Non-targeted plasma metabolomics of Adult B-cell Acute Lymphoblastic Leukemia

Fengmin Yang[a,b], Qian Li[a] , Junfeng Xiang[a], Hong Zhang[a], Hongxia Sun[a*],   Guorui Ruan[c*], Yalin Tang[a,b*]

1 National Laboratory for Molecular Sciences, Center for Molecular Sciences, State Key Laboratory for Structural Chemistry of Unstable and Stable Species, Institute of Chemistry, Chinese Academy of Sciences, Beijing, P. R. China, **2** University of the Chinese Academy of Sciences, Beijing, P. R. China, **3** Peking University People's Hospital and Institute of Hematology, Beijing, P. R. China.

* hongxsun@iccas.ac.cn (HS); tangyl@iccas.ac.cn (YT); ruanguorui@pkuph.edu.cn (RG)

Table S1 Plasma samples used in study

Sample group	Sample size	Gender	Age	Stratification
Control	19	Male = 10 Female = 9	18-58(39.8)	-
ALL	19	Male = 10 Female = 9	23-58(40.1)	Com-B-ALL(13) Pro-B-ALL(5) Pre-B-ALL(1)

Table S2 Assignments of plasma metabolites

Key	Metabolites	Moieties	δ^1H(ppm) and multiplicity*
1	Lipids(mainly LDL)	CH_3, $(CH_2)_n$	0.85(t),1.25(m)
2	Lipids(mainly VLDL)	CH_3, $(CH_2)_n$	0.88(t), 1.29(m)
3	Isoleucine	δCH_3, βCH_3, half-γCH_2, βCH,αCH	0.94(t),1.01(d)
4	Leucine	δCH_3, βCH_3	0.92(d), 0.95(d),
5	Valine	γ-CH_3,αCH, βCH	0.99(d), 1.05(d)
6	Lactate	βCH_3, αCH	1.33(d), 4.11(q)
7	Alanine	CH_3, αCH	1.49(d),3.78(q)
8	Citrulline	γCH_2, βCH_2	1.58(m)
9	Arginine	βCH_2,γCH_2	1.89(m),1.73(dd)
10	Acetate	CH_3	1.91(s)
11	Proline	γCH_2	1.99(m)
12	N-acetyl-glycoproteins	CH_3	2.04(s)
13	Glutamate	half-βCH_2, γCH_2	2.08(m), 2.34(m)
14	Glutmaine	βCH_2, γCH_2	2.13(m), 2.45(m)
15	Pyruvate	CH_3	2.41(s)
16	Citrate	half-CH_2,half-CH_2	2.55(d),2.65(d)
17	Unsaturated lipid	CH=CH	2.75(m),5.31(m)
18	α-Ketoglutarate	CH_2,CH_2	2.45(t),3.02(t)
19	Creatine	N-CH_3, CH_2	3.04(s), 3.93(s)
20	Creatinine	N-CH_3, CH_2	3.05(s),4.05(s)
21	Choline	N-$(CH_3)_3$,N-CH_2, $\mathbf{CH_2}$OH	3.2(s),3.66(m), 4.30(m)
22	Phosphorylcholine(PC)/Glycerophosphocholine(GPC)	N-$(CH_3)_3$, NCH_2 CH_2OH	3.22(s), 3.61(t),3.68(t) 4.63（m)
23	Glucose	β(H2, H3, H5), α(H2, H3, H6)	β(3.24(dd),3.48(t), 3.90(dd)),α(3.54(dd), 3.71(t), 3.72(dd),3.83(m))
24	Trimethylamine-N-oxide	N-$(CH_3)_3$	3.26(s)
25	Glycine	CH_2	3.56(s)
26	Glycerol	half-CH_2, C2-H	3.58(dd), 3.66(dd)
27	Serine		3.83(dd), 3.92-4.00(m)
28	Cholesterol	C18(in VLDL)	0.70(m)
29	β-glucose	1-CH	4.66(d)
30	Glycerol of lipids	CHOCOR	5.20(m)
31	α-glucose	1-CH	5.23(d)

