Sphingolipid, fatty acid and phospholipid metabolites are associated with disease severity and mTOR inhibition in lymphangioleiomyomatosis

Leonardo Bottolo[1,2,3], Suzanne Miller[4], Simon R. Johnson[4,5]

[1]Department of Medical Genetics, University of Cambridge, J. J. Thomson Avenue, Cambridge CB2 0QQ, UK
[2]The Alan Turing Institute, 96 Euston Road, London NW1 2DB, UK
[3]MRC Biostatistics Unit, University of Cambridge, Robinson Way, Cambridge CB2 0SR, UK
[4]Division of Respiratory Medicine, University of Nottingham, Nottingham Biodiscovery Institute, NIHR Biomedical Research Centre and Nottingham Molecular Pathology Node, UK
[5]National Centre for Lymphangioleiomyomatosis, Nottingham University NHS Trust. Nottingham UK

Corresponding Author:
Simon Johnson,
Room D209 Nottingham Biodiscovery Institute, Science Road, University Park, University of Nottingham. Nottingham NG7 2RD .
simon.johnson@nottingham.ac.uk

ABSTRACT

Background. Lymphangioleiomyomatosis (LAM) is a rare multisystem disease almost exclusively affecting women which causes loss of lung function, lymphatic abnormalities and angiomyolipomas. LAM occurs sporadically and in people with tuberous sclerosis complex (TSC). Loss of *TSC* gene function leads to dysregulated mTOR signalling in the affected cells. As mTOR signalling is a major determinant of lipid and nucleotide synthesis, we hypothesised that the serum metabolome would be altered in LAM and related to disease severity and activity.

Methods. Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy was used to examine the serum metabolome of 79 closely phenotyped women with LAM, including 29 receiving treatment with an mTOR inhibitor and 43 healthy control women.

Results. Sphingolipid, fatty acid and phospholipid metabolites were associated with percent predicted FEV_1 in women with LAM (e.g., behenoyl sphingomyelin adj. $p = 8.10 \times 10^{-3}$). Those with higher LAM disease-burden scores had abnormalities in fatty acid, phospholipid and lysolipids. Rate of loss of FEV_1 was associated with differences in acyl-carnitine, acyl-glycines, acyl-glutamine, fatty acids, endocanbinoids and sphingolipids (e.g., myristoleoylcarnitine adj. $p = 0.07$). In TSC-LAM, rapamycin affected modules of interrelated metabolites which comprised linoleic acid, the tricarboxylic acid cycle, aminoacyl-tRNA biosynthesis, cysteine, methionine, arginine and proline metabolism. Metabolomic pathway analysis within modules reiterated the importance of glycerophospholipid metabolites (adj. $p = 0.047$).

Conclusions. Our findings show that women with LAM have multiple abnormalities in lipid metabolism. The associations between these metabolites, multiple markers of disease activity and their potential biological roles in cell survival and signalling, suggest that lipid species may be both disease-relevant biomarkers and potential therapeutic targets for LAM.

KEY MESSAGES

What is the key question?

Lymphangioleiomyomatosis (LAM) and Tuberous sclerosis complex (TSC) are monogenic diseases resulting in mTOR activation which we predict would alter the metabolic profile of patients and confer metabolic advantages to LAM cells.

What is the bottom line?

Multiple sphingolipid, fatty acid and phospholipid metabolites were associated with disease severity, activity and, in subjects with TSC treatment with mTOR inhibitors: suggesting that lipid species may be both disease-relevant biomarkers and potential therapeutic targets for LAM.

Why read on?

This is the first study to comprehensively evaluate the serum metabolome in LAM and consistent with recent studies highlighting the role of mTOR in cellular lipid metabolism suggests deranged lipid metabolism is of potential relevance to disease pathogenesis in LAM and also highlights previously unrecognised differences between those with sporadic and TSC-LAM.

INTRODUCTION

Lymphangioleiomyomatosis (LAM) is a rare multisystem disease which almost exclusively affects women. Sporadic LAM results from sequential somatic mutations of either *TSC1*, or more commonly *TSC2*. In patients with tuberous sclerosis complex (TSC-LAM), loss of TSC gene function occurs due to a germline mutation followed by loss of heterozygosity. The TSC proteins are suppressors of the mechanistic target of rapamycin (mTOR). Loss of TSC function in an as yet unidentified precursor cell leads to dysregulated mTOR signalling and the emergence of a clone of 'LAM cells' (1). mTOR is a multiprotein cellular signalling node which assimilates inputs from growth factors and cellular nutrient status (2). Constitutive mTOR activation leads to unregulated cell proliferation, migration, invasion of the lungs and lymphatics and altered metabolism. In the lungs, LAM cells and wild type stromal cells form nodules which together result in the formation of lung cysts, recurrent pneumothorax and progressive loss of lung function (3). Involvement of the axial lymphatics can result in lymphadenopathy, fluid-filled lymphatic masses and chylous collections in the abdomen and thorax. LAM cells also are a component of angiomyolipoma, a benign tumour affecting around half of the women with LAM (4). Pharmacological suppression of mTOR stabilises lung function, reduces lymphatic complications and angiomyolipoma volume in the majority of those treated (5,6).

The rate of disease progression is highly variable. Loss of FEV_1 is perhaps the best-characterised marker of progression with a mean loss of FEV_1 around 80-140 ml/year in most populations studied (5,7). However, some patients remain stable for many years whilst others have a rapid clinical course requiring lung transplantation. LAM cells express oestrogen and progesterone receptors and proliferate and metastasise in response to oestrogen *in vitro* and *in vivo* (8,9). Consistent with this, loss of FEV_1 is accelerated by pregnancy and slows after the menopause (7,10).

mTOR signalling is a major regulator of lipid synthesis, nucleotide synthesis and glucose metabolism. By regulating the transcription factor sterol responsive element binding protein (SREBP) both via S6 kinase and through phosphorylation of Lipin1, the expression of genes involved in fatty acid and cholesterol biosynthesis are regulated. By increasing glycolytic enzyme expression, mTOR complex 1 (mTORC1) supports cellular growth by modulating the balance between oxidative phosphorylation and glycolysis, increasing the activity of the pentose phosphate pathway. mTORC1 also promotes the synthesis of purine and pyrimidine nucleotides (11,12).

We, therefore, reasoned that women with LAM would have alterations of the serum metabolome, due to the metabolic effects of mTOR activation. Moreover, these metabolomic changes would be

related to disease severity and activity. To answer these questions, we analysed the serum metabolome of a cohort of closely phenotyped women with LAM.

METHODS

SUBJECTS AND CLINICAL DATA

Women with LAM were recruited from the National Centre for LAM in Nottingham UK. All subjects had LAM defined by current ATS/JRS criteria (13). All had a clinical assessment, including drug history, clinical screening for TSC, lung function and CT scanning of the chest and abdomen to detect angiomyolipoma and lymphatic involvement. Menopausal status was recorded and subjects were considered post-menopausal if they had had a bilateral oophorectomy or were over 50 years of age and had not had a period for 24 months. None of the subjects were using systemic hormone therapy. The study was approved by the East Midlands Research Ethics Committee (13/EM/0264). Control subjects were 43 healthy women over the age of 18 with no prior history of lung disease from three metabolomic studies (supplementary table 1). The use of control samples was approved by Nottingham University Ethics Committee (approval BT A27 08 2009). All subjects gave written informed consent.

Tertiles of percent predicted FEV_1 and DL_{CO} at the time of assessment were generated with tertile 1, the highest, and tertile 3, the lowest, values. Lung function decline was calculated as the loss of FEV_1 (ΔFEV1, ml/yr) in the period preceding blood sampling. Rate of loss of lung function was based on a period of observation of greater than 1 year for all subjects.

To stratify the whole-body burden of LAM, a disease burden score between 0 and 3 was used. Subjects were assigned one point for each of (i) more severe lung disease defined by an FEV_1 or DL_{CO} of < 60% predicted, (ii) the presence of an angiomyolipoma at the time of assessment and (iii) the presence of lymphatic involvement visible on imaging (lymphadenopathy, cystic lymphatic mass or chylous effusion).

SERUM SAMPLES

Blood was collected in serum separator tubes and processed immediately from subjects without fasting. Samples were allowed to clot for 30 minutes at room temperature, centrifuged at 10,000 G for 10 minutes and stored in aliquots at −80°C until analysis. Serum VEGF-D was determined using Quantikine ELISA DVED00 (R&D Systems, Abingdon, UK).

SERUM METABOLOMICS

Serum metabolomics was performed by Metabolon© (Morrisville, NC). Samples were prepared using the MicroLab STAR® system (Hamilton Company). Proteins were precipitated with methanol and the resulting extracts analysed by two separate reverse phases (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, and one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI. Controls were analysed with the experimental samples including a pooled matrix sample generated from experimental samples, extracted water samples as process blanks and quality control standards selected not to interfere with endogenous compounds. Experimental samples were randomized across the platform run with quality control samples spaced evenly among the injections.

The analysis was performed using Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy (UPLC-MS/MS) using a Waters ACQUITY UPLC and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyser operated at 35,000 mass resolution. Sample extracts were dried and reconstituted in appropriate solvents. Reconstitution solvents contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. Four methods appropriate for different ion types and compounds were used, for all identified species including lipids. Raw data were extracted, peak-identified and quality control processed. Compounds were identified by comparison to a library of >3,300 commercially available purified standard compounds and recurrent unknown entities using the retention time/index, mass to charge ratio and chromatographic data. Further details are provided in the supplementary methods.

STATISTICAL AND BIOINFORMATIC ANALYSIS

Full details of the statistical and bioinformatics analysis are presented in the supplementary methods. Briefly, normalisation of case and control metabolite samples was performed following the workflow described by Shin et al (14) and Krumsiek et al. (15). Each raw metabolite value was rescaled to have median 1 to adjust for variation due to instrument run-day tuning differences. Missing values were imputed using the KNN-TN method (16). Differential analysis of serum metabolites was performed using Limma (17) controlling for the False Discovery Rate (FDR) (18). Differential co-expression network analysis was used to detect significant associations between the relative abundance of metabolites in differential networks, or modules, between conditions and their significance assessed through permutation tests (19). Annotation of metabolites was performed using HMDB 4.0 (20) release 2019-01-16. Finally, metabolomic pathway analysis was

performed by using MetaboAnalyst 4.0 (21) with significant differential metabolite modules detected in the differential network analysis mapped in KEGG pathways.

RESULTS

THE METABOLOME IN LAM

Seventy-nine women with LAM were studied. The subjects were typical of other reported series: having an average age of approximately 50 years at the time of study and had had LAM for an average of 13.5 years (table 1). Two-thirds had angiomyolipoma, a quarter had had pneumothorax in the past and their mean serum VEGF-D was 1,250 pg/ml. Fifteen per cent had TSC and a third were post-menopausal. Twenty-nine were being treated for LAM with rapamycin at the time of the study. Although of similar age and BMI (two-sided Wilcoxon signed-rank test $p = 0.29$ and $p = 0.92$, respectively), those treated with rapamycin tended to have lower values of FEV_1 and DL_{CO} (two-sided Wilcoxon signed-rank test $p = 3.9 \times 10^{-4}$ and $p = 4.63 \times 10^{-6}$, respectively). Further details of women with LAM and are given in supplementary tables 2 and 3.

After compound identification, quantification and data curation, 1,326 serum metabolites were identified from 122 individuals. Metabolites spanned a wide range of biochemical classes or super-pathways including 183 Amino Acid (13.8%), 24 Carbohydrate (1.8%), 21 Cofactors and Vitamins (1.6%), 10 Energy (0.8%), 471 Lipid (35.5%), 32 Nucleotide (2.4%), 79 Peptide (6%) and 115 Xenobiotics (8.7%) and a remaining group of 391 untargeted metabolites (29.5%). A heatmap of the log-transformed corrected metabolites levels with colour-coded sidebars that depict the groups considered in the analysis is presented in figure 1. Super-pathways are also highlighted and metabolites hierarchically clustered.

We examined how individual metabolites varied with the presence or absence of LAM. Using an FDR of 10%, there were no differences in the abundance of individual metabolites between women with LAM who had not received rapamycin and healthy control women (supplementary table 4). Control metabolomic profiles were quite variable. Hierarchical clustering showed that there was no segregation between control samples from the current study and companion metabolites studies (supplementary figure 1). Variation was lower in women with LAM and differences in the metabolic profile were observed when women with LAM were stratified according to menopausal status, lung function, disease burden and disease activity (table 2).

ASSOCIATION OF METABOLIC CHANGES WITH DISEASE SEVERITY AND ACTIVITY

To test the hypothesis that dysregulation of the metabolome was related to disease severity and activity, individuals with LAM were stratified according to the severity of their lung disease, the whole-body burden of LAM and rate of loss of FEV_1. As we suspected that mTOR inhibition would affect the metabolome, we restricted this analysis to the 50 subjects who had not received rapamycin. Subjects were grouped into tertiles of FEV_1 and DL_{CO} (tertile 1 highest values, tertile 3 lowest, see supplementary table 3). Comparing tertiles 1 and 2 applying a 10% FDR, reductions were seen in sphingolipids (behenoyl sphingomyelin (adj. p = 8.10×10^{-3}), lignoceroyl sphingomyelin (adj. p = 8.10×10^{-3}), lactosyl-N-behenoyl-sphingosine (adj. p = 3.06×10^{-2}), tricosanoyl sphingomyelin (adj. p = 6.23×10^{-2}) and palmitoyl dihydrosphingomyelin (adj. p = 9.75×10^{-2})), fatty acids (eicosenoylcarnitine (adj. p = 8.68×10^{-2})), phospholipids (1-palmityl-2-stearoyl-GPC (adj. p = 8.10×10^{-3})) and pantothenate (adj. pantothenate (adj. p =9.75×10^{-2})), see supplementary figure 2. Full details are presented in supplementary table 5. There were no significant associations between metabolites and DL_{CO} (table 2).

