A systematic survey of lipids across mouse tissues

Mohit Jain,1,2,3,4 Soeun Ngoy,4 Sunil A. Sheth,5 Raymond A. Swanson,5 Eugene P. Rhee,1 Ronglih Liao,4 Clary B. Clish,1 Vamsi K. Mothra,1,2,3 and Roland Nilsson6,7

1Broad Institute, Cambridge, Massachusetts; 2Department of Systems Biology, Harvard Medical School, Boston, Massachusetts; 3Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts; 4Division of Cardiovascular Medicine, Department of Medicine, Brigham and Women’s Hospital, Boston, Massachusetts; 5Department of Neurology, University of California San Francisco and Neurology Service, San Francisco Veterans Affairs Medical Center, San Francisco, California; 6Unit of Computational Medicine, Karolinska Institutet, Stockholm, Sweden; and 7Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden

Submitted 3 July 2013; accepted in final form 7 February 2014

Jain M, Ngoy S, Sheth SA, Swanson RA, Rhee EP, Liao R, Clish CB, Mothra VK, Nilsson R. A systematic survey of lipids across mouse tissues. Am J Physiol Endocrinol Metab 306: E854–E868, 2014. First published February 11, 2014; doi:10.1152/ajpendo.00371.2013.—Lipids are a diverse collection of macromolecules essential for normal physiology, but the tissue distribution and function for many individual lipid species remain unclear. Here, we report a mass spectrometry survey of lipid abundance across 18 mouse tissues, detecting ~1,000 mass spectrometry features, of which we identify 179 lipids from the glycerolipids, glycerophospholipids, lysophospholipids, acylcarnitines, sphingolipids, and cholesteryl ester classes. Our data reveal tissue-specific organization of lipids and can be used to generate testable hypotheses. For example, our data indicate that circulating triglycerides positively and negatively associated with future diabetes in humans are enriched in mouse adipose tissue and liver, respectively, raising hypotheses regarding the tissue origins of these diabetes-associated lipids. We also integrate our tissue lipid data with gene expression profiles to predict a number of substrates of lipid-metabolizing enzymes, highlighting choline phosphotransferases and sterol O-acyltransferases. Finally, we identify several tissue-specific lipids not present in plasma under normal conditions that may be of interest as biomarkers of tissue injury, and we show that two of these lipids are released into blood following ischemic brain injury in mice. This resource complements existing compendia of tissue gene expression and may be useful for integrative physiology and lipid biology.

Lipids are a family of naturally occurring hydrophobic or amphipathic macromolecules, ubiquitous across all living cells and second only to proteins in abundance (19, 45, 52). The lipid family is remarkably diverse, comprising thousands of individual species, and can be subdivided based on structural features into a number of classes, each comprised of a number of lipid species of varying fatty acyl chain lengths and degrees of saturation (45, 53). These lipid species have evolved to perform a wide variety of essential functions. In mammals, choline and ethanolamine lipids are major structural components of cell membranes and influence membrane fluidity and ion exchange as well as apoptotic signaling (51). Glycerolipids, particularly triacylglycerols, are the main long-term energy-storing molecules in mammalian cells, and subtle modifications of triacylglycerol composition can have a profound impact on energy expenditure (10). Many lipids have tissue-specific functions in mammals. For example, surfactant phospholipids in the lung stabilize alveolar structure to facilitate gas exchange, and sphingomyelins surround axons to promote nerve conduction. Disruption of lipid homeostasis underlies a number of rare monogenic disorders (23, 46) and is also associated with common diseases, including cardiovascular disease (28), cancer (20), and metabolic disorders such obesity and diabetes (15, 36). Circulating lipids also serve as diagnostic and prognostic biomarkers of human disease (2, 28, 43). Nutritional interventions, including the use of polyunsaturated fish oils for cardiovascular disease (30, 50), and pharmacologic targeting of lipid-metabolizing enzymes, such as HMG-CoA reductase inhibitors (statins) for hypercholesterolemia (3), are regularly employed therapies.

In mammals, lipids can be synthesized de novo or consumed from exogenous sources and are metabolized by a host of enzymes, giving rise to a remarkably diverse “lipidome” (53). Advances in mass spectrometry (MS) now offer the means to accurately monitor a large number of lipid species (45, 54), but there is still little data available on lipid abundance across multiple mammalian tissues, and it remains unclear to what extent lipids exhibit tissue specificity. Whereas proteins can been studied by genetic means, lipids can rarely be selectively deleted or increased in mammalian tissues, rendering systematic studies of lipid function difficult. Monitoring the relative abundance of lipids across a large collection of tissues may provide an alternative approach for studying lipids, as previously demonstrated for mRNA transcripts (49) and proteins (37), and may perhaps shed light on the tissue basis for observed associations between circulating lipids and human disease states.

