphatidylcholine, LPC (O-) - (Ether) Lysophophatidylcholine, PG - Phosphatidylglycerol, PI - Phosphatidylinositol, PS - Phosphatidylserine, Cer - Ceramide, HexCer - Hexosylceramide, SM - Sphingomyelin.

to PI, PG and PS lipids, of which only few were measured. This suggests a metabolic co-regulation of DG, PC and PE lipids in the white adipose tissue.

While visualization of lipidomic alterations on biochemical networks gives an overview, a dedicated algorithm that can reveal the center of dysregulation in the metabolic network provides an additional benefit to data analysis.

2.4 Inferring enzymatic dysregulation from lipidomics data

Network enrichment is a powerful methodology to find functionally dysregulated molecular modules. In the context of metabolic or lipid networks, such methods could reveal underlying changes of enzymatic activity. However, it is more more challenging than network enrichment of e.g. proteomics data on PPI networks where changes of protein amounts correspond directly to functional changes of the nodes in the network. Additionally, multispecific enzymes catalyze the majority of reactions in lipid networks. Therefore, one enzyme or a set of enzymes corresponds to multiple edges in the network.

We developed a network enrichment method that takes this multispecificity into account. Figure 4A shows the process of preparing a lipid network for the enrichment analysis. Starting with a lipid species network, in which each node represents a lipid species and each edge an enzymatic reaction, we introduce a second type of nodes, representing lipid class reactions. These reaction nodes are

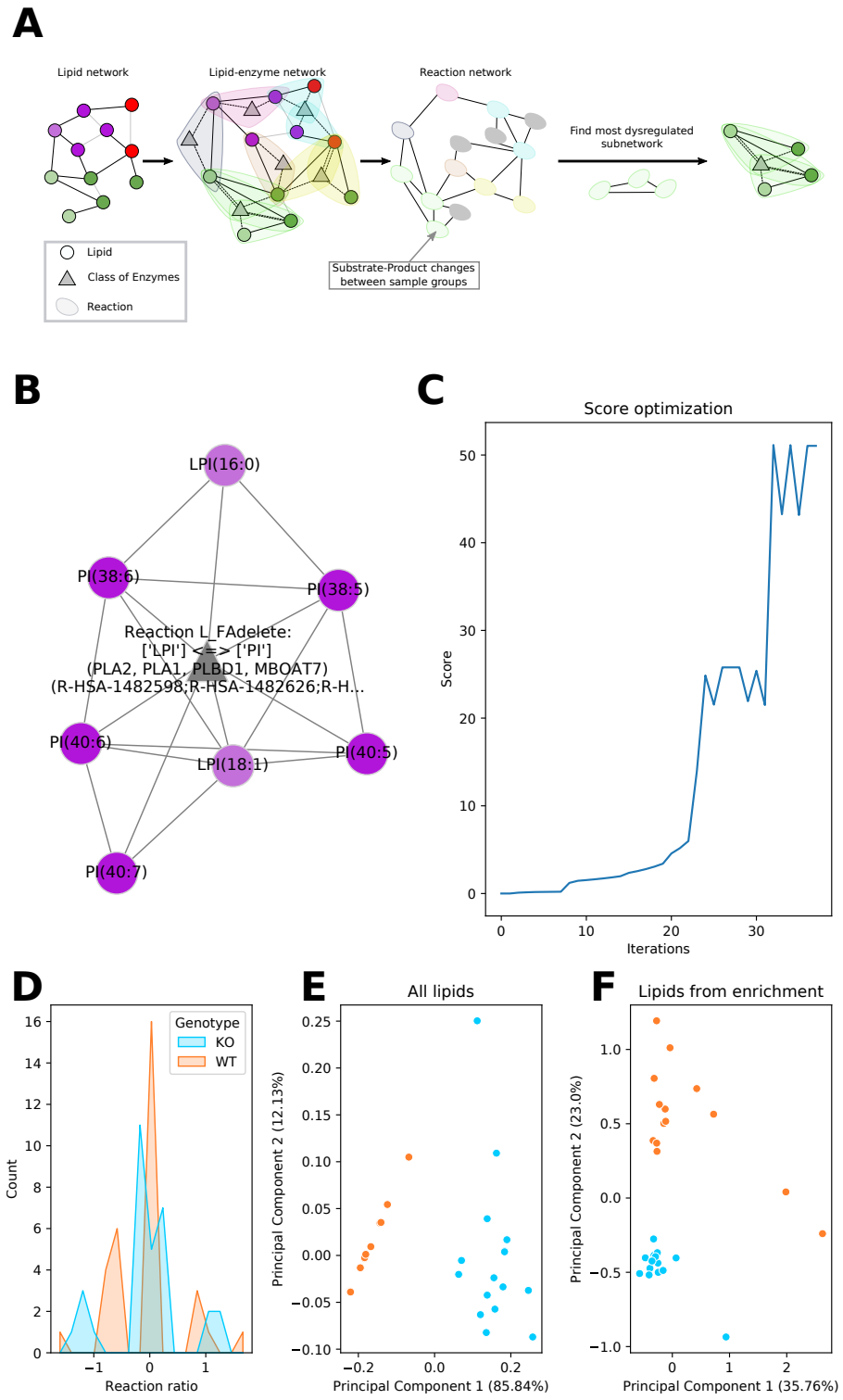


Figure 4: **A** Description of network preparation for the enrichment algorithm. The lipid network is converted into a reaction network via a hypergraph, to connect reactions catalyzed by the same enzymes. **B** Output of the enrichment for MBOAT7 knock-out data with the LPI to PI reaction at the center. **C** Trajectory of the optimization algorithm that yielded to the subnetwork in B. **D** Substrate-product ratio for the LPI to PI class reaction for all lipid species reaction pairs for the two experimental conditions MBOAT7 deficient (KO) and wild type (WT). **E** Principal component analysis of MBOAT7 data with all lipids and **F** lipids from the enriched subnetwork from B. Samples are colored by experimental condition.

then connected to all lipids, participating in a lipid species reaction it represents, via hyperedges. . For simplicity of the network analysis, the hypergraph is converted into a reaction network, where each node corresponds to a lipid species reaction (previous hyperedge) and reactions are connected if they share lipid species of enzyme sets. The advantage of this reaction network is that lipid species reactions extended from the same class reaction are now always neighboring nodes (See Figure 4A, nodes of the reaction network of one color are adjacent). The enrichment algorithm can therefore always easily connect solutions from the same class reaction, promoting more simple solutions with less dysregulated enzymes.

The reaction network is then used as an input for a simulated annealing-supported local search that finds a maximally dysregulated subnetwork. We quantify dysregulation by the relative change of the substrate-product ratio or difference between two experimental conditions. This is important especially for reversible reactions, whose rate depends on the substrates and products [37]. Heuristic reactions, such as ether conversions or fatty acid reactions are penalized in the objective function of the network enrichment and serve only to increase connectivity. Additionally, the number of class reactions in the network can be penalized to favor parsimonious solutions with a simple mechanistic explanation (For full description see Materials & Methods).

As a proof of principle for the method, we applied it to published lipidomics data, with a known alteration of lipid enzymatic activity. For that we selected the data from Thangapandi et al. [60]. In the paper, the authors compared liver lipidomics of mice with a hepatocyte-specific deletion of MBOAT7 (KO) to wild type (WT) mice under a high fat diet. MBOAT7 is known to catalyze a fatty acyl-CoA + LPI to PI + CoA with a specific preference for arachidonic acid (20:4(ω-6)) [21]. This study is well suited for validation of the method, since lipidomics data from tissue was acquired and a central lipid enzyme was knocked-out.