To examine the relationship between the metabolome and the whole body burden of LAM, we then stratified subjects by a score that reflected the presence of disease in the lungs, lymphatic tissue and angiomyolipomas (score 0 lowest disease burden, score 3 highest disease burden, see supplementary table 3). Those with the highest disease-burden score had abnormalities in four fatty acid, four phospholipid and one lysolipid metabolite (supplementary table 5). Stearoylcholine, lineloylcholine, palmitoylcholine, oleylcholine, docosahexaenoylcholine, arachadonycholine, eicosapentaenoylcholine, dihomo-linolenoyl-choline and 2-stearoyl-GPI (18:0) differed in abundance across levels of disease burden, tending to be lower in more extensive disease (figure 2 and supplementary figure 3). Although the same pathways were associated with both lung function and disease burden, the individual metabolites within these pathways differed (supplementary table 5).

Rate of loss of FEV_1 was used as a measure of disease activity. Subjects were divided into tertiles, with 1 being the lowest (best) rate of FEV_1 loss. Changes in abundance of 22 metabolites, at an FDR of 10% were observed between those with the lowest ΔFEV_1 in tertile 1 and those with the fastest ΔFEV_1 in tertile 3 (figure 3). Again, metabolites associated with ΔFEV_1 were almost all lipid-based comprising eight acyl-carnitine, one acyl-glycine, one acyl-glutamine, five fatty acid, three endocanbinoid and two sphingolipids (table 3 and supplementary table 5).

METABOLIC CHANGES POST-MENOPAUSE

Several studies have reported that the mean rate of loss in FEV_1 is slower in post-menopausal than pre-menopausal women with LAM (7,10). We, therefore, compared the metabolome of the 33

pre- and 17 post-menopausal women who had not been treated with rapamycin. In post-menopausal women, as expected, 18 sex steroid metabolites were altered (supplementary table 4 and supplementary figure 5). Metabolomic pathway analysis identified that menopausal status was associated with significant enrichment of the steroid hormone biosynthesis pathway (supplementary figure 6). Also, 14 acyl-carnitine, one acyl-glycine, one acyl-glutamine 12 phospholipids, five fatty acids, five diacylglycerols and two secondary bile acid metabolites were altered (supplementary table 4).

EFFECT OF RAPAMYCIN ON THE SERUM METABOLOME

Women with sporadic LAM treated with rapamycin had similar metabolomic profiles to those not treated with rapamycin. In contrast, in those with TSC-LAM, rapamycin affected a network of interrelated metabolites whose significance was assessed by empirical p-value < 0.1. Although no individual metabolites significantly differed in abundance (supplementary figure 4), changes in related groups of metabolites were identified as differential co-expression networks or 'modules'. The modules, arbitrarily described by colours, are described in detail in supplementary table 6. These modules were not detected if a co-expression network analysis was performed separately on TSC-LAM samples with and without rapamycin treatment (supplementary figure 7). The relationship between modules, suggests these metabolites are either interdependent or affected in a similar manner by mTOR inhibition (figure 4). Individual modules (blue, tan and magenta) were also significantly negatively associated with FEV_1, while ΔFEV_1, a measure of disease activity, was significantly positively correlated with the blue module. These associations were not seen in subjects treated with rapamycin (figure 5B), suggesting they are downstream of mTORC1. Metabolomic pathway analysis of the individual modules again shows a significant contribution from glycerophospholipid metabolites that are altered by treatment with rapamycin (figure 5A).

DISCUSSION

This study is the first to comprehensively examine the serum metabolome in a large cohort of women with LAM. Sphingolipid and fatty acid metabolism were associated with the severity of lung disease, total body burden of LAM and disease activity defined by the rate of loss of FEV_1. Comparison of mTOR inhibitor treated and untreated women showed changes in metabolite networks incorporating glycerophospholipid metabolites in subjects with TSC-LAM but not sporadic LAM. These findings are likely to reflect the altered metabolism of the LAM cell with metabolites related to disease activity and severity.

We observed that FEV_1, a measure of airflow limitation, was related to sphingolipid and fatty acid metabolism, although this was not a simple relationship with the greatest abnormalities seen in

those with moderate impairment (tertile 2). Airflow limitation in LAM is thought to be due to loss of elastic recoil resulting from parenchymal destruction and these metabolomic alterations may reflect the amount of parenchymal involvement. However, as we observed that the same metabolomic pathways were also related to the rate of loss of FEV_1, a measure of disease activity, it is also possible that FEV_1 tertile 2 contained those with the most active disease, both in terms of disease progression and mTOR driven metabolic changes. Subjects with better-preserved FEV_1 (tertile 1) may have less active disease and those with the most advanced disease (tertile 3) may have reached a 'burned out', less metabolically active stage. This idea is supported by the strong linear association between acyl-carnitine fatty acid and sphingolipid metabolism with the rate of FEV_1 loss, suggesting changes in the lipidome are associated with both disease extent and activity.

We reasoned that at least some of the metabolic changes observed would be related to the number of LAM cells in the body. The disease's clinical phenotype, and hence LAM cell burden, is variable with some having extensive lymphatic involvement or large angiomyolipomas and others having only lung disease. It is not currently possible to accurately determine the 'whole-body burden' of LAM cells, particularly as some organ involvement is likely to be sub-clinical (22). We therefore, used a simple score comprising mild or more advanced lung disease, the presence of angiomyolipoma and lymphatic involvement visible by clinical imaging. As we predicted, this 'disease burden' score was associated with the degree of metabolic abnormality in acyl-choline fatty acid metabolism and phospholipids.

The observations that LAM is one of the most female-specific diseases described, that rate of loss of FEV_1 is lower in postmenopausal women (7,10), can be exacerbated by pregnancy (23,24) and occurs earlier in those exposed to oestrogen containing contraceptives (25), suggest that LAM is likely to be oestrogen driven and we had expected the metabolome to reflect increased disease activity in pre-menopausal women. As expected, there were widespread changes in sex steroid metabolism between pre- and post-menopausal women with LAM. We also observed changes in phospholipid and acyl-carnitine fatty acid metabolism. Whilst it is tempting to speculate this change is related to reduced disease activity as a result of lower oestrogen levels, the menopause in healthy women is also associated with changes in acyl-carnitine, fatty acid and lysophosphatidylcholine metabolism (26) and we are unable to relate the changes observed here to a specific set of processes.

The molecular defect in the TSC proteins and the impressive reduction in lung function decline during mTOR inhibitor therapy suggests that mTOR activation is the key driver of LAM and TSC (1,5). In our study, the metabolic abnormalities associated with disease extent and activity are

almost exclusively comprised of lipids, particularly sphingolipids, phospholipids and acyl-carnitine fatty acids, reflecting metabolic processes downstream of mTOR. As an investigation in a monogenic disease characterised by mTOR dysregulation; our study provides a translational link between basic mTOR biology and disease. The changes observed represent a metabolic signature of increased fatty acid synthesis and uptake which is also a characteristic of cancer cells (27,28). Glycerophospholipid and sphingolipid metabolites are also associated with airflow obstruction in patients with COPD, whilst mTOR can be activated to some degree in COPD it is also possible these changes are a response to tissue damage (29). mTORC1 and 2 stimulate expression and processing of SREBP1 and suppress its negative regulator LIPIN1. SREBP1 is a transcription factor responsible for the expression of the fatty acid synthases, ATP citrate lyase, fatty acid synthase, acetyl-CoA carboxylase 1 and stearoyl-CoA desaturase. In addition, mTORC1 further stabilises these gene transcripts to enhance fatty acid synthesis and enhances fatty acid uptake via CD38. Lipidomic profiling of TSC2 null murine fibroblasts has similarly shown an increase in multiple lipid species including phosphatidylcholine, lysophosphatidylcholines, phospholipids, sphingomyelins and acylglycerols (30). Inhibition of mTOR signalling with rapamycin in TSC2 null murine fibroblasts did not suppress the overproduction of most lipid species, whereas downregulation of SREBP1 did, perhaps as SREBP1 is known to be only partially rapamycin sensitive (31). These observations are consistent with our findings showing that serum lipid abnormalities were only affected by rapamycin therapy in subjects with TSC where although groups of metabolites, or modules, were altered by rapamycin therapy, only glycerophospholids were clearly differently abundant in those receiving rapamycin. Also consistent with mTOR activity governing the metabolome in LAM was the observation that the associations between metabolite networks with lung function and disease activity were not seen in subjects who were being treated with rapamycin.

The LAM cell, analogous to a cancer cell, is likely to use altered lipid metabolism to support cell membrane synthesis: as constituents of membrane lipid rafts, saturated and monounsaturated phospholipids, potentially protect cancer cells from oxidative damage by reducing lipid peroxidation (32). As second messengers, lipids, including sphingosine 1-phosphate, promote cell proliferation and migration and in turn feedback to activate mTOR signalling (28). Depletion of SREBP1 and 2 reduces monounsaturated fatty acids, resulting in mitochondrial dysfunction, the accumulation of reactive oxygen species and endoplasmic reticulum stress (33). Inhibitors of various lipid synthetic enzymes are currently in clinical trials for cancer and may also have therapeutic potential in conjunction with mTOR inhibitors for LAM and TSC.

The metabolomic platform used identifies metabolites based upon a combination of three MS/MS criteria referenced against library data of over 3,300 purified standards. Whilst no targeted assays are performed the approach has been extensively applied and validated (34). We were surprised to observe that after statistical corrections, healthy women and those with LAM did not have different metabolic profiles. The healthy control samples were variable, with no apparent clustering dependent upon the source of the control sample. In addition, within women with LAM, significant metabolic changes in response to rapamycin were only observed in patients with TSC-LAM. Whilst TSC and sporadic LAM have the same underlying molecular abnormality (35), those with TSC-LAM may have larger changes in lipid metabolism, due to other TSC related tumours and perhaps widespread haplo-insufficiency for TSC.

LAM is a rare disease with a variable clinical phenotype and despite this being the largest metabolic analysis in LAM to date, the relatively small sample size reduces the statistical power to detect differential metabolites which may account for some of the unexpected results we have observed: in particular between healthy women and those with LAM not treated with rapamycin. Possible approaches would be to define a more liberal level of significance or FDR although this may increase false positive observations. Alternatively, pooling ‘similar’ metabolites (36), such as those that belong to the same KEGG pathway, could increase power.

In conclusion, we have performed the largest comprehensive metabolomic study in women with LAM to date and have demonstrated multiple abnormalities in lipid metabolism, particularly, sphingo- and phospho-lipids, and the fatty acids acyl-carnitine, acyl-choline and acyl-glycine. These changes are poorly responsive to rapamycin, the only proven effective therapy for LAM and are likely to contribute to LAM cell growth, survival and resulting lung damage. The associations between these metabolites, multiple markers of disease activity and their potential roles in cell survival and signalling, suggest that lipid species may be both disease-relevant biomarkers and potential therapeutic targets.

ACKNOWLEDGEMENTS

The authors are grateful to women with LAM who contributed to the study and to Paul Kirk and Angelos Alexopoulos for suggestions and insightful comments regarding the KNN-TN software.

