Plasma Metabolomics Implicate Modified Transfer RNAs and Altered Bioenergetics in the Outcome of Pulmonary Arterial Hypertension

Running title: *Rhodes et al.; Plasma metabolomics in pulmonary hypertension*

Christopher J. Rhodes, PhD[1]†; Pavandeep Ghataorhe, BM, BCh[1]†; John Wharton, PhD[1];
Kevin C. Rue-Albrecht, PhD[1]; Charaka Hadinnapola, BM, BCh[2]; Geoffrey Watson, BM, BCh[1];
Marta Bleda, PhD[2]; Matthias. Haimel, BSc[2]; Gerry Coghlan, MD[3]; Paul A. Corris, FRCP[4];
Luke. S. Howard, DPhil[5,6]; David G. Kiely, MD[7,8]; Andrew J. Peacock, MD[9];
Joanna Pepke-Zaba, PhD[10]; Mark R. Toshner, MD[2,10]; S. John Wort, PhD[11];
J. Simon R. Gibbs, MD[5,6]; Allan Lawrie, PhD[8]; Stefan Gräf, PhD[2,12];
Nicholas W. Morrell, MD, FMedSci[2]; Martin R. Wilkins, MD, FMedSci[1]‡

[1]Department of Medicine, Imperial College London, Hammersmith Campus, Du Cane Road, London, UK;[2]Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge, UK;[3]Cardiology Department, Royal Free Hospital, London, UK; [4]Institute of Cellular Medicine, Newcastle University and The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK;[5]National Pulmonary Hypertension Service, Imperial College Healthcare NHS Trust, Hammersmith Hospital, London, UK; [6]National Heart and Lung Institute (NHLI), Imperial College London, Hammersmith Campus, London, UK;[7]Sheffield Pulmonary Vascular Disease Unit, Royal Hallamshire Hospital, Sheffield, UK;[8]Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield, Sheffield, UK; [9]Scottish Pulmonary Vascular Unit, Golden Jubilee National Hospital, Glasgow, UK;[10]Pulmonary Vascular Disease Unit, Papworth Hospital, Cambridge, UK; [11]Pulmonary Hypertension Service, Royal Brompton Hospital, London, UK; [12]Department of Haematology, University of Cambridge, Cambridge, UK
†These authors contributed equally to this work
‡On behalf of the NIHR BioResource – Rare Diseases (BRIDGE) PAH Consortium
and the UK National PAH Cohort Study Consortium

Address for Correspondence:
Martin R. Wilkins, MD, FMedSci
Department of Medicine
Imperial College London
Hammersmith Campus
Du Cane Road
London, W12 0NN, UK
Tel: 020 3313 2049
Email: m.wilkins@imperial.ac.uk,

Journal Subject Terms: Biomarkers; Metabolism; Pulmonary Hypertension; Mortality/Survival; Vascular Disease

1

Abstract

Background— Pulmonary arterial hypertension (PAH) is a heterogeneous disorder with high mortality.

Methods—We conducted a comprehensive study of plasma metabolites using ultra-performance liquid chromatography mass-spectrometry to (1) identify patients at high risk of early death, (2) identify patients who respond well to treatment and (3) provide novel molecular insights into disease pathogenesis.

Results— 53 circulating metabolites distinguished well-phenotyped patients with idiopathic or heritable PAH (n=365) from healthy controls (n=121) following correction for multiple testing (p<7.3e-5) and confounding factors, including drug therapy, renal and hepatic impairment. A subset of 20/53 metabolites also discriminated PAH patients from disease controls (symptomatic patients without pulmonary hypertension, n=139). 62 metabolites were prognostic in PAH, with 36/62 independent of established prognostic markers. Increased levels of tRNA-specific modified nucleosides (N2,N2-dimethylguanosine, N1-methylinosine), TCA cycle intermediates (malate, fumarate), glutamate, fatty acid acylcarnitines, tryptophan and polyamine metabolites and decreased levels of steroids, sphingomyelins and phosphatidylcholines distinguished patients from controls. The largest differences correlated with increased risk of death and correction of several metabolites over time was associated with a better outcome. Patients who responded to calcium channel blocker therapy had metabolic profiles similar to healthy controls.

Conclusions— Metabolic profiles in PAH are strongly related to survival and should be considered part of the deep phenotypic characterisation of this disease. Our results support the investigation of targeted therapeutic strategies that seek to address the alterations in translational regulation and energy metabolism that characterise these patients.

Key-words: metabolome; metabolomics; metabolism; pulmonary circulation; pulmonary hypertension

Clinical Perspective

What is new?

- This study provides a comprehensive analysis of circulating metabolite levels in patients with pulmonary arterial hypertension (PAH) and controls.

- It is the first to relate metabolite levels to clinical outcomes in PAH.

- Increases in circulating modified nucleosides originating from transfer RNAs, energy metabolism intermediates, tryptophan and polyamine metabolites, and decreased steroids, sphingomyelins and phosphatidylcholines independently discriminate PAH patients from controls and predict survival.

- Correction of metabolite levels over time is linked to better clinical outcomes and patients who respond well to calcium-channel blocker therapy have metabolic profiles comparable with healthy controls.

What are the clinical implications?

- Energy metabolism and stress-response pathways are disturbed in PAH.

- Monitoring plasma metabolites that report on these pathways over time could be useful to assess disease progression and response to therapy.

- Therapeutic strategies targeted against metabolic disturbances in PAH, particularly translational regulation and energy metabolism, merit further investigation

Introduction

Pulmonary arterial hypertension (PAH) is a progressive vascular disorder that leads to increased pulmonary vascular resistance, right ventricular (RV) dysfunction [1] and premature death [2]. The pathogenesis of the vascular pathology is poorly understood [3, 4]. Several genetic mutations have been reported in hereditary and isolated idiopathic presentations of PAH, providing insight into perturbed signalling pathways [5, 6], and genome sequencing of clinically well characterised patient cohorts is underway in anticipation of finding new mutations. A complementary approach to identifying the molecular drivers of PAH is to conduct deep molecular phenotyping of patients beyond standard clinical tests.

Metabolomic technologies, such as ultra-performance liquid chromatography mass spectrometry, enable the detection and semi-quantitative measurement of hundreds of unique metabolites, representing a broad range of metabolic pathways, in small volumes of biofluids [7]. These approaches have identified differences in circulating metabolites that distinguish physiological and disease states, such as diabetes and systemic cardiovascular disorders, and predict clinical outcomes [8-10]. As yet, few metabolomics studies have been undertaken in patients with pulmonary vascular disease. Evidence of abnormal oxidation, arginine and sphingosine pathways have been found from mass spectrometry analysis of lung tissue from PAH patients [11, 12] and analysis of breath samples showed exhaled volatile compounds discriminate between severe idiopathic PAH and healthy volunteers [13]. A targeted analysis of 105 circulating plasma metabolites in PAH, primarily amino acids, nucleosides and their derivatives, showed abnormal levels of tryptophan, purine and tricarboxylic acid cycle metabolites correlated to haemodynamic measures [14].

In this study, we used a broad metabolomics platform to analyse 1416 metabolites in plasma from patients with idiopathic or heritable PAH (n=365) in three distinct cohorts and compared circulating levels with both healthy (n=121) and disease controls (n=139). We identified specific metabolites that both discriminate PAH patients from healthy and disease controls and independently predict survival. These metabolites included several modified nucleosides specific to transfer RNAs that indicate alterations in cell proliferation and translation of disease-related proteins as well as several constituents of energy metabolism.

Methods

Sample collection

Samples were obtained from patients with idiopathic or heritable PAH attending the National Pulmonary Hypertension Service at Hammersmith Hospital, London between 2002-2015 and from patients recruited from other UK national centres as part of the National Cohort Study of Idiopathic and Heritable Pulmonary Arterial Hypertension (ClinicalTrials.gov NCT01907295). Control plasma samples were obtained from healthy subjects and disease controls, the latter being symptomatic patients presenting to the service but in whom pulmonary hypertension was excluded by cardiac catheterisation. The diagnosis of PAH was based on standard criteria from the most recent guidelines [1]. Vasoresponders were defined as those who dropped their mean pulmonary artery pressure >10 mmHg to <40 mmHg, with preserved cardiac output, in response to an acute pulmonary vasodilator challenge and remained stable on calcium channel blocker therapy alone for at least 1 year [1]. Whole-genome sequencing data from the UK National Institute of Health Research Biomedical Research Centres Inherited Diseases Genetic Evaluation

(BRIDGE) consortium were used to determine which patients had known pathogenic mutations in the gene encoding the bone morphogenetic protein type II receptor (*BMPR2*) [5].

Venous blood samples were drawn from the antecubital fossa and collected in EDTA Vacutainer tubes (BD, Oxford, UK), immediately put on ice, centrifuged (1,300g, 15 minutes) within 30 minutes and stored at -80^{0}C until required. WHO functional class and six minute walk distance at sample date and clinical biochemical data (within 30 days) were recorded. All subjects provided informed written consent and local research ethics committees approved the study. A subset of patients consented to provide additional samples at later dates whilst attending follow-up clinical appointments.

Metabolomics

Metabolomic profiling by ultra-performance liquid chromatography mass spectrometry was conducted by Metabolon (Durham, NC, USA) [7], who provided semi-quantitative assessment of 949 named and 467 unnamed metabolite levels, annotated with pathways. Named compounds identified based on mass and fragmentation analysis but yet to be confirmed with standards are indicated by asterisks. Details can be found in the supplementary materials.

Angiogenin

Plasma angiogenin levels were determined by ELISA (Ref:DAN00, R&D Systems, Abingdon, UK) as per manufacturer's guidelines, with EDTA plasma diluted 1:800 before assay.

Statistical Analysis

To prevent skewing of results by outliers, initial group comparisons between controls and patients were performed using non-parametric Mann Whitney U tests. Prior to modelling, metabolites whose distribution was not normal were transformed either by $\log_{10}$ or power transformations (x^Y, with Y from -2 to 2 in 0.5 steps, as performed for Box-Cox transformations

[15]), whichever best normalised the data based on Kolmogorov-Smirnov tests, or ranked if no test met $p > 0.05$. Samples where metabolites were undetected were imputed with the minimum detected level for the metabolite. Data in all groups were z-score transformed based on the mean and standard deviation in all healthy controls for ease of comparisons. Data are presented as absolute numbers, percentages, mean or median (±standard deviation, SD) and percentile range. Linear regression analysis was conducted to assess the relationships between metabolite levels, diagnoses and potential confounders to determine whether differences in metabolite levels between groups was independent of age, gender, ethnicity, body mass index, drugs, renal and hepatic dysfunction. In the disease control and PAH cohorts, preserved renal function was defined as creatinine <75 µmol/L, and liver function as bilirubin <21 µmol/L. Logistic regression was conducted to determine metabolites that independently distinguished between diagnostic groups. Orthogonal partial least squares discriminant analysis (OPLS-DA) modelling was used to test the performance of these metabolites. R^2 scores indicate model performance and Q^2 scores estimate reproducibility, based on cross validation (subjects were divided into 7 groups and their diagnosis was predicted based on the other subjects in seven analyses). Pathway enrichment analysis was conducted on discriminating and prognostic metabolites using Fisher's exact test. All survival analyses were performed using time from sampling to death/census. Cox regression analysis was used to identify prognostic predictors, with proportional hazard assumptions tested and Kaplan-Meier plots used to illustrate events from time of sampling in relation to metabolite levels. Receiver Operating Characteristic (ROC) curves were used to assess discriminating and prognostic value of metabolites against diagnosis and all-cause mortality respectively. Hierarchical clustering based on Euclidean distances was used to assess if metabolites and patients clustered by functional pathways and phenotypes, respectively.

Network analysis was performed by calculating second order Spearman's rank correlations using ParCorA [16, 17] and visualised using Cytoscape [18]. 'Hub' nodes are metabolites with the most 'edges' (correlations) to other metabolites.

Statistical analysis was performed using IBM SPSS Statistics 22 (International Business Machines Corporation, New York, USA), Matlab (Matrix Laboratory, MathWorks, Natick, Massachusetts, USA), Microsoft Excel (Microsoft, Redmond, Washington, USA), SIMCA-P software, (Umetrics, Umea, Sweden) and R with RStudio and associated packages [19].

Results

Metabolites distinguishing between PAH and controls

We first compared plasma metabolite profiles from 116 consecutive patients with idiopathic or heritable PAH attending Hammersmith Hospital between November 2011 and August 2013 and 58 healthy controls (Table 1). To minimise confounding factors, only PAH patients aged 19-70 were compared with age- and sex-matched healthy controls in this analysis. Results were validated in 75 PAH patients recruited between 2002 and 2015 against a separate control group (n=63). A second validation analysis used 174 PAH patients recruited from other specialist centres in the UK from August 2013 to June 2015 and compared to all controls. Metabolites identified as xenobiotics or detected in less than 95% of samples were excluded from the analysis, leaving 686 well-quantified biological metabolites.

Circulating levels of 97 metabolites distinguished PAH from healthy subjects in all three analyses following Bonferroni correction (p<7.3e-5). Of these metabolites, 53 distinguished healthy and PAH subjects after correcting for potential confounders, including age, gender, ethnicity, body mass index, creatinine, bilirubin and drug therapies (p<0.05, Table S1, Fig. 1).

The most common confounders associated with metabolite levels were liver function (bilirubin) and renal function (creatinine). To determine whether these metabolite differences could be detected prior to initiation of PAH therapies, we performed a sub-analysis comparing 40 patients who were treatment naïve at the time of sampling and all 53 metabolites distinguished this group from healthy controls (p<0.05, Table S1). Patients with pathogenic *BMPR2* mutations (n=42) had similar metabolite levels to PAH patients without these mutations (Fig. S1, Table S1). Given that many metabolic alterations might occur in a chronic disease such as PAH, we set out to prioritise more disease-specific metabolites by comparing the PAH patients with disease controls, the latter comprising symptomatic patients in whom pulmonary hypertension had been excluded. We again adopted a discovery and validation design, with two groups of disease controls (n=70 and 69). A subset (20/53) of the metabolites distinguished PAH patients from disease controls after correcting for potential confounders (p<0.05, Table S1). These 'PAH-specific' differences in metabolites included increases in purine, polyamine and tricarboxylic acid (TCA) cycle metabolites, and decreases in phosphocholines and sphingomyelins (Fig. 2A), with network analysis showing the importance of 'hub' metabolites N2,N2-dimethylguanosine and malate (Fig. 2B).