We used our network enrichment algorithm to find a dysregulated subnetwork maximizing the reaction difference between the WT and KO condition. The LINEX results, including the full lipid-metabolic network, is available as a html file in the supplement. The optimal solution can be seen in Figure 4B together with the scores during the local search. The trajectory of the score shows the advantage of the simulated annealing, to overcome local maxima that are found during the search (Figure 4C). The resulting network consists only of PI and LPI lipid species, and the LPI to PI class reactions, which can be catalyzed by MBOAT7. Furthermore, only fatty acids of at least 20 carbon atoms and 4 double bonds are attached to the LPIs, to form PI species. The reaction, where arachidonic acid (AA) is attached to an LPI is also explicitly found by the algorithm (LPI(18:1) $\rightarrow$ PI(38:5)). The result confirms the experimental KO of the MBOAT7 gene in the liver. Additionally, our method pinpoints the already known preference of MBOAT7 towards AA [21], as well as other fatty acids as potential substrates of the enzyme, such as 20:5, 22:4, 22:5 or 22:6. These substrates are not surprising, because similar to AA they are also strongly bent, requiring a certain pocket shape in the enzyme for the substrate. The lipid network enrichment is therefore also suited to reveal probable preferences of multispecific lipid enzymes for subsequent validation. Additionally, we implemented an empirical p-value estimation procedure for the enrichment results (Detailed description in Materials & Methods section Network enrichment). The proposed center of dysregulation (Figure 4B), has a p-value of 0.0018, indicating a highly significant mechanistic solution over unconnected and therefore non-mechanistic solutions.

It is also important to discuss the advantage of the network enrichment over the independent analysis of the substrate-product ratios per class reaction. We plotted these ratios for all reaction pairs of the LPI to PI reaction for the WT and KO condition (Figure 4D). The distributions show a peak around zero, indicating that many reaction ratios are not influenced by the KO. However, two more peaks around 60 and -60 can be observed for both conditions, where the peaks of the KO are shifted slightly more towards absolutely higher values. The altered activity of the MBOAT7 enzyme is far less obvious here compared to the network enrichment that visualizes lipid species reactions directly as altered subnetworks and shows fatty acid specific effects. To show the separating capability of the network enrichment, we plotted the principal component analysis (PCA) of the full lipidomics data and the lipids identified by the enrichment (Figure 4E). Due to the controlled experimental conditions, the full lipidomics data already shows a good separation of the two groups on the second principal component, which corresponds to 23% of the explained variance. However, after selecting the LPI and PI species from the enrichment solution, a perfect separation along the first principal component can be observed covering 85.6% of the variance. This shows that the enrichment is able to find the lipids describing a majority of the variance between the KO and WTs.

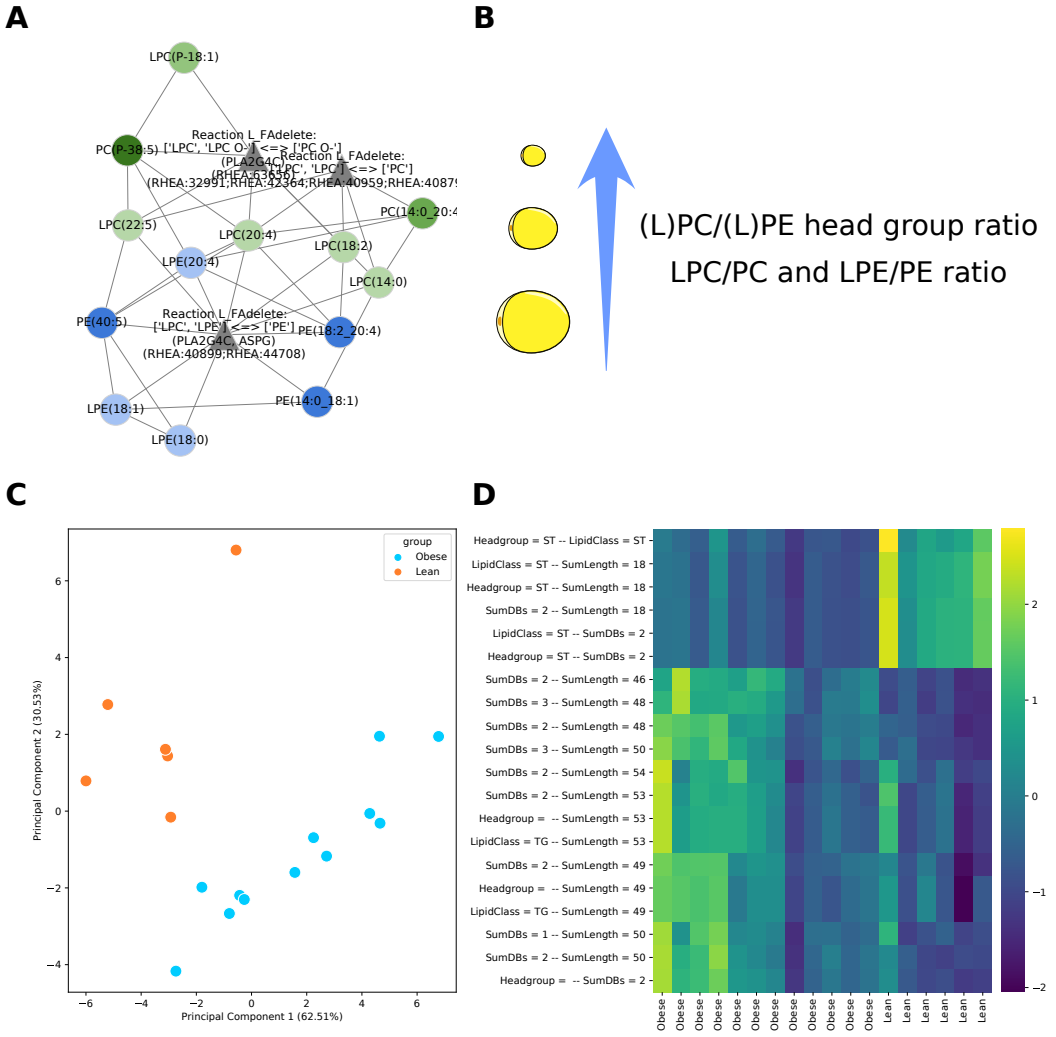


Figure 5: LINEX results from application on the Adipo Atlas data. **A** Subnetwork returned by the introduced enrichment algorithm. The enriched subnetwork contains three reaction nodes, all representing fatty acid transfer between lysophospho- and phospholipids. Furthermore, the network shows a preference towards long-chain polyunsaturated fatty acids. **B** Schematic of the relation between the assumed change of the ratio of LPC/LPE to PC/PE and the ratio of lipids with a PC and a PE headgroup as indicated by literature **C** Principal Component Analysis of the best 20 substructure combinations, selected via coefficients of a linear regression. **D** Heatmap with samples clustered by hierarchical clustering, based on the data from the top 20 combinations pairs.

With this experimental data, where a central lipid-metabolic enzyme was knocked-out, we could show that it is possible to recover enzymatic dysregulation from lipidomics data. In the next step, we aimed at applying our method to clinical lipidomics data, to derive hypotheses for regulatory changes.