FUNDING

Alan Turing Institute under the Engineering and Physical Sciences Research Council (EP/N510129/1 to L.B.). NIHR RD-TRC and MRC/EPSRC Nottingham Molecular Pathology Node (MR/N005953/1)

REFERENCES

1. Henske EP, McCormack FX. Lymphangioleiomyomatosis - A wolf in sheep's clothing. Vol. 122, Journal of Clinical Investigation. 2012. p. 3807–16.
2. Saxton RA, Sabatini DM. mTOR signaling in growth, metabolism, and disease. Vol. 168, Cell. 2017. p. 960–76.
3. Clements D, Dongre A, Krymskaya VP, Johnson SR. Wild type mesenchymal cells contribute to the lung pathology of lymphangioleiomyomatosis. PLoS One. 2015;10(5).
4. Johnson SR, Taveira-DaSilva AM, Moss J. Lymphangioleiomyomatosis. Vol. 37, Clinics in Chest Medicine. 2016. p. 389–403.
5. McCormack FX, Inoue Y, Moss J, Singer LG, Strange C, Nakata K, et al. Efficacy and safety of sirolimus in lymphangioleiomyomatosis. N Engl J Med. 2011;364(17):1595–606.
6. Bee J, Fuller S, Miller S, Johnson SR. Lung function response and side effects to rapamycin for lymphangioleiomyomatosis: A prospective national cohort study. Thorax. 2018;73(4):369–75.
7. Johnson SR, Tattersfield AE. Decline in lung function in lymphangioleiomyomatosis: Relation to menopause and progesterone treatment. Am J Respir Crit Care Med. 1999;160(2):628–33.
8. Yu JJ, Robb VA, Morrison TA, Ariazi EA, Karbowniczek M, Astrinidis A, et al. Estrogen promotes the survival and pulmonary metastasis of tuberin-null cells. Proc Natl Acad Sci U S A. 2009;106(8):2635–40.
9. Clements D, Asprey SL, McCulloch TA, Morris TA, Watson SA, Johnson SR. Analysis of the oestrogen response in an angiomyolipoma derived xenograft model. Endocr Relat Cancer. 2009;16(1):59–72.
10. Gupta N, Lee HS, Ryu JH, Taveira-DaSilva AM, Beck GJ, Lee JC, et al. The NHLBI LAM registry: Prognostic physiologic and radiologic biomarkers emerge from a 15-Year prospective longitudinal analysis. Chest. 2019;155(2):288–96.
11. Ben-Sahra I, Hoxhaj G, Ricoult SJH, Asara JM, Manning BD. mTORC1 induces purine synthesis through control of the mitochondrial tetrahydrofolate cycle. Science (80-). 2016;351(6274):728–33.
12. Robitaille AM, Christen S, Shimobayashi M, Cornu M, Fava LL, Moes S, et al. Quantitative phosphoproteomics reveal mTORC1 activates de novo pyrimidine synthesis. Science (80-). 2013;339(6125):1320–3.
13. Gupta N, Finlay GA, Kotloff RM, Strange C, Wilson KC, Young LR, et al. Lymphangioleiomyomatosis diagnosis and management: High-resolution chest computed tomography, transbronchial lung biopsy, and pleural disease management an Official American thoracic society/Japanese respiratory society clinical practice guideline. Am J

Respir Crit Care Med. 2017;196(10):1337–48.
14. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. Nat Genet. 2014;46(6):543–50.
15. Krumsiek J, Mittelstrass K, Do KT, Stückler F, Ried J, Adamski J, et al. Gender-specific pathway differences in the human serum metabolome. Metabolomics. 2015;11(6):1815–33.
16. Shah JS, Rai SN, DeFilippis AP, Hill BG, Bhatnagar A, Brock GN. Distribution based nearest neighbor imputation for truncated high dimensional data with applications to pre-clinical and clinical metabolomics studies. BMC Bioinformatics. 2017;18(1).
17. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
18. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. J R Stat Soc Ser B. 1995;57(1):289–300.
19. Tesson BM, Breitling R, Jansen RC. DiffCoEx: A simple and sensitive method to find differentially coexpressed gene modules. BMC Bioinformatics. 2010;11.
20. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, et al. HMDB 4.0: The human metabolome database for 2018. Nucleic Acids Res. 2018;46(D1):D608–17.
21. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, et al. MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. Nucleic Acids Res. 2018;46(W1):W486–94.
22. Hayashi T, Kumasaka T, Mitani K, Terao Y, Watanabe M, Oide T, et al. Prevalence of uterine and adnexal involvement in pulmonary lymphangioleiomyomatosis: A clinicopathologic study of 10 patients. Am J Surg Pathol. 2011;35(12):1776–85.
23. Johnson SR, Tattersfield AE. Clinical experience of lymphangioleiomyomatosis in the UK. Thorax. 2000;55(12):1052–7.
24. Cohen MM, Freyer AM, Johnson SR. Pregnancy experiences among women with lymphangioleiomyomatosis. Respir Med. 2009;103(5):766–72.
25. Oberstein EM, Fleming LE, Gómez-Marin O, Glassberg MK. Pulmonary lymphagioleiomyomatosis (LAM): Examining oral contraceptive pills and the onset of disease. J Women's Heal. 2003;12(1):81–5.
26. Ke C, Hou Y, Zhang H, Yang K, Wang J, Guo B, et al. Plasma metabolic profiles in women are menopause dependent. PLoS One. 2015;10(11).
27. Kim J, Guan KL. mTOR as a central hub of nutrient signalling and cell growth. Vol. 21, Nature Cell Biology. 2019. p. 63–71.
28. Mossmann D, Park S, Hall MN. mTOR signalling and cellular metabolism are mutual determinants in cancer. Vol. 18, Nature Reviews Cancer. 2018. p. 744–57.

29. Bowler RP, Jacobson S, Cruickshank C, Hughes GJ, Siska C, Ory DS, et al. Plasma sphingolipids associated with chronic obstructive pulmonary disease phenotypes. Am J Respir Crit Care Med. 2015;191(3):275–84.
30. Priolo C, Ricoult SJH, Khabibullin D, Filippakis H, Yu J, Manning BD, et al. Tuberous sclerosis complex 2 loss increases lysophosphatidylcholine synthesis in lymphangioleiomyomatosis. Am J Respir Cell Mol Biol. 2015;53(1):33–41.
31. Peterson TR, Sengupta SS, Harris TE, Carmack AE, Kang SA, Balderas E, et al. mTOR complex 1 regulates lipin 1 localization to control the SREBP pathway. Cell. 2011;146(3):408–20.
32. Rysman E, Brusselmans K, Scheys K, Timmermans L, Derua R, Munck S, et al. De novo lipogenesis protects cancer cells from free radicals and chemotherapeutics by promoting membrane lipid saturation. Cancer Res. 2010;70(20):8117–26.
33. Griffiths B, Lewis CA, Bensaad K, Ros S, Zhang Q, Ferber EC, et al. Sterol regulatory element binding protein-dependent regulation of lipid synthesis supports cell survival and tumor growth. Cancer Metab. 2013;1(1):3.
34. Thaiss CA, Itav S, Rothschild D, Meijer MT, Levy M, Moresi C, et al. Persistent microbiome alterations modulate the rate of post-dieting weight regain. Nature. 2016;540(7634):544–51.
35. Badri KR, Gao L, Hyjek E, Schuger N, Schuger L, Qin W, et al. Exonic mutations of TSC2/TSC1 are common but not seen in all sporadic pulmonary lymphangioleiomyomatosis. Vol. 187, American Journal of Respiratory and Critical Care Medicine. 2013. p. 663–5.
36. Stingo FC, Chen YA, Tadesse MG, Vannucci M. Incorporating biological information into linear models: A Bayesian approach to the selection of pathways and genes. Ann Appl Stat. 2011;5(3):1978–2002.

TABLES

Table 1. Patient characteristics.

	All subjects	No rapamycin treatment	Rapamycin treated
Sample size	79	50	29
Age (yrs)	49.6 (11.1)	50.6 (10.9)	48.0 (11.5)
BMI (kg/m^2)	25.6 (6.1)	25.6 (6.2)	25.5 (6.1)
Disease duration (yrs)	13.5 (11.0)	13.9 (11.1)	12.8 (11.0)
Post menopause (%)	33	22	11
Phenotype			
TSC (%)	15	14	17
Angiomyolipoma (%)	65	72	52
Lymphatic involvement (%)	18	16	21
Pneumothorax (%)	25	30	17
Serum VEGF-D (pg/ml)	1,250 (1,201)	1,287 (1,165)	1,182 (1,284)
Lung function			
FEV_1 (% predicted)	62 (21)	69 (21)	50 (17)
DL_{CO} (% predicted)	55 (16)	60 (16)	43 (11)
pre-treatment ΔFEV_1 (ml/yr)	-77 (168)	-54 (91)	-113 (241)

Group means (standard deviation) of subjects stratified by treatment with rapamycin. BMI = Body mass index; TSC = tuberous sclerosis complex; VEGF-D = vascular endothelial growth factor-D; FEV_1 = forced expiratory volume in one second (ml); DL_{CO} = diffusing capacity of the lung for carbon monoxide (ml/min/kPa); ΔFEV_1 = annual decline in forced expiratory volume (ml/yr).

Table 2. Number of differential metabolites identified.

Groups	Age covariate removed		Age covariate not removed	
	< 0.10% FDR	< 0.10% FDR & $\|\log_2 FC\| > 0.4$	< 0.10% FDR	< 0.10% FDR & $\|\log_2 FC\| > 0.4$
LAM untreated vs healthy women	0	0	-	-
LAM pre-menopause untreated vs LAM post-menopause untreated	-	-	68	40
mTOR inhibition				
TSC-LAM untreated vs TSC-LAM rapamycin treated	0	0	-	-
Sporadic LAM untreated vs sporadic LAM rapamycin treated	0	0	-	-
TSC-LAM untreated vs sporadic LAM untreated	1	0	-	-
TSC-LAM rapamycin treated vs sporadic LAM rapamycin treated	9	2	-	-
FEV_1				
T1 vs T2	11	0	-	-
T2 vs T3	0	0	-	-
T1 vs T3	0	0	-	-
T1 vs T2 & T3	0	0	-	-
DL_{CO}				
T1 vs T2	0	0	-	-
T2 vs T3	0	0	-	-
T1 vs T3	0	0	-	-
T1 vs T2 & T3	0	0	-	-
Disease burden				
S0 vs S1	1	0	-	-
S1 vs S2	0	0	-	-
S2 vs S3	32	20	-	-
S0 vs S1, S2 & S3	10	9	-	-
Disease activity				
ΔFEV_1 T1 vs T2	0	0	-	-
ΔFEV_1 T2 vs T3	3	2	-	-
ΔFEV_1 T1 vs T3	22	2	-	-
ΔFEV_1 T1 vs T2 & T3	3	0	-	-

Number of differential metabolites at 10% FDR threshold and at 10% FDR threshold with $|\log_2$ Fold Change (FC)| > 0.4 after removing the effect of sex, BMI, ethnicity, study covariate (only in 'LAM

untreated vs healthy women' contrast) and experimental run day and excluding metabolites of unknown identity and the 'Xenobiotics' super-pathway. TSC = tuberous sclerosis complex; VEGF-D = vascular endothelial growth factor-D (pg/ml); FEV_1 = forced expiratory volume in one second (ml); ΔFEV_1 = annual decline in forced expiratory volume (ml/yr). OBD = overall disease burden score; T = tertile, 1 highest, 3 lowest. For group definitions, see supplementary table 3. Full results are presented in supplementary tables 4 and 5.

Table 3. Lipid mediators associated with rate of loss of FEV_1.

Biochemical name	Sub-pathway	ΔFEV_1 T1 vs T2		ΔFEV_1 T2 vs T3		ΔFEV_1 T1 vs T3		ΔFEV_1 T1 vs T2 & T3	
		FC	Adj. p	FC	Adj. p	FC	Adj. p	FC	Adj. P
myristoleoylcarnitine (C14:1)*	Acyl carnitine	-0.107	0.998	-0.338	0.165	-0.445	0.070	-0.377	0.812
laurylcarnitine (C12)	Acyl carnitine	-0.149	0.998	-0.266	0.256	-0.415	0.070	-0.376	0.812
palmitoleoylcarnitine (C16:1)*	Acyl carnitine	-0.068	0.998	-0.243	0.160	-0.311	0.070	-0.253	0.812
oleoyl ethanolamide	Endocanbinoid	-0.022	0.998	-0.191	0.154	-0.213	0.070	-0.152	0.859
N-palmitoylglycine	Acyl glycine	-0.017	0.998	-0.164	0.154	-0.182	0.070	-0.107	0.859
palmitoyl ethanolamide	Endocanbinoid	-0.019	0.998	-0.151	0.154	-0.170	0.070	-0.103	0.859
sphingomyelin (d18:2/24:1, d18:1/24:2)*	Sphingolipid	0.018	0.998	-0.153	0.047	-0.135	0.070	-0.080	0.859
ximenoylcarnitine (C26:1)*	Acyl carnitine	-0.148	0.968	-0.056	0.705	-0.204	0.073	-0.294	0.039
adrenate (22:4n6)	Fatty acid	-0.036	0.998	-0.295	0.160	-0.331	0.078	-0.229	0.859
dihomolinoleate (20:2n6)	Fatty acid	-0.028	0.998	-0.293	0.160	-0.321	0.078	-0.220	0.859
docosapentaenoate (n6 DPA; 22:5n6)	Fatty acid	-0.006	0.998	-0.246	0.154	-0.252	0.078	-0.192	0.859
dihomolinolenate (20:3n3 or 3n6)	Fatty acid	-0.024	0.998	-0.177	0.160	-0.201	0.078	-0.153	0.859
oleoylcarnitine (C18)	Acyl carnitine	-0.076	0.998	-0.101	0.371	-0.176	0.078	-0.178	0.751
sphingomyelin (d18:1/24:1, d18:2/24:0)*	Sphingolipid	-0.034	0.998	-0.068	0.275	-0.102	0.078	-0.094	0.858
allantoin	Purine	0.190	0.263	-0.011	0.943	0.179	0.078	0.308	0.006
hexanoylglutamine	Acyl glutamine	-0.056	0.998	-0.328	0.175	-0.384	0.082	-0.226	0.859
suberoylcarnitine (C8-DC)	Acyl carnitine	-0.038	0.998	-0.323	0.160	-0.361	0.082	-0.206	0.859
docosadienoate (22:2n6)	Fatty acid	-0.028	0.998	-0.245	0.160	-0.272	0.082	-0.173	0.859
linoleoyl ethanolamide	Endocanbinoid	-0.031	0.998	-0.165	0.175	-0.196	0.082	-0.142	0.859
palmitoylcarnitine (C16)	Acyl carnitine	-0.062	0.998	-0.092	0.371	-0.154	0.085	-0.133	0.859
gamma-glutamylalanine	Gamma-glutamyl amino acid	0.071	0.998	0.106	0.371	0.178	0.085	0.121	0.859
eicosenoylcarnitine (C20:1)*	Acyl carnitine	-0.097	0.998	-0.068	0.565	-0.164	0.097	-0.179	0.751

Lipid mediators associated with rate of change of FEV_1 (ΔFEV_1) (ml/yr) at 10% FDR. Metabolites are ranked by significance and $-\log_2$ Fold Change comparing the largest difference, Tertile 1 (T1, lowest (best) ΔFEV_1) with T3 (greatest (worst) ΔFEV_1). Adj. p = adjusted p-value, FC = $-\log_2$ Fold Change. Red and blue cells indicate metabolites that are significantly different between conditions at 10% FDR with $|\log_2 FC| > 0.4$ and metabolites significant at 10% FDR with $|\log_2 FC| < 0.4$, respectively. *Compounds identified without confirmation by a standard. Volcano plot of ΔFEV_1 T1 vs T3 is shown in figure 3. Full results are presented in supplementary table 5.