Discriminant analyses to distinguish PAH and control groups

To identify a minimal set of metabolites which could in combination best distinguish PAH patients, we performed logistic regression analysis. We found 7/53 metabolites –
dehydroisoandrosterone sulfate (DHEA-S), methionine sulfone, N1-methylinosine, oleoylcarnitine, palmitoylcholine, sphingomyelin (d18:1/20:0, d16:1/22:0)* and X-24513 – independently distinguished PAH (aged 19-70) and healthy subjects in the discovery analysis, with 90% accuracy in an orthogonal partial least squares discriminant analysis (OPLS-DA,

R^2=0.64, Q^2=0.61). This model classified healthy and PAH subjects in the two validation

analyses with 89% and 84% accuracy, respectively. In addition, 90% (9/10) of PAH

vasoresponders in the discovery cohort had metabolite levels typical of healthy controls (Fig.

3A-B).

Out of the 20 'PAH-specific' metabolites, four – N-acetylaspartate, octadecanedioate,

palmitoylcholine and X-13737 – distinguished PAH patients and disease controls with 83%

accuracy in the discovery analysis (R^2=0.49, Q^2=0.47). This model classified disease controls

and PAH subjects in the two validation analyses with 69% and 67% accuracy, respectively (Fig.

3C-D).

Survival analysis of plasma peaks of interest in PAH

We hypothesised that metabolites most closely related to the disease pathobiology would be

associated with clinical outcomes. To identify metabolites associated with disease progression

and mortality, we performed survival analyses. 28/116 and 25/75 patients died in the discovery

and first validation PAH groups, with an average follow-up of 3.3±1.0 and 4.5±4.0 years,

respectively. The length of patient follow-up in the second validation cohort was insufficient to

permit analysis. After controlling for creatinine and diuretic use, no other potentially

confounding factor was associated with survival. Of the 686 well-quantified metabolites, 640

met the assumptions of Cox regression analysis, and 62 of these were prognostic after accounting

for creatinine and diuretic use in both analyses. ROC analysis at 3 years of follow-up confirmed

these metabolites were prognostic and identified optimal cut-offs (Fig. S2, $p<0.05$).

To identify metabolites that report on novel pathways independent of current prognostic

estimates, we compared the 62 prognostic metabolites with three markers previously found to

best predict survival in our patients – namely, N-terminal brain natriuretic peptide (NT-proBNP),

10

six minute walk distance (6MWD) and red cell distribution width (RDW) [20]. 36/62 of the metabolites were independent of these measures (p<0.05, Fig. 4A, Table S2) and network analysis indicated two main clusters with 'hub' metabolites including again, among others, N2,N2-dimethylguanosine (Fig. 4B).

Enrichment and clustering of metabolites of interest

The above analyses identified and validated, after controlling for confounders, a total of 100 metabolites that were either discriminating or prognostic in PAH, representing twenty five metabolic pathways. Six pathways in particular were enriched with metabolites of interest, including fatty acid (acyl carnitines), polyamine and nucleoside metabolism (Table S3, p<0.05). Sixteen of these metabolites both discriminated PAH and were prognostic; these, along with the three other metabolites which were selected by logistic regression modelling to best distinguish PAH and healthy subjects, clustered into defined metabolic pathways (Fig. 5).

Analysis of serial samples

Changes in metabolite levels in individuals over time may indicate pathways that report clinical improvement or whose correction itself leads to improved outcomes. We analysed serial samples from 86 patients who were followed up for a minimum of 1 year (median 1.50, interquartile range 1.33-2.95 years) after the second sample. Twenty nine patients died during follow-up. Changes in metabolite levels between the two samples (median time between samples 1.75, IQR 1.07-2.58 years) were compared between 'survivors' and 'non-survivors'.

Changes in 27/100 metabolites were significantly different between survivors and non-survivors (p<0.05), including several modified amino acids and nucleosides. ROC analysis confirmed these associations (Table S4) and identified prognostic cut-offs (Fig. 6).

11

Association of elevated modified nucleosides with elevated plasma angiogenin

Modified nucleosides can be released into the circulation during stress following cleavage of tRNAs by the ribonuclease angiogenin [21]. To determine whether this mechanism was relevant to PAH, we measured plasma angiogenin in a representative subset of age- and sex-matched healthy controls and PAH patients from the discovery analysis (Table S5). Angiogenin levels were elevated in plasma from PAH patients and correlated with N2,N2-dimethylguanosine levels (Rho:0.49, p<0.001, Fig. 7). The strength of correlation was similar in male and female subjects (data not shown).

Discussion

This study represents the most comprehensive analysis of circulating metabolites in patients with PAH to date. It is the first to robustly identify and validate differences in comparison to both healthy and symptomatic disease controls without pulmonary hypertension, and to associate metabolic profiles with outcomes in PAH, strengthening the evidence that the pathways identified could be important modifiers of disease progression. Changes in the levels of metabolites over time were associated with survival in a direction that suggests that correction of these disturbances is linked to improved outcomes. In agreement with this, patients defined as vasoresponders, who have excellent outcomes on calcium channel blocker therapies, demonstrated metabolic profiles more similar to healthy controls than other patients. Metabolic profiles seen in incident cases were similar to those with established PAH, emphasising that metabolic dysregulation is not corrected in the majority of cases by current therapy.

Modified nucleosides

Two of the most robust distinguishing and prognostic differences identified in PAH patients were increased levels of N1-methylinosine and N2,N2-dimethylguanosine. These are recognised epigenetic, post-transcriptional modifications of transfer RNA (tRNA) [21-23], and other tRNA modifications also found to be increased and prognostic included pseudouridine, N6-carbamoylthreonyladenosine and N1-methyladenosine. N2,N2-dimethylguanosine is found in the majority of tRNAs at position 26, upstream of the anticodon sequence at positions 34-36, and promotes the folding of tRNAs towards the classical clover-leaf structure [24]. N1-methylinosine is found 3' adjacent to the anticodon at position 37 of eukaryotic tRNAs and is formed from inosine by a specific S-adenosylmethionine-dependent methylase [25]. Increased serum and urine levels of N2,N2-dimethylguanosine, as well as pseudouridine and 1-methylinosine, have been observed in multiple solid tumour malignancies [26] and may reflect the general upregulation of the translational apparatus, including tRNA turnover, in hyperproliferative cancerous cells [27]. Increased circulating 1-methyladenosine has also been shown to be an early indicator of oxidative stress, cell damage and mortality in kidney disease [28].

Intracellular tRNA pools are dynamically regulated. For example under stress, tRNAs required for the translation of stress response proteins are preferentially expressed [21]. The altered levels of specific nucleoside modifications in PAH patients may reflect preferential expression of tRNAs that harbour them, as part of a switch towards translation of disease-related proteins. In addition, stress-induced cleavage of tRNA produces fragments that propagate the stress response and interfere with eukaryotic initiation factor (eIF)-4G and eIF4F[21]. Furthermore, eIF2α kinase-4 (GCN2), which prevents eIF2α interacting with the initiating Met-tRNA, suppressing general protein synthesis and activating stress-inducible transcription factors, is mutated and causally

implicated in some cases of pulmonary vascular disease [29]. Mutations in tRNA genes themselves have also been reported to cause pulmonary hypertension driven by mitochondrial dysfunction [30].

The initial cleavage of tRNAs is mediated by angiogenin [21], which we showed to be elevated in the plasma of PAH patients in concert with elevated levels of modified nucleosides. Angiogenin is also upregulated in cancer cells, mediating angiogenesis, cell proliferation and protection from apoptosis [31], and is increased in breath condensates from patients with pulmonary hypertension [32], indicating a possible pulmonary origin in this disease. Angioproliferative plexiform vascular lesions are characteristic of advanced PAH and the pro-angiogenic activity of angiogenin is inhibited by mutation of its ribonuclease active site [33], suggesting that elevated angiogenin and nucleoside levels may report patients developing this type of pulmonary vascular remodelling. Alterations in tRNA biology appear to be capable of driving the development of rare forms of pulmonary hypertension and are closely linked to the progression of PAH, and circulating levels of modified nucleosides may reflect increases in both pulmonary vascular cell proliferation and stress.

Energy metabolism

Significant alterations were observed in several pathways related to cellular energy production, with accumulation of multiple acylcarnitines, glutamate and TCA cycle intermediates. Their accumulation in PAH patients may represent a failed attempt to increase utility of fatty acids as an energy source, perhaps reflecting the inability of fatty acid beta-oxidation to keep pace with the demands of the overburdened right ventricle. Glutaminolysis is another alternative energy production pathway to glucose oxidation, with the product glutamate entering the TCA cycle as α-ketoglutarate. Inhibition of glutaminolysis and restoration of glucose oxidation has beneficial effects in rat models of right ventricular hypertrophy [34]. Increased circulating glutamate levels

14

have previously been seen in cancer patients [35], however, anti-glutaminolysis therapeutic targets have demonstrated toxic side effects [36]. The build-up of TCA intermediates and the precursors to the molecules that enter the cycle (acylcarnitines and glutamate) may indicate dysfunction of this cycle, or at least the inability to keep pace with the demands of the most active cells, such as proliferating pulmonary vascular cells. Increased levels of citrate, succinate and fatty acid metabolites have been demonstrated in lung tissue from PAH patients [11], suggesting dysfunctional energy metabolism is a feature of the diseased tissue. Restoration of glucose oxidation by dichloroacetate therapy is under investigation as a treatment for PAH [37], and maximising the capacity of the TCA cycle to process the acetyl-CoA produced may be a complementary therapeutic approach.

Consistent with previous reports we found a significant increase in the circulating levels of long-chain acylcarnitines (oleoylcarnitine) [38], and also short- (myristoylcarnitine, acetylcarnitine, hydroxbutyrylcarnitine) and medium- (adipoylcarnitine, suberoylcarnitine) chain products. The accumulation of acylcarnitines may itself be detrimental, effecting cardiac electrophysiological changes and arrhythmias [39]. There is also increasing evidence that accumulation of long chain acylcarnitines may contribute to insulin resistance [40], which is itself common and associated with prognosis in PAH [41].

Lipids, steroids, polyamines and tryptophan metabolites

Multiple sphingomyelin and phosphatidylcholine lipid species were significantly reduced in PAH patients, relating to increased mortality. Sphingomyelins are the most abundant subclass of sphingolipids, with other subclasses including sphinogosines, ceramides and glycophospholipids [42]. In patients with chronic obstructive pulmonary disease (COPD), low plasma levels of several sphingomyelins relate to disease severity [43]. As a membrane constituent, sphingomyelins are

15

implicated in trans-membrane signalling and are generated from phosphatidylcholine and ceramide by sphingomyelin synthase, knockout of which leads to mitochondrial dysfunction and reduced insulin release [44]. Sphingomyelins may also be considered a source of ceramide, which directly (and indirectly through other active lipid products) regulates cell proliferation, apoptosis, cell migration and autophagy [45].

Reduced lineoyl-GPC has been shown to be an early marker of insulin resistance in non-diabetics [46], and decreased circulating levels of several phosphatidylcholines were seen in patients with severe heart failure [47]. Phospholipids are also sources of multiple cellular signalling molecules including eicosanoids such as prostacyclin [48], levels of which are known to be reduced in pulmonary hypertension, with replacement an established treatment option. Circulating levels of DHEA-S and its metabolites (androsterone, epiandrosterone and androstenediol/4-androsten-3beta, 17beta-diol disulfate) were reduced in PAH patients compared to healthy controls, consistent with a recent report of reduced circulating levels of DHEA-S in a small cohort of 23 male PAH patients compared to healthy controls [49]. Differences in DHEA-S between PAH and controls were independent of the more subtle effects of both gender and age (Fig. S3), and lower DHEA-S levels were independently associated with mortality. Treatment with DHEA or DHEA-S has repeatedly been shown to prevent and reverse pulmonary hypertension in experimental rat models [50], with clinical trials ongoing in COPD-associated pulmonary hypertension (ClinicalTrials.gov NCT00581087).

We found increased levels of a breakdown product of N1-acetylspermidine, acisoga. Other metabolites of polyamine metabolism (4-acetamidobutanoate and N-acetylputrescine) were increased in PAH in relation to bilirubin levels and were prognostic in 2 distinct PAH cohorts, independent of established prognostic markers. Several animal models of pulmonary

hypertension have demonstrated evidence of increased polyamine levels and metabolism in lung tissue [51]. Administration of monocrotaline to rats led to significantly increased levels of polyamines and the development of pulmonary hypertension and right ventricular hypertrophy, which could be prevented by the administration of an inhibitor of polyamine biosynthesis [52], suggesting these molecules may be novel therapeutic targets.

In our study, we validated findings of elevated circulating tryptophan metabolites [14] with increased C-glycosyltryptophan and kynurenine in PAH patients compared to healthy controls, but changes in kynurenine were related to increased bilirubin levels and liver dysfunction. Levels of tryptophan and its other major metabolite, serotonin, were not significantly altered in our analysis.

Limitations

The majority of patients included in this study were prevalent cases on established treatments. A sub-analysis was conducted with 40 incident cases and showed similar results. Corrections were also made for potential treatment effects in the main analyses, including PAH-specific and comorbidity-related therapies, as well as demographics and renal/hepatic function. Patients and controls were sampled in the non-fasting state and information on insulin resistance was not available for all cases. Patients were also sampled from a peripheral vein. The stronger performance of discriminating models in discovery analyses suggests that optimisation could further improve their performance in distinct cohorts. Strongly correlated metabolites would not have been selected in the discriminant modelling, so each metabolite used in the final models may represent clusters of multiple metabolites. Evidence of tissue specificity and the source of circulating metabolites in pulmonary hypertension require further studies, for example, by transpulmonary sampling and direct measurements from tissue samples, to better localise the

source of informative metabolites. Plasma levels of metabolites may not reflect levels in the most important tissues, for example increased bile acid metabolites have been demonstrated in PAH lung tissue [53] but no differences were seen in circulating levels in our study.

Summary and conclusions

Increased circulating modified nucleosides (N2,N2-dimethylguanosine, N1-methylinosine), TCA cycle intermediates (malate, fumarate), glutamate, fatty acid acylcarnitines and polyamine metabolites and decreased levels of steroids, sphingomyelins and phosphatidylcholines are a characteristic of patients with PAH that distinguish them from symptomatic patients without pulmonary hypertension. Improvements in circulating metabolite levels are associated with a better prognosis and could be used to monitor response to PAH treatments. Indeed, our results support the investigation of therapeutic strategies targeted at alterations in energy metabolism in PAH and suggest options for correcting translational regulation also merit further study.