32	Urea	NH$_2$+NH$_2$	5.63
33	Tyrosine	CH,CH	6.90(dd), 7.20(dd)
34	Histidine	H4,H2	7.05(s),7.77(s)
35	Formate	CH	8.46(s)
36	Cholesterol	CH3(C18(in HDL), C18(in VLDL),C21),CH2COOR	0.66(m), 0.70(m), 0.91(m), 2.34(m)
37	Lipids (CH2)	CH3(CH2)n,CH2CH2CO,	0.94(m), 1.58(m)
38	Lipids (CH2CO)	CH2CO	2.25(m)
39	O-acetyl glycoproteins	CH3	2.06(s)
40	Polyunsaturated fatty acid	C=C-CH2-C=C	2.78(m)
41	Triglyceride	CH2OCOR	2.43(m), 3.53(m), 4.06(m), 4.28(m), 5.2(m) 5.23(m), 5.26(m), 5.27(m), 5.29(m),

s: single; d: doublet; t: triplet; q: quartet; m: multiplet; dd: doublet of doublet

Table S3 Differential metabolites between healthy controls and ALL patients in BPP-LED spectra.

m: multiplet

Metabolites	$\delta^1H(ppm)$	Multiplicity	P-value	Changes in the ALL patients against the healthy controls
LDL	1.25	m	0.0001	↑
Lipids (CH_2)	1.58	m	0.0004	↑
Lipids (CH_2CO)	2.25	m	0.0011	↑
PC/GPC	3.68	m	0.0005	↓
Unsaturated lipid	5.31	m	0.0007	↑

Table S4 The related parameters of logistic regression equation

Regression model	metabolites	Regression coefficients	p
The healthy controls vs. the ALL patients	Choline	-34.631	0.024
	Tyrosine	333.037	0.029
	Unsaturated liqid	49.748	0.046
	Constant	-27.649	0.087