FIGURE LEGENDS

Figure 1. Heatmap of log-transformed metabolite levels after correcting for confounding factors. Hierarchically clustered metabolites (horizontal) stratified by super-pathway and participants of the study (vertical) stratified by healthy controls and women with LAM pre- and post-menopause with and without rapamycin treatment (greyscale colour code); patients with TSC-LAM and sporadic LAM with and without rapamycin treatment (red scale); FEV_1 = forced expiratory volume in one second (ml) (blue scale); DL_{CO} = diffusing capacity of the lung for carbon monoxide (ml/min/kPa) (green scale); overall disease burden (orange scale); disease activity based on ΔFEV_1 = annual decline in forced expiratory volume (ml/yr); T = tertile, 1 highest, 3 lowest. S = score, 0 lowest, 3 highest. For group definitions, see supplementary table 3.

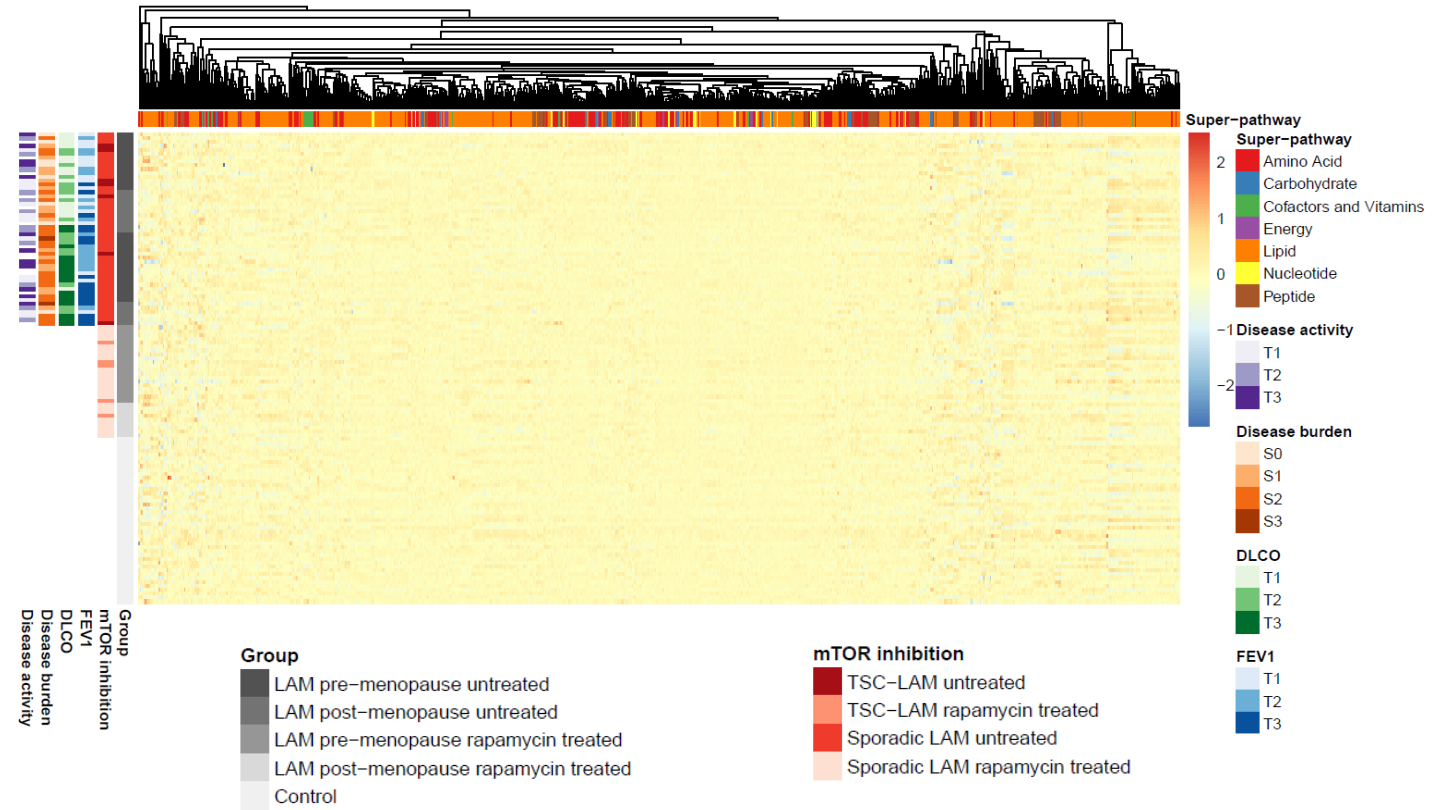
Figure 2. Volcano plot of fatty acids and phospholipids across overall disease burden. Significantly different metabolites at 10% False Discovery Rate (FDR) (*y*-axis) with $|\log_2$ Fold Change (FC)| > 0.4 (*x*-axis) in patients not treated with rapamycin between overall disease burden score S0 (lowest, n = 8) and scores S1 (n = 17), S2 (n = 23) and S3 (highest, n = 2) (for group definitions, see supplementary table 3). Red, blue, green and grey dots indicate metabolites that are significantly different between conditions at 10% FDR with $|\log_2$ FC| > 0.4, metabolites significant at 10% FDR with $|\log_2$ FC| < 0.4, non-significant at 10% FDR with $|\log_2$ FC| > 0.4 and overall NS = non-significant, respectively. *Compounds identified without confirmation by a standard. Full results are presented in supplementary table 5.

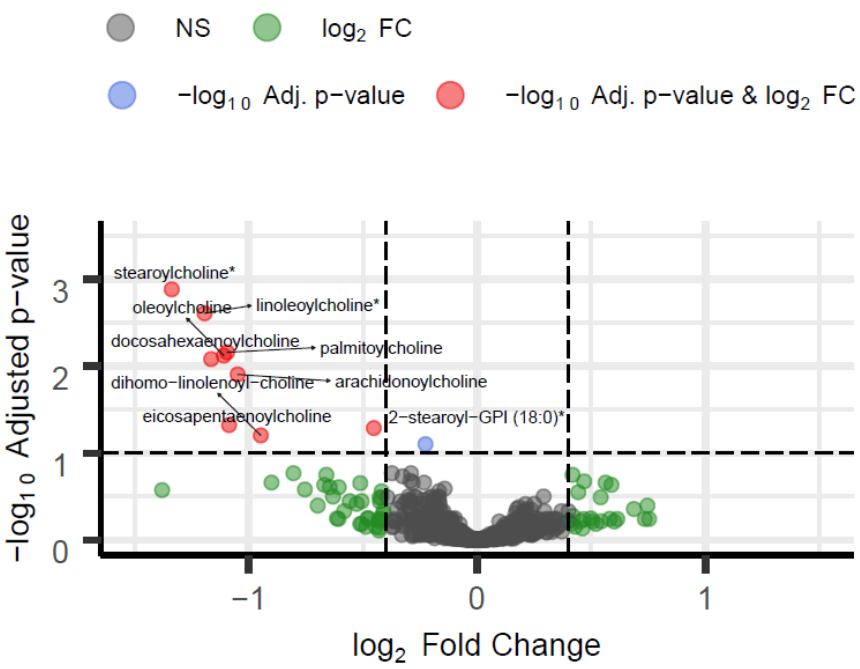
Figure 3. Volcano plot of metabolic changes in patients with rate loss of FEV_1. Significantly different metabolites at 10% False Discovery Rate (FDR) (*y*-axis) with $|\log_2$ Fold Change (FC)| > 0.4 (*x*-axis) in patients not treated with rapamycin with $\Delta FEV1_1$ T1 (n = 13) and ΔFEV_1 T3 (n = 12) (for group definitions, see supplementary table 3). Red, blue and grey dots indicate metabolites that are significantly different between conditions at 10% FDR with $|\log_2$ FC| > 0.4, metabolites significant at 10% FDR with $|\log_2$ FC| < 0.4 and overall NS = non-significant, respectively. *Compounds identified without confirmation by a standard. Full results are provided in supplementary table 5.

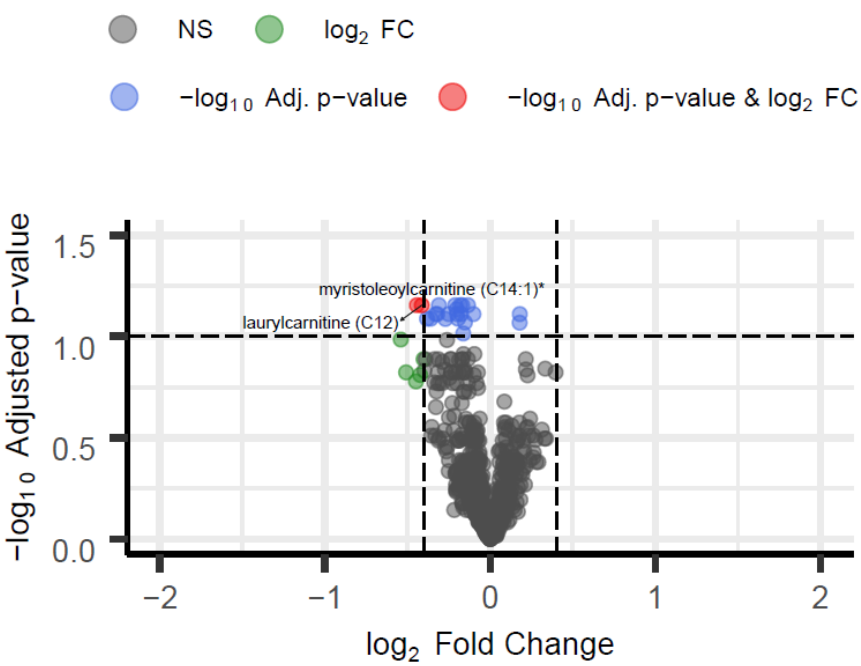
Figure 4. Differential co-expression network analysis after mTOR inhibitor treatment. (**A**) Dendrogram showing clustering of differential co-expressed metabolites (modules) based on TSC-LAM samples without (n = 5) and with (n = 7) mTOR inhibitor treatment (for group definitions, see supplementary table 3). Top bar: metabolites with $|\log_2$ FC| > 0.4 (green); bottom colour bar:

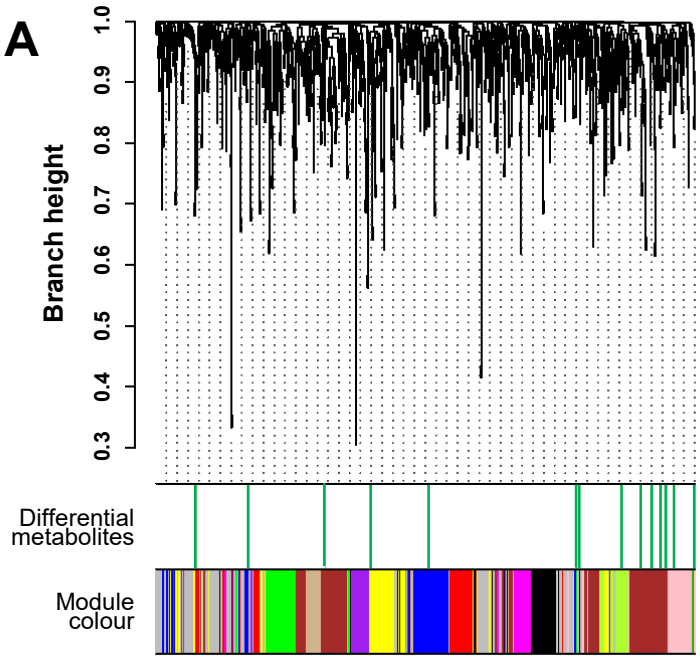
colours identify 15 modules generated by unsupervised hierarchical clustering. (**B**) Comparative correlation heat map (red corresponds to positive correlations, blue to negative correlations). The upper diagonal matrix shows the correlation between pairs of metabolites in the TSC-LAM rapamycin group while the lower diagonal matrix shows the correlation between pairs of metabolites in the TSC-LAM untreated group. Modules are identified in the heatmap by colour-coded squares and on the top by a colour bar. (**C**) Significance of the identified differential co-expressed metabolites and covariation between modules (empirical p-values inside each modules pair). Modules are identified on the top by a colour bar with number of metabolites in each module.