Acknowledgements

18

Disclosures

None

References

1.   Galie N, Humbert M, Vachiery JL, Gibbs S, Lang I, Torbicki A, Simonneau G, Peacock A, Vonk Noordegraaf A, Beghetti M, Ghofrani A, Gomez Sanchez MA, Hansmann G, Klepetko W, Lancellotti P, Matucci M, McDonagh T, Pierard LA, Trindade PT, Zompatori M and Hoeper M. 2015 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension: The Joint Task Force for the Diagnosis and Treatment of Pulmonary Hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS): Endorsed by: Association for European Paediatric and Congenital Cardiology (AEPC), International Society for Heart and Lung Transplantation (ISHLT). *European Respiratory Journal*. 2015;46:903-75.
2.   Benza RL, Miller DP, Barst RJ, Badesch DB, Frost AE and McGoon MD. An evaluation of long-term survival from time of diagnosis in pulmonary arterial hypertension from the REVEAL Registry. *Chest*. 2012;142:448-56.
3.   Ryan JJ and Archer SL. Emerging concepts in the molecular basis of pulmonary arterial hypertension: part I: metabolic plasticity and mitochondrial dynamics in the pulmonary circulation and right ventricle in pulmonary arterial hypertension. *Circulation*. 2015;131:1691-702.
4.   Maron BA and Leopold JA. Emerging Concepts in the Molecular Basis of Pulmonary Arterial Hypertension: Part II: Neurohormonal Signaling Contributes to the Pulmonary Vascular and Right Ventricular Pathophenotype of Pulmonary Arterial Hypertension. *Circulation*. 2015;131:2079-91.
5.   Machado RD, Eickelberg O, Elliott CG, Geraci MW, Hanaoka M, Loyd JE, Newman JH, Phillips JA, 3rd, Soubrier F, Trembath RC and Chung WK. Genetics and genomics of pulmonary arterial hypertension. *Journal of the American College of Cardiology*. 2009;54:S32-42.

6.   Ma L and Chung WK. The genetic basis of pulmonary arterial hypertension. *Human Genetics*. 2014;133:471-9.

7.   Evans AM, DeHaven CD, Barrett T, Mitchell M and Milgram E. Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Analytical Chemistry*. 2009;81:6656-67.

8.   Lewis GD, Farrell L, Wood MJ, Martinovic M, Arany Z, Rowe GC, Souza A, Cheng S, McCabe EL, Yang E, Shi X, Deo R, Roth FP, Asnani A, Rhee EP, Systrom DM, Semigran MJ, Vasan RS, Carr SA, Wang TJ, Sabatine MS, Clish CB and Gerszten RE. Metabolic signatures of exercise in human plasma. *Science Translational Medicine*. 2010;2:33ra37.

9.   Wang TJ, Larson MG, Vasan RS, Cheng S, Rhee EP, McCabe E, Lewis GD, Fox CS, Jacques PF, Fernandez C, O'Donnell CJ, Carr SA, Mootha VK, Florez JC, Souza A, Melander O, Clish CB and Gerszten RE. Metabolite profiles and the risk of developing diabetes. *Nature Medicine*. 2011;17:448-53.

10.   Shah SH, Kraus WE and Newgard CB. Metabolomic profiling for the identification of novel biomarkers and mechanisms related to common cardiovascular diseases: form and function. *Circulation*. 2012;126:1110-20.

11.   Zhao Y, Peng J, Lu C, Hsin M, Mura M, Wu L, Chu L, Zamel R, Machuca T, Waddell T, Liu M, Keshavjee S, Granton J and de Perrot M. Metabolomic heterogeneity of pulmonary arterial hypertension. *PLoS One*. 2014;9:e88727.

12.   Zhao YD, Chu L, Lin K, Granton E, Yin L, Peng J, Hsin M, Wu L, Yu A, Waddell T, Keshavjee S, Granton J and de Perrot M. A Biochemical Approach to Understand the Pathogenesis of Advanced Pulmonary Arterial Hypertension: Metabolomic Profiles of Arginine, Sphingosine-1-Phosphate, and Heme of Human Lung. *PLoS One*. 2015;10:e0134958.

13.   Mansoor JK, Schelegle ES, Davis CE, Walby WF, Zhao W, Aksenov AA, Pasamontes A, Figueroa J and Allen R. Analysis of volatile compounds in exhaled breath condensate in patients with severe pulmonary arterial hypertension. *PLoS One*. 2014;9:e95331.

14.   Lewis GD, Ngo D, Hemnes AR, Farrell L, Domos C, Pappagianopoulos PP, Dhakal BP, Souza A, Shi X, Pugh ME, Beloiartsev A, Sinha S, Clish CB and Gerszten RE. Metabolic Profiling of Right Ventricular-Pulmonary Vascular Function Reveals Circulating Biomarkers of Pulmonary Hypertension. *Journal of the American College of Cardiology*. 2016;67:174-89.

15.   Box G and Cox D. An Analysis of Transformations. *Journal of the Royal Statistical Society*. 1964;26:211-252.

16.   http://www.comp-sys-bio.org/software.html.

17.   de la Fuente A, Bing N, Hoeschele I and Mendes P. Discovery of meaningful associations in genomic data using partial correlation coefficients. *Bioinformatics*. 2004;20:3565-74.

18.   http://www.cytoscape.org/.

19.   http://CRAN.R-project.org/.

20.   Rhodes CJ, Wharton J, Howard LS, Gibbs JS and Wilkins MR. Red cell distribution width outperforms other potential circulating biomarkers in predicting survival in idiopathic pulmonary arterial hypertension. *Heart*. 2011;97:1054-60.

21.   Kirchner S and Ignatova Z. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nature Review Genetics*. 2015;16:98-112.

22.   Slotkin W and Nishikura K. Adenosine-to-inosine RNA editing and human disease. *Genome Medicine*. 2013;5:105.

23.   Torres AG, Batlle E and Ribas de Pouplana L. Role of tRNA modifications in human diseases. *Trends in Molecular Medicine*. 2014;20:306-14.

24.   Steinberg S and Cedergren R. A correlation between N2-dimethylguanosine presence and alternate tRNA conformers. *RNA*. 1995;1:886-91.

25.   Grosjean H, Auxilien S, Constantinesco F, Simon C, Corda Y, Becker HF, Foiret D, Morin A, Jin YX, Fournier M and Fourrey JL. Enzymatic conversion of adenosine to inosine and to N1-methylinosine in transfer RNAs: a review. *Biochimie*. 1996;78:488-501.

26.   Waalkes TP, Gehrke CW, Zumwalt RW, Chang SY, Lakings DB, Tormey DC, Ahmann DL and Moertel CG. The urinary excretion of nucleosides of ribonucleic acid by patients with advanced cancer. *Cancer*. 1975;36:390-8.

27.   Anderson P and Ivanov P. tRNA fragments in human health and disease. *FEBS letters*. 2014;588:4297-304.

28.   Mishima E, Inoue C, Saigusa D, Inoue R, Ito K, Suzuki Y, Jinno D, Tsukui Y, Akamatsu Y, Araki M, Araki K, Shimizu R, Shinke H, Suzuki T, Takeuchi Y, Shima H, Akiyama Y, Toyohara T, Suzuki C, Saiki Y, Tominaga T, Miyagi S, Kawagisihi N, Soga T, Ohkubo T, Yamamura K, Imai Y, Masuda S, Sabbisetti V, Ichimura T, Mount DB, Bonventre JV, Ito S, Tomioka Y, Itoh K and Abe T. Conformational change in transfer RNA is an early indicator of acute cellular damage. *Journal of the American Society of Nephrology*. 2014;25:2316-26.

29.   Eyries M, Montani D, Girerd B, Perret C, Leroy A, Lonjou C, Chelghoum N, Coulet F, Bonnet D, Dorfmuller P, Fadel E, Sitbon O, Simonneau G, Tregouet DA, Humbert M and Soubrier F. EIF2AK4 mutations cause pulmonary veno-occlusive disease, a recessive form of pulmonary hypertension. *Nature Genetics*. 2014;46:65-9.

30.   Sproule DM, Dyme J, Coku J, de Vinck D, Rosenzweig E, Chung WK and De Vivo DC. Pulmonary artery hypertension in a child with MELAS due to a point mutation of the mitochondrial tRNA((Leu)) gene (m.3243A>G). *Journal of Inherited Metabolic Disease*. 2008;31 Suppl 3:497-503.

31.   Saikia M and Hatzoglou M. The Many Virtues of tRNA-derived Stress-induced RNAs (tiRNAs): Discovering Novel Mechanisms of Stress Response and Effect on Human Health. *The Journal of Biological Chemistry*. 2015;290:29761-8.

32.   Seyfarth HJ, Sack U, Gessner C and Wirtz H. Angiogenin, bFGF and VEGF: angiogenic markers in breath condensate of patients with pulmonary hypertension. *Pneumologie*. 2015;69:207-11.

33.   Shapiro R and Vallee BL. Site-directed mutagenesis of histidine-13 and histidine-114 of human angiogenin. Alanine derivatives inhibit angiogenin-induced angiogenesis. *Biochemistry*. 1989;28:7401-8.

34.   Piao L, Fang YH, Parikh K, Ryan JJ, Toth PT and Archer SL. Cardiac glutaminolysis: a maladaptive cancer metabolism pathway in the right ventricle in pulmonary hypertension. *Journal of Molecular Medicine (Berl)*. 2013;91:1185-97.

35.   Koochekpour S, Majumdar S, Azabdaftari G, Attwood K, Scioneaux R, Subramani D, Manhardt C, Lorusso GD, Willard SS, Thompson H, Shourideh M, Rezaei K, Sartor O, Mohler JL and Vessella RL. Serum glutamate levels correlate with Gleason score and glutamate blockade decreases proliferation, migration, and invasion and induces apoptosis in prostate cancer cells. *Clinical Cancer Research*. 2012;18:5888-901.

36.   Tennant DA, Duran RV and Gottlieb E. Targeting metabolic transformation for cancer therapy. *Nature Reviews Cancer*. 2010;10:267-77.

37.   Piao L, Fang YH, Cadete VJ, Wietholt C, Urboniene D, Toth PT, Marsboom G, Zhang HJ, Haber I, Rehman J, Lopaschuk GD and Archer SL. The inhibition of pyruvate dehydrogenase kinase improves impaired cardiac function and electrical remodeling in two models of right ventricular hypertrophy: resuscitating the hibernating right ventricle. *Journal of Molecular Medicine (Berl)*. 2010;88:47-60.

38.   Brittain EL, Talati M, Fessel JP, Zhu H, Penner N, Calcutt MW, West JD, Funke M, Lewis GD, Gerszten RE, Hamid R, Pugh ME, Austin ED, Newman JH and Hemnes AR. Fatty Acid Metabolic Defects and Right Ventricular Lipotoxicity in Human Pulmonary Arterial Hypertension. *Circulation*. 2016;133:1936-44.

39.   DaTorre SD, Creer MH, Pogwizd SM and Corr PB. Amphipathic lipid metabolites and their relation to arrhythmogenesis in the ischemic heart. *Journal of Molecular and Cellular Cardiology*. 1991;23 Suppl 1:11-22.

40.   Schooneman MG, Vaz FM, Houten SM and Soeters MR. Acylcarnitines: reflecting or inflicting insulin resistance? *Diabetes*. 2013;62:1-8.

41.   Zamanian RT, Hansmann G, Snook S, Lilienfeld D, Rappaport KM, Reaven GM, Rabinovitch M and Doyle RL. Insulin resistance in pulmonary arterial hypertension. *Eur Respir J*. 2009;33:318-24.

42.   Hannun YA and Obeid LM. Principles of bioactive lipid signalling: lessons from sphingolipids. *Nature Reviews Molecular Cell Biology*. 2008;9:139-50.

43.   Bowler RP, Jacobson S, Cruickshank C, Hughes GJ, Siska C, Ory DS, Petrache I, Schaffer JE, Reisdorph N and Kechris K. Plasma sphingolipids associated with chronic obstructive pulmonary disease phenotypes. *American Journal of Respiratory and Critical Care Medicine*. 2015;191:275-84.

44.   Yano M, Watanabe K, Yamamoto T, Ikeda K, Senokuchi T, Lu M, Kadomatsu T, Tsukano H, Ikawa M, Okabe M, Yamaoka S, Okazaki T, Umehara H, Gotoh T, Song WJ, Node K, Taguchi R, Yamagata K and Oike Y. Mitochondrial dysfunction and increased reactive oxygen species impair insulin secretion in sphingomyelin synthase 1-null mice. *The Journal of Biological Chemistry*. 2011;286:3992-4002.

45.   Taniguchi M and Okazaki T. The role of sphingomyelin and sphingomyelin synthases in cell death, proliferation and migration-from cell and animal models to human disorders. *Biochim Biophys Acta*. 2014;1841:692-703.

46.   Ferrannini E, Natali A, Camastra S, Nannipieri M, Mari A, Adam KP, Milburn MV, Kastenmuller G, Adamski J, Tuomi T, Lyssenko V, Groop L and Gall WE. Early metabolic markers of the development of dysglycemia and type 2 diabetes and their physiological significance. *Diabetes*. 2013;62:1730-7.

47.   Cheng ML, Wang CH, Shiao MS, Liu MH, Huang YY, Huang CY, Mao CT, Lin JF, Ho HY and Yang NI. Metabolic Disturbances Identified in Plasma Are Associated With Outcomes in Patients With Heart Failure: Diagnostic and Prognostic Value of Metabolomics. *Journal of the American College of Cardiology*. 2015;65:1509-20.

48.   Wymann MP and Schneiter R. Lipid signalling in disease. *Nature Reviews Molecular Cell Biology*. 2008;9:162-76.

49.   Ventetuolo CE, Baird GL, Barr RG, Bluemke DA, Fritz JS, Hill NS, Klinger JR, Lima JA, Ouyang P, Palevsky HI, Palmisciano AJ, Krishnan I, Pinder D, Preston IR, Roberts KE and Kawut SM. Higher Estradiol and Lower Dehydroepiandrosterone-Sulfate Levels Are Associated with Pulmonary Arterial Hypertension in Men. *American Journal of Respiratory and Critical Care Medicine*. 2016;193:1168-75.

50.   Alzoubi A, Toba M, Abe K, O'Neill KD, Rocic P, Fagan KA, McMurtry IF and Oka M. Dehydroepiandrosterone restores right ventricular structure and function in rats with severe pulmonary arterial hypertension. *American Journal of Physiology: Heart and Circulatory Physiology*. 2013;304:H1708-18.

51.   Hoet PH and Nemery B. Polyamines in the lung: polyamine uptake and polyamine-linked pathological or toxicological conditions. *American Journal of Physiology: Heart and Circulatory Physiology*. 2000;278:L417-33.

52.   Olson JW, Atkinson JE, Hacker AD, Altiere RJ and Gillespie MN. Suppression of polyamine biosynthesis prevents monocrotaline-induced pulmonary edema and arterial medial thickening. *Toxicology and Applied Pharmacology*. 1985;81:91-9.

53.   Zhao YD, Yun HZ, Peng J, Yin L, Chu L, Wu L, Michalek R, Liu M, Keshavjee S, Waddell T, Granton J and de Perrot M. De novo synthesize of bile acids in pulmonary arterial hypertension lung. *Metabolomics*. 2014;10:1169-1175.