2.5 Understanding lipidomic alterations of adipocytes

To demonstrate how our analysis pipeline enables a functional analysis of clinically relevant patient data, we applied the presented framework to the Adipo Atlas data introduced before. An overview of the LINEX network is presented in fig. 3. An interactive version of the network as well as all other analyses conducted with LINEX are available in a html file in the supplement.Applying our network enrichment algorithm to this data, aiming at the separation of lean and obese sample, yielded the lipid metabolic subnetwork shown in Figure 5a. The solution's p-value of 0.0002 indicates a strong significance of the enrichment results. In contrast to the enrichment results on the

MBOAT7 knockout data, the enrichment network contains three different reaction nodes. However, these reactions represent very similar transformations, namely the transfer of an acyl-chain from a Lysophospholipid onto a LPC, resulting in an (Ether-)PC or the inverse reaction, the removal of a FA from a PC species. All of these connections include the same enzyme, the Phospholipase 2 Group IVC (PLA2G4C). The PLA2 enzyme family is generally known to preferably act on the sn-2 position and that polyunsaturated FAs, such as AA, are commonly cleaved by them. This preference is also reflected in the subnetwork, where all but one lipid, with both sn-positions esterified, have 4 to 5 double bonds.

Literature research shows that PLA2G4C has been reported to be differentially expressed in obese individuals [8, 27] and products of (c)PLA2 activity are known mediators of adipose tissue metabolism [2]. These insights further support the practical feasibility of our reaction enrichment approach.

Furthermore, the subnetwork suggests a transfer of FAs between lipids with a Phosphocholine and a Phosphoethanolamine headgroup and their respective Lyso-Phospholipid species. This finding is in line with the observation that the number of fat cells does not change, but rather fat cells expand [57], since expansion requires adipocytes to adjust their membrane curvature. This property is influenced by both the ratio of LPC/LPE to PC/PE as well as the ratio of lipids with a Phosphocholine headgroup to lipids with a Phosphoethanolamine headgroup [10, 17] (Figure 5B). Additionally, a change in this ratio is associated with altered membrane integrity and fluidity [36, 13], which has been observed in obesity - though not in adipose tissue yet [9]. We confirmed this with a LION enrichment analysis [43], where we used the lipids of the enriched subnetworks as a target list (Supplementary Figure S1A). The analysis resulted in significant membrane related terms.

In addition, we analyzed cell culture data from [34], that compares the lipidomes of mesenchymal stem cells (MSC), adipogenic and osteogenic cells under normal and docosahexaenoic acid (DHA) added conditions. We focused on the comparison of the plasma membrane lipidomes of MSC vs. adipogenic cells under normal conditions. Our network enrichment algorithm indicates a change in enzymes converting PC to LPC (and vice-versa) (Supplementary Figure S1C), such as the PLA2 family, but with different FA preferences than the ones observed in the lean to obese comparison on the WAT data. This finding supports the observation of adipocyte expansion in favor of increased adipogenesis and suggests that membrane remodeling of MSCs upon adipogenesis is distinct from membrane remodeling in adipocytes due to obesity.

Next to the inference of a hypothesis on enzymatic changes from the WAT lipid data, we utilized a substructure analysis approach, inspired by the glycan deconvolution of Bao et al. [5], to reveal functional changes on the lipidome level. Using all pairwise combinations of lipid moieties such as headgroups, lipid classes (i.e. the combination of headgroup, number of fatty acids and fatty acid linkage), sum lengths and the total number of double bonds as structural features, the resulting matrix contains 650 features. Using a principal component analysis (PCA) to visualize the major axes of variation shows that both the separation between healthy and obese samples as well as the fraction of explained variance are very similar between the full lipid species matrix (Supplementary Figure S4A) and the substructure matrix (Supplementary Figure S4B). Subsequently, we aimed at finding the most discriminative feature combinations between lean and obese samples. For this purpose we used slope coefficients from a linear regression model with the sample groups as the target variable and reported the top 20 moieties (i.e. those with the highest absolute coefficients). To show how well the variation of these features corresponds to variation between lean and obese samples, we again used PCA. Principal components 1 and 2 from the subset of substructural combinations separate obese and healthy similarly well, but explain a much higher portion of the contained variance (Figure 5C).

Looking into the features selected and their scaled abundances (Figure 5D), with mono- and di-unsaturated sum species with a chain length of 49 and 50 in healthy samples and with neutral lipids in obese samples. While the neutral lipid association is expected, the chain length is higher as expected, since the most common FAs in WAT have a length between 16 and 18. However, an increase of chain length as a result of obesity has been reported by Tan et. al [68]. Since both moieties with longer as well as shorter chain lengths are decreased, our results suggest a sum composition-specific change of Triacylglycerol (TG) levels. This is especially interesting in combination with the network enrichment results, since PLA2 activity can influence the activity of hormone-sensitive lipase (HSL), which hydrolyzes TGs into FAs and glycerol and has as high substrate specificity depending on the FA composition [51].

Another structural component enriched in the selected moieties set are steryl esters, mostly in combination with FAs with a chain length of 18 and/or 2 double bonds. Considering the aforementioned FA composition in WAT, these results could indicate both a specific rise in unsaturated C18 steryl esters levels, as well as a general steryl ester increase, with the observed specificity being induced by the WAT FA pool.

3 Discussion

We presented a method to visualize and analyze lipid-metabolic networks. With information from two reaction databases and a comprehensive list of lipid classes, we showed that a whole lipidome tissue atlas can be mapped on the network and global changes of the lipidome can be observed. However, specialized modifications are not (yet) included in the LINEX database, due to their more complicated generalizability. Furthermore, we developed a network enrichment algorithm, to estimate candidates for enzymatic dysregulation from lipidomics data. We validated this approach on liver lipidomics data, where the deficient MBOAT7 enzyme was successfully recovered from the data.

Network enrichment for molecular biological data analysis has first been applied in 2002 [26] and a variety of methods have been developed since then. The challenge in developing network enrichment methods for metabolomics or lipidomics data lies in the fact that dysregulation on the enzymatic level is not measured directly, but can only be observed in indirect changes of the metabolome. For lipid networks, only one approach has been developed so far [19]. In contrast to our proposed network enrichment algorithm, this method is only searching for activated reaction chains between lipids of the same fatty acid combination. Since dysregulation in enzymatic activity on a lipid class level impacts many lipids of one or multiple lipid classes, independent of the fatty acid combination, LINEX2 is able to generate novel hypotheses. Furthermore, it is able to cover fatty acid addition/removal, enabling insights such as the MBOAT7 example we show in this work. To illustrate how LINEX2 compares to BioPAN [19] , we computed the BioPAN network (Supplementary Figure S2A) as well as the predicted list of active reactions, which does not include lyso-glycerophospholipids (LPI) species, and only one reaction chain with a PI species (Supplementary Figure S2B). The dysregulation of the MBOAT7 reaction can not be observed with this method.

Hence, LINEX lipid network enrichment is the only available method that aims at inferring enzymatic dysregulation from lipidomics data. Another novelty of the method is the usage of hypergraphs, to take multispecificity of lipid enzymes into account and put them into the center of the search for dysregulation. Compared to previous network enrichment methods, our approach profits from multiple reactions that correspond to the same enzyme. This increases confidence for the retrieved mechanism. Multispecificity also plays a role beyond lipids. For instance in the glycan metabolism, where enzymes extend various branched glycan structures [5], or DNA methylation [64]. Our principle of network analysis and enrichment can probably be extended into these fields and help to discover underlying dysregulation.

A limitation of our enrichment algorithm is that it computes substrate-product ratios independent from each other. In reality, however, reactions are linked through shared substrates or products and metabolic changes are propagated through the network. These effects can be due to, e.g. metabolic self-regulation [22] and structural or signaling functions. Since each lipid species takes part in a plethora of reactions, results of altered enzymatic activity might not be observed directly for the substrates and products of that reaction. This is also the case for multiple reactions, which form a consecutive transformation sequence that change at the same time. However, most disordered conditions are caused by only a few enzymatic steps, making the settings for such inaccurate approximations rare cases. Another limitation is the missing knowledge about enzymatic specificity. Therefore, our method is constrained to returning a set of candidate enzymes, which are attributed to the same type of reaction, without pinpointing specific variants. With more data available, such as the work from Hayashi et al. [24], better estimates for fatty acid-specific subnetworks can be made.

