Electronic Supplementary Material (ESI) for Molecular Omics.
This journal is © The Royal Society of Chemistry 2021

Supplementary information

NMR metabolome and microarray-based transcriptome data integration identifies unique molecular signatures of hypersensitivity pneumonitis

NMR measurements

Serum: 200 µL of serum and 400 µL of saline solution (prepared from 0.9% NaCl, 15% D_2O and 3 mM trimethylsilyl propanoic acid (TSP) were mixed together. Following centrifugation at 1500×g for 10 min, an aliquot of 600 µL of each sample supernatant was transferred into 5 mm NMR tubes.
EBC: 70 µL D_2O solution, containing 1mM TSP and 3 mM sodium azide, was added to 630 µL of EBC. 600 µL of the sample was transferred into 5 mm NMR tubes.
BALF:400 µL D_2O with 0.1 mM TSP was mixed with 200 µL of BALF. 600 µL of the sample was used for NMR experiments.

Carr–Purcell–Meiboom–Gill (CPMG) spin-echo spectra of serum samples were recorded with spectral width of 14 000 Hz, 256 transients, 4.0 s relaxation delay, 32 k data points each and an acquisition time of 0.58 s. For EBC samples, the carrier frequency (O1) value was set on the water resonance, 4 s relaxation delay, 0.82 s acquisition time, 10000 Hz spectral width, 16K time domain was and 160 number of transients. For processing, a shifted sine bell (SSB) of 2 Hz was applied prior to Fourier transformation (FT) and a real spectrum size of 16 K was used. For BALF samples, one-dimensional spectra were recorded using the excitation sculpting pulse sequence for suppression of water signal. The relaxation delay was 4 sec and 160 number of scans were recorded for each sample.

0.3 Hz line broadening was applied prior to Fourier transformation to all the free induction decays (FIDs). All spectra were manually phased and baseline corrected, and referenced to TSP (δ = 0.0 ppm) using MestReNova version 12.0.2 (Mestrelab Research, Santiago de Compostela, Spain). Normalization of the spectrum to its total area was performed using the same tool to limit the effect of concentration differences between samples. Due to poor signal to noise ratio in the aromatic region, univariate analysis was carried out in this region. Human Metabolome Database (HMDB 3.6) was used to identify each metabolite in the spectra. Furthermore, TOCSY (total correlation spectroscopy) along with previously published literature, was used to validate the assignment of peaks. While performing the metabolomics experiments, metabolomics standards initiative (MSI) guidelines were followed. Quality control (QC) samples were run throughout to examine the analytical variance. Equivalent volumes of each biofluid

(50 μL) were aliquoted from all subjects and respectively pooled together to prepare the QC samples. Coefficient of variation (CVs) for each of the identified and quantified metabolites were calculated based on the QC samples (~ 40) and was found to be <15%.

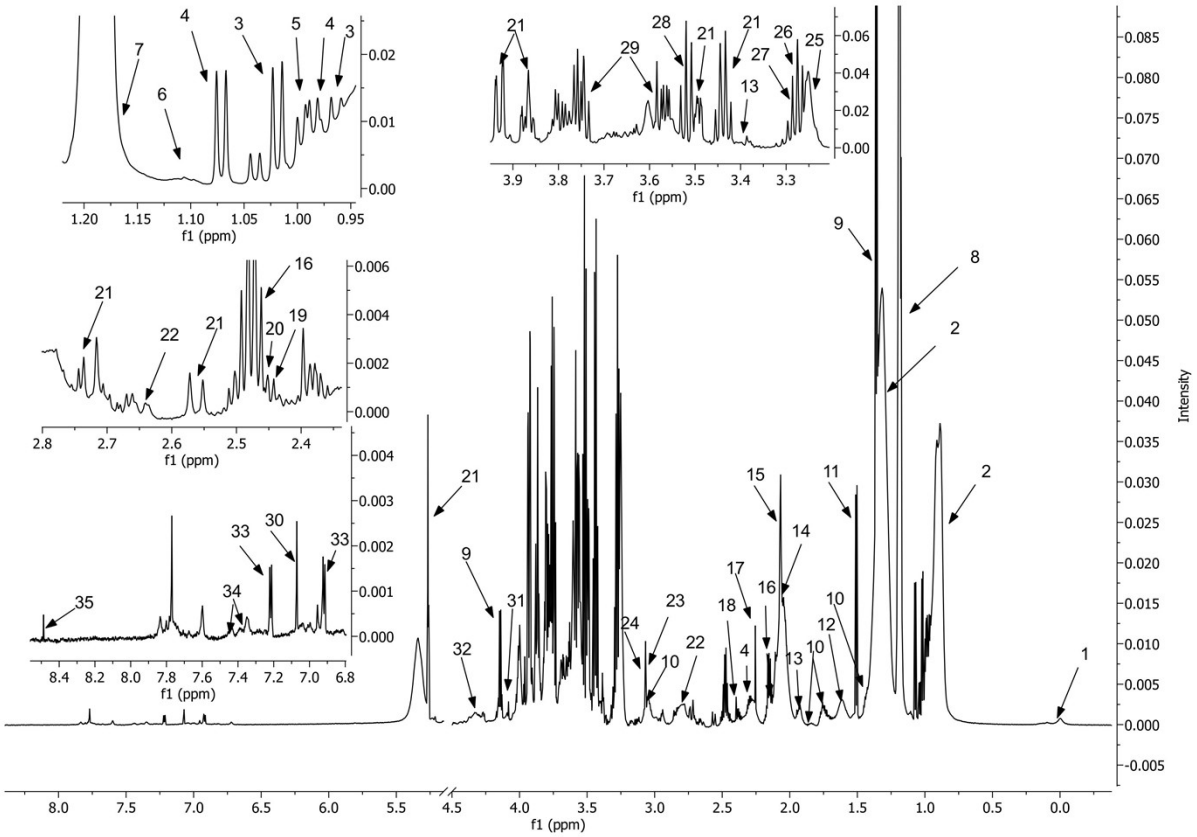
Spectral pre-processing

Using MestReNova, spectra was binned into equal frequency windows of 0.003 ppm, thereby resulting in 5200 integrated segments [1]. Multivariate analysis was performed to the spectral region of 0–5.5 ppm for serum and BALF and to 0-5 ppm region for EBC. Normalization (by sum) and unit variance scaling were performed on the working region of the binned data matrix using Metaboanalyst 4.0 [2]. Following preprocessing and missing value imputation, the data was subjected to multivariate analysis using SIMCA 13.0.2 (Umetrics, Sweden).