Table 1. Cohort Characteristics. Means and standard deviations or counts are given. BMI, body mass index; NT-proBNP, N-terminal brain natriuretic peptide at time of sample; RDW, red cell distribution width; HC, healthy controls; DC, disease controls; COPD, chronic obstructive pulmonary disease; CAD, coronary artery disease; IHD, ischaemic heart disease; AF, atrial fibrillations; PDE5, phosphodiesterase 5; ERA, endothelin receptor antagonists; CCB, calcium channel blocker; ACE, angiotensin converting enzyme. Ethnicity is shown for subjects who self-declared.

	Discovery				Validation 1			Validation 2
	HC (n=58)	DC (n=70)	PAH (19-70) (n=88)	PAH (>70) (n=28)	HC (n=63)	DC (n=69)	PAH (n=75)	PAH-2 (n=174)
Age at sampling, years	48+/-13.5	56.5+/-15.9	48.1+/-13.8	75.9+/-4.8	49.1+/-16.1	55.9+/-18.5	52+/-16.3	52+/-15
Sex, Female:Male (ratio)	38:20 (1.9:1)	48:22 (2.2:1)	61:27 (2.3:1)	19:9 (2.1:1)	40:23 (1.7:1)	48:21 (2.3:1)	52:23 (2.3:1)	127:47 (2.7:1)
Ethnicity, % non-Caucasian	32.5	56.9	17.0	14.8	38.1	44.1	18.6	8.1
BMI, kg/m²	30.5+/-10.5	27.9+/-6	28.7+/-8.2	27.9+/-5.7	26+/-4.1	27.4+/-5.9	30.1+/-7.2	29.2+/-7
Baseline haemodynamics at diagnosis								
Pulmonary capillary wedge pressure, mmHg		11.6+/-4.5	11.9+/-6	12.1+/-4.6		10.8+/-3.5	10.8+/-4.9	9.4+/-3.8
Mean pulmonary artery pressure, mmHg		19.3+/-4.1	53.1+/-14.3	46.3+/-14.6		19.1+/-4.7	53.7+/-11.1	56.2+/-15.2
Pulmonary vascular resistance, Woods units			12.3+/-5.7	8.6+/-4.7			12.4+/-5.8	13+/-6.4
Mean right atrial pressure, mmHg		7.3+/-3.4	10.0+/-5.7	9.3+/-4.9		6.3+/-3.2	11.8+/-5.8	9.2+/-5.3
Cardiac output, L/min			4.4+/-1.7	4.5+/-1.8			4.0+/-1.7	4.0+/-1.3
Functional status and pathology								
Six minute walk distance, m			279.4+/-153.4	197.9+/-158.7			271.3+/-169.5	334.8+/-119.0
WHO Functional Class, I/II/III/IV			2/11/65/10	0/2/23/1			0/7/52/9	2/24/119/18
RDW, %			14.8+/-2.1	15.0+/-1.1			16.0+/-2.9	14.7+/-3.3
NT-proBNP, pmol/L			735+/-882	1137+/-1123				895+/-1244
Creatinine, umol/L		76.3+/-21.8	81.2+/-29.1	107.6+/-35.7		87+/-40.6	92.7+/-32.6	89.9+/-25.1
Bilirubin, umol/L		13.9+/-13.9	15.6+/-10.7	12.4+/-8.5		16.3+/-22.8	17.2+/-10.4	12.1+/-9.0
Comorbidities								
Asthma/COPD		17.1	11.4	11.1		11.6	17.1	17.6
Diabetes		8.6	17.0	48.1		13.0	24.3	14.2
CAD/IHD		10.0	12.5	37.0		11.6	12.9	8.3
AF/flutter		14.3	18.2	25.9		17.4	8.6	5.9
Systemic hypertension		22.9	19.3	77.8		36.2	22.9	17.2
Hypercholesterolaemia/lipidaemia		12.9	12.5	22.2		10.1	15.7	6.5
Drug therapy								
Anticoagulation		32.9	71.6	78.6		34.8	65.3	66.7
PDE5 inhibitors		0.0	65.9	67.9		0.0	48.0	77.8
ERA		0.0	48.0	30.8		0.0	33.8	64.3
Diuretics		12.9	37.5	75.0		21.7	53.3	54.4

Aldosterone antagonists		7.1	31.8	21.4		1.4	34.7	8.2
Statins/lipid lowering drugs		25.7	22.7	57.1		33.3	24.0	24.0
CCB		10.0	14.8	28.6		24.6	14.7	21.1
Cardiac glycosides		7.1	17.0	14.3		8.7	21.3	9.4
Antidiabetic drugs		8.6	12.5	35.7		10.1	18.7	9.4
Prostanoids		0.0	9.1	0.0		0.0	6.7	33.3
Iron replacement therapy		4.3	13.6	25.0		5.8	4.0	10.5
ACE inhibitors		27.1	17.0	60.7		37.7	28.0	16.4

Figure Legends

Figure 1. Analysis flow chart. Summary of analytical workflow showing numbers of metabolites that distinguish PAH from controls, and/or are prognostic in PAH.

Figure 2. Metabolites which discriminate PAH and control subjects. (A) Average metabolite levels in PAH and control subjects for 20 metabolites found to significantly distinguish PAH and both healthy and disease controls, independent of potential confounders. Values plotted are z-scores calculated based on mean and standard deviation of all healthy volunteers in study - negative values indicate metabolites at lower levels in patients versus healthy controls and positive values indicate higher levels of metabolites in patients. For the discovery analysis only data from PAH patients aged 19-70 is plotted and for the validation analysis all patients data are shown. (B) Network analysis of the same 20 metabolites based on second order correlations. Line thickness indicates strength of correlations (all p<0.0001). *probable metabolite identity, but unconfirmed (see methods). EPE,eicosapentaenoyl; DHE,docosahexaenoyl; DPE,docosapentaenoyl; DHEA-S,dehydroisoandrosterone sulphate; GPC, glycerophosphocholine; SM, sphingomyelin.

Figure 3. Discriminant analysis models based on low numbers of metabolites distinguish PAH patients from controls. (A&C) Dotplots showing individual subjects' model scores in healthy controls (HC), PAH patients, vasoresponders and disease controls (DC) in discovery and validation analyses. Metabolites were selected by logistic regression of PAH-HC (A) and PAH-

26

DC (C) comparisons, respectively. (B&D) ROC curves showing performance of models in distinguishing PAH and HC (B) and DC (D) subjects.

Figure 4. Prognostic metabolites independent of established risk factors. (A) Hazard ratios after correcting for creatinine and diuretic use of 36 metabolites which were prognostic in PAH patients independent of RDW, NT-proBNP and six minute walk distance. Hazard ratios indicate the risk of a change in each metabolite of 1 standard deviation, for ease of comparison. Patients of all ages were included in both discovery and validation survival analyses. (B) Network analysis of the same 36 metabolites based on second order correlations. Line thickness indicates strength of correlations (all p<0.0001). Red lines indicate negative correlations. *probable metabolite identity, but unconfirmed (see methods). EPE,eicosapentaenoyl; DHE,docosahexaenoyl; DPE,docosapentaenoyl; DHEA-S,dehydroisoandrosterone sulphate; GPC, glycerophosphocholine; GPE, glycerophosphoethanolamine.

Figure 5. Hierarchical clustering of 19 discriminating and prognostic metabolites in PAH patients. (A) Venn diagram shows overlap between metabolites that discriminate PAH from healthy controls in all 3 cohorts, from logistic regression between PAH and healthy controls and prognostic metabolites in the discovery and first validation cohorts. (B) Clustering of the 19 overlapping metabolites from A. is shown between healthy controls (n=58), PAH survivors (n=110, alive at 3 years post-sample) and non-survivors (n=24) in the discovery analysis. Red indicates metabolite levels that are increased (and blue levels that are decreased) in PAH versus controls. *probable metabolite identity, but unconfirmed (see methods), ‡metabolites also

distinguish PAH from disease controls. GPC, glycerophosphocholine; EPE, eicosapentaenoyl; SM, sphingomyelin.

Figure 6. Analysis of serial samples. (A) ROC analysis of changes in metabolite levels and survival status at last follow-up. (B&D) Changes in individual patient metabolite levels, grouped by survival status at last follow-up. (C) Kaplan-Meier analysis illustrating survival over time in PAH patients divided into groups according to the changes in N-acetyl-methionine levels between serial samples.

Figure 7. Circulating angiogenin levels. (A) Plasma angiogenin levels determined by ELISA in healthy controls and PAH patients. (B) Scatter-plot of plasma N2,N2-dimethylguanosine versus plasma angiogenin in controls and PAH patients. Statistics shown are from Spearman's Rank test.

1416 metabolites quantified, 187
xenobiotics excluded, 686 metabolites
quantified in >95% of subjects

97 distinguish IPAH and
healthy subjects in 3
analyses (p<7.3e-5)

53 independent of
confounders – age, gender,
ethnicity, BMI, liver/renal
function, therapies (p<0.05)

62 prognostic, independent
of creatinine and diuretic
use, in PAH patients in 2
analyses (p<0.05)

*100 either distinguish
or prognostic*

*16 both distinguishing
and prognostic*

20 'PAH-specific' –
distinguish PAH from both
healthy and disease
controls, independent of
confounders (p<0.05)

36 independent of
established prognostic
markers NT-proBNP, 6
minute walk and RDW
(p<0.05)

A

Discriminant score based on 7 metabolites

| HC | PAH 19-70 | Vaso-responders | DC | HC | PAH | HC | PAH |

Discovery — *Validation1* — *Validation2*

B

Sensitivity vs 1 - Specificity

- Discovery
- Validation
- Validation-2

Analysis	AUC	Sig.	95% CI
Discovery	0.990	2E-22	0.979 - 1.000
Validation-1	0.950	2E-19	0.912 - 0.989
Validation-2	0.939	1E-37	0.914 - 0.964

C

Discriminant score based on 4 metabolites

| DC | PAH 19-70 | HC | DC | PAH | DC | PAH |

Discovery — *Validation1* — *Validation2*

D

Sensitivity vs 1 - Specificity

- Discovery
- Validation
- Validation-2

Analysis	AUC	Sig.	95% CI
Discovery	0.914	4E-18	0.869 - 0.958
Validation-1	0.723	2E-06	0.647 - 0.816
Validation-2	0.747	6E-14	0.692 - 0.801

A

Discovery | Validation

N6-succinyladenosine
N6-carbamoylthreonyladenosine
N1-methylinosine
N2,N2-dimethylguanosine
X - 24020
N-acetylmethionine
X - 24513
1-methylimidazoleacetate
X - 12472
4-acetamidobutanoate
pimeloylcarnitine/3-methyladipoylcarnitine
X - 12739
N-acetylalanine
N1-methyladenosine
X - 24527
X - 12688
N-formylmethionine
pseudouridine
N-acetylputrescine
X - 24728
X - 15503
X - 11564
urate
X - 24411
X - 11429
dehydroisoandrosterone sulfate (DHEA-S)
1-linoleoyl-2-docosahexaenoyl-GPC (18:2/22:6)*
1-oleoyl-2-docosapentaenoyl-GPC (18:1/22:5n6)*
1-eicosapentaenoyl-GPE (20:5)*
1-myristoyl-2-arachidonoyl-GPC (14:0/20:4)*
phosphatidylcholine (18:0/20:5, 16:0/22:5n6)*
1-palmitoyl-2-eicosapentaenoyl-GPC (16:0/20:5)*
1-linoleoyl-2-eicosapentaenoyl-GPC (18:2/20:5)*
1-eicosapentaenoyl-GPC (20:5)*
1-myristoyl-2-docosahexaenoyl-GPC (14:0/22:6)*
X - 24041

0.25 0.5 1 2 4 8 16   0.25 0.5 1 2 4 16 64

Hazard ratios, after correcting for creatinine and diuretic use

B

Amino acid metabolism
- ■ Alanine and Aspartate
- ■ Histidine
- ■ Met, Cys, SAM and Taur
- ■ Polyamine

Nucleoside metabolism
- ■ Purine, Adenine
- ■ Pyrimidine, Uracil

Unidentified metabolites
- ■ Unknown

Lipid metabolism
- ■ Acyl carnitine
- ■ Lysolipid
- ■ Phospholipid
- ■ Steroid

1: N1-methylinosine
2: 4-acetamidobutanoate
3: N-acetylalanine
4: X - 11564
5: X - 24513

X - 24728
N6-succinyl adenosine
N6-carbamoyl threonyladenosine
N-acetyl putrescine
N-acetyl methionine
N-formylmethionine
N1-methyl adenosine
X - 15503
X - 24020
X - 12688
pseudouridine
N2,N2-dimethyl-guanosine
urate
DHEA-S
1-methylimida-zoleacetate
X - 11429
pimeloyl/3-methyl adipoyl-carnitine
X - 12472
X - 24527
X - 12739
1-oleoyl-2-DPE-GPC (18:1/22:5n6)*
phosphatidylcholine (18:0/20:5, 16:0/22:5n6)*
1-myristoyl-2-arachidonoyl-GPC (14:0/20:4)*
1-myristoyl-2-DHE-GPC (14:0/22:6)*
X - 24041
1-palmitoyl-2-EPE-GPC (16:0/20:5)*
1-linoleoyl-2-EPE-GPC (18:2/20:5)*
1-EPE-GPE (20:5)*
1-linoleoyl-2-DHE-GPC (18:2/22:6)*

Discriminating metabolites

Metabolites from logistic regression

34

3

0

4

12

0

46

Prognostic metabolites

z-score

Survivor
Non-survivor
HC

Met/Cys/SAM/Tau — N-formylmethionine
N-acetylmethionine
Mevalonate — 3-hydroxy-3-methylglutarate ‡
Met/Cys/SAM/Tau — N-acetyltaurine
X – 24527
Unknown — X – 12739
X – 12472
Fatty Acid — oleoylcarnitine
TCA cycle — fumarate
Purine — N2,N2-dimethylguanosine ‡
N1-methylinosine ‡
Unknown — X – 24513
X – 12688 ‡
Met/Cys/SAM/Tau — methionine sulfone
Steroid — DHEA-S
Fatty Acid — palmitoylcholine
Phospholipid — 1-linoleoyl-2-EPE-GPC (18:2/20:5)* ‡
Sphingolipid — SM (d18:1/20:0, d16:1/22:0)* ‡
SM (d18:1/22:1, d18:2/22:0, d16:1/24:1)* ‡