In our analysis, we also employed the lipid substructure analysis to provide insights based on lipid moiety occurrence/abundances. This analysis is not considering biosynthetic relations. While this is a clear limitation, we showed on the AdipoAtlas lipidome that in a combined analysis with the lipid-metabolic network based enrichment, can lead to new insights.Therefore, the combination of both types of analyses in our tool allows lipidomics researchers to generate hypotheses on causal mechanisms behind lipidome changes as well as how these changes affect the lipidome.

With the ability to connect enzymatic activity to lipidomics data, LINEX provides the basis for a knowledge-driven integration of lipidomics with proteomics data. The inclusion of quantitative proteome information could further improve the performance of our enrichment algorithm presented in this paper and open up the possibility to directly identify causal proteins. This but could be of great value for the causal interpretation of lipidome changes, which would directly translate into relevance for clinical applications, due to the many associations of lipids with various disorders [27, 38, 29, 59, 63].

With our new LINEX web service, we offer new analysis methods for lipidomic data. Ranging from network visualization to generating hypotheses for dysregulation. Available through a user-friendly interface, lipidomics researchers do not need to be experts in bioinformatics to perform sophisticated analysis of the lipidome in a metabolic context. Moreover, LINEX networks can be the basis for further methodological developments that help to enhance biological interpretability of lipidomics experiments by enabling inference of metabolic regulation from lipid data.

4 Materials & Methods

4.1 Webtool and data upload

LINEX is available as a web service at `https://exbio.wzw.tum.de/linex/`. The new LINEX2 web interface is an updated version of the previously published LINEX software [31]. It is built with the Django web framework (`https://www.djangoproject.com/`) in the python programming language (version 3.8, `https://www.python.org/`). PostgreSQL (`https://www.postgresql.org/`) is used as a back-end database for Django, to store data, networks and all computed attributes. Cookies are used to connect a browser session to uploaded data, their corresponding computed networks and analyses. For interactive network visualizations vis-network [48] is used and other interactive plots are done with Plotly [50]. All other user-site functionalities are implemented in plain JavaScript. PDF versions of networks are generated with the NetworkX package [23] in conjunction with the matplotlib library [25]. The backend was implemented in python. To achieve compatibility across operating systems, LINEX2 can be built in a Docker environment. In the public LINEX2 version, uploaded user data is temporarily stored on our server for a certain time or can be deleted manually by the user (for further information see `https://exbio.wzw.tum.de/linex/request-data-delete`). However, using the provided Dockerfiles LINEX can also be easily run locally on any computer (for instructions check the source code repository). LINEX2 is free software, published under the aGPLv3 license. The source code is available at `https://gitlab.lrz.de/lipitum-projects/linex`. While we adapted the procedure to generate lipid species networks, the original LINEX version can still be accessed through the website (marked as version 1).

Identified and quantified lipidomics data with optional sample labels can be uploaded to LINEX. Lipidomics data must be uploaded as a table with samples, lipids and their corresponding concentrations/amounts. To convert lipids into our internal programming model, we recommend the LIPID MAPS nomenclature [15]. However, we integrated the LipidLynxX [44] software, which is able to convert multiple lipid nomenclatures, increasing the compatibility of LINEX with multiple formats.

4.2 Database parsing & curation

We obtained lipid related reactions from the Rhea [39] and Reactome [28] databases. From Rhea, all reactions involving lipids were parsed (based on ChEBI ontology, subclass of CHEBI:18059). All reactions included in the category "Metabolism of Lipids" for all available organisms (e.g. R-HSA-556833 for Homo sapiens) were parsed from Reactome.

After parsing, all lipids and reactions were manually curated. Lipids were annotated and assigned to classes according to an updated version of lipid nomenclature from Pauling et al. [46] (Supplementary Table 1). Lipids that could not be annotated were not considered. Lipids are commonly composed of a headgroup, a backbone, and a set of attached fatty acids. From the databases, we extracted reactions showing conversions between common lipid classes, which are usually based on changes in one of these three attributes of lipids. We classified these lipid class reactions with at least one annotated lipid available into 5 categories: deadgroup modification (e.g. PC $\leftrightarrow$ PE), deadgroup addition/removal (e.g. DG $\leftrightarrow$ PA), fatty acid addition/removal (e.g. LPC $\leftrightarrow$ PC), fatty acid ether

exchange (ester to ether bond, e.g. LPA ↔ LPA O-), lipid merging (e.g. PA + PG ↔ CL) (see next section and Supplementary Figure S3 for more detailed descriptions). Reactions of the fatty acid metabolism can be manually added or banned by the user. Default available reactions are fatty acid elongation (increasing the chain length by 2), fatty acid desaturation (adding one double bond), hydroxylation/oxidation (adding one hydroxylation/oxidation to a fatty acid).

4.3 Network extension to species level

The network extension in LINEX2 is fundamentally different compared to the first LINEX version. Reactions from databases are used instead of user-entered conversions This allows more complex reactions. All steps for the network extension are explained below. To properly evaluate the reactions, molecular lipid species are required. This means that for each lipid the attached fatty acid must be available. Therefore, all lipid species, which are only available as sum species, are converted into a set of possible molecular species. As an example a PC(40:2) has to be converted into possible molecular species such as PC(20:0_20:2) or PC(22:2_18:0). For this, possible common (class-specific) fatty acids can be added by the user. Only if at least one molecular species can be generated that has the same sum formula as the original sum species, it is considered for the network extension.

Extension of lipid class metabolic networks to lipid species networks can be divided into two steps: extension of the class metabolism and fatty acid metabolism.

Extension of class metabolism Lipid class reactions are evaluated using the five defined reaction categories (headgroup removal/addition, headgroup modification, fatty acid addition/removal, fatty acid ether exchange and lipid merging). For each reaction, all lipids from the user data, which match the lipid classes that participate in a reaction are selected. Reactions with more than one lipid class as substrate and product are only possible or available for certain reaction categories. If possible, these are explicitly mentioned. The reaction evaluations are under the condition that a lipid class reaction for the substrate-product set exists. For a "Headgroup modification" reaction the substrate and product lipid require the same set of fatty acids, e.g. PC(18:0_16:0) → PE(18:0_16:0, (Supplementary Figure S3A). A "Headgroup addition/removal" also requires the substrate and product lipid to have the same set of fatty acids, e.g. DG(18:0_18:1) → PA(18:0_18:1) (Supplementary Figure S3B). The reaction is also possible for two lipids as substrates and two lipids as products, e.g. PC + Cer ↔ DG + SM (Supplementary Figure S3C). In this case, the headgroup is shifted from one lipid to another. For this evaluation, two substrate product pairs are matched of the lipid donating the headgroup (PC → DG) and the lipid accepting it (Cer → SM). These are then evaluated independently if at least one reaction per pair can be found. The "Fatty acid addition/removal" reactions in the case of one lipid as substrate and product require one lipid with one less fatty acid and the fatty acids of the lipid with fewer fatty acids to be contained in the other lipid, e.g. DG(18:0_18:1) → TG(18:0_18:1_16:0) (Supplementary Figure S3D). For two substrates and two products, a fatty acid is shifted from one lipid to another, e.g. PC + Chol ↔ CE + LPC. Again, two substrate product pairs are matched of the lipid donating the fatty acid (PC → LPC) and the lipid accepting it (Chol → CE). They are evaluated independently and the edges added to the network if two pairs can be found which donate/accept the same fatty acid. Another case exists for reactions with two substrates and one product (e.g. LPC + LPC ↔ PC). Also here, a fatty acid is shifted from one lipid to another, however, the donor is not considered a lipid, after the fatty acid is removed. Similarly, a pair of lipids accepting the fatty acid is formed (LPC → PC). Edges are then added to the network if the accepting and donating lipids have combined the same fatty acids as the resulting lipid. The reaction type "Lipid merging" describes two lipids that are bound together by a reaction, e.g. PG + PA ↔ CL (Supplementary Figure S3E). The molecular species of the substrates require the same combined fatty acids as the resulting lipid for this reaction to occur and be added to the network. Finally, "Fatty acid ether exchange" reactions transform lipids into ether lipids. In the cell, they occur by replacing an acyl bound fatty acid with fatty alcohol, e.g. RHEA:36171. However, fatty alcohols are usually not measured in mass spectrometry experiments and these reactions can therefore not be evaluated. We therefore optionally consider edges between lipid classes and their corresponding ether classes if they share the same set of fatty acids, e.g. LPA(18:1) → LPA(O-18:1), to improve network connectivity and stress fatty acid-specific effects in the network. Edges in the network are undirected, since we cannot conclude the net flux of a reaction from the lipidomics data, especially since for most reactions, counterparts in the opposite direction exist.