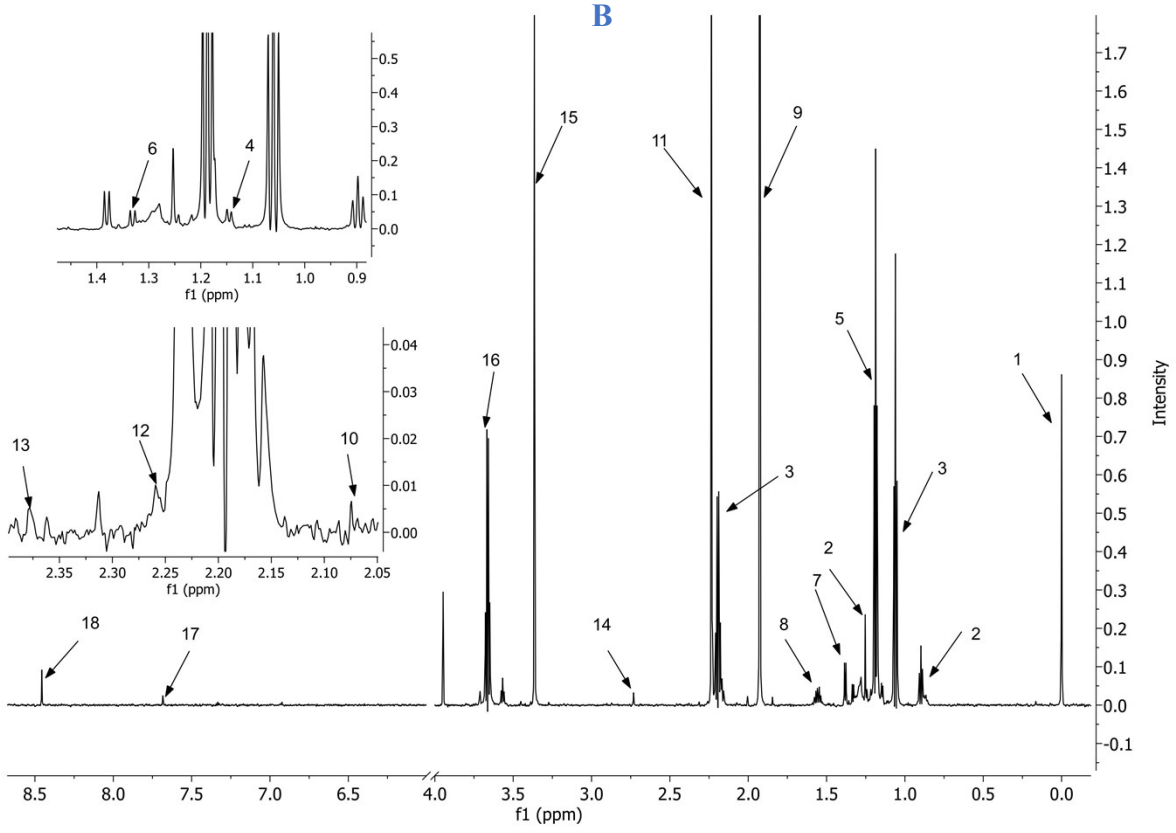
References

1. A. Hajduk, J. Mrochem-Kwarciak, A. Skorupa, M. Ciszek, A. Heyda, K. Składowski and M. Sokół, *Metabolomics*, 2016, **12**, 102.
2. J. Chong, O. Soufan, C. Li, I. Caraus, S. Li, G. Bourque, DS. Wishart, J. Xia, *Nucleic Acids Res*., 2018, **46**, W486-94.

A



B



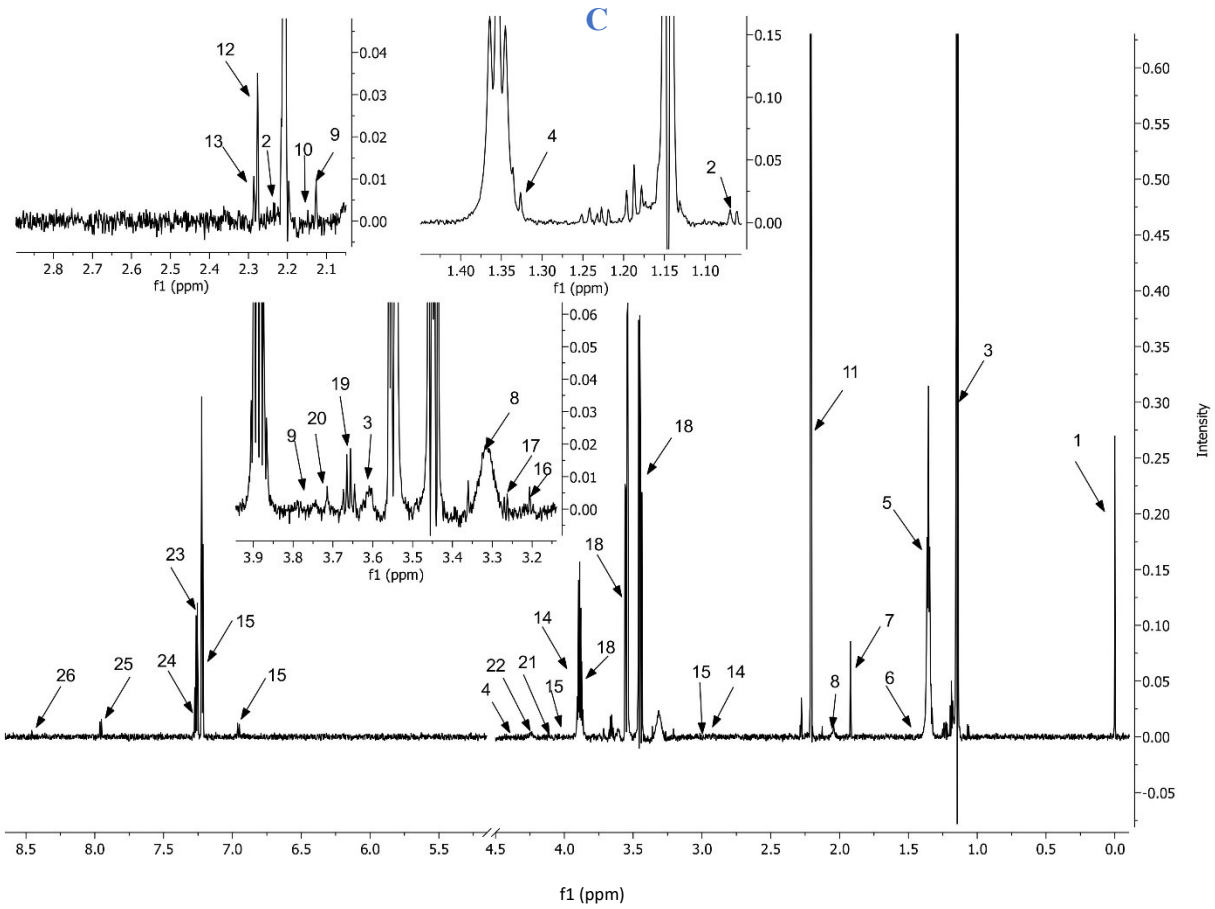


Fig. S1 Representative spectra of human (A) serum (B) exhaled breath condensate (EBC) and (C) bronchoalveolar lavage fluid (BALF) measured at 300 K, 800 MHz. Assigned numbers indicate the following metabolites:

(A) 1, TSP; 2, lipid; 3, isoleucine; 4, valine; 5, L-leucine; 6, propylene glycol; 7, isopropyl alcohol; 8, 3-hydroxybutyric acid; 9, lactic acid; 10, L-lysine; 11, L-alanine; 12, adipic acid; 13, acetic acid; 14, proline; 15, N-acetyl glycoproteins; 16, glutamine; 17, acetone; 18, glutamate; 19, pyruvate; 20, succinate; 21, citrate; 22, asparagine; 23, creatine; 24, ornithine; 25, carnitine; 26, glucose; 27, betaine; 28, glycine; 29, glycerol; 30, histidine; 31, creatinine; 32, threonine 33, tyrosine; 34, phenylalanine; 35, formate.

(B) 1, TSP; 2, fatty acid; 3, propionate; 4, isopropanol; 5, ethanol; 6, lactate; 7, threonine; 8, alanine; 9, acetate; 10, proline; 11, acetone; 12, valine; 13, pyruvate; 14, trimethylamine; 15, methanol; 16, glycerol; 17, phenylalanine; 18, formate

(C) 1, TSP; 2, fatty acid; 3, valine; 4, ethanol; 5, threonine; 6, lactate; 7, alanine; 8, acetate; 9, proline; 10, glutamate; 11, methionine; 12, acetone; 13, pyruvate; 14, succinate; 15, ornithine; 16, arginine; 17, betaine; 18, glycerol; 19, creatine; 20, choline; 21, gluconic acid; 22, tyrosine; 23, tryptophan; 24, phenylalanine; 25, histidine; 26, formate