A — ROC curve plot. Sensitivity vs 1 - Specificity with legend:
- 3-hydroxy-3-methylglutarate
- 3-hydroxybutyrylcarnitine (2)
- 4-acetamidobutanoate
- acetylcarnitine
- C-glycosyltryptophan
- hexadecanedioate
- malate
- N-acetylalanine
- N-acetylmethionine
- N-acetylputrescine
- N-acetyltaurine
- N-formylmethionine
- N1-methyladenosine
- N2,N2-dimethylguanosine
- 5,6-dihydrothymine
- fumarate
- Reference Line

B — Change in N2,N2-dimethylguanosine, Survivors vs Non-survivors, P<0.0001

C — Kaplan-Meier: Cum Survival vs Years survived since sample 2. "Decrease more than -0.2 standard deviations", "Increase more than -0.2 standard deviations", Change in N-acetyl-methionine, p<0.001

D — Change in N-acetylmethionine, Survivors vs Non-survivors, P<0.0001

Circulation

American Heart Association®

Plasma Metabolomics Implicate Modified Transfer RNAs and Altered Bioenergetics in the Outcome of Pulmonary Arterial Hypertension

Christopher J. Rhodes, Pavandeep Ghataorhe, John Wharton, Kevin C. Rue-Albrecht, Charaka Hadinnapola, Geoffrey Watson, Marta Bleda, Matthias Haimel, Gerry Coghlan, Paul A. Corris, Luke S. Howard, David G. Kiely, Andrew J. Peacock, Joanna Pepke-Zaba, Mark Toshner, Stephen J. Wort, J. Simon R. Gibbs, Allan Lawrie, Stefan Gräf, Nicholas W. Morrell and Martin R. Wilkins

The online version of this article, along with updated information and services, is located on the World Wide Web at:

http://circ.ahajournals.org/content/early/2016/11/21/CIRCULATIONAHA.116.024602

Free via Open Access

Data Supplement (unedited) at:

http://circ.ahajournals.org/content/suppl/2016/11/21/CIRCULATIONAHA.116.024602.DC1.html

SUPPLEMENTAL MATERIAL

Title: Plasma metabolomics implicate modified transfer RNAs and altered bioenergetics in the outcome of pulmonary arterial hypertension

Authors: Christopher J. Rhodes[1]†, PhD; Pavandeep Ghataorhe[1]†, BM BCh; John Wharton[1], PhD; Kevin C. Rue-Albrecht[1], PhD; Charaka Hadinnapola[2], BM BCh; Geoffrey Watson[1], BM BCh; Marta Bleda[2], PhD; Matthias. Haimel[2], BSc; Gerry Coghlan[3], MD; Paul A. Corris[4], FRCP; Luke. S. Howard[5,6], DPhil; David G. Kiely[7,8], MD; Andrew J. Peacock[9], MD; Joanna Pepke-Zaba[10], PhD; Mark R. Toshner[2,10], MD; S. John Wort[11], PhD; J. Simon R. Gibbs[5,6], MD; Allan Lawrie[8], PhD; Stefan Gräf[2,12], PhD; Nicholas W. Morrell[2], MD FMedSci; Martin R. Wilkins[1]*‡, MD FMedSci

Affiliations:

[1]Department of Medicine, Imperial College London, Hammersmith Campus, Du Cane Road, London, W12 0NN, UK

[2]Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge, CB2 OQQ, UK

[3]Cardiology Department, Royal Free Hospital, London, NW3 2QG, UK

[4]Institute of Cellular Medicine, Newcastle University and The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, NE7 7DN, UK

[5]National Pulmonary Hypertension Service, Imperial College Healthcare NHS Trust, Hammersmith Hospital, Du Cane Road, London, W12 0HS, UK

[6]National Heart and Lung Institute (NHLI), Imperial College London, Hammersmith Campus, Du Cane Road, London, W12 0NN, UK

[7]Sheffield Pulmonary Vascular Disease Unit, Royal Hallamshire Hospital, Sheffield, S10 2JF, UK

[8]Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield, Sheffield, S10 2RX, UK

[9]Scottish Pulmonary Vascular Unit, Golden Jubilee National Hospital, Glasgow, G81 4HX, UK

[10]Pulmonary Vascular Disease Unit, Papworth Hospital, Cambridge, CB23 3RE, UK

[11]Pulmonary Hypertension Service, Royal Brompton Hospital, London, SW3 6NP, UK

[12]Department of Haematology, University of Cambridge, Cambridge, CB2 0PT, UK

*Corresponding author: Dr Martin R Wilkins, MD FMedSci

Department of Medicine, Imperial College London, Hammersmith Campus, Du Cane Road,

London, W12 0NN, UK.

Email: m.wilkins@imperial.ac.uk,

Tel: 020 3313 2049

†These authors contributed equally to this work

‡On behalf of the NIHR BioResource – Rare Diseases (BRIDGE) PAH Consortium and the UK National PAH Cohort Study Consortium

Supplementary Materials

Metabolomic profiling methodology

Samples were prepared with use of an automated MicroLab STAR system (Hamilton

Company, Reno, NV, USA). For quality control (QC), a pooled sample from all experimental

samples was used throughout the experiment, and a mixture of Metabolon QC standards were

spiked into all experimental samples to monitor instrument performance and chromatographic

alignment. Samples were randomised prior to experimentation.

Experiments were conducted on Waters Acuity ultra-performance liquid chromatography

(UPLC) systems (Waters Corporation, Milford, MA, USA) using Thermo Scientific Q-

Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray

ionization (HESI-II) source and Orbitrap mass analyser (Thermo Fisher Scientific, MA, USA).

The analysis platform used four methods for Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy (UPLC-MS/MS) including a) positive ion mode electrospray ionisation (ESI), b) positive ion mode optimised for hydrophobic compounds, c) negative ion mode ESI and d) negative ionisation following elution from a hydrophilic interaction chromatography (HILIC) column. Scan time varied between methods and covered 70-1000m/z.

The resulting spectra were compared to the in-house Metabolon standard library using retention time, mass (m/z), adducts and MS/MS spectra. Analysis using this platform has been applied to measure metabolite levels in human plasma in control [1, 2] and disease populations [3, 4]. All experiments and runs were conducted on the same day.

Supplementary Materials

Table S1 (below). Metabolites distinguishing pulmonary arterial hypertension (PAH) from healthy (HC) and disease controls (DC). 97 metabolites that are significantly different between PAH and healthy controls in 3 cohorts (p<7.3e-5) are shown. Mean values are given and the data is scaled to the healthy control group. Significance from linear regression is shown (p value), and for metabolites with p>0.05 in PAH HC linear regression, the significant confounder is shown. Significance is also shown for Mann Whitney U test between PAH treatment naïve patients versus all HC, and PAH bone morphogenetic protein type 2 receptor (*BMPR2*) mutation carriers versus patients with no *BMPR2* mutation. GPC, glycerophosphocholine. *probable metabolite identity, but unconfirmed (see methods).

Metabolite	Metabolic Pathway	Group averages (z-score relative to healthy control levels)									Statistics (p-values) and major confounders				
		Discovery			Validation1			Validation2	PAH	BMPR2	Linear Regression		Confounder	Sub-analyses	
		HC	DC	PAH (19-70)	HC	DC	PAH	PAH2	Naïve	Mutation	HC vs PAH	HC+DC vs PAH	HC vs PAH	PAH Naïve vs HC	BMPR2 vs PAH
Increased in PAH vs HC and DC (independent of confounders)															
N-acetylaspartate (NAA)	Alanine & Aspartate	0.12	0.21	1.13	-0.11	0.64	1.32	0.87	1.47	0.93	2.3E-04	4.6E-03		3.0E-09	0.46
octadecanedioate	Fatty Acid, Dicarboxylate	0.07	0.12	0.82	-0.07	0.34	0.93	0.35	0.95	0.43	1.5E-02	1.3E-02		7.3E-06	0.25
3-hydroxy-3-methylglutarate	Mevalonate	-0.07	0.57	0.89	0.07	0.57	1.15	0.76	1.20	0.75	3.5E-02	3.0E-02		1.6E-09	0.12
acisoga	Polyamine	-0.10	1.07	1.42	0.09	1.14	1.95	1.22	1.93	1.04	1.9E-04	1.5E-02		2.7E-11	2.9E-02
N1-methylinosine	Purine, Hypo-Xanthine/Inosine	-0.22	0.58	1.50	0.20	0.86	1.85	1.29	1.78	1.50	8.1E-04	2.2E-02		1.5E-13	0.48
xanthine	Purine, Hypo-Xanthine/Inosine	0.11	0.56	0.98	-0.10	0.62	1.06	1.00	1.35	1.16	2.0E-07	2.8E-03		3.4E-09	0.19
N2,N2-dimethylguanosine	Purine, Guanine	-0.11	0.83	1.56	0.10	1.15	2.00	1.39	1.95	1.36	1.1E-02	3.9E-02		2.2E-13	0.11
3-ureidopropionate	Pyrimidine, Uracil	-0.14	0.22	0.62	0.13	0.09	1.16	0.54	1.16	0.47	1.7E-02	9.0E-04		4.0E-09	0.36
malate	TCA Cycle	-0.18	0.60	1.17	0.17	0.69	1.64	1.02	1.81	1.11	8.8E-04	9.1E-03		1.5E-14	0.35
X - 12688	Unknown	-0.11	1.01	1.50	0.10	1.14	1.70	1.26	1.76	0.99	4.9E-05	3.2E-02		6.4E-13	1.6E-02
X - 13737	Unknown	-0.05	0.19	0.97	0.05	0.38	1.03	0.91	1.15	0.94	1.8E-02	6.1E-03		1.9E-08	0.70
X - 21796	Unknown	0.07	0.19	0.96	-0.07	0.44	1.06	0.73	1.33	0.83	5.0E-03	1.3E-02		1.9E-12	0.83
Decreased in PAH vs HC and DC (independent of confounders)															
palmitoylcholine	Fatty Acid (Acyl Choline)	-0.05	-0.14	-1.33	0.05	-0.35	-0.73	-1.10	-0.94	-1.35	1.7E-03	8.3E-03		3.9E-05	4.8E-02
1-arachidonoyl-GPC (20:4n6)*	Lysolipid	-0.08	-0.14	-1.09	0.08	-0.25	-0.74	-0.81	-0.81	-1.00	1.2E-02	3.2E-02		5.9E-05	0.27
1-docosapentaenoyl-GPC (22:5n3)*	Lysolipid	-0.03	-0.17	-1.03	0.03	-0.21	-0.80	-0.84	-0.79	-1.02	9.0E-03	6.7E-03		1.6E-05	0.23
1-linoleoyl-2-eicosapentaenoyl-GPC (18:2/20:5)*	Phospholipid	0.03	-0.39	-0.81	-0.03	-0.13	-1.13	-0.81	-1.16	-0.77	2.4E-02	1.2E-02		3.4E-08	0.68
sphingomyelin (d18:1/20:0, d16:1/22:0)*	Sphingolipid	0.05	-0.43	-1.02	-0.05	-0.38	-1.09	-1.05	-1.24	-0.93	1.3E-02	7.7E-03		1.2E-08	0.49
sphingomyelin (d18:1/21:0, d17:1/22:0, d16:1/23:0)*	Sphingolipid	0.03	-0.44	-0.89	-0.03	-0.32	-0.93	-0.76	-1.04	-0.83	2.7E-03	4.0E-02		1.4E-08	0.78
sphingomyelin (d18:1/22:1, d18:2/22:0, d16:1/24:1)*	Sphingolipid	0.07	-0.37	-1.09	-0.06	-0.32	-1.16	-1.03	-1.21	-1.12	2.6E-03	7.2E-04		1.8E-07	0.95
sphingomyelin (d18:2/23:0, d18:1/23:1, d17:1/24:1)*	Sphingolipid	0.04	-0.27	-0.63	-0.04	-0.17	-0.60	-0.54	-0.68	-0.66	4.9E-04	4.4E-02		1.0E-06	0.40

Supplementary Materials

Increased in PAH vs HC (independent of confounders)

oleoyl ethanolamide	Endocannabinoid	0.10	1.02	0.88	-0.09	1.15	1.02	0.52	0.84	0.29	1.3E-03	2.2E-01	5.6E-05	5.4E-03
3-hydroxybutyrylcarnitine (1)	Fatty Acid (Acyl Carnitine)	0.02	0.52	0.91	-0.01	0.81	0.97	0.60	1.11	0.62	6.1E-05	1.1E-01	7.3E-09	4.5E-02
3-hydroxybutyrylcarnitine (2)	Fatty Acid (Acyl Carnitine)	0.07	0.76	1.26	-0.06	1.02	1.24	0.90	1.49	0.81	4.7E-03	6.8E-01	3.4E-09	3.8E-02
acetylcarnitine	Fatty Acid (Acyl Carnitine)	0.11	0.43	1.02	-0.10	0.77	0.88	0.59	0.94	0.56	1.2E-02	2.9E-01	5.7E-06	3.1E-02
adipoylcarnitine	Fatty Acid (Acyl Carnitine)	-0.04	0.41	1.44	0.04	0.94	1.69	1.25	1.77	1.20	4.6E-02	3.9E-01	5.8E-10	0.09
myristoleoylcarnitine*	Fatty Acid (Acyl Carnitine)	0.04	0.33	0.75	-0.04	0.64	0.85	0.53	0.91	0.37	1.6E-02	4.6E-01	6.2E-05	1.0E-02
myristoylcarnitine	Fatty Acid (Acyl Carnitine)	0.01	0.46	0.81	-0.01	0.87	1.06	0.60	1.07	0.44	4.1E-02	9.5E-01	9.4E-06	1.1E-02
oleoylcarnitine	Fatty Acid (Acyl Carnitine)	0.03	0.53	0.96	-0.03	0.83	1.39	1.20	1.22	0.87	1.8E-02	6.7E-01	3.2E-07	0.10
palmitoleoylcarnitine*	Fatty Acid (Acyl Carnitine)	0.08	0.56	0.91	-0.07	0.90	1.23	0.85	1.14	0.69	8.5E-03	8.9E-01	7.9E-07	4.1E-02
suberoylcarnitine	Fatty Acid (Acyl Carnitine)	0.10	0.55	1.40	-0.09	0.98	1.61	1.23	1.66	1.03	7.1E-03	3.5E-01	1.3E-09	2.9E-02
glutamate	Glutamate	0.16	0.47	1.13	-0.14	0.41	0.98	0.71	1.12	0.83	2.9E-02	1.4E-01	2.7E-07	0.98
methionine sulfone	Met, Cys, SAM & Tau	-0.03	0.49	1.57	0.03	0.98	1.79	1.40	1.56	1.29	3.6E-02	4.9E-01	1.8E-09	0.14
N-acetylmethionine	Met, Cys, SAM & Tau	-0.04	0.48	1.08	0.04	0.63	1.27	0.60	1.30	0.80	3.3E-03	1.6E-01	1.3E-09	0.22
N-acetyltaurine	Met, Cys, SAM & Tau	0.07	0.56	1.02	-0.06	0.71	1.15	0.63	1.15	0.37	1.0E-03	7.1E-02	2.7E-07	9.7E-04
N-formylmethionine	Met, Cys, SAM & Tau	-0.17	0.49	1.08	0.16	0.69	1.53	0.78	1.59	0.85	4.8E-02	3.0E-01	5.8E-10	0.10
5,6-dihydrothymine	Pyrimidine, Thymine	-0.22	0.65	1.33	0.20	0.96	1.37	0.62	1.38	0.71	2.0E-04	6.7E-02	1.2E-08	0.15
alpha-ketoglutarate	TCA Cycle	-0.04	0.45	0.68	0.04	0.71	1.57	1.13	1.61	1.14	2.2E-02	2.4E-01	2.6E-13	1.00
fumarate	TCA Cycle	-0.15	0.30	0.65	0.13	0.45	1.14	0.56	1.23	0.31	2.1E-02	9.8E-02	3.5E-16	0.28
C-glycosyltryptophan	Tryptophan	-0.05	0.55	1.11	0.05	0.84	1.37	0.76	1.30	0.59	4.8E-02	3.9E-01	5.9E-07	4.7E-03
X - 12127	Unknown	0.06	0.63	0.85	-0.06	0.61	1.05	0.67	1.07	0.73	1.6E-02	4.4E-01	1.6E-08	0.53
X - 12472	Unknown	-0.04	0.71	1.03	0.04	0.82	1.05	0.96	1.15	0.93	3.4E-03	8.2E-01	2.5E-07	0.59
X - 12739	Unknown	-0.09	0.87	1.02	0.08	1.11	1.12	0.99	1.18	0.97	2.4E-02	1.8E-01	6.8E-07	0.35
X - 12824	Unknown	-0.02	0.54	0.85	0.02	0.74	1.08	0.79	1.16	0.78	2.0E-02	6.8E-01	9.5E-08	0.44
X - 17327	Unknown	0.03	0.66	0.86	-0.03	0.81	0.97	0.82	1.02	0.78	2.6E-02	7.0E-01	2.4E-06	0.34
X - 21829	Unknown	-0.07	0.44	0.82	0.06	0.65	1.24	0.66	1.41	0.62	2.5E-02	5.6E-01	2.4E-10	0.10
X - 24307	Unknown	0.04	0.68	0.70	-0.04	0.52	1.18	0.97	1.09	0.94	6.9E-05	7.2E-02	1.2E-07	0.88