Recent studies have employed multiple complementary chromatographic and MS methods to comprehensively catalog a wide range of lipids present in a single cellular model such as yeast (12) or mammalian cell lines (11) or in isolated tissues including brain (17), adipose tissue (8), or human plasma (42). Here, we employ a simple lipid extraction and LC-MS approach that samples only a fraction of the lipidome but facilitates the study of a larger number of mouse tissues with multiple replicates to provide insight into tissue diversity. The present study provides data on the distribution of lipid species from nine lipid classes across 17 mouse tissues and plasma. The full dataset is available as an online resource for the scientific community (see supplementary dataset, linked to this article). Here, we describe this dataset and explore its use for generating hypotheses about the physiological functions of lipids and the use of lipids as markers for local tissue injury.
MATERIALS AND METHODS

An extended experimental procedure is available in the APPENDIX. Tissues and plasma were obtained from C57BL/6J male mice at 10–12 wk of age following a 12-h fast to remove any contamination from recent ingested chow. All animal procedures and handling were performed using protocols approved by the Harvard Medical School and Longwood Medical Area Institutional Animal Care and Use Committee. For lipid extraction, we employed a single-step protocol using isopropanol in excess, modified from previous studies of plasma and cells (1, 39, 43). Following anesthesia, tissues were rapidly dissected, washed, placed in 19× (wt/wt) HPLC grade 100% isopropanol, and rapidly snap-frozen. Lipids were extracted by mechanical homogenization of tissues in isopropanol, and the homogenate was centrifuged for collection of supernatants for LC-MS analysis. For each tissue, samples from at least five animals were analyzed. Prior to LC-MS, all samples were diluted with additional isopropanol as needed to ensure that total lipids in each tissue were within MS dynamic range. Internal standards were added to extracted samples after dilution. Lipid separation was achieved by reverse-phase chromatography using a Prophere HP C4 column (150 × 3.0 mm, Grace). MS data were acquired using a 4000 QTRAP triple quadrupole/linear ion trap mass spectrometer (AB SCIEX, Foster City, CA) using electrospray ionization and Q1 scans (full scan MS) in the positive ion mode. Aliquots of a pooled plasma sample were analyzed after every 20 samples and used to determine technical variation for each of the lipid species monitored.

Raw LC-MS data were binned into 0.1 m/z units and median filtered to remove “speckle” noise. Integrated peak areas were normalized to the internal standards. To determine the molecular identity of specific lipids, retention times and m/z values for 80 total lipid standards (Avanti Polar Lipids) were determined, including 57 qualitative standards and 23 quantitative Lipid MAPS MS standards (45) (see APPENDIX). To account for variation in ionization efficiency across lipid species, a set of quantitative Lipid MAPS MS standards (see APPENDIX) covering all lipid classes except LPE and FAC (for which no quantitative standards were available) were measured at equimolar concentrations, and ionization efficiency was estimated from the resulting intensity values and normalized out, as in previous reports (7).

For cluster analysis, pairwise distances between samples were computed as 1 minus the Pearson correlation coefficient. Data were projected into two-dimensional space using a nonlinear algorithm “SMACOF,” which searches for a configuration of points in 2-D whose pairwise distances closely approximate the original correlation distances. For the heatmap visualization (see Fig. 2A), species that did not reach a Z-score over zero of at least 2 in at least one tissue were discarded, and the remaining data were converted to fold changes over the overall mean. For the tissue composition analysis (see Fig. 3), relative abundance was calculated by normalizing each value to the total abundance of all lipids within a lipid class. For the lipid-lipid correlation analysis (see Fig. 5), we used 1 minus the Spearman correlation coefficient of the fold change data to avoid single tissues dominating the result. For the untargeted analysis, Z-scores were defined for every 2-D matrix element as the difference of tissue averages \( \bar{x}, \bar{y} \) divided by the pooled standard deviation. The generation and preprocessing of the mouse tissue gene expression data set has been previously described (29). False discovery rates were determined by the Benjamini-Hochberg procedure (4).

An experimental model of stroke, modified from prior reports (31), was induced through transient 1-h occlusion of the middle cerebral artery (MCA) with collection of blood by cardiac puncture at designated time points. Sham-operated mice underwent an identical surgical and collection procedure without MCA occlusion.

RESULTS

Assessing lipids across mammalian tissues. We analyzed the relative abundance of lipids from samples of 17 solid tissues and plasma isolated from 10- to 12-wk-old male C57BL/6 mice fasted for 12 h (Fig. 1A). For each tissue, we obtained samples from 5–6 individual mice for a total of 104 samples. For lipid extraction, we employed a single-step isopropanol-based protocol, which allows for capture of a fraction of the lipidome (1, 39, 43), including glycerolipids, glycerophospholipids, lyso-phospholipids, acylcarcinines, sphingolipids, and cholesteryl esters, without requiring additional wash steps, phase separation, or drying and resuspension of lipid extracts prior to analysis. Lipid extracts were diluted when necessary, separated by reverse-phase high-performance liquid chromatography (LC), and quantified using electrospray ionization full-scan mass spectrometry (MS) in positive ion mode, covering a mass/charge (m/z) ratio of 400–1,100 atomic mass units, resulting in a two-dimensional (retention time vs. m/z) full-scan data matrix for each tissue sample (Fig. 1A). More complex extraction methods (5, 14), use of negative ion mode, and scanning over a greater m/z range would be required to monitor additional lipid classes including negatively charged phospholipids such as phosphatidic acids, phosphatidylerines, and nonesterified sterols, as well as free fatty acids and their derivatives, among other lipid species. We estimate that the resulting data set contains >1,000 individual LC-MS features (peaks at distinct retention time and m/z coordinates) likely representing hundreds of distinct lipid species.