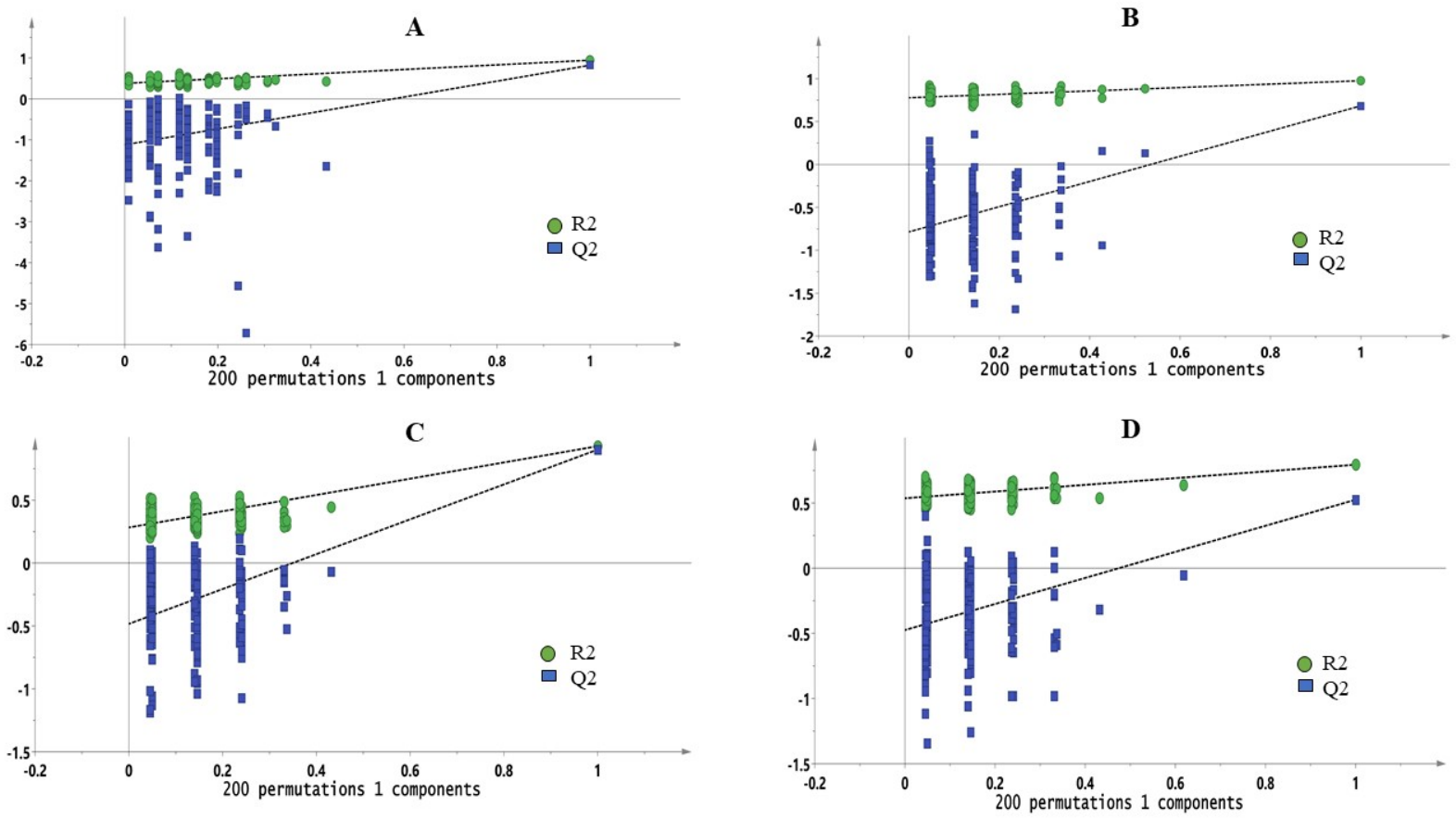


Fig. S2 Permutation tests indicating predictive ability of orthogonal partial least squares discriminant analysis (OPLS-DA) model for hypersensitivity pneumonitis (HP) and healthy controls. (A) indicates serum and (B) represents exhaled breath condensate (EBC) of discovery cohort. (C) indicates serum and (D) represents EBC of the validation cohort.

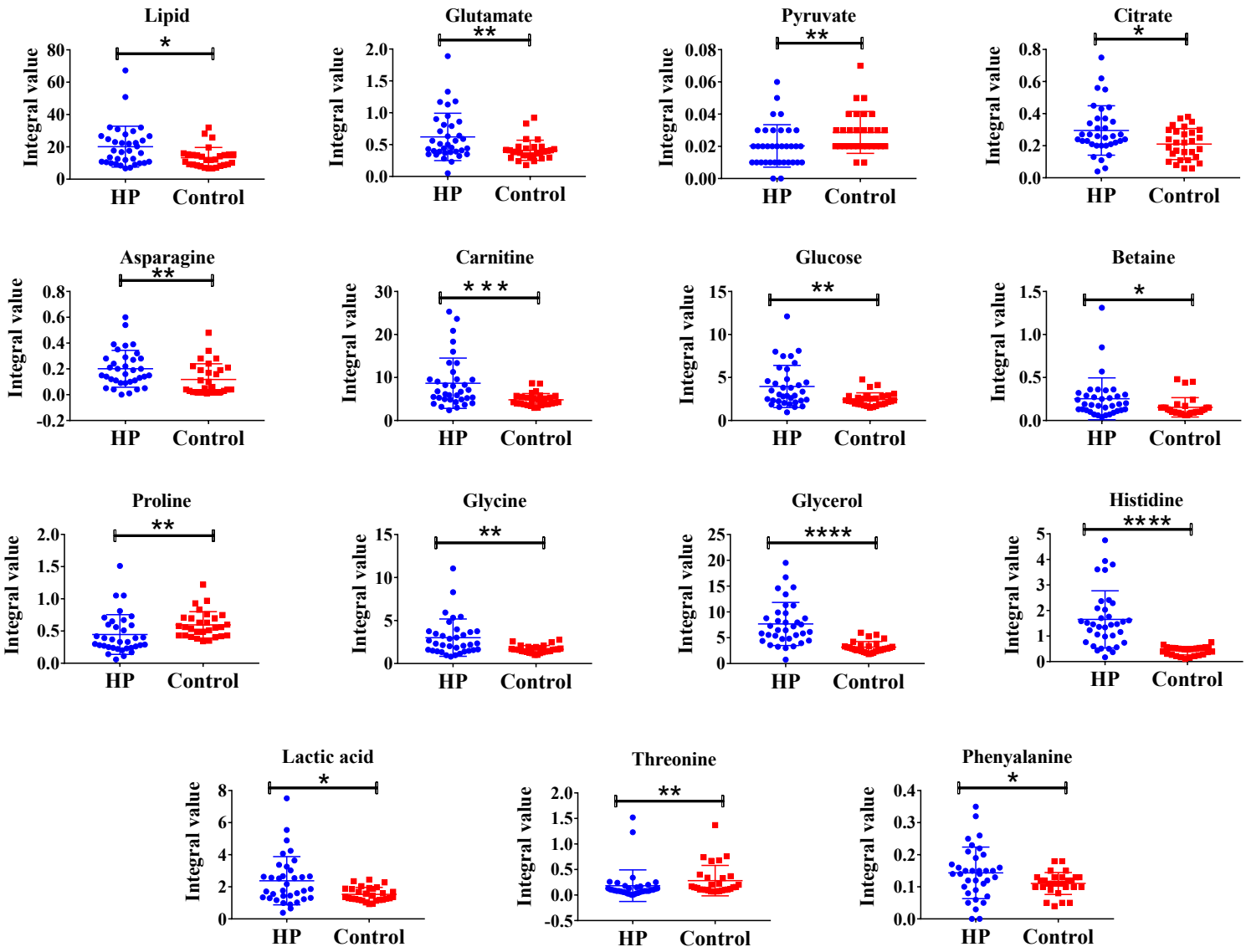


Fig. S3 Scatter plots of significantly altered metabolites in serum of HP patients as compared with controls. *p value < 0.05; **p value< 0.01; ***p value < 0.001