Supplementary Materials

X - 24513	Unknown	-0.10	0.67	1.05	0.09	0.92	1.44	0.98	1.30	0.77	2.4E-02	4.8E-01		1.4E-07	9.1E-03
X - 24527	Unknown	-0.08	0.86	1.11	0.08	1.17	1.22	1.06	1.22	0.94	3.5E-02	1.9E-01		2.2E-06	0.25
X - 24678	Unknown	0.21	0.74	1.20	-0.19	0.62	1.34	0.70	1.33	0.83	4.9E-03	1.2E-01		3.6E-11	0.42
X - 24766	Unknown	-0.31	0.45	0.82	0.28	0.60	1.07	0.85	0.89	0.92	1.9E-02	3.9E-01		5.1E-07	0.78
Decreased in PAH vs HC (independent of confounders)															
asparagine	Alanine & Aspartate	-0.06	-0.58	-0.88	0.05	-0.55	-1.00	-0.90	-1.22	-0.95	8.7E-04	8.5E-02		2.3E-07	0.64
dehydroisoandrosterone sulfate (DHEA-S)	Steroid	-0.05	-1.17	-1.53	0.04	-0.95	-1.74	-1.77	-1.53	-1.56	2.5E-02	4.1E-01		1.3E-09	0.46
X - 23765	Unknown	0.09	-0.50	-0.72	-0.08	-0.50	-0.90	-0.90	-1.02	-1.08	8.1E-04	2.0E-01		9.7E-07	0.15
Increased in PAH vs HC															
N-acetylalanine	Alanine & Aspartate	-0.11	0.55	0.76	0.10	0.88	1.15	0.66	1.04	0.50	9.9E-01	2.3E-01	Diuretics	3.2E-05	1.6E-02
N-acetylneuraminate	Aminosugar	-0.11	0.50	0.93	0.11	0.63	0.88	0.47	0.76	0.28	7.4E-02	2.4E-01	Age	1.8E-04	1.0E-03
erythronate*	Aminosugar	-0.11	0.45	0.71	0.10	0.68	1.16	0.59	1.10	0.37	7.1E-01	6.5E-01	Bilirubin	5.4E-06	6.0E-03
N-acetylglucosamine/N-acetylgalactosamine	Aminosugar	-0.14	0.50	0.72	0.13	0.89	0.95	0.58	0.62	0.54	1.0E-01	7.3E-01	Gender	1.8E-03	4.4E-02
gulonic acid*	Ascorbate & Aldarate	-0.08	0.49	0.61	0.07	0.57	0.79	0.61	0.81	0.47	8.5E-01	5.3E-01	Age	4.5E-05	2.8E-02
palmitoylcarnitine	Fatty Acid (Acyl Carnitine)	-0.03	0.49	0.91	0.03	1.03	1.28	0.95	1.23	0.90	6.0E-02	9.2E-01	Bilirubin	1.6E-07	0.43
malonylcarnitine	Fatty Acid Synthesis	0.00	0.20	0.56	0.00	0.29	0.94	0.48	0.96	0.57	1.3E-01	3.8E-02	Bilirubin	3.8E-09	0.53
N-acetylserine	Glycine, Serine & Threonine	-0.16	0.63	1.00	0.15	1.04	1.49	0.97	1.38	0.85	5.2E-01	4.0E-01	Bilirubin	6.9E-08	0.05
N-acetylthreonine	Glycine, Serine & Threonine	-0.22	0.31	0.68	0.20	0.64	0.99	0.55	1.10	0.47	6.5E-01	9.7E-01	Bilirubin	4.4E-07	2.0E-02
1-methylimidazoleacetate	Histidine	-0.05	0.52	0.86	0.04	0.50	1.32	0.86	1.02	0.74	6.0E-02	9.5E-02	Age	7.9E-07	0.08
imidazole propionate	Histidine	-0.21	0.44	0.84	0.19	0.71	0.99	0.79	0.98	0.84	1.4E-01	8.1E-01	Bilirubin	1.1E-07	0.54
quinolinate	Nicotinate & Nicotinamide	-0.11	0.46	0.76	0.10	0.68	1.19	0.86	0.95	0.58	6.0E-01	5.0E-01	Bilirubin	1.8E-04	3.8E-02
vanillylmandelate (VMA)	Phenylalanine & Tyrosine	-0.21	0.35	1.00	0.19	0.37	1.52	0.83	1.75	0.96	2.3E-01	1.5E-02	Age	6.3E-12	0.37
4-acetamidobutanoate	Polyamine	-0.10	0.72	1.22	0.09	0.90	1.93	1.26	1.83	1.19	8.6E-01	7.0E-01	Bilirubin	1.0E-11	0.15
N-acetylputrescine	Polyamine	-0.16	0.66	0.84	0.15	0.72	1.32	0.66	0.93	0.53	5.8E-01	6.2E-01	Bilirubin	3.0E-06	3.8E-02
N6-carbamoylthreonyl adenosine	Purine, Adenine	-0.12	0.58	1.04	0.11	0.90	1.49	1.06	1.31	0.92	1.7E-01	4.5E-01	Age	8.4E-09	4.6E-02
N1-methyladenosine	Purine, Adenine	-0.21	0.70	0.94	0.19	0.91	1.26	0.65	1.03	0.46	8.6E-01	1.4E-01	Bilirubin	2.4E-05	4.4E-03
N6-succinyladenosine	Purine, Adenine	-0.06	0.38	0.63	0.05	0.43	0.86	0.50	0.94	0.44	1.5E-01	2.7E-01	Bilirubin	1.2E-09	2.7E-03
N4-acetylcytidine	Pyrimidine, Cytidine	0.07	0.55	1.31	-0.06	0.77	1.50	1.22	1.33	1.19	2.7E-01	6.7E-01	Bilirubin	1.3E-08	0.47
orotidine	Pyrimidine, Orotate	-0.11	0.58	1.07	0.10	0.95	1.27	0.96	1.27	0.83	2.8E-01	5.3E-01	Bilirubin	1.4E-08	0.08
pseudouridine	Pyrimidine, Uracil	-0.08	0.77	1.22	0.08	1.05	1.66	1.17	1.68	0.97	1.5E-01	8.6E-01	Bilirubin	5.2E-11	1.4E-02

Supplementary Materials

Metabolite	Pathway													Factor		
kynurenine	Tryptophan	-0.14	0.29	1.00	0.13	0.57	1.37	0.77	1.44	0.68	5.6E-01	3.3E-01	Bilirubin	1.2E-08	0.07
X - 12100	Unknown	-0.13	0.49	1.04	0.12	0.85	1.26	0.88	1.28	0.75	1.3E-01	4.6E-01	Age	9.8E-08	0.06
X - 11564	Unknown	-0.13	0.51	0.94	0.12	0.90	1.48	0.83	1.43	0.73	9.6E-01	8.8E-01	Bilirubin	3.5E-07	0.06
X - 12026	Unknown	-0.16	0.69	1.58	0.14	1.01	2.01	1.36	2.20	1.26	6.9E-01	7.7E-01	Bilirubin	3.8E-12	0.07
X - 12117	Unknown	-0.12	0.66	1.18	0.11	0.98	1.55	1.35	1.50	1.20	6.7E-01	3.6E-01	Bilirubin	7.7E-09	0.19
X - 15503	Unknown	-0.17	0.28	1.04	0.15	0.70	1.64	1.06	1.68	1.33	6.3E-01	5.6E-02	Bilirubin	2.5E-10	0.63
X - 11429	Unknown	-0.07	0.88	1.52	0.07	1.31	1.98	1.45	1.88	1.18	9.1E-02	8.4E-01	Diuretics	3.4E-10	2.0E-02
X - 21736	Unknown	-0.06	0.33	1.04	0.06	0.59	1.46	0.92	1.57	0.95	2.7E-01	3.5E-01	Diuretics	1.1E-10	0.42
pro-hydroxy-pro	Urea cycle; Arginine & Proline	-0.03	0.32	0.87	0.03	0.51	1.24	0.90	1.07	0.98	9.6E-01	6.9E-01	Diuretics	2.9E-06	0.95
Decreased in PAH vs HC																
histidine	Histidine	-0.07	-0.74	-1.57	0.06	-0.91	-1.89	-1.71	-1.78	-1.60	6.1E-02	8.2E-03	Prostanoids	1.2E-11	4.2E-01
1-linoleoyl-GPC (18:2)	Lysolipid	-0.10	-0.59	-1.22	0.09	-0.51	-1.14	-1.12	-1.33	-1.10	5.8E-02	2.4E-02	Bilirubin	1.3E-07	5.2E-01
2-linoleoyl-GPC (18:2)*	Lysolipid	-0.11	-0.55	-1.23	0.10	-0.42	-0.97	-1.09	-1.35	-1.15	9.4E-02	2.9E-02	Bilirubin	4.5E-08	9.6E-01
1-dihomo-linoleoyl-GPC (20:2)*	Lysolipid	0.00	-0.39	-0.95	0.00	-0.49	-0.86	-1.14	-0.98	-1.30	2.4E-01	4.3E-01	Statins	5.5E-05	1.2E-01
1-(1-enyl-palmitoyl)-2-linoleoyl-GPC (P-16:0/18:2)*	Plasmalogen	0.01	-0.38	-1.06	-0.01	-0.17	-1.13	-0.98	-1.28	-1.08	2.3E-01	1.5E-01	PDE5 inhib	5.1E-08	9.4E-01
behenoyl sphingomyelin (d18:1/22:0)*	Sphingolipid	0.02	-0.22	-0.75	-0.01	-0.20	-0.71	-0.92	-0.78	-0.77	1.4E-01	3.0E-02	Gender	3.6E-06	8.4E-01
4-androsten-3beta,17beta-diol disulfate (1)	Steroid	-0.04	-0.64	-0.87	0.03	-0.47	-0.81	-1.14	-0.55	-1.06	2.4E-01	5.4E-01	Gender	4.0E-03	6.8E-01
4-androsten-3beta,17beta-diol monosulfate (1)	Steroid	-0.01	-0.85	-1.15	0.01	-0.68	-1.30	-1.40	-1.18	-1.32	5.6E-02	3.7E-01	Gender	5.8E-07	9.3E-01
androsterone sulfate	Steroid	-0.02	-0.70	-1.33	0.02	-0.69	-1.40	-1.40	-1.14	-1.33	1.5E-01	2.0E-01	Gender	4.6E-06	7.3E-01
epiandrosterone sulfate	Steroid	-0.02	-0.86	-1.42	0.01	-0.72	-1.50	-1.49	-1.23	-1.40	1.3E-01	3.9E-01	Gender	1.1E-06	6.4E-01
pregn steroid monosulfate*	Steroid	-0.08	-0.83	-0.88	0.07	-0.68	-0.93	-1.04	-0.71	-0.84	1.0E-01	6.7E-01	Gender	3.3E-04	2.6E-01
X - 23749	Unknown	0.04	-0.16	-1.10	-0.03	-0.34	-1.05	-1.12	-0.94	-1.41	2.0E-01	2.7E-01	BMI	5.4E-05	0.02
arginine	Urea cycle; Arginine & Proline	-0.05	-0.67	-1.05	0.04	-0.79	-1.40	-1.47	-1.80	-1.32	1.1E-01	3.1E-01	DM drugs	2.1E-11	6.5E-01
homoarginine	Urea cycle; Arginine & Proline	-0.05	-0.70	-0.91	0.05	-0.59	-1.13	-1.06	-1.11	-0.86	5.1E-02	2.1E-01	Ethnicity	4.8E-07	7.6E-02

Supplementary Materials

Table S1. Metabolites distinguishing pulmonary arterial hypertension (PAH) from healthy (HC) and disease controls (DC). 97 metabolites that are significantly different between PAH and healthy controls in 3 cohorts (p<7.3e-5) are shown. Mean values are given and the data is scaled to the healthy control group. Significance from linear regression is shown (p value), and for metabolites with p>0.05 in PAH HC linear regression, the significant confounder is shown. Significance is also shown for Mann Whitney U test between PAH treatment naïve patients versus all HC, and PAH bone morphogenetic protein type 2 receptor (*BMPR2*) mutation carriers versus patients with no *BMPR2* mutation. GPC, glycerophosphocholine. *probable metabolite identity, but unconfirmed (see methods).

Table S2. (below) Survival analysis.