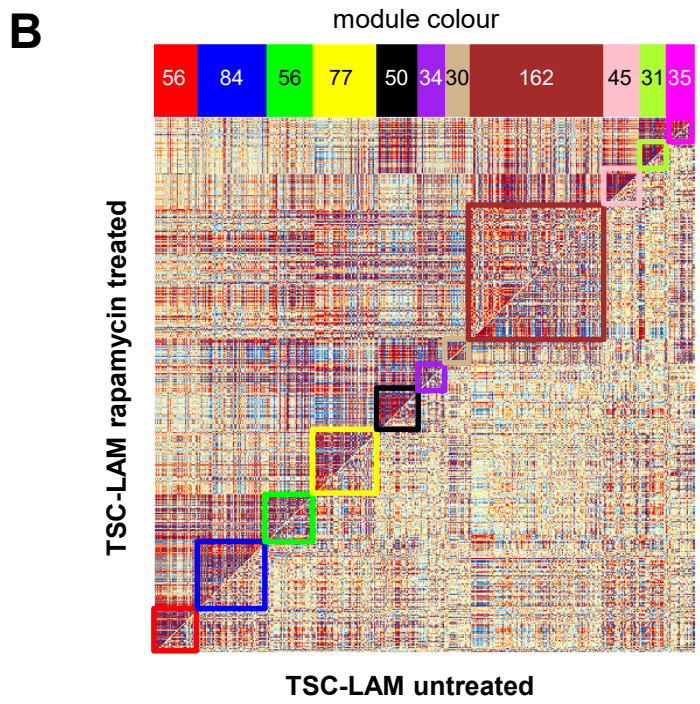
Figure 5. Pathway analysis of modules affected by mTOR inhibition. Differential co-expressed metabolites (modules) pathway analysis and correlation with clinical phenotypes in the TSC-LAM samples without (n = 5) and with (n = 7) rapamycin treatment (for group definitions, see supplementary table 3). (**A**) Scatterplots depicting $-\log_{10}$ p-values of the metabolomic pathway analysis (*y*-axis) and impact values from topology pathway analysis (*x*-axis). For each significant module (empirical p-value < 0.1) with at least one significant pathway, metabolomic pathways detected at 10% Holm–Bonferroni correction are highlighted (full results are presented in supplementary table 6). Modules are identified by the colour bar, dot sizes are proportional to the number of altered metabolites within the pathway. (**B**) Table of correlations and p-values for studying the relationships between modules 'eigen-metabolites' (rows) and selected clinical traits (columns) in TSC-LAM samples without (left) and with (right) rapamycin treatment. Significant Spearman correlations at 10% nominal level of significance and corresponding p-values are reported inside the tables.

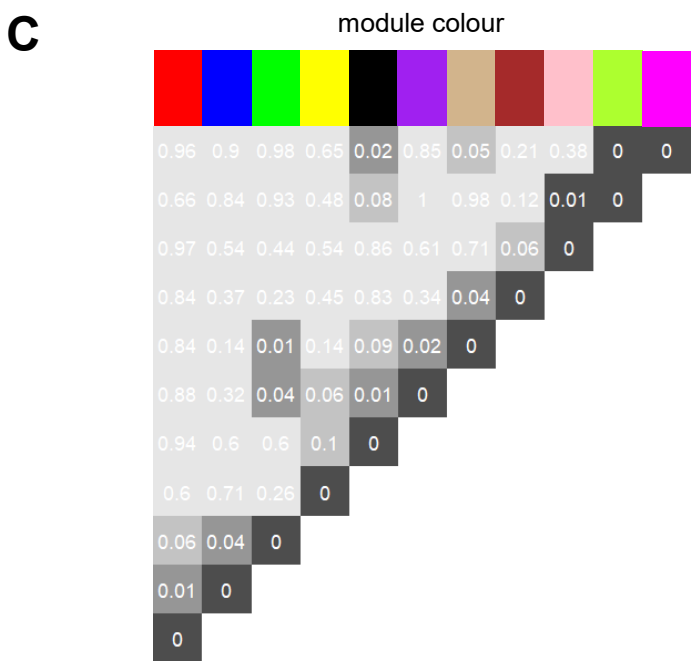

Super−pathway
Super−pathway
Amino Acid
Carbohydrate
Cofactors and Vitamins
Energy
Lipid
Nucleotide
Peptide
2
1
0
−1
−2
Disease activity
T1
T2
T3
Disease burden
S0
S1
S2
S3
DLCO
T1
T2
T3
FEV1
T1
T2
T3
Disease activity
Disease burden
DLCO
FEV1
mTOR inhibition
Group
Group
LAM pre−menopause untreated
LAM post−menopause untreated
LAM pre−menopause rapamycin treated
LAM post−menopause rapamycin treated
Control
mTOR inhibition
TSC−LAM untreated
TSC−LAM rapamycin treated
Sporadic LAM untreated
Sporadic LAM rapamycin treated

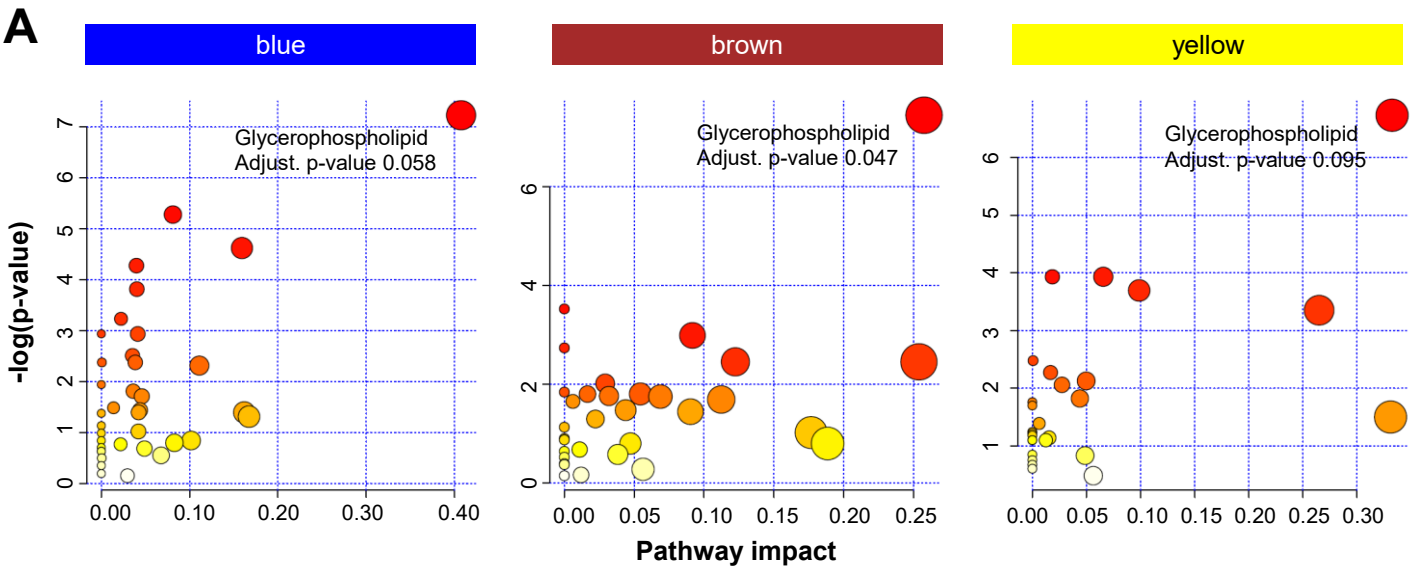

NS
log2 FC
−log10 Adj. p−value
−log10 Adj. p−value & log2 FC
stearoylcholine*
oleoylcholine
linoleoylcholine*
docosahexaenoylcholine
palmitoylcholine
dihomo−linolenoyl−choline
arachidonoylcholine
eicosapentaenoylcholine
2−stearoyl−GPI (18:0)*
−log10 Adjusted p−value
log2 Fold Change
0
1
2
3
−1
0
1

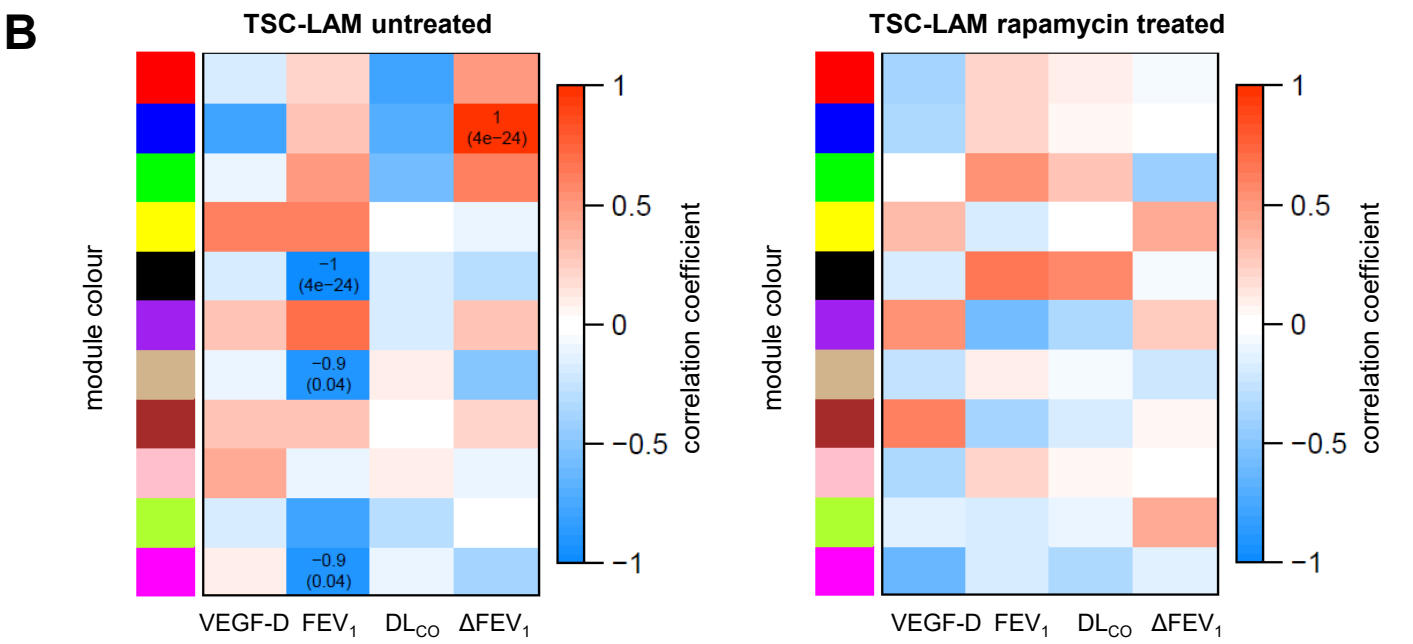

NS
log2 FC
−log10 Adj. p−value
−log10 Adj. p−value & log2 FC
myristoleoylcarnitine (C14:1)*
laurylcarnitine (C12)*
−log10 Adjusted p−value
1.5
1.0
0.5
0.0
−2
−1
0
1
2
log2 Fold Change


A
Branch height
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Differential metabolites
Module colour


B
module colour
56
84
56
77
50
34
30
162
45
31
35
TSC-LAM rapamycin treated
TSC-LAM untreated


C
module colour
0.96 0.9 0.98 0.65 0.02 0.85 0.05 0.21 0.38 0 0
0.66 0.84 0.93 0.48 0.08 1 0.98 0.12 0.01 0
0.97 0.54 0.44 0.54 0.86 0.61 0.71 0.06 0
0.84 0.37 0.23 0.45 0.83 0.34 0.04 0
0.84 0.14 0.01 0.14 0.09 0.02 0
0.88 0.32 0.04 0.06 0.01 0
0.94 0.6 0.6 0.1 0
0.6 0.71 0.26 0
0.06 0.04 0
0.01 0
0


A
blue
brown
yellow
Glycerophospholipid
Adjust. p-value 0.058
Glycerophospholipid
Adjust. p-value 0.047
Glycerophospholipid
Adjust. p-value 0.095
-log(p-value)
Pathway impact


B
TSC-LAM untreated
TSC-LAM rapamycin treated
module colour
1
(4e−24)
−1
(4e−24)
−0.9
(0.04)
−0.9
(0.04)
correlation coefficient
VEGF-D
FEV1
DLCO
ΔFEV1

Sphingolipid, fatty acid and phospholipid metabolites are associated with disease severity and mTOR inhibition in lymphangioleiomyomatosis

Leonardo Bottolo, Suzanne Miller, Simon R. Johnson

SUPPLEMENTARY METHODS

SAMPLE PROCESSING AND ANALYSIS

All samples were maintained at -80^{o}C until processed. Quality control recovery standards were added prior to extraction for QC purposes. Samples were prepared using the MicroLab STAR® system (Hamilton Company). Proteins were precipitated with methanol by shaking for 2 minutes followed by centrifugation. The resulting extract was divided into five fractions: two for analysis by two separate reverse phase (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI, and one backup. The organic solvent was removed using a TurboVap® (Zymark). Sample extracts were stored overnight under nitrogen before preparation for analysis. Controls analysed with the experimental samples comprised a pooled matrix sample generated by taking a small volume of each experimental sample as a technical replicate throughout the data set, extracted water samples as process blanks and various QC standards selected not to interfere with the endogenous compounds were spiked into every analysed sample, allowed instrument performance monitoring and aided chromatographic alignment. Instrument variability was determined by calculating the median relative standard deviation (RSD) for the standards that were added to each sample prior to injection into the mass spectrometers. Overall process variability was determined by calculating the median RSD for all endogenous metabolites present in all pooled matrix samples. Experimental samples were randomized across the platform run with QC samples spaced evenly among the injections.

ULTRAHIGH PERFORMANCE LIQUID CHROMATOGRAPHY-TANDEM MASS SPECTROSCOPY (UPLC-MS/MS)

All methods utilized a Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyser operated at 35,000 mass resolution. Sample extracts were dried and reconstituted in solvents compatible with each of the four methods. Reconstitution solvents contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency.