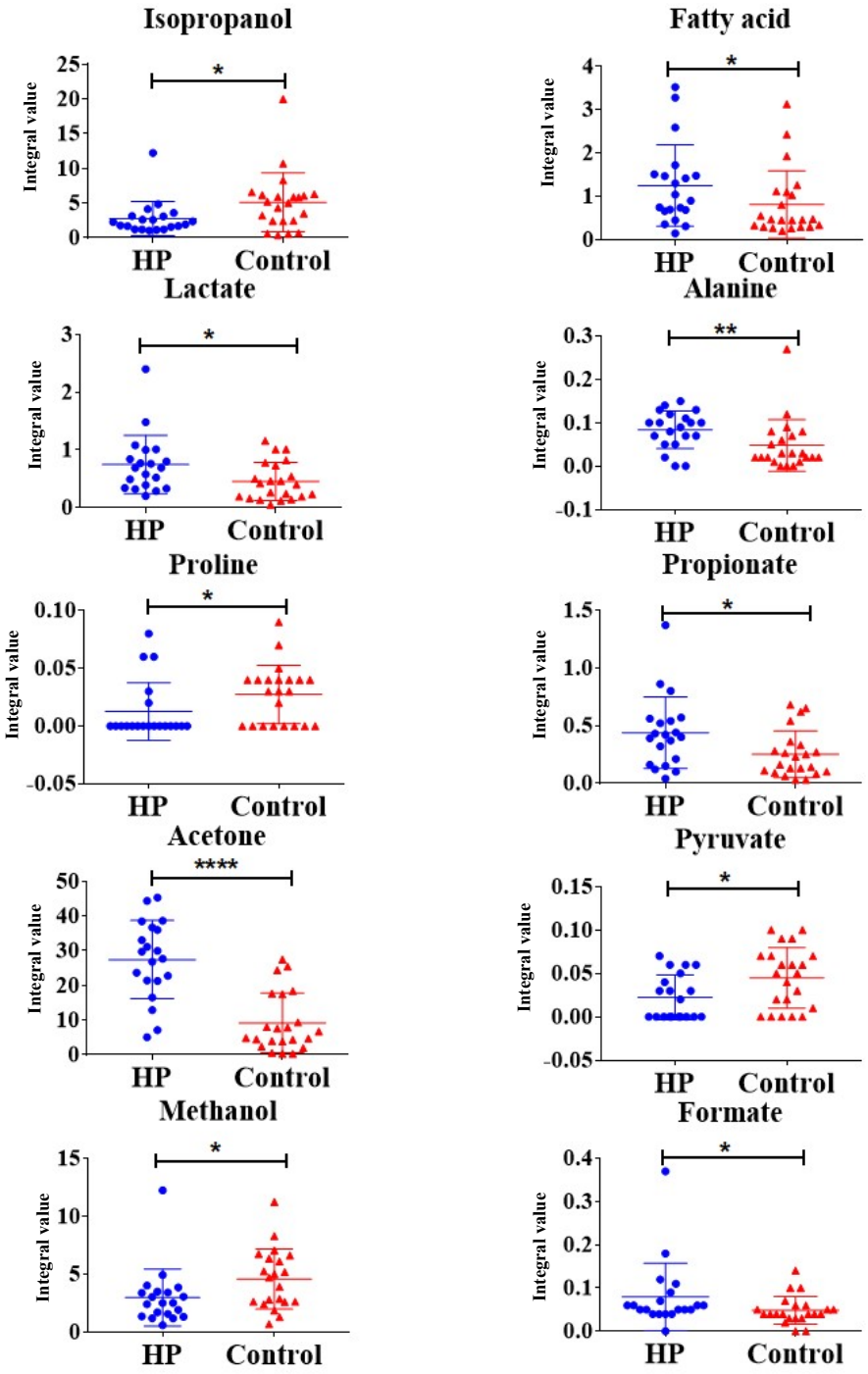


Fig. S4 Scatter plots of significantly altered metabolites in exhaled breath condensate (EBC) of HP patients as compared with controls.

*p value < 0.05; **p value< 0.01; ***p value < 0.001

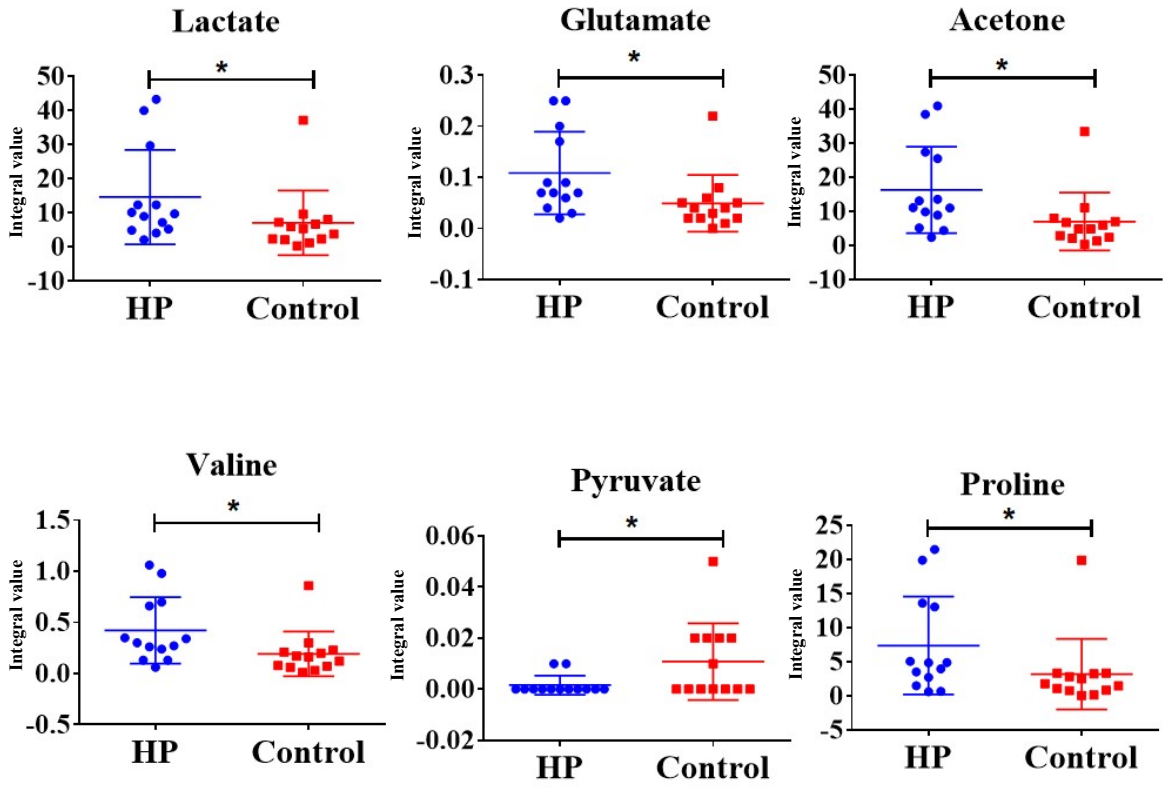


Fig. S5 Scatter plots of significantly altered metabolites in bronchoalveolar lavage fluid (BALF) of HP patients as compared with controls.
*p value < 0.05