Metabolite	Metabolic Pathway	Discovery		Validation1	
		Hazard Ratio	Sig	Hazard Ratio	Sig
Higher value indicates mortality. Independent of established prognostic markers					
N-acetylalanine	Alanine and Aspartate	2.02 (1.22-3.36)	6.43E-03	2.08 (1.12-3.86)	2.04E-02
pimeloylcarnitine/3-methyladipoylcarnitine	Fatty Acid (Acyl Carnitine)	2.16 (1.28-3.66)	4.03E-03	2.52 (1.24-5.10)	1.04E-02
1-methylimidazoleacetate	Histidine	2.26 (1.37-3.73)	1.43E-03	1.74 (1.03-2.93)	3.93E-02
N-acetylmethionine	Methionine, Cysteine, SAM and Taurine	2.36 (1.41-3.96)	1.16E-03	2.29 (1.18-4.43)	1.44E-02
N-formylmethionine	Methionine, Cysteine, SAM and Taurine	1.79 (1.20-2.68)	4.50E-03	1.98 (1.20-3.25)	7.15E-03
4-acetamidobutanoate	Polyamine	2.20 (1.45-3.35)	2.19E-04	2.02 (1.30-3.14)	1.83E-03
N-acetylputrescine	Polyamine	1.74 (1.04-2.91)	3.54E-02	2.92 (1.51-5.66)	1.50E-03
N1-methylinosine	Purine , (Hypo)Xanthine/Inosine	2.82 (1.74-4.57)	2.42E-05	1.73 (1.09-2.77)	2.11E-02
urate	Purine , (Hypo)Xanthine/Inosine	1.61 (1.06-2.42)	2.43E-02	2.14 (1.26-3.64)	4.73E-03
N6-succinyladenosine	Purine , Adenine	3.89 (1.40-10.82)	9.18E-03	8.31 (1.94-35.54)	4.29E-03
N6-carbamoylthreonyladenosine	Purine , Adenine	3.10 (1.60-6.00)	8.04E-04	2.08 (1.08-4.00)	2.81E-02
N1-methyladenosine	Purine , Adenine	1.94 (1.25-3.01)	2.92E-03	1.93 (1.12-3.32)	1.75E-02
N2,N2-dimethylguanosine	Purine , Guanine	2.53 (1.57-4.08)	1.35E-04	1.86 (1.14-3.03)	1.25E-02
pseudouridine	Pyrimidine , Uracil	1.78 (1.07-2.94)	2.54E-02	2.75 (1.48-5.12)	1.45E-03
X - 24020	Unknown	2.47 (1.42-4.30)	1.36E-03	1.84 (1.00-3.39)	4.94E-02
X - 24513	Unknown	2.34 (1.28-4.29)	5.92E-03	2.18 (1.11-4.29)	2.34E-02
X - 12472	Unknown	2.24 (1.44-3.47)	3.11E-04	1.57 (1.00-2.46)	4.97E-02
X - 12739	Unknown	2.07 (1.41-3.04)	2.07E-04	1.56 (1.06-2.28)	2.28E-02
X - 24527	Unknown	1.85 (1.38-2.48)	4.09E-05	1.43 (1.03-1.99)	3.12E-02
X - 12688	Unknown	1.81 (1.23-2.68)	2.83E-03	2.02 (1.27-3.19)	2.80E-03
X - 24728	Unknown	1.70 (1.09-2.64)	1.96E-02	2.04 (1.09-3.79)	2.50E-02
X - 15503	Unknown	1.67 (1.11-2.52)	1.38E-02	1.51 (1.03-2.22)	3.48E-02
X - 11564	Unknown	1.62 (1.10-2.38)	1.49E-02	1.60 (1.03-2.47)	3.51E-02

Supplementary Materials

X - 24411	Unknown	1.54 (1.04-2.27)	3.05E-02	2.16 (1.32-3.54)	2.17E-03
X - 11429	Unknown	1.47 (1.05-2.06)	2.65E-02	1.79 (1.18-2.71)	5.76E-03
Lower value indicates mortality. Independent of established prognostic markers					
1-eicosapentaenoyl-GPE (20:5)*	Lysolipid	0.60 (0.43-0.84)	2.65E-03	0.62 (0.39-1.00)	4.98E-02
1-eicosapentaenoyl-GPC (20:5)*	Lysolipid	0.47 (0.31-0.73)	6.30E-04	0.50 (0.31-0.82)	6.05E-03
1-linoleoyl-2-docosahexaenoyl-GPC (18:2/22:6)*	Phospholipid	0.64 (0.46-0.89)	7.40E-03	0.65 (0.47-0.91)	1.27E-02
1-oleoyl-2-docosapentaenoyl-GPC (18:1/22:5n6)*	Phospholipid	0.64 (0.47-0.87)	4.11E-03	0.62 (0.43-0.90)	1.06E-02
1-myristoyl-2-arachidonoyl-GPC (14:0/20:4)*	Phospholipid	0.57 (0.39-0.83)	3.68E-03	0.68 (0.48-0.96)	2.66E-02
phosphatidylcholine (18:0/20:5, 16:0/22:5n6)*	Phospholipid	0.54 (0.37-0.77)	7.30E-04	0.42 (0.25-0.72)	1.44E-03
1-palmitoyl-2-eicosapentaenoyl-GPC (16:0/20:5)*	Phospholipid	0.53 (0.37-0.75)	4.11E-04	0.43 (0.26-0.74)	1.96E-03
1-linoleoyl-2-eicosapentaenoyl-GPC (18:2/20:5)*	Phospholipid	0.48 (0.33-0.70)	1.38E-04	0.60 (0.41-0.88)	8.59E-03
1-myristoyl-2-docosahexaenoyl-GPC (14:0/22:6)*	Phospholipid	0.47 (0.31-0.72)	4.84E-04	0.49 (0.29-0.81)	5.62E-03
dehydroisoandrosterone sulfate (DHEA-S)	Steroid	0.71 (0.52-0.97)	3.19E-02	0.55 (0.33-0.91)	2.06E-02
X - 24041	Unknown	0.46 (0.29-0.72)	6.88E-04	0.61 (0.38-0.99)	4.45E-02
Higher value indicates mortality					
hexadecanedioate	Fatty Acid, Dicarboxylate	1.53 (1.09-2.14)	1.40E-02	1.67 (1.16-2.41)	5.58E-03
N-acetyltaurine	Methionine, Cysteine, SAM and Taurine	1.97 (1.16-3.35)	1.19E-02	2.24 (1.13-4.43)	2.05E-02
3-hydroxy-3-methylglutarate	Mevalonate	1.68 (1.03-2.74)	3.78E-02	2.08 (1.02-4.23)	4.40E-02
vanillylmandelate (VMA)	Phenylalanine and Tyrosine	1.48 (1.06-2.06)	2.02E-02	1.58 (1.06-2.34)	2.38E-02
N-acetyl-beta-alanine	Pyrimidine , Uracil	1.54 (1.08-2.21)	1.80E-02	1.65 (1.18-2.32)	3.65E-03
fumarate	TCA Cycle	1.70 (1.00-2.88)	4.97E-02	2.57 (1.09-6.06)	3.09E-02
N2,N5-diacetylornithine	Urea cycle; Arginine and Proline	2.15 (1.27-3.63)	4.37E-03	2.15 (1.13-4.09)	1.95E-02
Lower value indicates mortality					
valine	Leucine, Isoleucine and Valine	0.66 (0.49-0.89)	6.80E-03	0.68 (0.48-0.95)	2.21E-02
1-docosahexaenoyl-GPE (22:6)*	Lysolipid	0.68 (0.49-0.94)	1.86E-02	0.54 (0.34-0.85)	8.27E-03
1-dihomo-linolenoyl-GPC (20:3n3 or 6)*	Lysolipid	0.57 (0.41-0.79)	7.43E-04	0.71 (0.53-0.96)	2.82E-02
1-docosahexaenoyl-GPC (22:6)*	Lysolipid	0.54 (0.35-0.81)	3.17E-03	0.57 (0.35-0.93)	2.32E-02
phosphatidylcholine (16:0/22:5n3, 18:1/20:4)*	Phospholipid	0.68 (0.48-0.96)	2.64E-02	0.59 (0.36-0.98)	4.16E-02
1-stearoyl-2-docosahexaenoyl-GPE (18:0/22:6)*	Phospholipid	0.64 (0.47-0.88)	5.99E-03	0.52 (0.32-0.85)	8.73E-03
1-palmitoleoyl-2-docosahexaenoyl-GPC (16:1/22:6)*	Phospholipid	0.62 (0.43-0.90)	1.09E-02	0.50 (0.28-0.88)	1.67E-02

Supplementary Materials

1-pentadecanoyl-2-docosahexaenoyl-GPC (15:0/22:6)*	Phospholipid	0.62 (0.42-0.90)	1.30E-02	0.34 (0.17-0.68)	1.96E-03
1-stearoyl-2-docosahexaenoyl-GPC (18:0/22:6)	Phospholipid	0.60 (0.42-0.87)	7.16E-03	0.45 (0.27-0.76)	2.76E-03
1-stearoyl-2-docosapentaenoyl-GPC (18:0/22:5n3)*	Phospholipid	0.58 (0.41-0.82)	2.11E-03	0.50 (0.30-0.83)	8.18E-03
1-palmitoyl-2-dihomo-linolenoyl-GPC (16:0/20:3n3 or 6)*	Phospholipid	0.58 (0.38-0.88)	9.67E-03	0.54 (0.35-0.84)	6.47E-03
1-(1-enyl-stearoyl)-2-docosahexaenoyl-GPC (P-18:0/22:6)*	Plasmalogen	0.56 (0.37-0.85)	6.20E-03	0.57 (0.33-0.98)	4.06E-02
1-(1-enyl-stearoyl)-2-docosahexaenoyl-GPE (P-18:0/22:6)*	Plasmalogen	0.55 (0.37-0.82)	3.45E-03	0.49 (0.27-0.91)	2.31E-02
1-(1-enyl-palmitoyl)-2-docosahexaenoyl-GPC (P-16:0/22:6)*	Plasmalogen	0.54 (0.37-0.79)	1.31E-03	0.53 (0.32-0.87)	1.30E-02
1-(1-enyl-palmitoyl)-2-docosahexaenoyl-GPE (P-16:0/22:6)*	Plasmalogen	0.49 (0.34-0.71)	1.63E-04	0.47 (0.26-0.82)	8.30E-03
sphingomyelin (d18:2/14:0, d18:1/14:1)*	Sphingolipid	0.73 (0.53-1.00)	4.83E-02	0.48 (0.30-0.77)	2.36E-03
sphingomyelin (d18:1/20:0, d16:1/22:0)*	Sphingolipid	0.60 (0.41-0.87)	6.60E-03	0.57 (0.35-0.93)	2.32E-02
sphingomyelin (d18:1/22:1, d18:2/22:0, d16:1/24:1)*	Sphingolipid	0.59 (0.42-0.83)	2.51E-03	0.53 (0.33-0.85)	7.63E-03
behenoyl sphingomyelin (d18:1/22:0)*	Sphingolipid	0.42 (0.27-0.64)	5.94E-05	0.47 (0.27-0.82)	7.42E-03

Table S2. Survival analysis in PAH. 62 metabolites significantly different between PAH survivors and non-survivors in discovery and validation1 cohorts (p<0.05) are shown. Hazard ratio and significance (Sig) is shown from Cox regression analysis. Hazard ratios shown are for 1 standard deviation change in each metabolite for ease of comparison. Metabolites which are also independent of established prognostic markers in the discovery cohort are also shown. *probable metabolite identity, but unconfirmed (see methods). GPC, glycerophosphocholine; GPE, glycerophosphoethanolamine.

Pathway	Result	Metabolite	HC v PAH	Prog-nosis
Fatty Acid Metabolism(Acyl Carnitine)	4.5E-05	3-hydroxybutyrylcarnitine (1)	*6E-05*	0.3499
		pimeloylcarnitine/3-methyladipoylcarnitine	NA	*0.0006*
		suberoylcarnitine	*0.0071*	*0.0044*
		3-hydroxybutyrylcarnitine (2)	*0.0047*	*0.0046*
		palmitoleoylcarnitine*	*0.0085*	0.3042
		adipoylcarnitine	*0.0464*	*0.0106*
		acetylcarnitine	*0.0118*	0.141
		myristoleoylcarnitine*	*0.0165*	0.6142
		oleoylcarnitine	*0.0176*	0.0598
		myristoylcarnitine	*0.0412*	0.2631
		palmitoylcarnitine	0.0604	0.1959
		linoleoylcarnitine*	NA	0.0615
		hexanoylcarnitine	NA	0.5981
		stearoylcarnitine	NA	0.6611
		octanoylcarnitine	NA	0.8074
		laurylcarnitine	NA	0.8641
		decanoylcarnitine	NA	0.912
		cis-4-decenoyl carnitine	NA	0.9541
Polyamine Metabolism	0.01079	N-acetylputrescine	0.5825	*2E-05*
		4-acetamidobutanoate	0.8618	*3E-05*
		acisoga	*0.0002*	*0.0003*
		5-methylthioadenosine (MTA)	NA	*0.0433*
Alanine and Aspartate Metabolism	0.02412	N-acetylaspartate (NAA)	*0.0002*	0.2634
		asparagine	*0.0009*	0.9422
		N-acetylalanine	0.9901	*0.009*
		alanine	NA	0.6529
		aspartate	NA	0.9677
Purine Metabolism, (Hypo)Xanthine/Inosine containing	0.04317	xanthine	*2E-07*	0.438
		urate	NA	*0.0004*
		N1-methylinosine	*0.0008*	*0.0144*
		AICA ribonucleotide	NA	*0.0066*
		allantoin	NA	0.0934
		hypoxanthine	NA	0.6652
Purine Metabolism, Adenine containing	0.04317	N6-carbamoylthreonyladenosine	0.1735	*7E-05*
		N6-succinyladenosine	0.1452	*0.0005*
		N1-methyladenosine	0.8632	*0.0006*
		adenine	NA	0.4835
		adenosine	NA	0.5677
		adenosine 5'-monophosphate (AMP)	NA	0.5988
Pyrimidine Metabolism, Uracil containing	0.04317	pseudouridine	0.1506	*6E-05*
		N-acetyl-beta-alanine	NA	*0.0002*
		3-ureidopropionate	*0.0175*	*0.0024*
		uridine	NA	0.2684
		5-methyluridine (ribothymidine)	NA	0.5196
		beta-alanine	NA	0.8665

Table S3. Pathway enrichment analysis results. Pathways analysed and enrichment p-values are given, as well as metabolites within each pathway and significance values from tests used to select metabolites considered to be disease-associated. HC v PAH, significance of PAH on metabolite levels after controlling for potential confounders by linear regression; Prognosis, weakest significance of metabolite association with survival by Cox analysis in discovery or validation cohorts. *probable metabolite identity, but unconfirmed (see methods).