1. Using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. Extracts were gradient eluted from a C18 column (Waters UPLC BEH C18-2.1x100 mm, 1.7 µm) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA).
2. Using acidic positive ion conditions, optimised for more hydrophobic compounds at a higher organic content.

3. Using basic negative ion optimised conditions on a separate C18 column. The basic extracts were gradient eluted using methanol and water, with 6.5mM Ammonium Bicarbonate at pH 8.
4. Via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1x150 mm, 1.7 µm) using a water/acetonitrile with 10mM Ammonium Formate, pH 10.8 gradient. The MS analysis alternated between MS and data-dependent MS^n scans using dynamic exclusion. The scan range varied slighted between methods but covered 70-1000 m/z.

DATA EXTRACTION AND COMPOUND IDENTIFICATION

Raw data were extracted, peak-identified and QC processed at Metabolon© as described previously (1). The platform and compound identification algorithm uses biochemical identifications based on three criteria: 1) the retention index within a narrow RI window of the proposed identification, 2) mass match to the library +/- 10 ppm, and MS/MS forward and reverse scores between experimental data and standards. The MS/MS scores are based on a comparison of ions present in the experimental spectrum to those present in the library spectrum. The combined use of all three data points is used to distinguish and differentiate compounds. Over 3300 purified standards are registered in LIMS for analysis for determination of their identity.

CONTROL GROUP

To increase study power, 21 control samples from the current study were merged with female controls from two companion metabolites studies available from the NIHR BioResource Rare Diseases, University of Cambridge, resulting in 43 controls subjects. All were healthy women over the age of 18 with no prior history of lung disease (supplementary table 1).

LAM GROUP

The LAM subjects comprised 79 women recruited from the National Centre for LAM in Nottingham UK between 2011 and 2018. All subj

ects had LAM defined by current ATS/JRS criteria (2). Subjects had a clinical assessment, comprising CT of the chest, abdomen and pelvis, screening for TSC, full lung function. At follow up visits, lung function FEV_1 and DL_{CO} were measured. The study was approved by the East Midlands Research Ethics Committee (13/EM/0264) and all participants gave written informed consent.

Prospective change in FEV_1 was calculated by the regression slope of all FEV_1 values (ΔFEV_1) and expressed as change in ml/year. Only subjects with greater than one year of observations were included for calculation of ΔFEV_1.

Serum VEGF-D was determined using Quantikine ELISA DVED00, (R&D Systems, Abingdon, UK).

For exploratory analyses, subjects were categorised into those with mild and more severe disease based upon lung function and disease activity defined by ΔFEV_1. Subjects were also segregated by menopausal status and treatment with rapamycin.

DATA PRE-PROCESSING

Normalisation and imputation of case and control samples serum metabolites were performed following the workflow presented in (1) and (3). Relevant normalisation (N) steps can be summarised as follows: (N.1) Untargeted metabolites and metabolites belonging to 'Xenobiotics' biochemical class were removed from the analysis, reducing the number of targeted metabolites to 820; (N.2) After checking the proportion of missing values across samples and metabolites, no metabolites were removed. (N.3) Each metabolite raw value was rescaled to have median 1 to adjust for variation due to instrument run-day tuning differences; (N.4) A log transformation with base 10 was applied to all the metabolites; (N.4) After transformation, data points lying more than 4 standard deviations from the mean of each metabolite concentration were excluded. For the imputation (I) of missing values, we employed the KNN-TN method of (4) which consists of the following steps: (I.1) Estimation of the detection level (DL) of the machine to be the minimum observed value for the whole dataset; (I.2) Maximum Likelihood Estimation (MLE) of μ_m and σ_m, assuming that each metabolite m (m = 1,…,820) follows a left-truncated (on the DL) Gaussian distribution with mean μ_m and standard deviation σ_m; (I.3) Standardisation of each metabolite using the MLEs of μ_m and σ_m; (I.4) For each metabolite m with a missing value in sample i, detection of its K = 10 closest metabolites (which have an observed value for their ith sample) using the k-nearest neighbours algorithm; (I.5) Imputation of the missing value with a weighted average of the K values found in (I.4). The weights are functions of the Pearson correlations between the metabolite with missing values and its K closest metabolites; (I.6) Transformation of each metabolite back to the original scale as it was before step (I.3).

DIFFERENTIAL ANALYSIS

Differential analysis of 820 targeted serum metabolites was performed by using Limma (5) after correcting for BMI, ethnicity and run day (recording which samples were run on which days relative to each other) using a linear mixed model (6) with study and run day as crossed random effects. We also corrected for age when the hypothesis to test did require the assessment of its effect on the metabolites' levels, and 'study' covariate as a third crossed random effect when we tested differences between LAM women not treated with rapamycin and healthy controls (since we added two extra control groups from the companion metabolites studies). Significant differential metabolites were declared at 10% FDR (7).

DIFFERENTIAL NETWORK ANALYSIS

Based on the WGCNA package (8), Differential Network Analysis (9) allows the detection of differential networks (modules) between conditions. Relevant steps of this method include: (S.1) Build correlation matrix *C* within each condition. We used robust Spearman's rank correlation coefficient to calculate the correlation between any pair of metabolites in each condition. (S.2) Compute matrix of adjacent (powered correlation) differences. The soft-threshold power parameter β is chosen such that it is the lowest power for which the scale-free topology R^2 fit between the degree of connectivity *k* and the proportion of metabolites that have connectivity *k* exceeds 0.85. In the real data analysis, this automatic procedure leads to estimated value of β ranging between 5 and 7. (S.3) Derive the Topological Overlap Measure (TOM distance) based on the dissimilarity matrix *T* in order to identify metabolites that share the same metabolites' neighbours in the graph obtained from the matrix of adjacency differences (S.2); (S.4) Hierarchical clustering of dissimilarity matrix *T* based on TOM distance allows partitioning the metabolites into modules that share similar metabolites' neighbours. Thresholding of hierarchical clustering is obtained by using the Dynamic Tree Cut R package (10). (S.5) The permutation-based procedure is employed to assess the statistical significance of the modules detected in (S.4), with the number of permutations = 1,000. The permutation consists in shuffling observations between conditions and, for a given partition obtained in (S.4), the empirical p-value is obtained by calculating how many time the observed average powered correlation difference in a module is greater than the one obtained by shuffling the observations. Finally, (S.5) for each identified module, principal component analysis is performed and the first eigenvalue ('eigen-metabolites') is correlated (Spearman correlation) with selected clinical traits.

BIOINFORMATIC ANALYSIS

Metabolomic pathway analysis was performed by using MetaboAnalyst 4.0 (11) with both significant differential metabolites and metabolites modules detected in the differential network analysis mapped in KEGG pathways. Given the lack of pathways annotation for a large fraction of metabolites, significant metabolomic pathways were declared at a conservative 10% Holm–Bonferroni correction. Finally, topology pathway analysis was performed by selecting the relative-betweenness centrality measure (ranging between 0 and 1) which quantifies the importance of a subgroup of metabolites in a given metabolomic pathway.

REFERENCES

1. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. Nat Genet. 2014;46(6):543–50.
2. McCormack FX, Gupta N, Finlay GR, Young LR, Taveira-Da Silva AM, Glasgow CG, et al.

Official American thoracic society/Japanese respiratory society clinical practice guidelines: Lymphangioleiomyomatosis diagnosis and management. Am J Respir Crit Care Med. 2016;194(6):748–61.

3. Krumsiek J, Mittelstrass K, Do KT, Stückler F, Ried J, Adamski J, et al. Gender-specific pathway differences in the human serum metabolome. Metabolomics. 2015;11(6):1815–33.
4. Shah JS, Rai SN, DeFilippis AP, Hill BG, Bhatnagar A, Brock GN. Distribution based nearest neighbor imputation for truncated high dimensional data with applications to pre-clinical and clinical metabolomics studies. BMC Bioinformatics. 2017;18(1).
5. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
6. Pinheiro J, Bates D, editors. Linear Mixed-Effects Models: Basic Concepts and Examples. In: Mixed-Effects Models in S and S-PLUS. Springer-Verlag; 2006. p. 3–56.
7. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. J R Stat Soc Ser B. 1995;57(1):289–300.
8. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9.
9. Tesson BM, Breitling R, Jansen RC. DiffCoEx: A simple and sensitive method to find differentially coexpressed gene modules. BMC Bioinformatics. 2010;11.
10. Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: The Dynamic Tree Cut package for R. Bioinformatics. 2008;24(5):719–20.
11. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, et al. MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. Nucleic Acids Res. 2018;46(W1):W486–94.

Sphingolipid, fatty acid and phospholipid metabolites are associated with disease severity and mTOR inhibition in lymphangioleiomyomatosis

Leonardo Bottolo, Suzanne Miller, Simon R Johnson.

Supplementary tables 1 - 3

Supplementary table 1. Demographics of control subjects.

Group	Sample size	**Age** (yrs)	Description
G1*	21	35.4 (11.7)	Healthy individuals University of Nottingham
G2*	13	48 (15)	Healthy individuals from companion metabolites study
G3*	9	46 (8)	Healthy individuals from companion metabolites study

*Control group is G1+G2+G3, n = 43. Standard deviation in brackets.

Supplementary table 2. Stratification of women with LAM by age, BMI and disease duration.

Group	Sample size	Age (yrs)	BMI	Disease duration (yrs)
TSC-LAM untreated	7	41.1 (9.8)	23.7 (3.2)	10.4 (9.7)
TSC-LAM rapamycin treated	5	40.8 (8.3)	26.8 (3.9)	6.4 (3.8)
Sporadic LAM untreated	43	52.1 (10.3)	26.0 (6.5)	14.4 (11.3)
Sporadic LAM rapamycin treated	24	49.5 (11.6)	25.2 (6.5)	14.1 (11.6)
FEV_1 T1	16	51.8 (10.9)	24.3 (5.2)	9.6 (10.8)
FEV_1 T2	17	46.3 (8.7)	27.0 (7.3)	11.4 (7.6)
FEV_1 T3	17	53.7 (12.0)	25.4 (5.9)	20.3 (12.0)
DL_{CO} T1	16	51.6 (11.9)	25.7 (6.2)	11.4 (9.7)
DL_{CO} T2	17	48.1 (10.2)	24.4 (4.3)	14.3 (9.9)
DL_{CO} T3	17	52.2 (10.6)	26.7 (7.7)	15.8 (13.5)
OBD S0	8	53.1 (11.4)	25.8 (6.9)	13.2 (14.3)
OBD S1	17	46.9 (10.1)	24.4 (4.0)	9.3 (6.6)
OBD S2	23	51.9 (11.0)	26.6 (7.3)	16.4 (11.4)
OBD S3	2	56.0 (14.1)	23.8 (7.1)	25.5 (17.0)
ΔFEV_1 T1	13	52.0 (13.1)	25.9 (7.0)	15.6 (11.1)
ΔFEV_1 T2	12	55.6 (9.0)	26.0 (5.7)	17.2 (11.1)
ΔFEV_1 T3	12	47.1 (9.5)	24.5 (5.0)	14.0 (11.9)

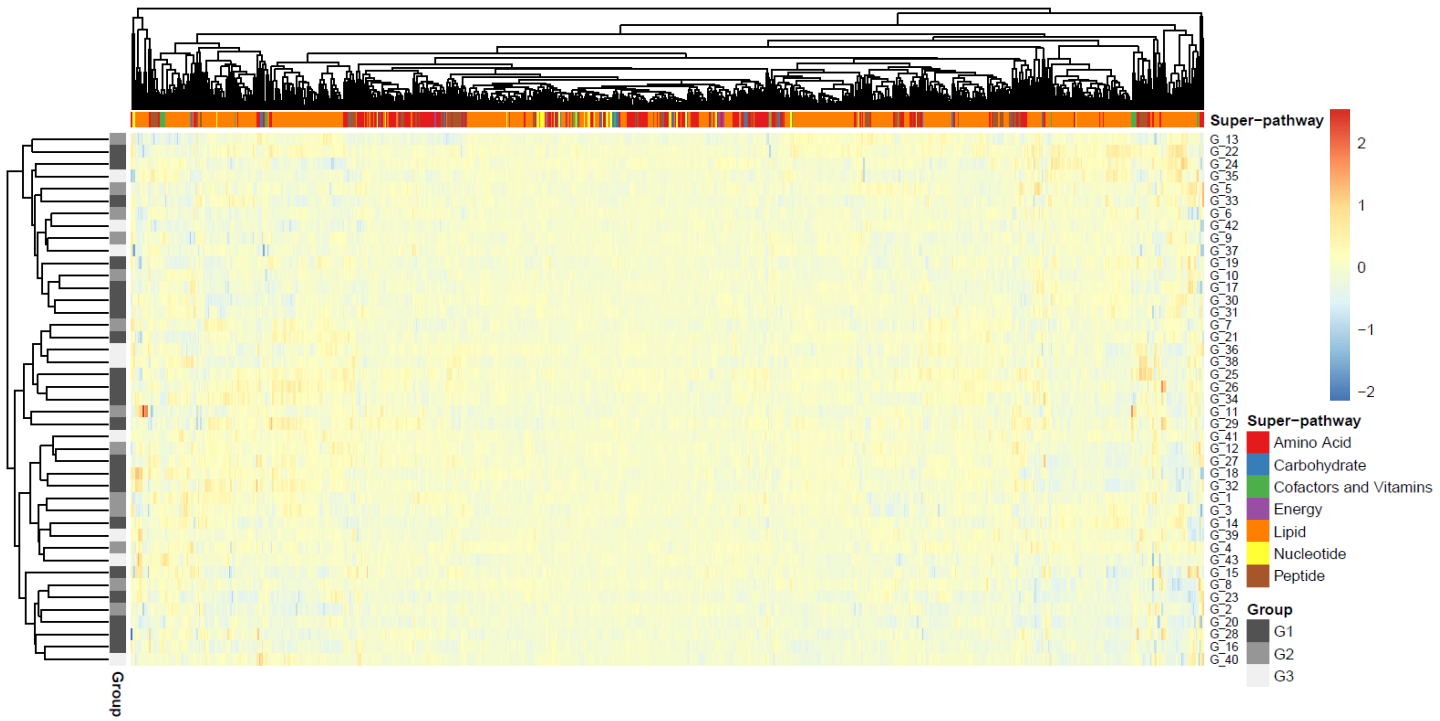
TSC = tuberous sclerosis complex; sporadic LAM = patients with LAM not associated with TSC; FEV_1 = forced expiratory volume in one second (ml); DL_{CO} = diffusing capacity of the lung for carbon monoxide (ml/min/kPa); ODB = overall disease burden score; ΔFEV_1 = annual decline in forced expiratory volume (ml/yr); T = tertile, 1 highest, 3 lowest; S = score, 0 lowest, 3 highest. Boxes inside the table indicate the characteristic used to stratify the patients. Standard deviation in brackets.