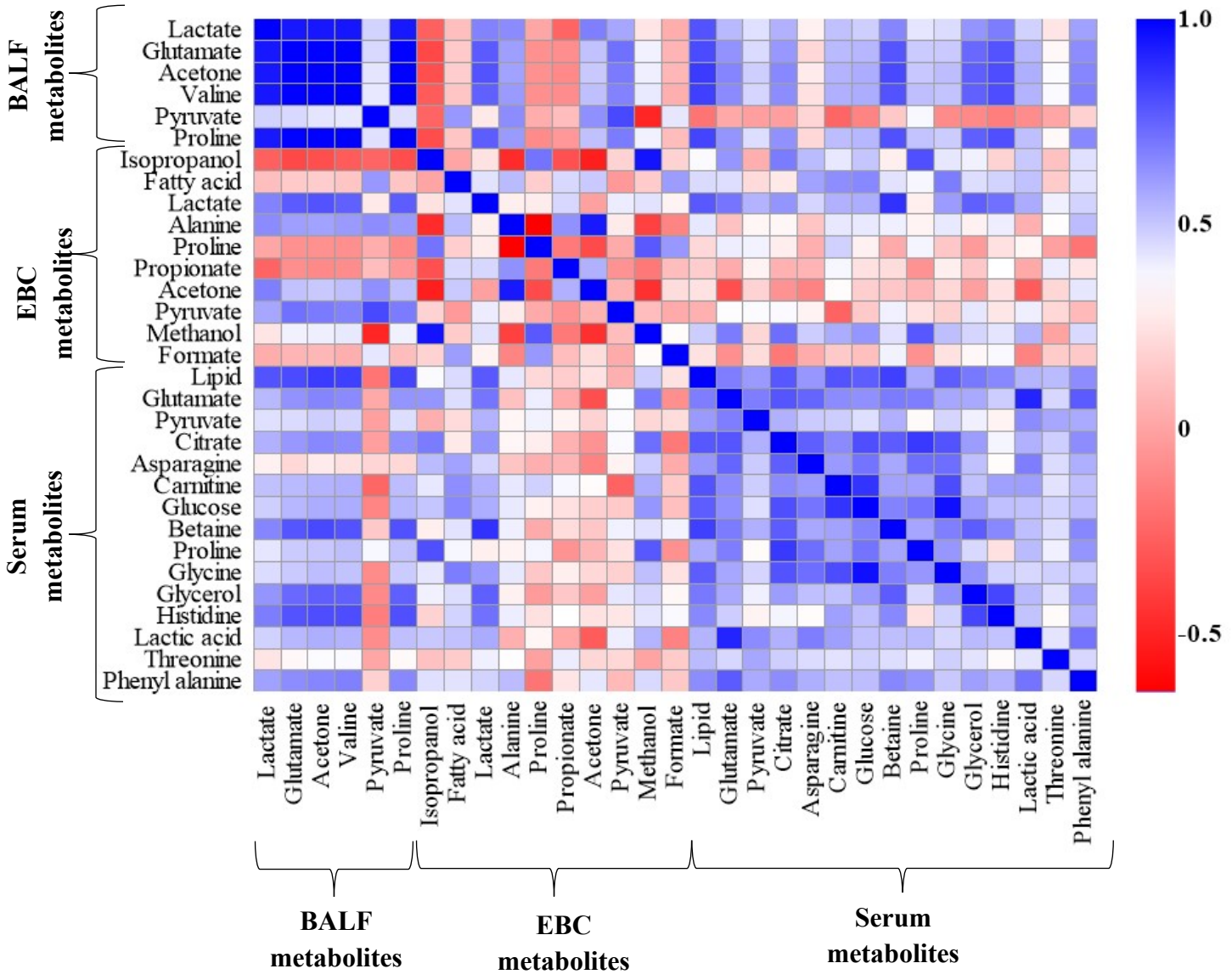


Fig. S6 Pearson's correlation heatmap shows association between dysregulated serum, exhaled breath condensate (EBC) and bronchoalveolar lavage fluid (BALF) metabolites in patients with HP as compared with controls.

Blue squares—positive correlation, red squares—negative correlation (p value<0.05)

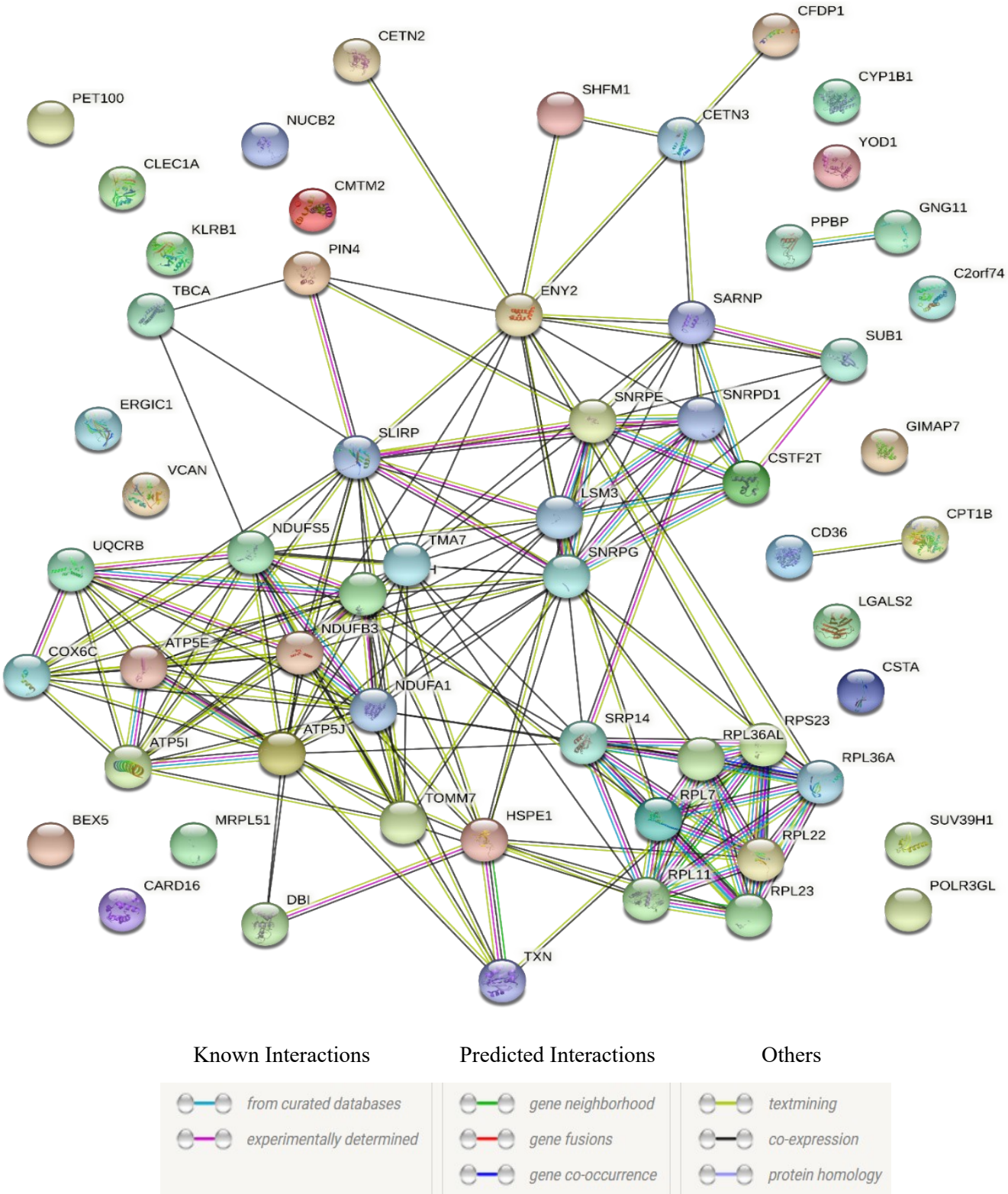


Fig. S7 A total of 59 significantly altered genes identified in patients with HP were used for protein interaction network (PPI) using STRING database. Different colored edges show protein-protein interactions. The network has 60 nodes and 179 edges

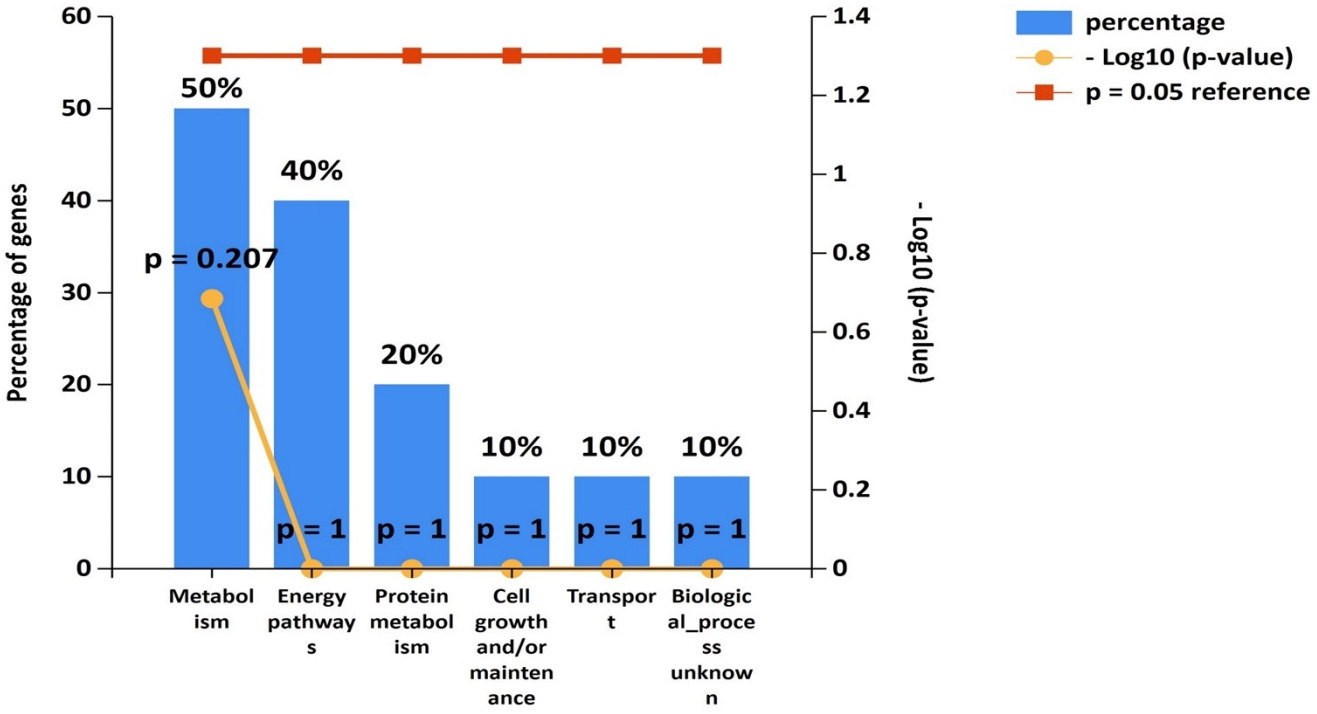


Fig. S8 Ten most significantly altered genes identified in patients with HP were used for functional enrichment analysis with a stand-alone software tool FunRich (Functional Enrichment analysis tool) version 3.1.3. Biological process enrichment analysis of the dysregulated genes is shown.