To identify specific lipids from the LC-MS features, we measured retention times and m/z values for a total of 57 qualitative and 23 quantitative standards (see APPENDIX), representing nine lipid classes. Once these representative members of a lipid class were identified, related species were annotated based on m/z values from the Lipid MAPS database (http://www.lipidmaps.org/data). Notably, in two-dimensional full-scan MS data, m/z and retention time follow a predictable pattern with respect to acyl chain length and saturation within a given lipid class (Fig. 1A), which facilitates identification of lipids. In this manner, we identified 179 species spanning nine major lipid classes, ranging in size from fatty acylcarcinines to triacylglycerides (Fig. 1B and supplementary dataset). Throughout, we denote these lipids as C:D, where C represents the total number of carbons and D the number of unsaturated double bonds in all acyl chains (R-groups in Fig. 1B), respectively. For example, TAG 56:4 represents a triacylglycerol lipid with a total of 56 carbons and 4 double bonds among the three acyl chains. We note that our method does not distinguish between closely related or isobaric species that differ only in acyl chain configuration but have equal total acyl chain carbon content and total degree of saturation. Hence, the species reported here may be composed of multiple, closely related chemical structures.

Although lack of quantitative stable-isotope standards precludes absolute quantitation of all lipid species, our data provide information on relative lipid abundance between tissues. Within a single sample, repeated lipid extraction and LC-MS analysis resulted in a median coefficient of variation (CV) of 9% across measured lipids, with more abundant lipids generally exhibiting lower CV, suggesting good technical reproducibility, comparable to other large-scale technologies.
such as microarrays and gel proteomics (32, 47). When variation across tissue samples from individual mice was assessed, median CV of all lipids was 25%; glycerolipids and acylcarnitines were the most variable (CV 32 and 54%, respectively), whereas structural lipids classes exhibited less variability (CV from 7 to 25%).

Relative abundance of classes of lipids across tissues. To first obtain a high-level view of lipid tissue distribution, we
estimated the total abundance of each lipid class as the sum of abundances of its member species (Fig. 1C). While we cannot account for varying extraction efficiency of lipids from different tissues, by this measure, cholesteryl esters (CE) were most abundant in plasma, likely reflecting CE in circulating lipoproteins (52). Total triacylglycerol (TAG) stores were 10-fold higher in white and brown adipose tissue than in any other tissue and were lowest in brain, consistent with the absence of TAG oxidation or stores in brain tissue (56). Diacylglycerides (DAGs) were also primarily found in adipose tissues, as well as in intestines and liver, mirroring TAG abundance. Fatty acyl-carnitines (FAC), substrates for mitochondrial β-oxidation (6), were abundant in heart and muscle but also in thymus and spleen (38). Total phosphatidylcholine (PC), a major constituent of cell membranes, was evenly distributed across tissues, whereas the membrane phospholipid phosphatidylethanolamine (PE) was more variable. The single acyl chain forms lysophosphatidylcholine (LPC) and lysophosphatidylethanolamine (LPE) generally correlated with their respective precursors PC and PE across solid tissues.

**Clustering tissues based of lipid content.** To systematically investigate the similarities between tissues with respect to lipid content, we projected abundance data for each tissue sample based on the 179 identified lipids into two dimensions using nonlinear dimensionality reduction (Fig. 1D and Appendix). In this analysis, most tissues were well separated based on lipid content, and functionally related tissues generally clustered together. Liver and kidney were closely aligned, as were immune organs (spleen and thymus), contractile tissues (cardiac and skeletal muscle), epithelium-rich tissues (aorta and epidermis), and steroidogenic glands (adrenals and testes) (Fig. 1D). Small and large intestine samples were more scattered, reflecting variation between individual mice (Fig. 1D). Similar results were obtained using hierarchical clustering analysis (Fig. 1E). Although many of these tissue-tissue relationships were recapitulated when clustering by whole genome expression profiles (29), others were unique to lipid profiles. For example, brown adipose tissue was most similar to muscle by gene expression, consistent with a common developmental origin (26), but was most similar to white adipose tissue with respect to lipid content (Fig. 1E). Thus, similar lipid composition may arise in cell types from distinct developmental lineages.

**Patterns of lipid species across tissues.** We next examined the detailed abundance patterns of the 179 identified lipid species across all tissues. Almost all lipids exhibited distinctive, highly reproducible patterns of tissue abundance (Fig. 2A and supplementary dataset), which were not evident when lipid classes as a whole were examined (Fig. 1C). For example, liver contained relatively little total CE (Fig. 1C) but exhibited high levels specifically of CE 18:1 (Fig. 2, A and B), in agreement with previous reports (34). Adrenal glands, testes, spleen, and thymus contained other long-chain CEs not found in the liver, such as CE 20:1 (Fig. 2, A and B). In adrenal glands, CE is a source of steroid hormone synthesis (16, 18), and since lymphocytes have been found to exhibit steroid synthesis activity (58), we hypothesize that these CEs may be precursors for steroids in lymphocytes as well.

Approximately two-thirds of the TAG species monitored were predominantly found in the fat-storing adipose tissues; the remaining TAGs exhibited distinct abundance patterns. For example, C60:2–3 TAGs were mainly found in adrenal glands, and C60:6–8 TAGs in heart and skeletal muscle (Figs. 2, A and B), suggesting a role for these specific TAGs outside adipose tissue. While overall PC abundance was similar across tissues (Fig. 1C), individual PC species again exhibited striking tissue abundance patterns (Fig. 2A). As expected, the PC 32:0–2 lipids were abundant in lung tissue (Fig. 2A); these lipids are known components of the alveolar surfactant (13). We also noted a group of polyunsaturated PC 40:6–8 species highly abundant in heart and muscle tissue (Fig. 2, A and B). Among sphingomyelin (SM) species, SM 36:1–2, a known constituent of neuronal myelin (44, 48), was specific to brain tissue (Fig. 2A), whereas other SM species exhibited completely different abundance patterns (Fig. 2, A and B). Taken together, these data are in good agreement with previously known tissue specificity of lipids and provides information on the tissue distribution of a number of additional lipid species.