Supplementary table 3. Characterisation of women with LAM stratified by age and disease characteristics.

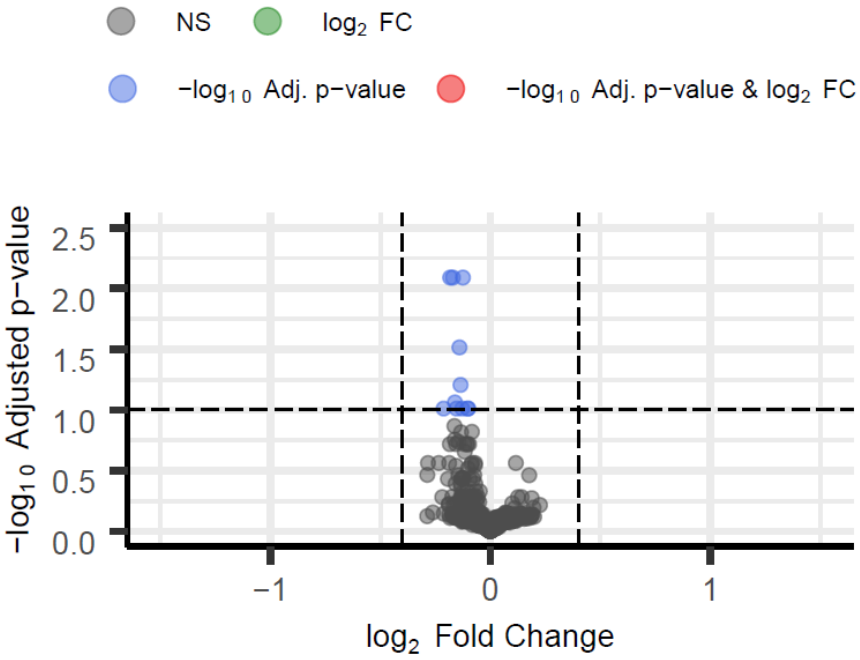
Group	TSC (%)	AML (%)	Lymph (%)	PTX (%)	VEGF-D (pg/ml)	FEV_1 (% pr.)	DL_{CO} (% pr.)	ΔFEV_1 (ml/yr)
TSC-LAM untreated	100	86	0	29	1,021	72	61	5
TSC-LAM rapamycin treated	100	100	20	20	2,159	50	47	-19
Sporadic LAM untreated	0	70	19	30	1,332	68	60	-61
Sporadic LAM rapamycin treated	0	42	21	17	961	50	42	-131
FEV_1 T1	19	62	19	12	1,085	92	72	-32
FEV_1 T2	12	65	12	41	1,598	69	59	-102
FEV_1 T3	12	88	18	35	1,186	46	50	-31
DL_{CO} T1	19	62	25	25	865	81	76	-40
DL_{CO} T2	18	76	12	35	1,472	68	62	-43
DL_{CO} T3	6	76	12	29	1,513	57	44	-84
OBD S0	0	50	25	12	634	93	74	-60
OBD S1	18	47	6	29	1,282	76	65	-67
OBD S2	17	96	13	35	1,549	57	52	-52
OBD S3	0	100	100	50	943	52	60	24
ΔFEV_1 T1	23	85	15	23	1,529	66	59	15
ΔFEV_1 T3	0	100	8	50	1,231	62	62	-50
ΔFEV_1 T3	8	50	25	25	1,291	74	62	-132

TSC = tuberous sclerosis complex; AML = angiomyolipoma; PTX = Pneumothorax; % = percentage of group with this characteristic; % pr. = percentage of predicted value; FEV_1 = forced expiratory volume in one second (ml); DL_{CO} = diffusing capacity of the lung for carbon monoxide (ml/min/kPa); ODB = overall disease burden; ΔFEV_1 = annual decline in forced expiratory volume (ml/yr); T = tertile, 1 highest, 3 lowest; S = score, 0 highest, 3 lowest. Boxes inside the table indicate the characteristic used to stratify the patients.

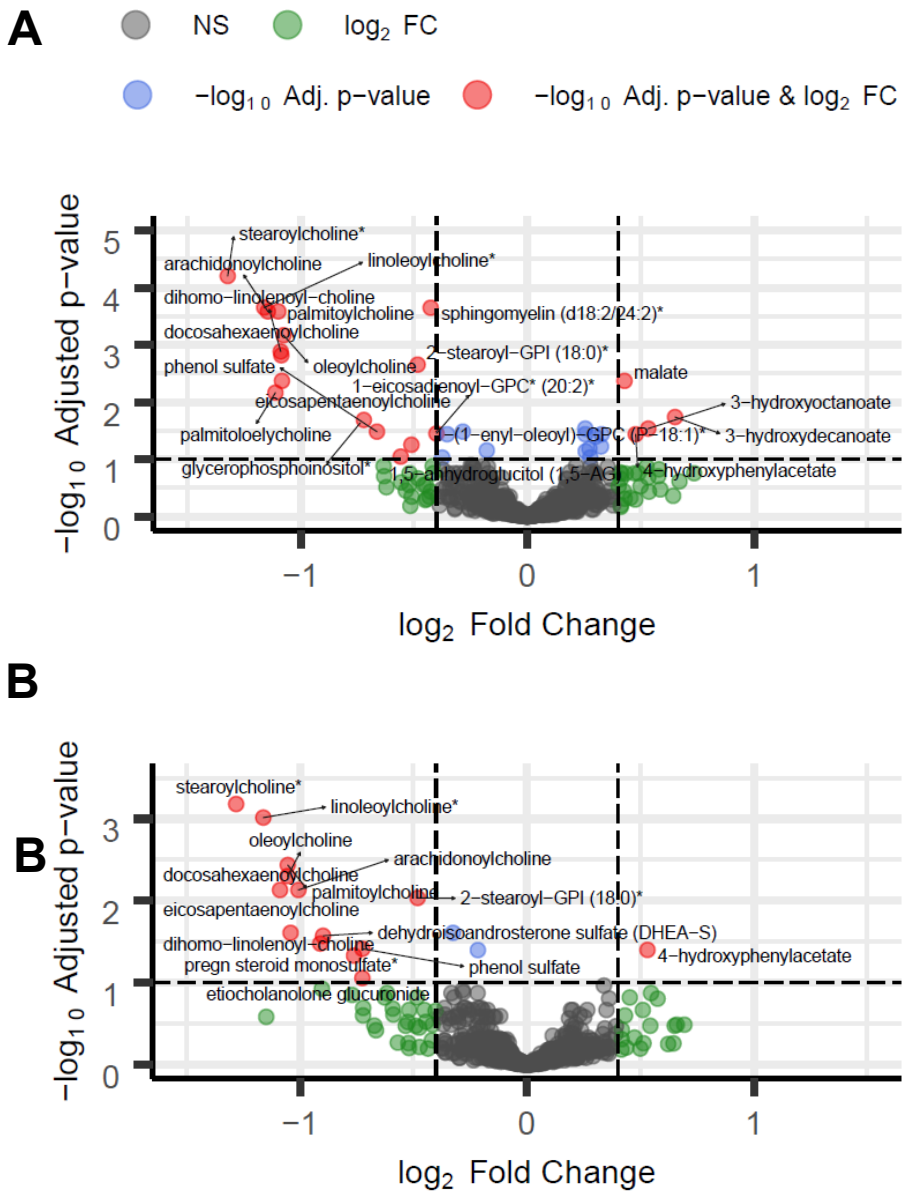
Supplementary figure 1. Hierarchical clustering of healthy controls based on log-transformed metabolites levels after correcting for confounding effects. Bi-clustering plot of the controls groups (horizontal) and metabolites levels (vertical) stratified by super-pathway. Control group G1 (n = 21), G2 (n = 13) and G3 (n = 9) (for controls groups definition, see supplementary table 1) are colour coded in dark grey, grey and light grey, respectively.



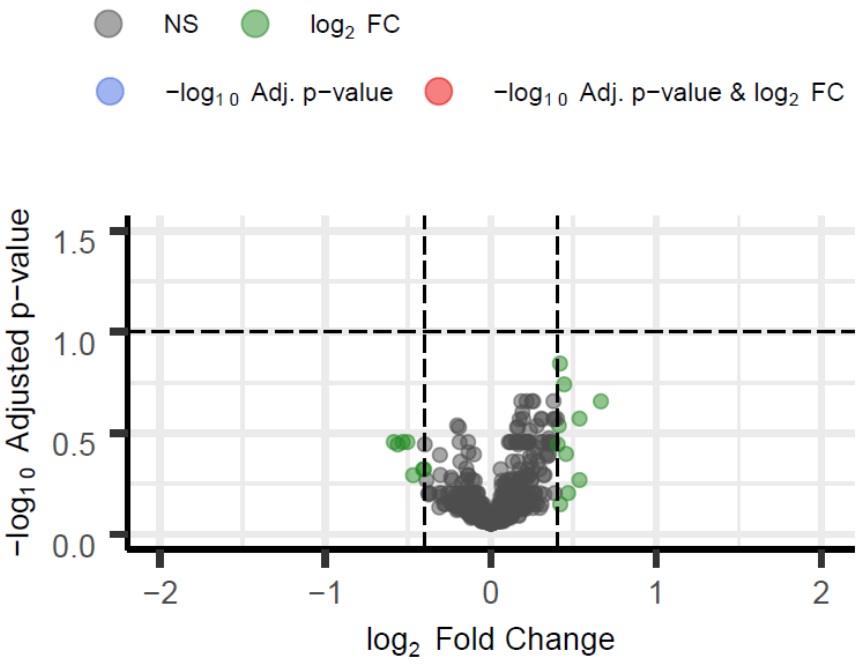
Supplementary figure 2. Volcano plot of metabolomic changes in patients with decrease of forced expiratory volume. Significantly different metabolites at 10% False Discovery Rate (FDR) (*y*-axis) with |$\log_2$ Fold Change (FC)| < 0.4 (*x*-axis) between FEV_1 T1 (n = 16) and T2 (n = 17) both not treated with rapamycin (for group definitions, see supplementary table 3). Blue and grey dots indicate metabolites that are significantly different between conditions at 10% FDR with |$\log_2$ FC| < 0.4 and overall NS = non-significant, respectively. Full results are provided in supplementary table 5.