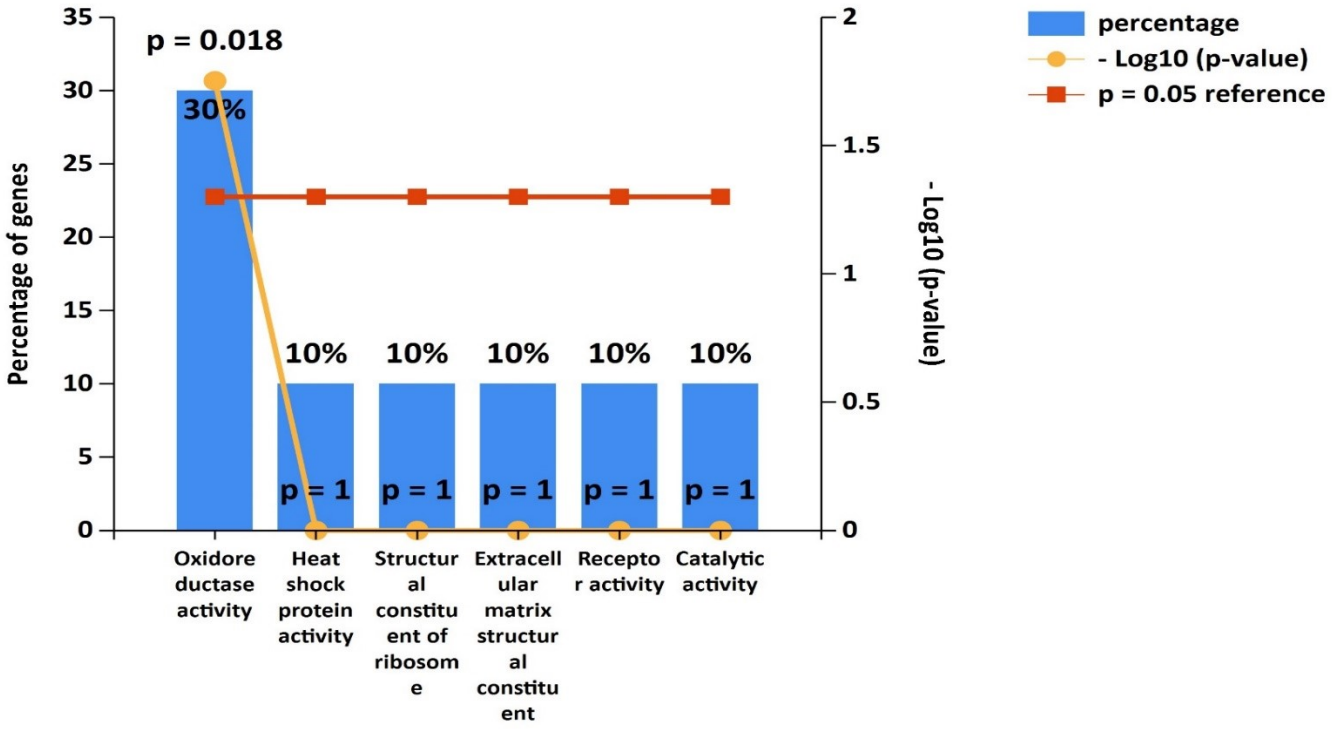


Fig. S9 Ten most significantly altered genes identified in patients with HP were used for functional enrichment analysis with a stand-alone software tool FunRich (Functional Enrichment analysis tool) version 3.1.3. Molecular function enrichment analysis of the dysregulated genes is shown

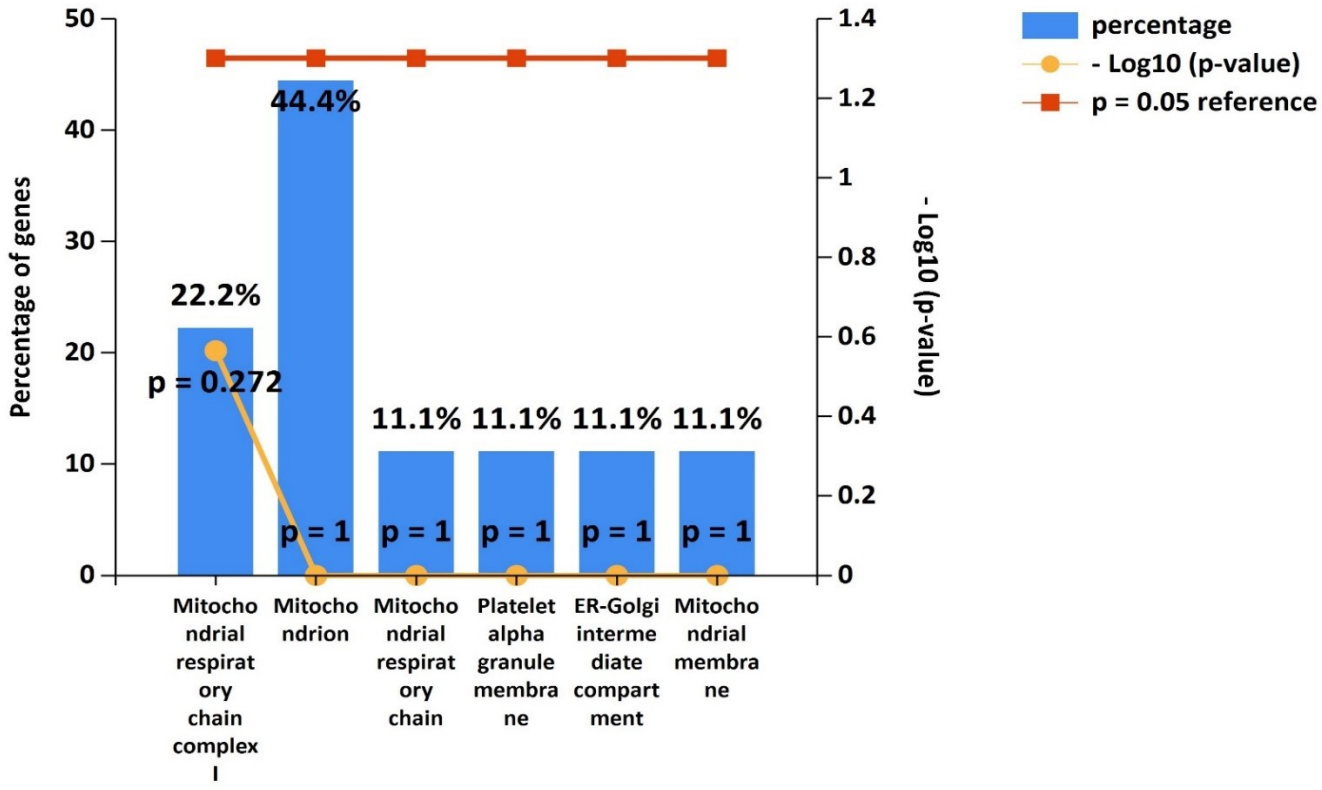


Fig. S10 Ten most significantly altered genes identified in patients with HP were used for functional enrichment analysis with a stand-alone software tool FunRich (Functional Enrichment analysis tool) version 3.1.3. Cellular component enrichment analysis of the dysregulated genes is shown

Table S1 Integral values of all identified metabolites in serum of hypersensitivity pneumonitis (HP) subjects as compared to controls (discovery cohort)