**Distribution of lipid species within tissues.** We next examined the relative abundance of lipids within each tissue. This analysis is complementary to that shown in Fig. 2A, where the lipid composition of a given tissue is obscured by the large variation in total lipids (Fig. 1C); for example, it is difficult to discern the composition of brain TAGs in Fig. 2A, since total TAG content of brain is very low relative to other tissues. Comparing relative abundances of lipids is complicated by variation in ionization efficiency, which typically is inversely proportional to lipid mass within a lipid class (7). To correct for this effect, for each lipid class, quantitative standards for lipids of different masses were analyzed at equimolar concentrations, their ionization efficiencies as a function of mass were estimated from the resulting ion intensities, and relative abundance estimates were adjusted accordingly, as in previous reports (7) (see Appendix). Relative abundances should be considered valid only within lipid classes; we here focus on TAG, DAG, and PC lipids, for which multiple quantitative standards across a range of masses were available (Fig. 3). Differences in relative abundances were much larger than any observed ionization effects, which varied at most threefold for TAGs, twofold for PCs, and up to fourfold for DAGs (data not shown). Moreover, the observed relative abundances varied markedly from tissue to tissue, suggesting that they were not simply due to differences in ionization efficiency, and were highly reproducible across mice (Fig. 3). Certain lipids were relatively abundant in all solid tissues, such as C50–54:3–4 TAGs (Fig. 3A), which were also underrepresented in plasma (Fig. 3A). Some tissues exhibited distinct patterns in TAG distributions: liver and kidneys were rich in highly unsaturated C56–60 TAGs, which were also found in plasma but not in intestines or adipose tissue. While brain contained very little total TAG (Figs. 1C, 2A), C56:6–7 TAGs were specifically enriched in brain but underrepresented in other tissues (Fig. 3A). The tissue composition of DAG was more diverse (Fig. 3B), and tissues with specialized TAG composition, such as brain (Fig. 3A), also had unique DAG distributions (Fig. 3B). Long-chain DAG species (C40) in adrenal gland and testes (Fig. 3B) mirrored the presence of C60 TAG species (Figs. 2B), and DAG 32:0 in lung (Fig. 3B) correlated with the lung surfactant PC 32:0 (Fig. 2, A and B) (40). PC composition also differed markedly among tissues (Fig. 3C).

**Differential display of full-scan lipid profiles.** To expand our analysis beyond the 179 identified lipids, we developed a
computational method for differential display of full scan 2-D LC-MS spectra, visualizing abundance differences between two tissues for all lipid species monitored. We calculated z-scores between two sample groups for each of the image elements, resulting in a plot highlighting any features that differed between two groups (Fig. 4). This approach may help spotlight more subtle but systematic differences between tissues with related lipid profiles (Fig. 1D). For example, comparison of continuously contracting heart vs. fast-twitch gastrocnemius muscle (Fig. 4, A and B) revealed a pattern where heart was enriched in longer (C56–60) TAGs whereas skeletal muscle preferred shorter (C48–52) TAGs (Fig. 4A). A wealth of other lipids exhibited differences between heart and skeletal muscle as well, the majority of which remain to be identified. A similar comparison between brown vs. white adipose tissue revealed another pattern, where brown adipose tissue exhibited higher levels of saturated TAGs whereas white adipose tissue was enriched for more unsaturated species (Fig. 4C). This trend has been noted before on the level on total fatty acyl saturation (57). These patterns consist of relatively small fold differences for each lipid, which would be difficult to detect by analyzing individual species. Hence, this strategy may afford increased power to detect trends in lipidomic data, analogous to enrichment methods used in gene expression analysis (33).

Fig. 2. Lipid abundance patterns across 18 mammalian tissues. A: heat map of abundance of 179 distinct lipid species across the 8 lipid classes, presented as fold change above average for each lipid; for highlighted species, see main text. B: selected lipid abundance patterns; for discussion see text. Error bars denote SD across 5–6 replicates.
Fig. 3. Lipid composition within mammalian tissues. A: relative abundance of triacylglycerols (TAGs), normalized per sample to the total abundance within this lipid class to obtain molar fractions. Each solid line indicates tissue from an individual mouse. Gray vertical lines separate TAGs by total number of acyl chain carbons. The number of double bonds increase from left to right within each group, as indicated by triangles. B and C: relative abundance of diacylglycerol (DAG) and phosphatidylcholine (PC) lipids, as described for A.
Fig. 4. Abundance patterns from 2-D LC-MS data. A: z-Scores computed directly from 2-D LC-MS data of heart vs. gastrocnemius muscle. A region covering ~1/3 of the full matrix is shown. B: magnified view of the region marked by a white rectangle in A. Left and middle: images of representative samples from heart and gastrocnemius tissue; dark color indicates high abundance. Right: corresponding z-scores. C: intensity and z-scores from white vs. brown adipose tissue, as in B. For highlighted lipids, see main text.
Cluster analysis of lipid species. Whole genome gene expression patterns across multiple tissues have been tremendously valuable for pinpointing the function of previously uncharacterized genes (49), and a similar approach may help elucidate the physiological functions of lipids. In this “guilt-by-association” approach, similarities between tissue abundance patterns are used to generate hypotheses about physiological function. We jointly analyzed all 179 identified lipids by computing all pairwise correlation coefficients and organizing the resulting matrix by hierarchical clustering (Fig. 5). This analysis revealed several clusters of lipids with similar tissue abundance patterns, often spanning multiple lipid classes. One such cluster (cluster c) consisted of FACs and the highly unsaturated PC and PE species present in muscle tissues (Fig. 5). These FAC and PC species have been implicated in the metabolism and contraction of muscle tissue (24, 25), leading to the hypothesis that PEs in this cluster could play a similar role. Similarly, another lipid cluster (cluster b) consisted of short-chain PC and PE species, including the 32-carbon PC that functions as lung surfactant, suggesting that the remaining phospholipids in this cluster may have a related function. Interestingly, this cluster was also abundant in brain, indicating that these lipids may also be important in the central nervous system.