Supplementary Materials

Metabolite	Area	Sig.	95% CI
3-hydroxy-3-methylglutarate	0.75	2.05E-04	0.63 - 0.86
3-hydroxybutyrylcarnitine (2)	0.67	1.10E-02	0.54 - 0.79
4-acetamidobutanoate	0.67	8.40E-03	0.56 - 0.79
5,6-dihydrothymine	0.67	1.04E-02	0.55 - 0.79
acetylcarnitine	0.65	2.80E-02	0.52 - 0.77
C-glycosyltryptophan	0.66	1.82E-02	0.53 - 0.78
fumarate	0.67	1.01E-02	0.55 - 0.79
hexadecanedioate	0.65	2.32E-02	0.52 - 0.78
malate	0.69	3.32E-03	0.58 - 0.81
N1-methyladenosine	0.69	3.62E-03	0.57 - 0.81
N2,N2-dimethylguanosine	0.69	4.70E-03	0.57 - 0.80
N-acetylalanine	0.67	1.01E-02	0.55 - 0.79
N-acetylmethionine	0.76	6.94E-05	0.66 - 0.87
N-acetylputrescine	0.69	4.07E-03	0.57 - 0.81
N-acetyltaurine	0.68	6.40E-03	0.56 - 0.80
N-formylmethionine	0.67	8.63E-03	0.56 - 0.79
X - 11564	0.65	2.16E-02	0.53 - 0.77
X - 12127	0.65	2.43E-02	0.53 - 0.77
X - 12472	0.65	2.32E-02	0.53 - 0.77
X - 12688	0.68	5.26E-03	0.57 - 0.80
X - 12739	0.68	7.75E-03	0.56 - 0.79
X - 13737	0.66	1.45E-02	0.54 - 0.78
X - 15503	0.68	5.89E-03	0.56 - 0.80
X - 21796	0.70	3.03E-03	0.59 - 0.80
X - 24020	0.67	1.15E-02	0.55 - 0.78
X - 24527	0.67	8.63E-03	0.55 - 0.79
X - 24766	0.67	1.07E-02	0.55 - 0.78

Table S4. ROC analysis of serial metabolite measurements. Area under the curve values for the association between metabolite level changes (i.e. sample 1 subtracted from sample 2) and survival during follow-up are shown for significantly associated metabolites. *probable metabolite identity, but unconfirmed (see methods).

	Controls	PAH (19-70)	PAH (>70)
	n=30	n=69	n=8
Female:Male ratio	2.0	2.3	3.0
Age	48.5 +/- 13.1	48.3 +/- 14.1	75.6 +/- 8.1
Angiogenin conc., ng/ml	360 +/- 110.3	479.7 +/- 176.6	554.9 +/- 234.2
N2,N2-dimethylguanosine	-0.1 +/- 1	1.6 +/- 1.2	2.7 +/- 1.4

Table S5. Demographics and circulating factor levels in subjects used for angiogenin study.

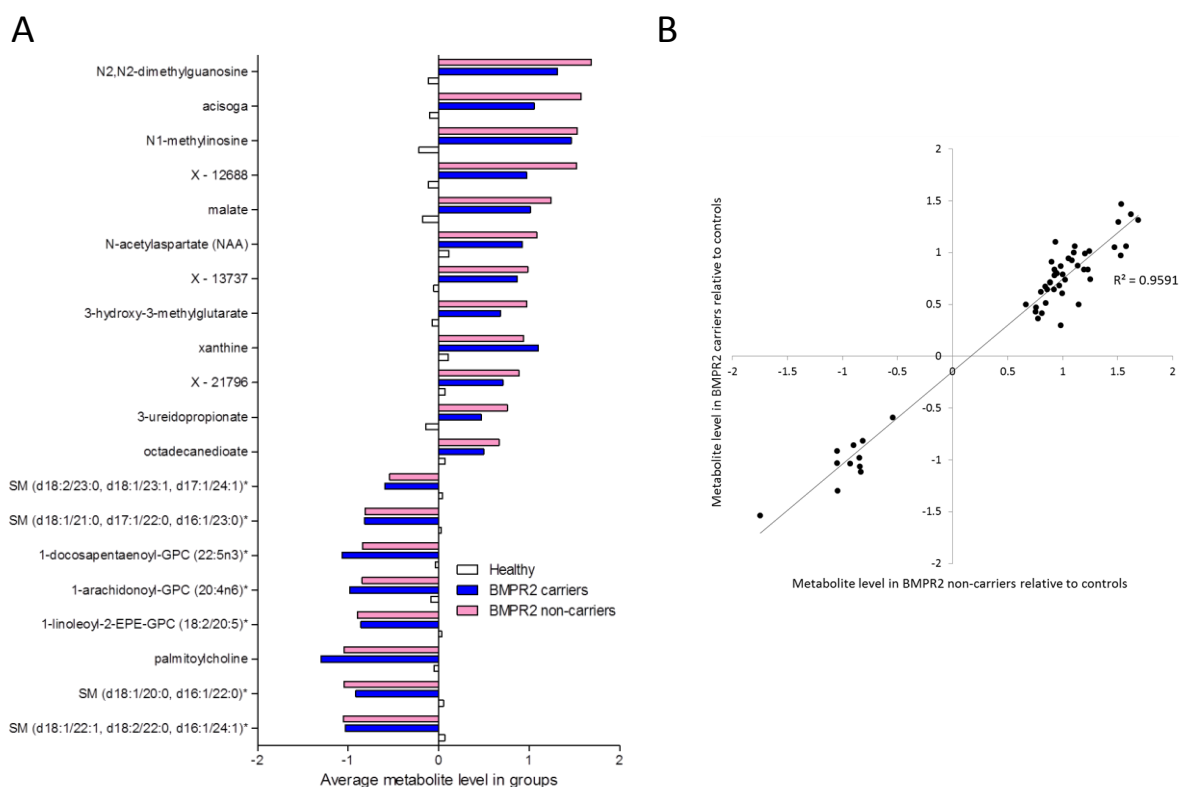Mean +/- standard deviation is shown for continuous variables.

Fig. S1. Metabolite levels in BMRP2 mutation carriers. (**A**) Average metabolite levels in PAH BMPR2 mutation carriers, non-carriers and control subjects for 20 metabolites found to significantly distinguish PAH and both healthy and disease controls, independent of potential confounders. (**B**) Correlation of average metabolite levels in BMPR2 mutation carriers and non-carriers relative to controls for 53 metabolites that distinguish PAH from healthy controls, independent of potential

confounders. Values plotted are z-scores calculated based on mean and standard deviation of all healthy volunteers in study - negative values indicate metabolites at lower levels in patients versus healthy controls and positive values indicate higher levels of metabolites in patients. *probable metabolite identity, but unconfirmed (see methods).
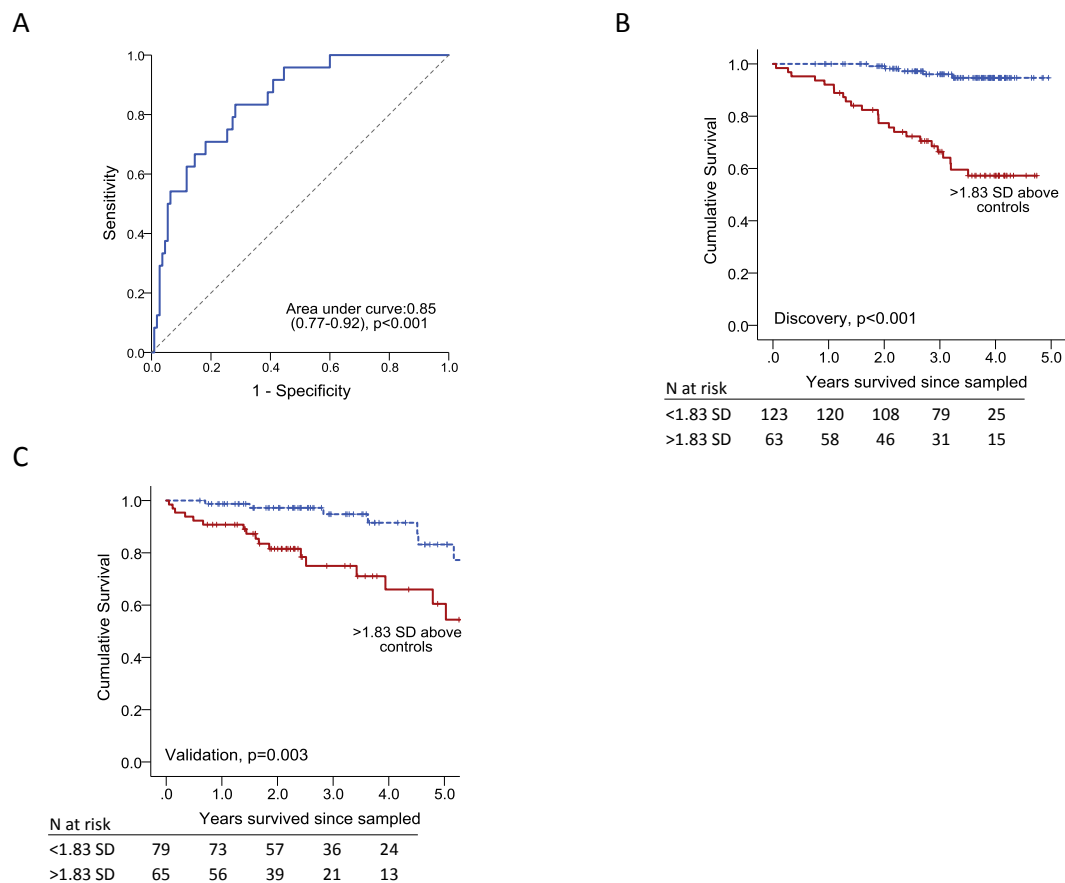
A

Area under curve:0.85
(0.77-0.92), p<0.001

Sensitivity

1 - Specificity

B

Cumulative Survival

>1.83 SD above
controls

Discovery, p<0.001

Years survived since sampled

N at risk					
<1.83 SD	123	120	108	79	25
>1.83 SD	63	58	46	31	15

C

Cumulative Survival

>1.83 SD above
controls

Validation, p=0.003

Years survived since sampled

N at risk					
<1.83 SD	79	73	57	36	24
>1.83 SD	65	56	39	21	13

Fig. S2– Survival analysis of PAH patients. (**A**) Receiver operating characteristic (ROC) curve for N2,N2-dimethylguanosine in the discovery cohort at 3 years of follow-up. The optimal cut-off for high/low risk levels of N2,N2-dimethylguanosine was derived from this for **B&C.** Kaplan Meier survival estimates in PAH patients in the discovery (**B**) and first validation (**C**) cohorts.
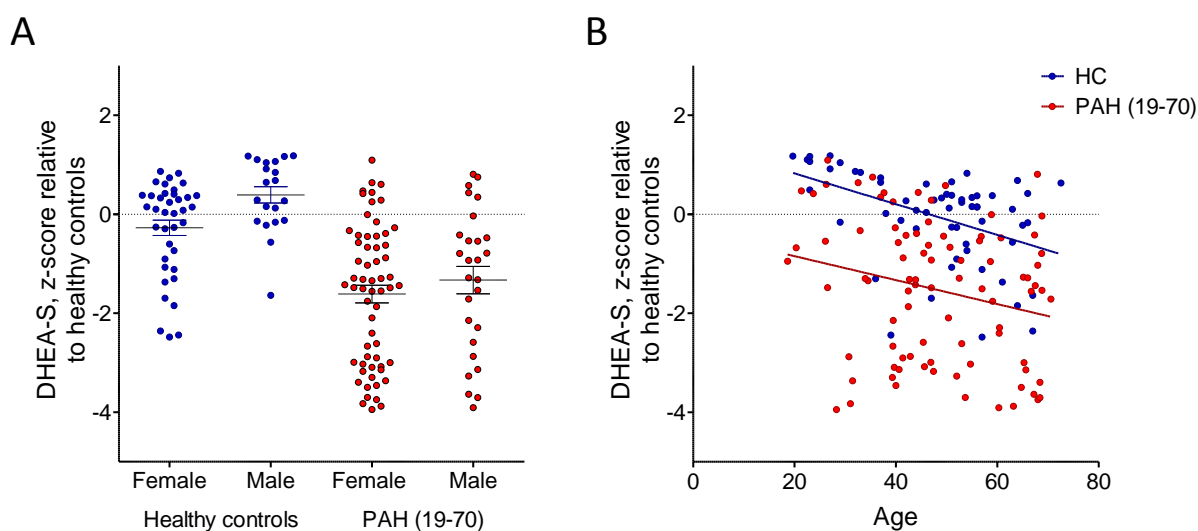
Fig. S3. Dehydroisoandrosterone-sulphate (DHEA-S). Plasma DHEA-S levels in the discovery cohort are shown for healthy controls (HC) and PAH (19-70) with (**A**) separation by gender and (**B**) against age.

Supplementary References

1. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, Arnold M, Erte I, Forgetta V, Yang TP, Walter K, Menni C, Chen L, Vasquez L, Valdes AM, Hyde CL, Wang V, Ziemek D, Roberts P, Xi L, Grundberg E, Multiple Tissue Human Expression Resource C, Waldenberger M, Richards JB, Mohney RP, Milburn MV, John SL, Trimmer J, Theis FJ, Overington JP, Suhre K, Brosnan MJ, Gieger C, Kastenmuller G, Spector TD and Soranzo N. An atlas of genetic influences on human blood metabolites. *Nature genetics*. 2014;46:543-50.
2. Guo L, Milburn MV, Ryals JA, Lonergan SC, Mitchell MW, Wulff JE, Alexander DC, Evans AM, Bridgewater B, Miller L, Gonzalez-Garay ML and Caskey CT. Plasma metabolomic profiles enhance precision medicine for volunteers of normal health. *Proceedings of the National Academy of Sciences of the United States of America*. 2015;112:E4901-10.
3. Lawton KA, Brown MV, Alexander D, Li Z, Wulff JE, Lawson R, Jaffa M, Milburn MV, Ryals JA, Bowser R, Cudkowicz ME, Berry JD and Northeast ALSC. Plasma metabolomic biomarker panel to distinguish patients with amyotrophic lateral sclerosis from disease mimics. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*. 2014;15:362-70.
4. Miller MJ, Kennedy AD, Eckhart AD, Burrage LC, Wulff JE, Miller LA, Milburn MV, Ryals JA, Beaudet AL, Sun Q, Sutton VR and Elsea SH. Untargeted metabolomic analysis for the clinical screening of inborn errors of metabolism. *Journal of Inherited Metabolic Disease*. 2015;38:1029-39.