Metabolites	Chemical shift (ppm)	Controls	HP
Lipid	0.90	13.31±6.35	20.11±12.64
Isoleucine	0.96	0.20±0.06	0.49±1.43
Valine	0.99	0.59±0.21	0.66±0.46
L-leucine	1.01	0.76±0.23	0.81±0.55
Isoleucine	1.04	0.27±0.10	0.30±0.17
Valine	1.07	0.76±0.20	0.85±0.51
Propylene glycol	1.13	0.01±0.008	0.01±0.01
Isopropyl alcohol	1.17	1.17±3.31	2.62±2.24
3 hydroxybutyric acid	1.19	0.03±0.02	0.07±0.09
Lipid	1.32	28.61±18.83	36.82±25.88
Lactic acid	1.36	7.74±2.18	11.40±8.72
Lysine	1.42	0.20±0.17	0.24±0.14
L-alanine	1.5	1.38±0.42	1.62±0.82
Adipic acid	1.60	0.90±0.89	1.14±1.29
Lysine	1.75	0.85±0.21	0.89±0.47
Lysine	1.90	0.33±0.11	0.42±0.29
Acetic acid	1.94	0.08±0.04	0.11±0.11
Proline	2.04	1.15±1.21	1.19±1.35
N- acetyl glycoprotein	2.07	5.03±3.30	8.01±6.82
L-glutamine	2.16	1.46±0.40	1.46±0.83
Acetone	2.26	0.84±0.42	0.77±0.50
Valine	2.29	0.76±0.85	0.44±0.57
Glutamate	2.38	0.41±0.15	0.62±0.37
Pyruvate	2.42	0.03±0.01	0.02±0.01
Succinate	2.43	0.04±0.02	0.04±0.03
Glutamine	2.46	1.03±0.28	1.02±0.54
Citrate	2.55	0.20±0.06	0.25±0.14
Asparagine	2.67	0.14±0.10	0.11±0.08
Citrate	2.71	0.21±0.09	0.29±0.15
Asparagine	2.93	0.11±0.12	0.20±0.14
Lysine	3.04	0.55±0.26	0.66±0.51
Creatine	3.06	0.26±0.06	0.32±0.17
Ornithine	3.07	0.07±0.06	0.08±0.07
Carnitine	3.24	4.82±1.44	8.66±5.84
Glucose	3.29	2.44±0.77	3.95±2.45
Betaine	3.31	0.15±0.11	0.25±0.24
Proline	3.35	0.59±0.20	0.44±0.30

Glucose	3.44	4.08±1.07	7.23±5.19
Glucose	3.50	3.03±0.82	5.83±4.12
Glycine	3.52	1.66±0.43	3.01±2.16
Glycerol	3.57	3.17±1.07	7.67±4.18
Glycerol	3.77	7.46±2.14	13.51±8.35
Glucose	3.86	2.98±0.85	5.40±3.85
Glucose	3.93	2.92±0.89	5.43±3.42
Histidine	4.00	0.42±0.16	1.65±1.12
Creatinine	4.08	0.14±0.04	0.19±0.10
Lactic acid	4.15	1.53±0.42	2.37±1.50
Threonine	4.26	0.28±0.29	0.18±0.30
Glucose	5.27	1.22±0.33	2.25±1.61
Tyrosine	6.91	0.10±0.03	0.14±0.07
Histidine	7.06	0.06±0.01	0.07±0.04
Tyrosine	7.21	0.11±0.03	0.14±0.07
Phenylalanine	7.35	0.11±0.03	0.14±0.08
Phenylalanine	7.39	0.03±0.02	0.06±0.06
Phenylalanine	7.44	0.09±0.03	0.10±0.07
Histidine	7.76	0.05±0.01	0.06±0.04
Formate	8.49	0.01±0.003	0.01±0.008

Table S2 Integral values of all identified metabolites in serum of HP subjects as compared to controls (validation cohort)

Metabolites	Chemical shift (ppm)	Controls	HP
Lipid	0.90	14.33±7.65	36.46±20.62
Isoleucine	0.96	0.36±0.52	0.38±0.23
Valine	0.99	0.56±0.22	1.37±1.52
L-leucine	1.01	0.73±0.22	1.26±0.65
Isoleucine	1.04	0.25±0.08	0.40±0.21
Valine	1.07	0.69±0.21	1.19±0.63
Propylene glycol	1.13	0.015±0.01	0.03±0.02
Isopropyl alcohol	1.17	0.39±0.43	0.65±1.10
3 hydroxybutyric acid	1.19	0.03±0.03	0.09±0.08
Lipid	1.32	30.39±21.90	67.16±43.20
Lactic acid	1.36	8.99±5.38	14.07±7.93
Lysine	1.42	0.23±0.26	0.53±0.45
L-alanine	1.50	1.35±0.39	2.24±1.18
Adipic acid	1.60	1.14±1.22	2.54±2.04
Lysine	1.75	0.76±0.24	1.25±0.60
Lysine	1.90	0.36±0.18	0.57±0.42
Acetic acid	1.94	0.08±0.06	0.11±0.10
Proline	2.04	1.15±1.36	3.41±5.51
N- acetyl glycoprotein	2.07	6.67±3.94	10.76±11.12
L-glutamine	2.16	1.31±0.49	2.31±1.04
Acetone	2.26	0.80±0.47	0.80±0.31
Valine	2.29	1.02±0.97	1.13±1.38
Glutamate	2.38	0.38±0.13	0.93±0.63
Pyruvate	2.42	0.04±0.07	0.02±0.01
Succinate	2.43	0.06±0.07	0.06±0.03
Glutamine	2.46	1.00±0.33	1.61±0.77
Citrate	2.55	0.18±0.05	0.31±0.17
Asparagine	2.67	0.09±0.05	0.16±0.15
Citrate	2.71	0.22±0.09	0.38±0.21
Asparagine	2.93	0.09±0.09	0.30±0.18
Lysine	3.04	0.60±0.31	1.24±0.95
Creatine	3.06	0.32±0.26	0.44±0.20
Ornithine	3.07	0.07±0.05	0.12±0.11
Carnitine	3.24	5.19±3.13	12.50±6.53
Glucose	3.29	2.37±0.80	5.08±2.69

Betaine	3.31	0.13±0.10	0.27±0.18
Proline	3.35	0.59±0.22	0.29±0.27
Glucose	3.44	4.14±1.64	9.36±4.94
Glucose	3.50	2.98±1.26	7.69±4.15
Glycine	3.52	1.64±0.66	3.99±2.23
Glycerol	3.57	2.60±0.92	11.68±7.24
Glycerol	3.77	7.21±2.77	18.29±9.43
Glucose	3.86	2.97±1.35	7.25±4.17
Glucose	3.93	2.91±1.19	7.21±3.77
Histidine	4.00	0.40±0.20	2.36±1.64
Creatine	4.08	0.12±0.03	0.22±0.11
Lactic acid	4.15	1.47±0.41	3.10±1.95
Threonine	4.26	0.36±0.42	0.15±0.08
Glucose	5.27	1.19±0.60	2.80±1.79
Tyrosine	6.91	0.10±0.03	0.17±0.08
Histidine	7.06	0.06±0.01	0.10±0.04
Tyrosine	7.21	0.10±0.03	0.18±0.07
Phenylalanine	7.35	0.10±0.05	0.17±0.11
Phenylalanine	7.39	0.02±0.02	0.04±0.06
Phenylalanine	7.44	0.08±0.04	0.11±0.11
Histidine	7.76	0.05±0.01	0.09±0.04
Formate	8.49	0.01±0.004	0.01±0.01

Table S3 Integral values of all identified metabolites in exhaled breath condensate (EBC) of HP subjects as compared to controls (discovery cohort)

Metabolites	Chemical shift (ppm)	Controls	HP
Fatty acid	0.90	0.27±0.33	0.31±0.22
Propionate	1.06	0.78±1.25	0.73±0.90
Isopropanol	1.18	5.11±4.26	2.76±2.46
Ethanol	1.2	1.65±2.45	0.64±0.48
Fatty acid	1.25	0.81±0.77	1.25±0.93
Lactate	1.33	0.45±0.32	0.74±0.50
Threonine	1.38	0.16±0.41	0.32±0.65
Alanine	1.48	0.04±0.05	0.08±0.04
Acetate	1.93	1.63±1.70	2.36±2.85
Proline	2.07	0.02±0.02	0.01±0.02
Propionate	2.18	0.25±0.20	0.43±0.31
Acetone	2.23	9.18±8.56	27.45±11.31
Valine	2.24	0.07±0.32	0.34±1.22
Pyruvate	2.39	0.04±0.03	0.02±0.02
Trimethylamine	2.9	0.02±0.02	0.02±0.05
Methanol	3.36	4.58±2.60	2.98±2.46
Glycerol	3.66	2.29±1.72	1.78±1.49
Phenylalanine	7.33	0.01±0.02	0.001±0.004
Formate	8.46	0.04±0.03	0.07±0.07