Extension of fatty acid metabolism Fatty acid synthesis and modification occurs commonly on activated fatty acids and they are not bound to lipids. However, to increase network connectivity and visualize fatty acid-specific effects on the lipid class level, fatty acid reactions on the lipid class level can be added to the network. As described earlier, this is done through user-defined reactions. A fatty acid reaction, e.g. PC(18:0_16:0) → PC(18:1_16:0), here fatty acid desaturation for the fatty acid 18:0 to 18:1, requires two lipids of the same lipid class and all but one identical fatty acid. Only one fatty acid modification is considered per reaction. In the case of elongation, the non-identical fatty acids require the same amount of double bonds and other modifications and have to differ by the length of two carbon atoms, e.g. 18:0 - 20:0. A desaturation requires fatty acids, which differ by a double bond, with all other attributes being the same. Custom fatty acid metabolism rules can be added, by providing the numeric changes of fatty acid attributes, such as length, double bonds, or modifications. Additionally, reactions between two specific fatty acids can be excluded. For example, the desaturation of fatty acid 18:2 to 18:3 is not possible in humans.

The network, with lipids as nodes is constructed with all edges between lipids from class, and fatty acid reactions. Class reactions can be added as nodes to the network with enzyme annotations. These are connected to all lipids that participate in these reactions.

Lipidomics data uploaded to LINEX can consist of sum and molecular species. In the network representation lipids are shown in the provided resolution. If provided as sum species, lipid nodes can also be shown as molecular species (based on the possible molecular species, as explained earlier). Sum-species including their statistical properties are then projected onto multiple potential molecular species.

4.4 Network enrichment

We developed a novel network enrichment algorithm for lipid networks. It aims to find the most dysregulated lipid subnetwork between two experimental conditions providing a hypothesis for enzymatic alteration/dysregulation. The methodology involves: 1. building a reaction network from a standard LINEX network and calculation of substrate-product changes per reaction. 2. Utilization of a local search algorithm to find the heaviest connected subgraph (i.e. the subgraph with the largest average substrate-product change) and 3. an empirical p-value estimation. All steps are described below.

Reaction network building To convert the lipid network to a reaction network, we generate a unique reaction identifier for each reaction (edge) in the network extension. This is especially important for reactions with more than one substrate and product, with multiple edges corresponding to one lipid species reaction. In the next step, all lipid species reactions are converted to a new network representation with reactions as nodes. Edges between two reaction nodes are drawn, if the reaction belongs to the same lipid class reaction or at least one lipid species can be found in both reactions.

Substrate-product change calculation The Substrate-product change is calculated independently for each reaction r_i of the set of all reaction nodes R in the network. It describes the relative substrate to product change between two experimental conditions. It can be calculated using absolute or relative substrate-product change. The data for the calculation consists of a set of measured samples N. With C denoting the subset of samples belonging to the control condition and D the samples belonging to the disease condition. The absolute substrate product difference Γ^a for reaction r_i for of the disease samples D is calculated as

$$\Gamma_{r_i}^{a,D} = \frac{\sum_{n \in D} \left(\sum_{p \in r_i} p_n - \sum_{s \in r_i} s_n \right)}{|D|}$$

with $\sum_{p \in r_i} p_n$ as the sum of all lipid products concentrations from reaction r_i in sample n and $\sum_{s \in r_i} s_n$ corresponding for substrates. The relative substrate-product difference Γ^a is calculated by

$$\Gamma_{r_i}^{r,D} = \frac{\sum_{n \in D} \left(\sum_{p \in r_i} p_n / \sum_{s \in r_i} s_n \right)}{|D|}$$

From the user, the absolute or relative score can be used to compute the final reaction score that compares both experimental sets C and D. It is calculated as follows:

$$\text{Score}^{r_i} = \frac{\left|\Gamma^D_{r_i} - \Gamma^C_{r_i i}\right|}{\Gamma^C_{r_i}}$$

This can be done with the relative or absolute substrate-product change. As previously explained, reactions of the fatty acid metabolism or ether lipid conversions are heuristic, to improve network connectivity. They do not occur directly on the lipid level. For that reason, They are also considered in the network enrichment but penalized (default = -1) to favor selection of the non-heuristic reactions.

Local search and simulated annealing Local search is a heuristic approach that is usually applied to hard optimization problems [1]. Local search investigates the search space by applying local changes to candidate solutions, such that the objective function value is increasing. The changes are applied until no more local improvements can be made. To avoid stagnation in a local maximum, the simulated annealing procedure [61] allows non-optimal solutions and thus increases the exploration space. The probability of accepting a suboptimal solution depends on the temperature parameter T, which decreases over time at rate α:

$$T = T_0 \cdot \alpha^n$$

where T_0 is the initial temperature, α is the rate of decrease and n is the iteration number. If no more local improvements are possible, a random solution is accepted under the following condition:

$$e^{\frac{o_{n-1} - o_n}{-T}} > \text{uniform}(0, 1)$$

where and are objective function scores at iterations n-1 and n correspondingly.

We employ local search on the reaction network $G = (V, E)$. Starting from a (random) set of connected starting nodes, also called seed, the local search can perform three actions for improvement in the objective function scores: node addition, node deletion, and node substitution. A minimum and maximum size for the subnetwork have to be entered as parameters, preventing the algorithm from selecting too small or big solutions. The action that allows improving the current value of the objective function is accepted, and thus a candidate solution is modified at each iteration. The algorithm terminates when a) no further improvements are possible, b) the simulated annealing condition is not satisfied, or c) the number of maximum iterations is reached. The best identified subnetwork is returned. The objective function score of a reaction subnetwork $G^* = (V^*, E^*)$ is computed as follows:

$$o = \frac{\sum_{v_i \in V^*} \text{Score}(v_i)}{|V^*| \times (p \times |\text{CR}(V^*)|)}$$

with a user defined penalty p for the number of different lipid class reactions in the subnetwork and $CR(V^*)$, the set of different lipid class reactions in the set nodes V^*. If the reaction network consists of unconnected components, the local search is run for each component independently and a subgraph for each component is returned.