The correlation matrix representation can also reveal higher-order relationships not immediately evident from the cluster tree. For example, two distinct clusters of triacylglycerides enriched in plasma (cluster a) and adipose tissue (cluster d) exhibited an off-diagonal correlation (ad) indicating that a secondary relationship existed (Fig. 5). Closer inspection revealed that these two lipid groups were both present in liver and kidney as well, explaining the off-diagonal correlation. Another lipid cluster (cluster e) was exclusively present in adipose tissues.

Using tissue lipids to interpret circulating plasma lipid profiles. Lipid profiling of plasma is now widely applied to discover associations between circulating lipid levels and human diseases (27, 43). Our data set of tissue lipids could be useful to generate hypotheses about the potential origin or fate of such disease-associated lipid species. For example, a recent study found that circulating shorter-chain, saturated TAGs are elevated in the plasma of individuals at risk for future diabetes, whereas longer-chain, unsaturated TAGs are lower in such individuals (43). However, the physiological basis for these observations remains obscure. In our data, the shorter, saturated TAGs were enriched in mouse adipose tissue (cluster e, Fig. 5) whereas longer, unsaturated triglycerides were more abundant in liver and kidney (cluster a, Fig. 5), raising hypotheses about specific organs of relevance to the observed human diabetes TAG signature.

Exploring lipid-enzyme relationships. We next sought to uncover associations between lipids and lipid-metabolizing enzymes by integrating our tissue lipid data with existing tissue mRNA expression data (29). Reasoning that lipids that co-occur with an enzyme across tissues are more likely to be products of that enzyme, we calculated the pairwise correla-
tions between abundance profiles for lipids in a given class and gene expression profiles for enzymes known to produce lipids of that class (supplementary dataset). We identified 121 significant gene-lipid relationships at a 20% false discovery rate (supplementary dataset). For example, we found that PCs segregated into two groups, one associated with the choline/ethanolamine phosphotransferase 1 (CEPT1) enzyme, and another associated with choline phosphotransferase 1 (CHPT1) (supplementary dataset). Whereas CEPT1 synthesizes both PC and PE species, the CHPT1 enzyme is PC specific and is believed to account for most PC synthesis in mammals (52). The correlation profiles of CEPT1 and CHPT1 (Fig. 6A) predict that CEPT1 should preferentially produce short-chain PC whereas CHPT1 should prefer longer-chain PCs and that neither enzyme should be active with saturated PCs. These predictions are in overall agreement with the lipid specificity observed experimentally for these two enzymes (21, 22).

CEs are synthesized from free cholesterol and fatty acyls in mammals by two enzymes, sterol O-acyltransferase 1/2 (SOAT1/2, also known as ACAT1/2). Of these, SOAT2 is thought to be the main enzyme for esterification of dietary cholesterol (52). In our analysis, SOAT1 expression was correlated with long-chain (C20–22) esters, whereas SOAT2 was associated with short-chain esters (Fig. 6B). This analysis agrees with biochemical evidence demonstrating a preference of SOAT2 for C16 and C18 fatty acyls (9). Our data also predict that SOAT1 activity extends to very long-chain acyls, whereas SOAT2 is less active on saturated acyl chains than SOAT1. Taken together, these findings indicate that this integrative approach can be used to generate hypotheses about enzyme specificity.

**Tissue-specific lipids as plasma biomarkers of organ injury.** Given the marked tissue specificity of several lipid species (Fig. 2A), we reasoned that lipids abundant in a specific tissue but undetected in blood might be released into the circulation upon tissue injury and could serve as biomarkers for clinical conditions involving specific tissue damage, similar to currently used protein markers of tissue injury. A search of our data set using a stringent criterion for tissue specificity revealed eight lipids that were specific to five tissues (Table 1) and low in plasma. Three brain-specific lipids (SM 36:1, SM 36:2, and PC 34:0) were virtually absent from plasma under normal conditions, and we hypothesized that these lipid species could be released into circulation with local brain injury, as occurs during ischemic stroke. We confirmed that these species were low or absent in plasma sampled from the jugular vein of mice, indicating that limited release occurs from brain tissue under normal conditions (Fig. 7A; PC 34:0 was undetectable in jugular vein samples, not shown). In contrast, other SM species were present in plasma from the jugular vein (Fig. 7B). We then experimentally induced cerebral ischemia in mice by transient 1-h occlusion of the middle cerebral artery and