Table S4 Integral values of all identified metabolites in exhaled breath condensate (EBC) of HP subjects as compared to controls (validation cohort)

Metabolites	Chemical shift (ppm)	Controls	HP
Fatty acid	0.90	0.44±0.60	0.58±0.40
Propionate	1.06	0.35±0.42	0.87±0.91
Isopropanol	1.18	4.47±3.39	1.62±3.07
Ethanol	1.2	4.91±4.27	1.35±1.12
Fatty acid	1.25	1.48±1.94	1.28±1.38
Lactate	1.33	0.42±0.63	0.73±0.81
Threonine	1.38	0.06±0.06	0.19±0.23
Alanine	1.48	0.01±0.02	0.02±0.02
Acetate	1.93	1.29±1.23	2.57±2.08
Proline	2.07	0.03±0.02	0.01±0.01
Propionate	2.18	0.22±0.26	0.63±0.65
Acetone	2.23	3.28±3.59	2.20±0.53
Valine	2.24	0.06±0.18	0.04±0.04
Pyruvate	2.39	0.03±0.01	0.01±0.01
Trimethylamine	2.9	0.03±0.02	0.02±0.02
Methanol	3.36	5.70±3.70	2.05±0.84
Glycerol	3.66	4.43±4.07	1.11±0.91
Phenylalanine	7.33	0.02±0.07	0.05±0.03
Formate	8.46	0.06±0.09	0.08±0.05

Table S5 Integral values of all identified metabolites in bronchoalveolar lavage fluid (BALF) of HP subjects as compared to controls (non-HP)

Metabolites	Chemical shift (ppm)	Controls (non-HP)	HP
Valine	1.07	0.12±0.14	0.10±0.05
Ethanol	1.15	19.41±23.85	15.08±8.43
Threonine	1.33	0.19±0.24	0.10±0.05
Lactate	1.36	7.07±9.48	14.62±13.82
Alanine	1.48	0.03±0.03	0.01±0.01
Acetate	1.92	0.14±0.09	0.17±0.08
Proline	2.04	0.39±0.46	0.26±0.29
Glutamate	2.13	0.04±0.05	0.10±0.08
Methionine	2.16	0.01±0.01	0.01±0.01
Acetone	2.21	7.02±8.49	16.31±12.70
Valine	2.28	0.19±0.21	0.42±0.32
Pyruvate	2.37	0.01±0.01	0.001±0.003
Succinate	2.41	0.009±0.01	0.004±0.008
Creatine	3.04	0.01±0.009	0.009±0.006
Tyrosine	3.15	0.01±0.02	0.006±0.01
Arginine	3.22	0.07±0.06	0.04±0.04
Betaine	3.27	0.02±0.02	0.03±0.03
Proline	3.35	3.16±5.16	7.37±7.18
Glucose	3.45	7.29±9.21	5.33±2.84
Ethanol	3.62	0.35±0.45	0.22±0.13
Glycerol	3.66	0.32±0.19	0.33±0.19
Glucose	3.71	0.06±0.07	0.08±0.05
Glutamate	3.79	0.01±0.01	0.006±0.01
Glucose	3.89	4.98±6.28	3.84±1.89
Tyrosine	4.00	0.01±0.02	0.01±0.03
Gluconic acid	4.11	0.03±0.03	0.02±0.01
Threonine	4.24	0.92±1.90	2.32±2.50
Tyrosine	6.96	0.07±0.10	0.16±0.13
Tryptophan	7.26	1.25±1.49	1.57±0.73
Phenylalanine	7.3	0.02±0.03	0.04±0.05
Histidine	7.82	0.009±0.01	0.008±0.01
Formate	8.46	0.01±0.007	0.01±0.008

Table S6 Differentially expressed genes (DEGs) of HP subjects as compared to controls

Downregulated				
ID	**Average expression**	**P value**	**Adjusted p value**	**Log FC**
CARD16	10.24789	8.22E-05	0.019838	-1.03802
C2orf74	7.210272	0.000626	0.074903	-1.33064
SRP14	10.99894	2.88E-10	1.31E-06	-1.15516
RPL11	10.39968	0.000364	0.052817	-1.18164
RPL7	12.02689	1.11E-05	0.00511	-1.14287
RPL23	11.73195	0.000139	0.027651	-1.31674
RPS23	12.63192	2.76E-11	1.67E-07	-1.19123
RPL36A	10.34827	1.53E-06	0.001229	-1.99295
COX6C	9.987171	0.000169	0.031569	-1.70249
NDUFS5	10.32554	3.37E-10	1.42E-06	-1.94756
TOMM7	10.66048	2.25E-05	0.008142	-1.39207
LSM3	8.083385	6.91E-06	0.003631	-1.59503
UQCRH	10.01942	1.65E-06	0.001268	-1.13621
SEM1	9.749123	6.53E-09	1.98E-05	-1.6944
NDUFA1	10.30143	1.55E-14	4.24E-10	-2.11598
ATP5PF	10.51079	6.24E-08	0.000103	-1.33031
DBI	9.374304	1.15E-05	0.005214	-1.41926
SNRPD1	8.347551	2.61E-08	5.95E-05	-1.54438
CFDP1	8.393305	3.75E-10	1.46E-06	-1.16125
SNRPE	9.800555	7.16E-08	0.000112	-1.52537
NDUFB3	7.950478	1.62E-06	0.001266	-1.16262
TBCA	8.443308	4.03E-07	0.000408	-1.24206
NUCB2	9.760718	4.00E-06	0.002456	-1.22225
GNG11	8.036379	6.01E-05	0.016111	-1.31566
PIN4	8.964413	0.000127	0.026462	-1.09043
CSTA	10.57433	9.68E-07	0.000853	-1.336
HSPE1	9.755914	5.66E-09	1.82E-05	-2.01688
SNRPG	10.17753	8.38E-05	0.020058	-1.13197
UQCRB	9.74728	1.90E-12	1.87E-08	-2.57168
ATP5ME	9.074964	4.32E-08	8.14E-05	-1.3424
RPL36AL	11.19874	4.10E-16	2.24E-11	-1.58784
LGALS2	9.071441	0.000933	0.096583	-1.79732
RPL22	12.38711	1.03E-13	1.87E-09	-1.50538
TXN	9.294564	1.77E-05	0.006882	-1.01231
CETN2	8.76321	2.45E-08	5.92E-05	-1.01747
CETN3	6.948075	0.000899	0.094322	-1.20798
SUB1	10.37239	5.35E-08	9.44E-05	-1.4188
PPBP	11.39095	4.93E-05	0.013896	-1.19989
KLRB1	8.251271	0.000176	0.032513	-1.16656
ATP5F1E	11.22985	6.48E-07	0.000621	-1.15667
ENY2	10.31544	1.00E-06	0.000869	-1.21679

SUV39H1	6.920705	3.37E-07	0.000368	-1.08737
SLIRP	7.58295	0.000831	0.09017	-1.05784
TMA7	10.84398	0.000131	0.026946	-1.1557
MRPL51	8.579389	1.73E-10	8.62E-07	-1.21598
POLR3GL	9.546271	1.62E-07	0.00021	-1.12614
SARNP	7.817862	1.37E-05	0.005754	-1.16995
PET100	10.90771	0.000112	0.024617	-1.20543
GIMAP7	10.4828	8.81E-07	0.000789	-1.14331
BEX5	6.415435	0.000231	0.037585	-1.20644
CMTM2	10.03815	7.27E-05	0.017908	-1.023
Upregulated				
CYP1B1	7.127202	0.000229	0.037585	1.10756
VCAN	7.147938	5.82E-05	0.015901	1.179384
CPT1B	8.152227	1.64E-07	0.00021	1.037231
CSTF2T	5.449478	0.000109	0.024292	1.029825
YOD1	6.412912	2.63E-06	0.001842	1.745776
CLEC1A	3.087444	0.000125	0.026462	1.019138
ERGIC1	9.151039	3.36E-05	0.010677	1.333477
CD36	7.194779	0.000337	0.049781	1.132183