Subnetwork p-value The network enrichment algorithm results in a subnetwork with a score for each run. To indicate if this subnetwork/score provides a significant insight comparing to equally sized random set of reactions, we compute an empirical p-value. For that, we sample reactions in the range of the minimum and maximum subnetwork size. These reactions are not connected, as in the subnetwork of the enrichment. This creates a distribution of scores. The distribution is then used to estimate a p-value for the solution found by the enrichment. The number of samples can be decided by the user, with more samples giving a better estimate of the distribution at increased runtime. The rationale behind sampling unconnected solutions is to estimate how much the connected (mechanistic) subnetwork scores compared to unconnected (non-mechanistic) solutions.

In the implementation for the LINEX2 web service, the local search is run multiple times (default=5), each with random seeds. The best result can then be optionally used as seed for another local search run, to improve this result. The best score achieved in all local search runs is then returned to the user.

4.5 Reaction ratio plots

Visualizations of reaction ratios were performed for each lipid class reaction individually. Similarly to the local search substrate/product ratios are computed for each lipid species reaction in the network

belonging to one lipid class reaction R_i and each sample n:

$$\frac{\prod P_{R_i}^{D,n}}{\prod S_{R_i}^{D,n}}$$

All ratios per experimental condition are compiled into a list and the density for each considered experimental condition is plotted.

4.6 Lipid substructure analysis

(Combined) Abundance of lipid features were implemented inspired by the glycan substructure method by Bao et al. [5]. We used the same vectorization and weighting as the authors, but with lipid substructures as features. These were: headgroup, backbone, independent fatty acids, sum length of fatty acids, sum double bonds of fatty acids and fatty acid hydroxylations. The features were weighted independently or in combination of pairs by occurrence in each lipid per sample. Lipid substructure analysis is also available through the LINEX2 webtool.

4.7 Lipid chain analysis

LINEX also offers lipid chain analysis. We implemented it in python according to the proposed method [42]. For each lipid class, lipid species with the same sum length of fatty acids are summed up per sample and a mean over all samples of one experimental condition is calculated. After that, the fold change between a selected control and e.g. a disease condition is calculated for each sum length per lipid class. The result is then plotted with an ascending fatty acid length on the x-axis, showing class specific fatty acid length fold changes between conditions.

4.8 Statistical measures implemented in LINEX2

To enable a combined visualization of the biochemical connections between lipid species and quantitative lipidomics measurements, LINEX2 offers the possibility to project different statistical metrics onto the species networks.

To characterize the changes of lipid levels between different experimental conditions, we provide unpaired parametric (t-test), non-parametric (Wilcoxon rank-sum test) test options, and a paired parametric test (Wilcoxon signed-rank test). All resulting p-values are automatically false-discovery rate (FDR) corrected using the Benjamini-Hochberg procedure [6]. Furthermore, fold-changes are computed to showcase effect size. For the computation of these metrics we used the scipy package [6, 62] in conjunction with the statsmodels package [55]. All measures can be visualized either as node sizes or node colors.

Measures for lipid connections (i.e. edges in the network) are correlation-based. Specifically, the options provided are spearman's correlation and partial correlation. All correlations with a significance above a user-specified threshold (default = 0.05) are set to 0 automatically. Correlation values can be visualized in the network representation through edge colors.

4.9 Principal component analysis

Feature reduction of lipidomics data was performed using principal component analysis from the python Scikit-learn package [47].

4.10 Analyzed data sets

The lipidomics data for MOBAT7 WT and knockout mice were taken from Thangapandi et al. [60]. No further processing was done and the data was analyzed as provided by the authors. Data for the Adipo Atlas was used as provided in the supplement of Lange et al. [32]. The comparison of MCS to adipogenic cells is coming from the supplement of Levental et al. [34]. Lipid species measured in less than 50% of all samples were removed prior to analysis with LINEX. For all data sets analyzed with LINEX, html files with the LINEX output are available in the supplement.

Data Availability Statement

LINEX is free software. Source code: GitLab (aGPLv3 License): `https://gitlab.lrz.de/lipitum-projects/linex`

Author Contributions

JKP supervised the project and secured the funding. NK, TDR, and JKP planned and conceptualized the work. NK and TDR developed the web service. NK, OL, and TDR designed and implemented the network enrichment procedure. LF, LK, and TDR parsed and curated the reaction databases, and implemented the network extension. NK and TDR applied, validated, and interpreted the approach on public lipidomics data. NK, OL, TDR, and JKP wrote the manuscript. All authors read, reviewed, and accepted the manuscript in its final form.

Acknowledgments

This project was funded by the Bavarian State Ministry of Science and the Arts in the framework of the Bavarian Research Institute for Digital Transformation (bidt; JKP, NK, TDR: Junior Research Group LipiTUM; O.L.: Doctoral Fellow).

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Emile Aarts and Jan Karel Lenstra. *Local Search in Combinatorial Optimization*. Princeton University Press, June 2018.

[2] Marcia J Abbott, Tianyi Tang, and Hei Sook Sul. The role of phospholipase a2-derived mediators in obesity, 2010.

[3] Nicolas Alcaraz, Tobias Friedrich, Timo Kötzing, Anton Krohmer, Joachim Müller, Josch Pauling, and Jan Baumbach. Efficient key pathway mining: combining networks and OMICS data. *Integr. Biol.*, 4(7):756–764, July 2012.

[4] John A Allen, Robyn A Halverson-Tamboli, and Mark M Rasenick. Lipid raft microdomains and neurotransmitter signalling. *Nat. Rev. Neurosci.*, 8(2):128–140, February 2007.

[5] Bokan Bao, Benjamin P Kellman, Austin W T Chiang, Yujie Zhang, James T Sorrentino, Austin K York, Mahmoud A Mohammad, Morey W Haymond, Lars Bode, and Nathan E Lewis. Correcting for sparsity and interdependence in glycomics by accounting for glycan biosynthesis. *Nat. Commun.*, 12(1):4988, August 2021.

[6] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.*, 57(1):289–300, January 1995.

[7] Stella Bernardi, Annalisa Marcuzzi, Elisa Piscianz, Alberto Tommasini, and Bruno Fabris. The complex interplay between lipids, immune system and interleukins in Cardio-Metabolic diseases. *Int. J. Mol. Sci.*, 19(12), December 2018.

[8] Nicholas J Carruthers, Clarissa Strieder-Barboza, Joseph A Caruso, Carmen G Flesher, Nicki A Baker, Samuel A Kerk, Alexander Ky, Anne P Ehlers, Oliver A Varban, Costas A Lyssiotis, Carey N Lumeng, Paul M Stemmer, and Robert W O'Rourke. The human type 2 diabetes-specific visceral adipose tissue proteome and transcriptome in obesity, 2021.

[9] Roberta Cazzola, Mariangela Rondanelli, Samantha Russo-Volpe, Ettore Ferrari, and Benvenuto Cestaro. Decreased membrane fluidity and altered susceptibility to peroxidation and lipid composition in overweight and obese female erythrocytes. *Journal of lipid research*, 45(10):1846–1851, 2004.

[10] Leonid Chernomordik. Non-bilayer lipids and biological fusion intermediates. *Chemistry and physics of lipids*, 81(2):203–213, 1996.

[11] Valerio Chiurchiù, Alessandro Leuti, and Mauro Maccarrone. Bioactive lipids and chronic inflammation: Managing the fire within. *Front. Immunol.*, 9:38, January 2018.

[12] Ludovic Cottret, Clément Frainay, Maxime Chazalviel, Floréal Cabanettes, Yoann Gloaguen, Etienne Camenen, Benjamin Merlet, Stéphanie Heux, Jean-Charles Portais, Nathalie Poupin, Florence Vinson, and Fabien Jourdan. MetExplore: collaborative edition and exploration of metabolic networks. *Nucleic Acids Res.*, 46(W1):W495–W502, July 2018.