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Lipid</th>
<th>vs. Second Tissue</th>
<th>vs. Plasma</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ratio</td>
<td>Z-Score</td>
<td>Ratio</td>
</tr>
<tr>
<td>Adrenal</td>
<td>TAG 58:2</td>
<td>3.3</td>
<td>2.8</td>
</tr>
<tr>
<td>Adrenal</td>
<td>TAG 60:3</td>
<td>4.8</td>
<td>4.8</td>
</tr>
<tr>
<td>Brain</td>
<td>PC 34:0</td>
<td>2.9</td>
<td>15.1</td>
</tr>
<tr>
<td>Brain</td>
<td>SM 36:1</td>
<td>2.4</td>
<td>12.0</td>
</tr>
<tr>
<td>Brain</td>
<td>SM 36:2</td>
<td>12.2</td>
<td>11.8</td>
</tr>
<tr>
<td>Kidney</td>
<td>SM 34:2</td>
<td>3.7</td>
<td>9.9</td>
</tr>
<tr>
<td>Lung</td>
<td>PC 32:2</td>
<td>2.4</td>
<td>7.5</td>
</tr>
<tr>
<td>Spleen</td>
<td>PE 34:0</td>
<td>2.0</td>
<td>4.9</td>
</tr>
</tbody>
</table>

Shown are lipid species with ratio and z-score of at least 2 vs. the second most abundant tissue, as well as ratio and z-score of at least 5 vs. plasma; ∞ indicates not detectable in plasma.
measured lipid abundance in peripheral blood at 24 h following tissue injury. The 36:1 and 36:2 SMs were markedly increased in the circulation of animals at 24 h after tissue injury (Fig. 7C; fold increase 5.4 ± 1.8 and 6.4 ± 2.2, respectively; mean ± SD, P < 0.01), but not in sham-operated animals. We did not observe any increase in blood at 6 h (data not shown). In contrast to SM lipids, the PC 34:0 species was not found to be released into circulation following cerebral ischemia, even at 24 h (data not shown).

DISCUSSION

Tissue atlases of gene expression have proved to be useful resources for understanding gene function and tissue physiology (49). In the same fashion, to study mammalian lipids, we have generated an initial survey of major lipid species across mouse tissues. Our data reveal an intricate tissue distribution of lipids such that each tissue appears to express a unique lipid profile. While the influence of diet on these profiles is not clear from our data, these complex lipid abundance patterns suggest that lipid biology is specialized in mammalian tissues, with dedicated physiological roles for individual lipids. In this paper, we have focused on the 179 individual lipids. In this context, the presence of a lipid in a specific tissue may provide an inroad to studying its physiological function. For instance, we note the specific polyunsaturated PC 40:6–8 species to be relatively abundant in heart and muscle tissue (Fig. 2A and B). Interestingly, dietary polyunsaturated fatty acids are beneficial in patients with impaired heart function (50), and polyunsaturated PC species have been proposed to modulate calcium transport at the sarcoplasmic reticulum to support rapid muscle contraction and metabolic rate (24, 25). Similarities between tissues in terms of their lipid profile may suggest similarities in lipid metabolism. For example, the strikingly similar TAG profiles of liver and kidney (Fig. 3A) raise the possibility that the kidneys have a liver-like TAG metabolism, as has previously been suggested (55). Computational techniques utilizing tissue abundance profiles as “fingerprints” of function have long been used to discover functions for mRNA (49) and protein (37) species and may prove valuable for lipids as well. These techniques require large sample collections from diverse conditions to generate profiles with high information content (35, 37, 49), and the resource described here may provide an opportunity to study lipid using such approaches.

Our analysis reveals clusters that could represent functionally related lipids, which warrant further examination and validation in focused studies. For example, the observed distinct clusters of TAGs (Fig. 5) might represent specialized modes of TAG metabolism. In addition, one cluster of short, saturated TAGs that we found to be stored exclusively in adipose tissue in mice under fasting conditions (cluster e, Fig. 5) has been associated with increased risk for diabetes in humans (43). One hypothesis is that these tissue-specific TAGs

Fig. 7. Release of tissue-specific lipids with tissue injury. A: abundance of sphingomyelin (SM) 36:1 and 36:2 in brain tissue relative to plasma isolated from the jugular vein. B: abundance of SM species in jugular vein plasma. C: release of SM 36:1 and 36:2 into peripheral blood of mice 24 h after ischemic stroke (●) compared with control animals undergoing sham surgery (○), expressed as fold change over time 0. Horizontal lines denote group means.
may represent early insulin resistance in adipose tissue and impaired TAG storage, which may be further evaluated in samples from animal models and humans with diabetes. Similarly, our data may help in identifying tissues of origin and focus follow-up studies on cell types of interest for other disease-associated lipids.

Finally, we find in proof-of-concept animal experiments that tissue-specific lipids in the brain, identified in this resource, may be released into peripheral circulation with tissue injury. Such lipids could potentially be exploited for diagnostic or prognostic biomarkers of tissue injury in a range of medical conditions for a number of organ systems, similar to clinically used protein biomarkers such as cardiac troponin I for heart injury or alanine transaminase for liver injury. Although additional studies are certainly needed to evaluate the potential relevance and clinical utility of such biomarkers, lipids may hold potential as biomarkers for disease and tissue injury.