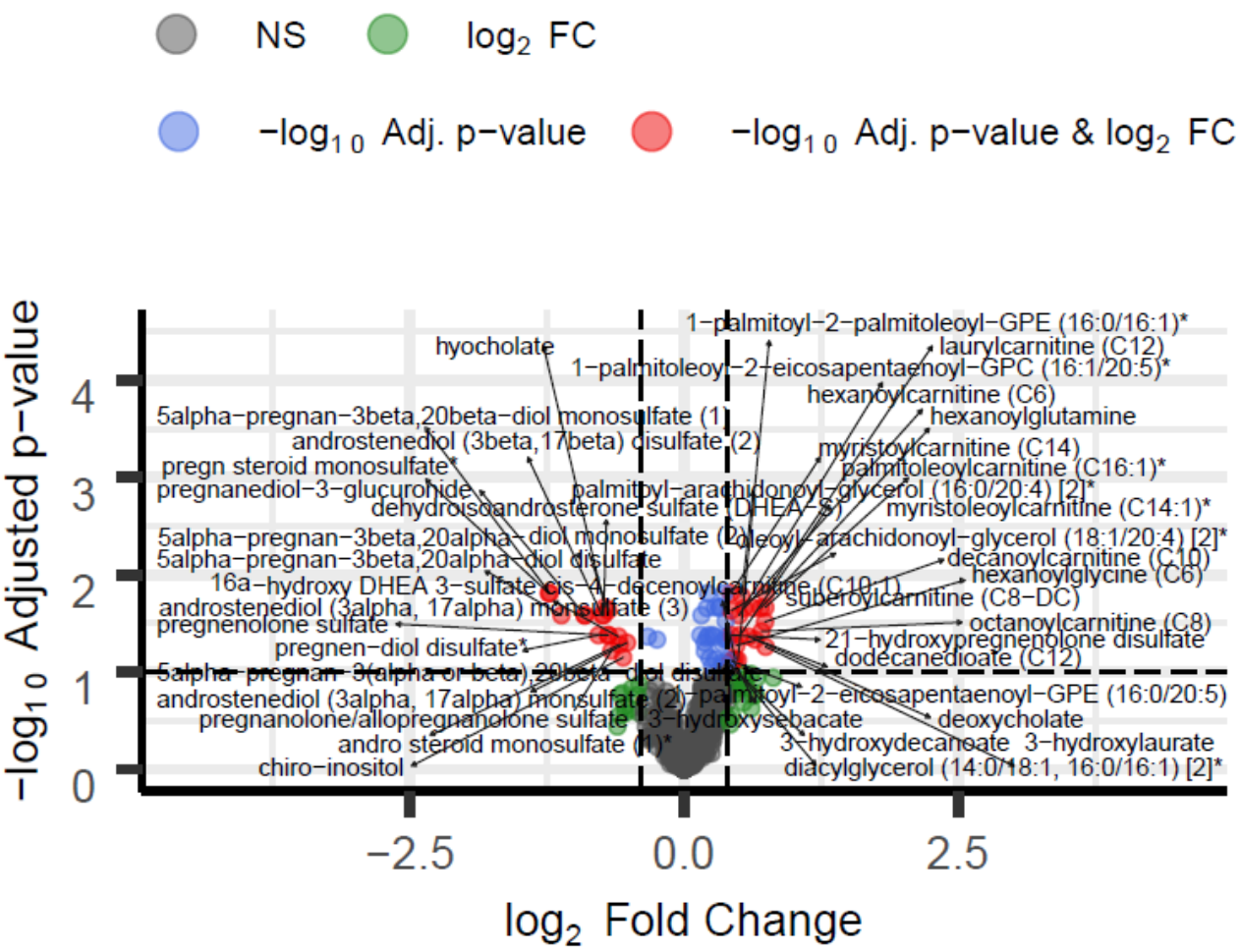
Supplementary figure 3. Volcano plot of fatty acids and phospholipids across overall disease burden. Significantly different metabolites at 10% False Discovery Rate (FDR) (*y*-axis) with |$\log_2$ Fold Change (FC)| > 0.4 (*x*-axis) in patients not treated with rapamycin between overall disease burden score (**A**) S2 (n = 23) and S3 (highest, n = 2) and (**B**) S0 (n = 8) and S3 (highest, n = 2) (for group definitions, see supplementary table 3). Red, blue, green and grey dots indicate metabolites that are significantly different between conditions at 10% FDR with |$\log_2$ FC| > 0.4, metabolites significant at 10% FDR with |$\log_2$ FC| < 0.4, non-significant at 10% FDR with |$\log_2$ FC| > 0.4 and overall NS = non-significant, respectively. *Compounds identified without confirmation by a standard. Full results are presented in supplementary table 5.



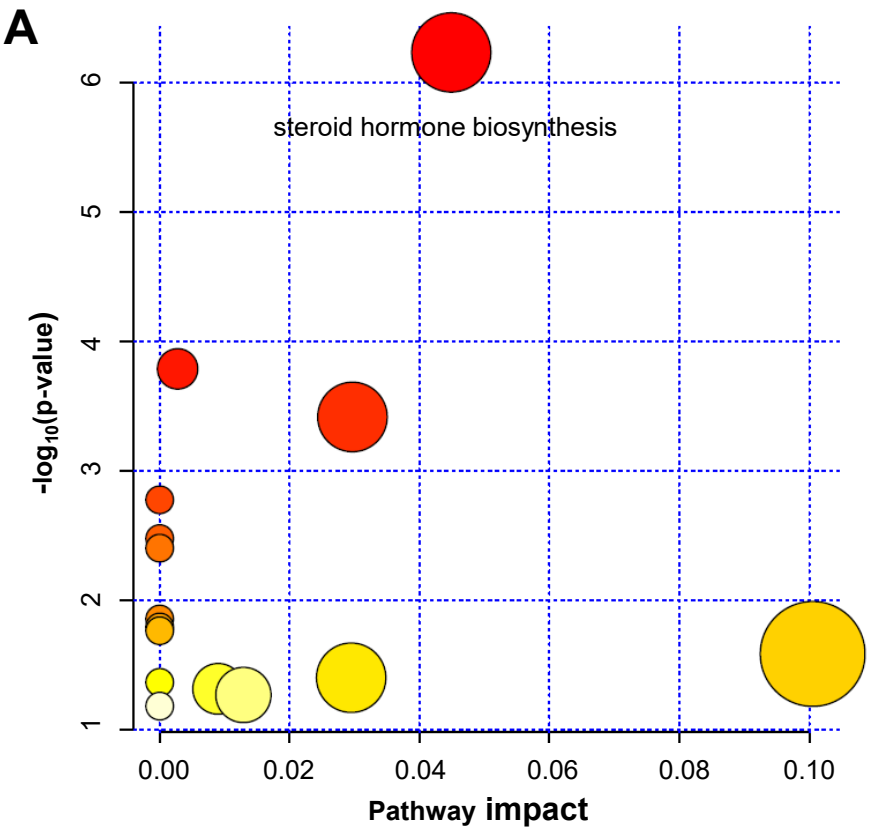
Supplementary figure 4. Volcano plot of mTOR inhibition on the TSC-LAM. Differential metabolites at 10% False Discovery Rate (FDR) (*y*-axis) with |$\log_2$ Fold Change (FC)| > 0.4 (*x*-axis) between TSC-LAM without (n = 7) and with (n = 5) mTOR inhibitor treatment. Green and grey dots indicate metabolites that are non-significant at 10% FDR with |$\log_2$ FC| > 0.4 and overall NS = non-significant, respectively. Full results are presented in supplementary table 5.

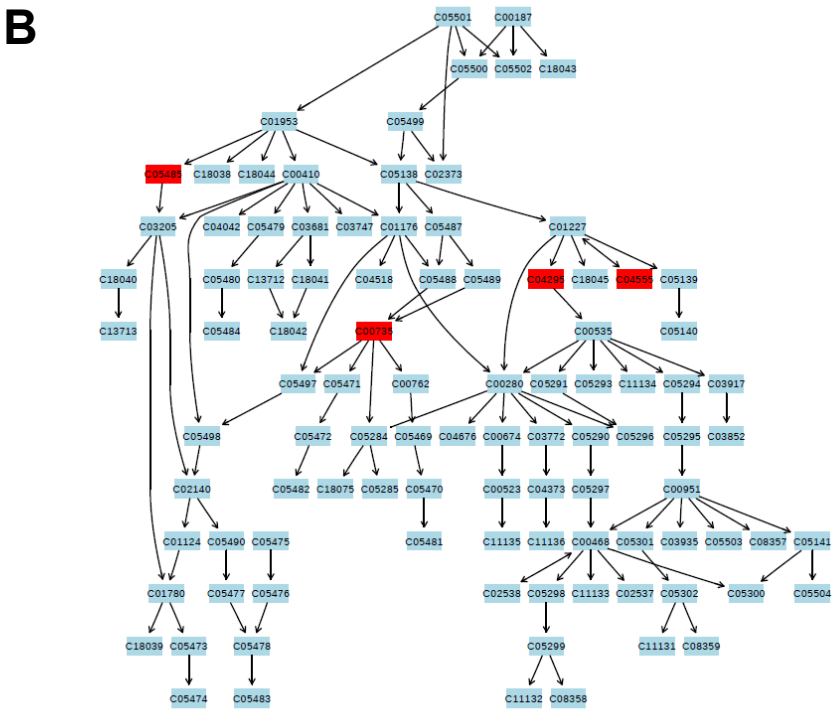


Supplementary figure 5. Volcano plot of metabolomic changes in post-menopause. Significantly different metabolites at 10% False Discovery Rate (FDR) (*y*-axis) with $|\log_2$ Fold Change (FC)| > 0.4 (*x*-axis) between pre- (n = 33) and post-menopause (n = 17) women with LAM not treated with rapamycin. Red, blue, green and grey dots indicate metabolites that are significantly different between conditions at 10% FDR with $|\log_2$ FC| > 0.4, metabolites significant at 10% FDR with $|\log_2$ FC| < 0.4, non-significant at 10% FDR with $|\log_2$ FC| > 0.4 and overall NS = non-significant, respectively. All downregulated metabolites, but 'chiro-inositol' belongs to the 'steroid' sub-pathway. *Compounds identified without confirmation by a standard. Full results are provided in supplementary table 4.

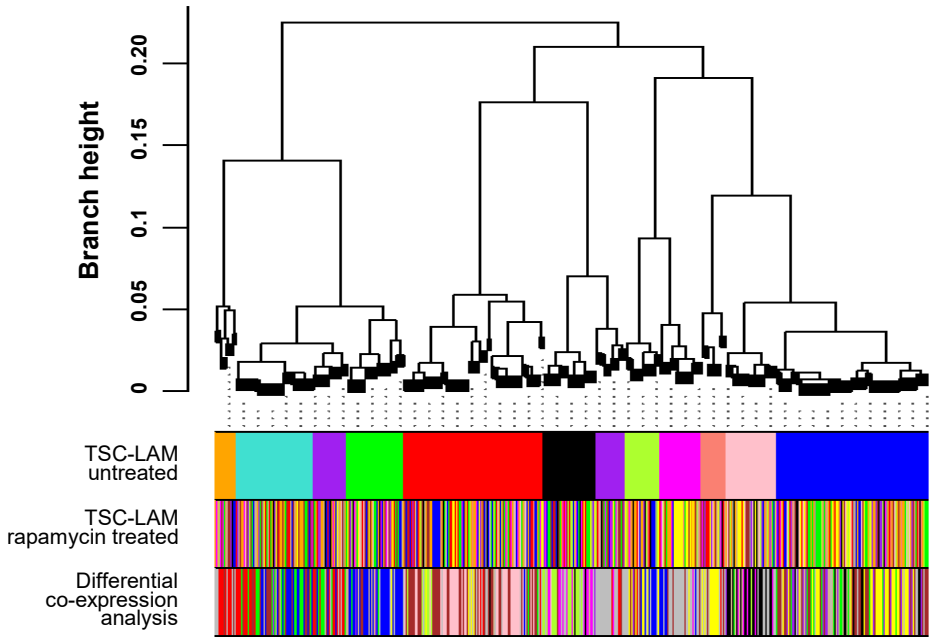


Supplementary figure 6. Metabolomic pathway analysis of differential metabolites in the post-menopause analysis (see supplementary table 4). (**A**) Scatterplot depicting -$\log_{10}$ p-values of the metabolomic pathway analysis (*y*-axis) and impact values from topology pathway analysis (*x*-axis). metabolomic pathway detected at 10% Holm–Bonferroni correction is highlighted. (**B**) Steroid hormone biosynthesis KEGG pathway. Red boxes indicate significant differential metabolites present in the KEGG pathway.





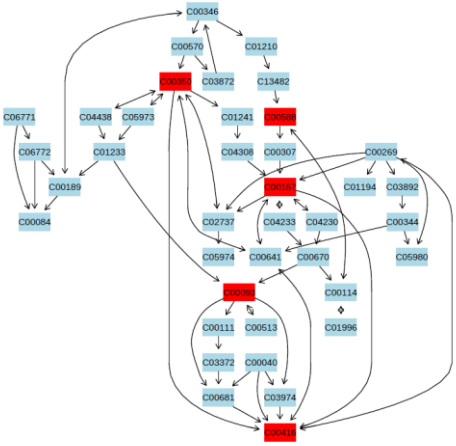
Supplementary figure 7. Co-expression network analysis performed separately on TSC-LAM samples treated and not treated with rapamycin. Dendrogram showing clustering of co-expressed metabolites (modules) for TSC-LAM samples not treated with rapamycin (n = 5) (for group definitions, see supplementary table 3). Top colour bar: colours identify 12 modules generated by unsupervised hierarchical clustering based on untreated TSC-LAM samples; middle colour bar: colours identify 16 modules generated by unsupervised hierarchical clustering based on TSC-LAM samples with rapamycin treatment (n = 7) and rearranged based on untreated TSC-LAM modules ordering; bottom colour bar: colours identify 11 modules generated by differential co-expression analysis and rearranged based on untreated TSC-LAM modules ordering.



Supplementary figure 8. Metabolomic pathways analysis of the differential co-expression networks (modules) based on TSC-LAM samples with and without rapamycin treatment. For each significant module (empirical p-value < 0.1), metabolomic KEGG pathways detected at 10% Holm–Bonferroni correction are plotted (see supplementary table 6). Modules are identified on the top by a colour bar. Within each metabolomic pathway, red boxes indicate metabolites present in the module and in the KEGG pathway.

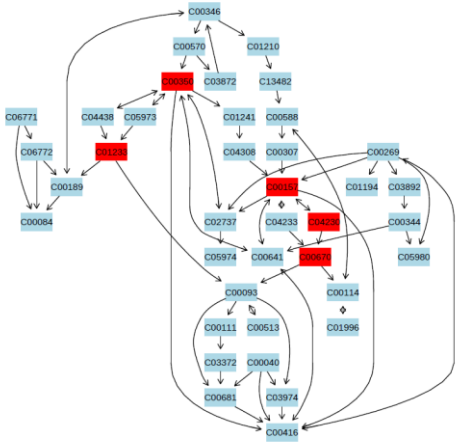
blue

Glycerophospholipid
KEGG pathway



brown

Glycerophospholipid
KEGG pathway



yellow

Glycerophospholipid
KEGG pathway