[13] Rosie Dawaliby, Cataldo Trubbia, Cédric Delporte, Caroline Noyon, Jean-Marie Ruysschaert, Pierre Van Antwerpen, and Cédric Govaerts. Phosphatidylethanolamine is a key regulator of membrane fluidity in eukaryotic cells. *Journal of Biological Chemistry*, 291(7):3658–3667, 2016.

[14] Zijian Ding, Wenbo Guo, and Jin Gu. ClustEx2: Gene module identification using Density-Based network hierarchical clustering, 2018.

[15] Eoin Fahy, Manish Sud, Dawn Cotter, and Shankar Subramaniam. LIPID MAPS online tools for lipid research. *Nucleic Acids Res.*, 35(Web Server issue):W606–12, July 2007.

[16] Clément Frainay, Sandrine Aros, Maxime Chazalviel, Thomas Garcia, Florence Vinson, Nicolas Weiss, Benoit Colsch, Frédéric Sedel, Dominique Thabut, Christophe Junot, and Fabien Jourdan. MetaboRank: network-based recommendation system to interpret and enrich metabolomics results. *Bioinformatics*, 35(2):274–283, January 2019.

[17] N Fuller and RP Rand. The influence of lysolipids on the spontaneous curvature and bending elasticity of phospholipid membranes. *Biophysical journal*, 81(1):243–254, 2001.

[18] S Gatt and Y Barenholz. Enzymes of complex lipid metabolism. *Annual Review of Biochemistry*, 42(1):61–90, 1973.

[19] Caroline Gaud, Bebiana C Sousa, An Nguyen, Maria Fedorova, Zhixu Ni, Valerie B O'Donnell, Michael J O Wakelam, Simon Andrews, and Andrea F Lopez-Clavijo. BioPAN: a web-based tool to explore mammalian lipidome metabolic pathways on LIPID MAPS. *F1000Res.*, 10:4, January 2021.

[20] Susan Dina Ghiassian, Jörg Menche, and Albert-László Barabási. A DIseAse MOdule detection (DIAMOnD) algorithm derived from a systematic analysis of connectivity patterns of disease proteins in the human interactome. *PLoS Comput. Biol.*, 11(4):e1004120, April 2015.

[21] Miguel A Gijón, Wayne R Riekhof, Simona Zarini, Robert C Murphy, and Dennis R Voelker. Lysophospholipid acyltransferases and arachidonate recycling in human neutrophils. *J. Biol. Chem.*, 283(44):30235–30245, October 2008.

[22] SR Hackett, VRT Zanotelli, W Xu, J Goya, JO Park, DH Perlman, PA Gibney, D Botstein, JD Storey, and JD Rabinowitz. Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science*, 354(6311):aaf2786, 2016.

[23] Aric A Hagberg, Daniel A Schult, and Pieter J Swart. Exploring network structure, dynamics, and function using networkx. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pages 11–15, 2008.

[24] Daiki Hayashi, Varnavas D Mouchlis, and Edward A Dennis. Omega-3 versus omega-6 fatty acid availability is controlled by hydrophobic site geometries of phospholipase as. *J. Lipid Res.*, page 100113, August 2021.

[25] John D Hunter. Matplotlib: A 2D graphics environment, 2007.

[26] Trey Ideker, Owen Ozier, Benno Schwikowski, and Andrew F Siegel. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*, 18 Suppl 1:S233–40, 2002.

[27] Laura Jackisch, Warunee Kumsaiyai, Jonathan D Moore, Nasser Al-Daghri, Ioannis Kyrou, Thomas M Barber, Harpal Randeva, Sudhesh Kumar, Gyanendra Tripathi, and Philip G McTernan. Differential expression of Lp-PLA2 in obesity and type 2 diabetes and the influence of lipids. *Diabetologia*, 61(5):1155–1166, May 2018.

[28] Bijay Jassal, Lisa Matthews, Guilherme Viteri, Chuqiao Gong, Pascual Lorente, Antonio Fabregat, Konstantinos Sidiropoulos, Justin Cook, Marc Gillespie, Robin Haw, Fred Loney, Bruce May, Marija Milacic, Karen Rothfels, Cristoffer Sevilla, Veronica Shamovsky, Solomon Shorser, Thawfeek Varusai, Joel Weiser, Guanming Wu, Lincoln Stein, Henning Hermjakob, and

Peter D'Eustachio. The reactome pathway knowledgebase. *Nucleic Acids Res.*, 48(D1):D498–D503, January 2020.

[29] Jingting Jiang, Peter Nilsson-Ehle, and Ning Xu. Influence of liver cancer on lipid and lipoprotein metabolism. *Lipids Health Dis.*, 5(1):1–7, March 2006.

[30] Minoru Kanehisa and Susumu Goto. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, 28(1):27–30, January 2000.

[31] Nikolai Köhler, Tim Daniel Rose, Lisa Falk, and Josch Konstantin Pauling. Investigating global lipidome alterations with the lipid network explorer. *Metabolites*, 11(8), July 2021.

[32] Mike Lange, Georgia Angelidou, Zhixu Ni, Angela Criscuolo, Jürgen Schiller, Matthias Blüher, and Maria Fedorova. AdipoAtlas: A reference lipidome for human white adipose tissue. *Cell Rep Med*, 2(10):100407, October 2021.

[33] Chih-Hao Lee, Peter Olson, and Ronald M Evans. Minireview: lipid metabolism, metabolic diseases, and peroxisome proliferator-activated receptors. *Endocrinology*, 144(6):2201–2207, June 2003.

[34] Kandice R Levental, Michal A Surma, Allison D Skinkle, Joseph H Lorent, Yong Zhou, Christian Klose, Jeffrey T Chang, John F Hancock, and Ilya Levental. ω-3 polyunsaturated fatty acids direct differentiation of the membrane phenotype in mesenchymal stem cells to potentiate osteogenesis. *Sci Adv*, 3(11):eaao1193, November 2017.

[35] Hagai Levi, Ran Elkon, and Ron Shamir. DOMINO: a network-based active module identification algorithm with reduced rate of false calls, 2021.

[36] Zhaoyu Li, Luis B Agellon, Theresa M Allen, Masato Umeda, Larry Jewell, Andrew Mason, and Dennis E Vance. The ratio of phosphatidylcholine to phosphatidylethanolamine influences membrane integrity and steatohepatitis. *Cell metabolism*, 3(5):321–331, 2006.

[37] Wolfram Liebermeister and Edda Klipp. Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor. Biol. Med. Model.*, 3:41, December 2006.

[38] Ding Liu, Maureen Meister, Shiying Zhang, Chi-In Vong, Shuaishuai Wang, Ruixie Fang, Lei Li, Peng George Wang, Pierre Massion, and Xiangming Ji. Identification of lipid biomarker from serum in patients with chronic obstructive pulmonary disease. *Respir. Res.*, 21(1):242, September 2020.

[39] Thierry Lombardot, Anne Morgat, Kristian B Axelsen, Lucila Aimo, Nevila Hyka-Nouspikel, Anne Niknejad, Alex Ignatchenko, Ioannis Xenarios, Elisabeth Coudert, Nicole Redaschi, and Alan Bridge. Updates in rhea: SPARQLing biochemical reaction data. *Nucleic Acids Res.*, 47(D1):D596–D600, January 2019.

[40] Jiapei Lv, Linlin Zhang, Furong Yan, and Xiangdong Wang. Clinical lipidomics: a new way to diagnose human diseases. *Clin. Transl. Med.*, 7(1):12, April 2018.