APPENDIX

Tissue isolation and lipid extraction. Adult C57BL/6j male mice 10–12 wk of age were obtained from The Jackson Laboratory. Mice were housed under alternating light-day cycles with standard mouse chow (Pico-Vac Mouse Diet 20) available ad libitum. All animal procedures and handling were performed under the guidelines of Harvard Medical School and Longwood Medical Area Institutional Animal Care and Use Committee, and the National Society for Medical Research. Twelve hours prior to tissue harvest, food was removed from cages to ensure no contamination of gut tissue lipids from incompletely digested chow. Mice were anesthetized using intraperitoneal pentobarbital, intubated, and ventilated to maintain tissue perfusion and oxygenation during harvest. All murine tissues were rapidly extracted in under 1 min, washed to remove any contaminating blood, and weighed in ice-cold saline. Tissues were placed in 19× (wt/wt) HPLC grade 100% isopropanol and rapidly snap-frozen in liquid nitrogen. Circulating blood was collected and centrifuged at 1,200 rpm for 10 min at 4°C, and the plasma fraction was collected, mixed with 19× (vol/vol/vol) HPLC grade 100% isopropanol, and snap-frozen. Tissues were disrupted in isopropanol using a mechanical homogenizer set at maximum speed while submerged in a cooling bath of ethanol and dry ice to maintain temperature of approximately –80°C. Tissue homogenates and plasma samples were vortexed for 60 s, and all samples centrifuged at 10,000 g for 10 min at 4°C. The supernatants with extracted lipid species were collected for LC-MS analysis. For each tissue, samples from at least five animals were analyzed with LC-MS.

LC-MS. LC-MS data were acquired on a 4000 QTRAP Triple Quadrupole/Linear Ion trap mass spectrometer (AB SCIEX, Foster City, CA) equipped with an HTS PAL autosampler (Leap Technologies, Carrboro, NC) and an Agilent 1200 Series binary HPLC pump (Santa Clara, CA). Four internal standards: PC 12:0/13:0, LPC 17:1, TAG 51:1 (17:0–17:1–17:0 D5), and CE 19:0 (Avanti Polar Lipids) (Santa Clara, CA) equipped with an HTS PAL autosampler (Leap Technologies, Carrboro, NC) and an Agilent 1200 Series binary HPLC pump (Santa Clara, CA) were added to all samples to monitor instrument drift. Lipid separation was achieved in samples by reverse-phase chromatography using a Prosphere HP C4 column (150× 3.0 mm, Grace) eluted isocratically at 1,200 rpm × 10 min at 4°C, and the plasma fraction was collected, mixed with 19× (vol/vol/vol) HPLC grade 100% isopropanol, and snap-frozen. Tissues were disrupted in isopropanol using a mechanical homogenizer set at maximum speed while submerged in a cooling bath of ethanol and dry ice to maintain temperature of approximately –80°C. Tissue homogenates and plasma samples were vortexed for 60 s, and all samples centrifuged at 10,000 g for 10 min at 4°C. The supernatants with extracted lipid species were collected for LC-MS analysis. For each tissue, samples from at least five animals were analyzed with LC-MS.

LC-MS data processing. Raw LC-MS data were binned into 0.1 m/z units to produce 2-D data matrices and preprocessed using a background cutoff at 10 intensity units followed by a 3 × 3 median filter to remove “speckle” noise. Data matrices were aligned against an arbitrarily chosen reference sample by computing for each m/z an optimal retention time shift, defined as the mode of the autocorrelation function, and averaging these shift values across the m/z range. For targeted analysis, the resulting data were converted to an intensity value for each targeted species by integrating the area of the peak given by the total chromatogram across 1 m/z unit, centered on the corresponding ion mass.

To determine the molecular identity of specific lipids, retention times and m/z values for 80 total lipid standards (Avanti Polar Lipids) were determined, including 57 qualitative standards and 23 quantitative Lipid MAPS MS standards (45) (see tables below). Integrated peak areas were first normalized to the internal standards LPC 17:1 and PC 12:0/13:0 to correct for instrument drift, so that the mean intensity of these two standards was equal across all samples. Note that this first normalization step does not alter relative abundance measures between lipids in a given sample.

To account for variation in ionization efficiency across lipid species with each class, a set of 23 quantitative Lipid MAPS MS standards (see tables above), covering all lipid classes except LPE and FAC (for which no quantitative standards were available), were measured at equimolar concentrations, and ionization efficiency was estimated from the resulting intensity values and used to normalize peak measures, as in previous reports (7). As we observed that, within each lipid class, ionization efficiency decreased with increasing ion mass, we fit a curve

\[ I(m) = a + bm \]

to describe the relationship between intensity I and lipid mass m. We then normalized the observed intensities of tissues lipids by dividing each value by the \( I(m) \) for that lipid. Similar correction methods have been described previously (7). In cases where only one standard was available for a lipid class (CE, LPC, and SM), \( I(m) \) was set to the corresponding constant value. For LPE and FAC, no standards were available.

Throughout, we identified lipid species as C:D, where C is the total number of carbons in all acyl chains, and D is the total number of unsaturated double bonds in all acyl chains. For example, TAG 56:4 represents a triacylglycerol with a total of 56 carbons and 4 double bonds in all three acyl chains. For SMs, the sphingoid base was included in the C:D numbers as the separation between distinct sphingoid bases was uncertain.

Cluster analysis. Pairwise distances between samples were computed as 1 minus the Pearson correlation coefficient for the lipid data expressed as fold changes, as described above. The data were projected into 2-D space by minimizing the difference between the true pairwise distances and 2-D Euclidean distances, using the SMACOF algorithm implemented in R.