[41] Haisu Ma, Eric E Schadt, Lee M Kaplan, and Hongyu Zhao. COSINE: COndition-SpecIfic sub-NEtwork identification using a global optimization method. *Bioinformatics*, 27(9):1290–1298, May 2011.

[42] Ahmed Mohamed, Jeffrey Molendijk, and Michelle M Hill. lipidr: A software tool for data mining and analysis of lipidomics datasets. *J. Proteome Res.*, 19(7):2890–2897, July 2020.

[43] Martijn R Molenaar, Aike Jeucken, Tsjerk A Wassenaar, Chris H A van de Lest, Jos F Brouwers, and J Bernd Helms. LION/web: a web-based ontology enrichment tool for lipidomic data analysis. *Gigascience*, 8(6), June 2019.

[44] Zhixu Ni and Maria Fedorova. LipidLynxX: a data transfer hub to support integration of large scale lipidomics datasets.

[45] Alicja Pakiet, Jarosław Kobiela, Piotr Stepnowski, Tomasz Sledzinski, and Adriana Mika. Changes in lipids composition and metabolism in colorectal cancer: a review. *Lipids Health Dis.*, 18(1):1–21, January 2019.

[46] Josch K Pauling, Martin Hermansson, Jürgen Hartler, Klaus Christiansen, Sandra F Gallego, Bing Peng, Robert Ahrends, and Christer S Ejsing. Proposal for a common nomenclature for fragment ions in mass spectra of lipids. *PLoS One*, 12(11):e0188394, November 2017.

[47] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.*, 12(85):2825–2830, 2011.

[48] Giancarlo Perrone, Jose Unpingco, and Haw-Minn Lu. Network visualizations with pyvis and VisJS. *arXiv preprint arXiv:1802.03426*, June 2020.

[49] Francesca Perrotti, Consuelo Rosa, Ilaria Cicalini, Paolo Sacchetta, Piero Del Boccio, Domenico Genovesi, and Damiana Pieragostino. Advances in lipidomics for cancer biomarkers discovery. *Int. J. Mol. Sci.*, 17(12), November 2016.

[50] Plotly Technologies Inc. Collaborative data science. *Montréal, QC*, 2015.

[51] T Raclot, C Holm, and D Langin. Fatty acid specificity of hormone-sensitive lipase. implication in the selective hydrolysis of triacylglycerols. *J. Lipid Res.*, 42(12):2049–2057, December 2001.

[52] Amit Rai and Kazuki Saito. Omics data input for metabolic modeling, 2016.

[53] Claudio R Santos and Almut Schulze. Lipid metabolism in cancer. *FEBS J.*, 279(15):2610–2623, August 2012.

[54] Jean E Schaffer. Lipotoxicity: when tissues overeat. *Curr. Opin. Lipidol.*, 14(3):281, June 2003.

[55] Skipper Seabold and Josef Perktold. Statsmodels: Econometric and statistical modeling with python, 2010.

[56] Charles N Serhan, Nan Chiang, and Thomas E Van Dyke. Resolving inflammation: dual anti-inflammatory and pro-resolution lipid mediators. *Nat. Rev. Immunol.*, 8(5):349–361, May 2008.

[57] Kirsty L Spalding, Erik Arner, Pål O Westermark, Samuel Bernard, Bruce A Buchholz, Olaf Bergmann, Lennart Blomqvist, Johan Hoffstedt, Erik Näslund, Tom Britton, Hernan Concha, Moustapha Hassan, Mikael Rydén, Jonas Frisén, and Peter Arner. Dynamics of fat cell turnover in humans. *Nature*, 453(7196):783–787, June 2008.

[58] Adam Stefanko, Christian Thiede, Gerhard Ehninger, Kai Simons, and Michal Grzybek. Lipidomic approach for stratification of acute myeloid leukemia patients. *PLoS One*, 12(2):e0168781, February 2017.

[59] Janel Suburu and Vong Q Chen. Lipids and prostate cancer. *Prostaglandins Other Lipid Mediat.*, 98(1-2):1–10, May 2012.

[60] Veera Raghavan Thangapandi, Oskar Knittelfelder, Mario Brosch, Eleonora Patsenker, Olga Vvedenskaya, Stephan Buch, Sebastian Hinz, Alexander Hendricks, Marina Nati, Alexander Herrmann, Devavrat Ravindra Rekhade, Thomas Berg, Madlen Matz-Soja, Klaus Huse, Edda Klipp, Josch K Pauling, Judith Ah Wodke, Jacobo Miranda Ackerman, Malte von Bonin, Elmar Aigner, Christian Datz, Witigo von Schönfels, Sophie Nehring, Sebastian Zeissig, Christoph Röcken, Andreas Dahl, Triantafyllos Chavakis, Felix Stickel, Andrej Shevchenko, Clemens Schafmayer, Jochen Hampe, and Pallavi Subramanian. Loss of hepatic mboat7 leads to liver fibrosis. *Gut*, 70(5):940–950, May 2021.

[61] P J van Laarhoven and E H Aarts. *Simulated Annealing: Theory and Applications*. Springer Science & Business Media, March 2013.

[62] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J van der Walt, Matthew Brett, Joshua Wilson, K Jarrod Millman, Nikolay Mayorov, Andrew R J Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E A Quintero, Charles R Harris, Anne M Archibald, Antônio H Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in python. *Nat. Methods*, 17(3):261–272, March 2020.

[63] Olga Vvedenskaya, Tim Daniel Rose, Oskar Knittelfelder, Alessandra Palladini, Judith Andrea Heidrun Wodke, Kai Schuhmann, Jacobo Miranda Ackerman, Yuting Wang, Canan Has, Mario Brosch, Veera Raghavan Thangapandi, Stephan Buch, Thomas Züllig, Jürgen Hartler, Harald C Köfeler, Christoph Röcken, Ünal Coskun, Edda Klipp, Witigo von Schoenfels, Justus

Gross, Clemens Schafmayer, Jochen Hampe, Josch Konstantin Pauling, and Andrej Shevchenko. Nonalcoholic fatty liver disease stratification by liver lipidomics. *J. Lipid Res.*, 62:100104, August 2021.

[64] J Walter, T A Trautner, and M Noyer-Weidner. High plasticity of multispecific DNA methyltransferases in the region carrying DNA target recognizing enzyme modules. *EMBO J.*, 11(12):4445–4450, December 1992.

[65] J M Weinberg. Lipotoxicity. *Kidney Int.*, 70(9):1560–1566, November 2006.

[66] Markus R Wenk. The emerging field of lipidomics. *Nat. Rev. Drug Discov.*, 4(7):594–610, July 2005.

[67] Furong Yan, Hong Zhao, and Yiming Zeng. Lipidomics: a promising cancer biomarker. *Clin. Transl. Med.*, 7(1):21, July 2018.

[68] Chong Yew Tan, Samuel Virtue, Steven Murfitt, Lee D Roberts, Yi Hui Phua, Martin Dale, Julian L Griffin, Francisco Tinahones, Philipp E Scherer, and Antonio Vidal-Puig. Adipose tissue fatty acid chain length and mono-unsaturation increases with obesity and insulin resistance. *Sci. Rep.*, 5:18366, December 2015.

[69] Linlin Zhang, Xianlin Han, and Xiangdong Wang. Is the clinical lipidomics a potential goldmine? *Cell Biol. Toxicol.*, 34(6):421–423, December 2018.

[70] Thomas Züllig, Martin Trötzmüller, and Harald C Köfeler. Lipidomics from sample preparation to data analysis: a primer. *Anal. Bioanal. Chem.*, 412(10):2191–2209, April 2020.