Targeted data analysis. For the heat map visualization (Fig. 2A), species that did not reach a z-score over 0 of at least 2 in at least one tissue were discarded, and the remaining data were converted to fold changes by dividing data on each lipid by its mean across all samples. Within each lipid class, lipids were ordered by hierarchical clustering. For the tissue composition plots (Fig. 3), relative abundance was calculated for each lipid class by normalizing each sample to the total abundance of all lipids within that class (so that each profile sums to 1). For the lipid-lipid correlation analysis (Fig. 5), we used 1 minus the Spearman correlation coefficient of the fold change data to avoid single tissues dominating the result. Centroids were computed as grand averages of fold change data per tissue.
Untargeted analysis. z-Scores were defined for every 2-D matrix element as the difference of tissue averages $\frac{\bar{x}}{\text{H20849}}$, $\frac{\bar{y}}{\text{H6126}}$ divided by the pooled standard deviation $\frac{\sqrt{s_x}}{\text{H11005}}$, $\frac{\sqrt{s_y}}{\text{H6126}}$ where $s_x$ and $s_y$ are the variance of the sample groups.

Tissue gene expression data. Whole genome mouse tissue gene expression data were obtained from the GNF Mouse GeneAtlas V3.
stored at 80°C until lipid extraction and analysis. Samples with evidence of hemolysis were discarded. Four animals were used per condition (4 stroke animals and 4 sham animals).

ACKNOWLEDGMENTS

We thank Laura Strittmatter, Alex Soukas, Scott Vafai, Sarah Calvo, Vishal Gohil, and Monty Krieger for scientific discussions and feedback, as well as Anna Conti and Nikhil Madhusudhan for assistance with obtaining and processing of stroke samples.

Present address for M. Jain: Depts. of Medicine and Pharmacology, University of California San Diego, La Jolla, CA.

GRANTS

M. Jain was supported by NIH GRANT K08 HL-107451. R. N. was supported by grants from the Swedish Research Council (K2012-99X-22007-01-3), the Foundation for Strategic Research (ICA10-0023), and the Knut & Alice Wallenberg Foundation. This work was supported by a gift from the Nestle Research Center to the Broad Institute (V. K. M.).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS


ENDNOTE

Please note that a full dataset is available at www.nilssonlab.se/resources.php.

REFERENCES


<table>
<thead>
<tr>
<th>Lipid</th>
<th>Ion</th>
<th>Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE 19:0</td>
<td>NH4</td>
<td>666.30</td>
</tr>
<tr>
<td>C16:1,C16:1 DAG-d5</td>
<td>NH4</td>
<td>569.51</td>
</tr>
<tr>
<td>C18:2,C18:2 DAG-d5</td>
<td>NH4</td>
<td>621.54</td>
</tr>
<tr>
<td>C18:1,C18:1 DAG-d5</td>
<td>NH4</td>
<td>625.54</td>
</tr>
<tr>
<td>C18:0,C18:0 DAG-d5</td>
<td>NH4</td>
<td>629.54</td>
</tr>
<tr>
<td>C13:0 LPC</td>
<td>H</td>
<td>453.29</td>
</tr>
<tr>
<td>C17:1 LPC</td>
<td>H</td>
<td>507.33</td>
</tr>
<tr>
<td>C12:0,13:0 PC</td>
<td>H</td>
<td>635.45</td>
</tr>
<tr>
<td>C17:0,C14:1 PC</td>
<td>H</td>
<td>717.53</td>
</tr>
<tr>
<td>C17:0,C20:4 PC</td>
<td>H</td>
<td>795.58</td>
</tr>
<tr>
<td>C21:0,C22:6 PC</td>
<td>H</td>
<td>875.64</td>
</tr>
<tr>
<td>C12:0,C13:0 PC</td>
<td>H</td>
<td>593.41</td>
</tr>
<tr>
<td>C17:0,C14:1 PE</td>
<td>H</td>
<td>675.5</td>
</tr>
<tr>
<td>C17:0,C20:4 PE</td>
<td>H</td>
<td>753.53</td>
</tr>
<tr>
<td>C21:0,C22:6 PE</td>
<td>H</td>
<td>833.59</td>
</tr>
<tr>
<td>d18:1,12:0 SM</td>
<td>H</td>
<td>646.92</td>
</tr>
<tr>
<td>C14:0,C16:1,C14:0 TAG-d5</td>
<td>NH4</td>
<td>753.69</td>
</tr>
<tr>
<td>C15:0,C18:1,C15:0 TAG-d5</td>
<td>NH4</td>
<td>809.75</td>
</tr>
<tr>
<td>C19:0,C12:0,C19:0 TAG-d5</td>
<td>NH4</td>
<td>839.8</td>
</tr>
<tr>
<td>C16:0,C18:0,C16:0 TAG-d5</td>
<td>NH4</td>
<td>839.8</td>
</tr>
<tr>
<td>C17:0,C17:1,C17:0 TAG-d5</td>
<td>NH4</td>
<td>851.80</td>
</tr>
<tr>
<td>C20:2,C18:3,C20:2 TAG-d5</td>
<td>NH4</td>
<td>937.81</td>
</tr>
<tr>
<td>C20:5,C22:6,C20:5 TAG-d5</td>
<td>NH4</td>
<td>975.74</td>
</tr>
</tbody>
</table>

Fig. 9. Quantitative MS lipid standards used to assess ionization efficiency.