Table S5 The evaluation of diagnostic efficacy of Logistic regression model

Positive	Negative	Specificity(true negative/negative)	Sensitivity(true positive/ positive)	The accuracy of the logistic regression model
17(true positive)	2(false negative)	94.7%	89.5%	92.1%
1(false positive)	18(true negative)			
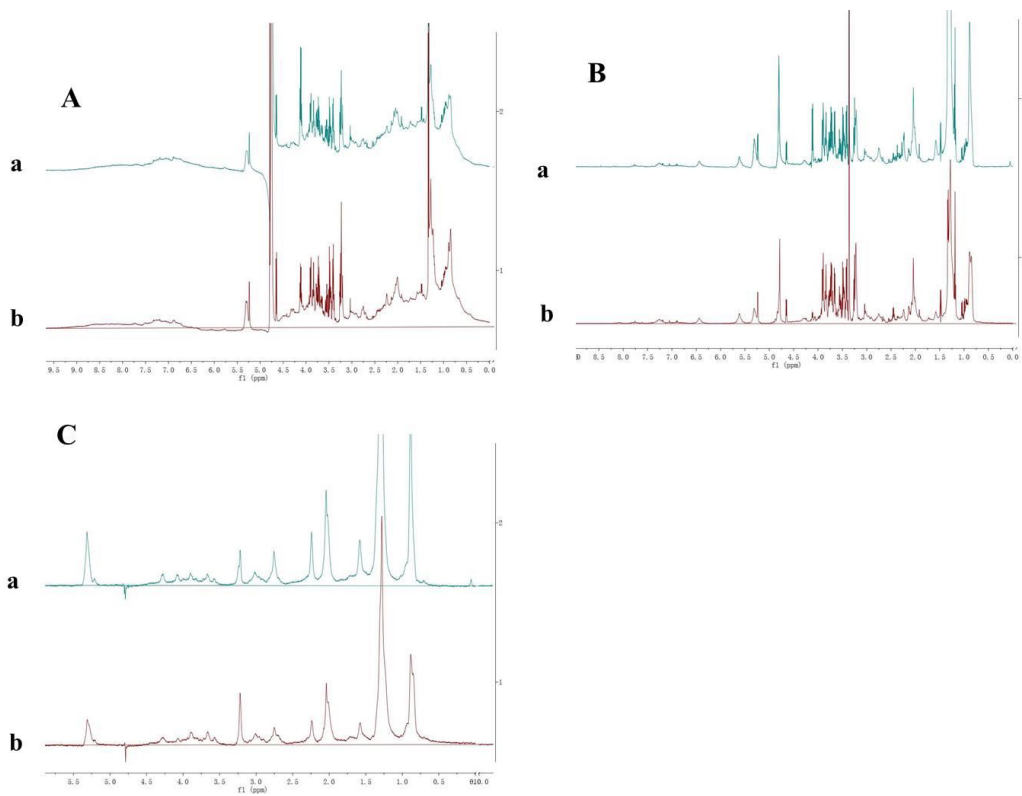
Figure S1 ^{1}H NMR spectra of plasma from healthy control (a) and ALL (b) (A) NOESY, (B) CPMG, (C) BPP-LED.
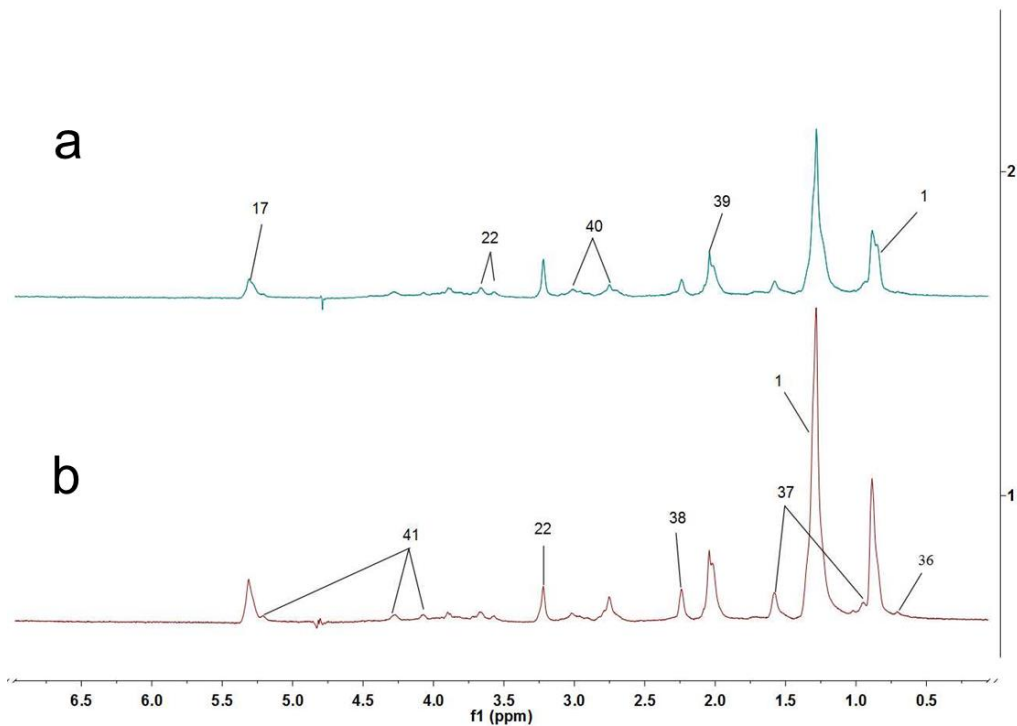
Figure S2 Typical BPP-LED ^{1}H NMR spectra of plasma for healthy control (a) and ALL patient (b).1. Lipids(mainly LDL); 17. Unsaturated lipid; 22. PC/ GPC; 36. Cholesterol; 37. Lipids (CH2); 38. Lipids (CH2CO); 39. O-acetyl glycoproteins; 40. Polyunsaturated fatty acid; 41. Triglyceride.
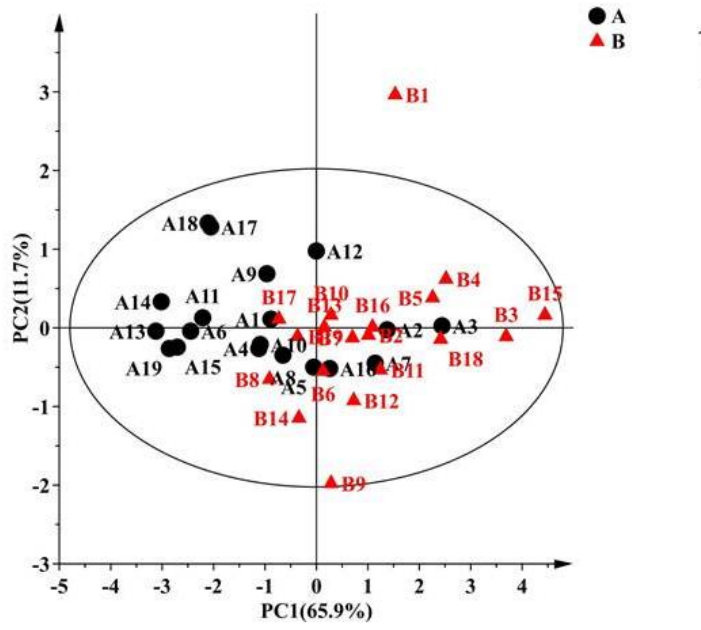
Figure S3 The scores of the PCA for the controls (green circle) and the ALL patients (blue box)
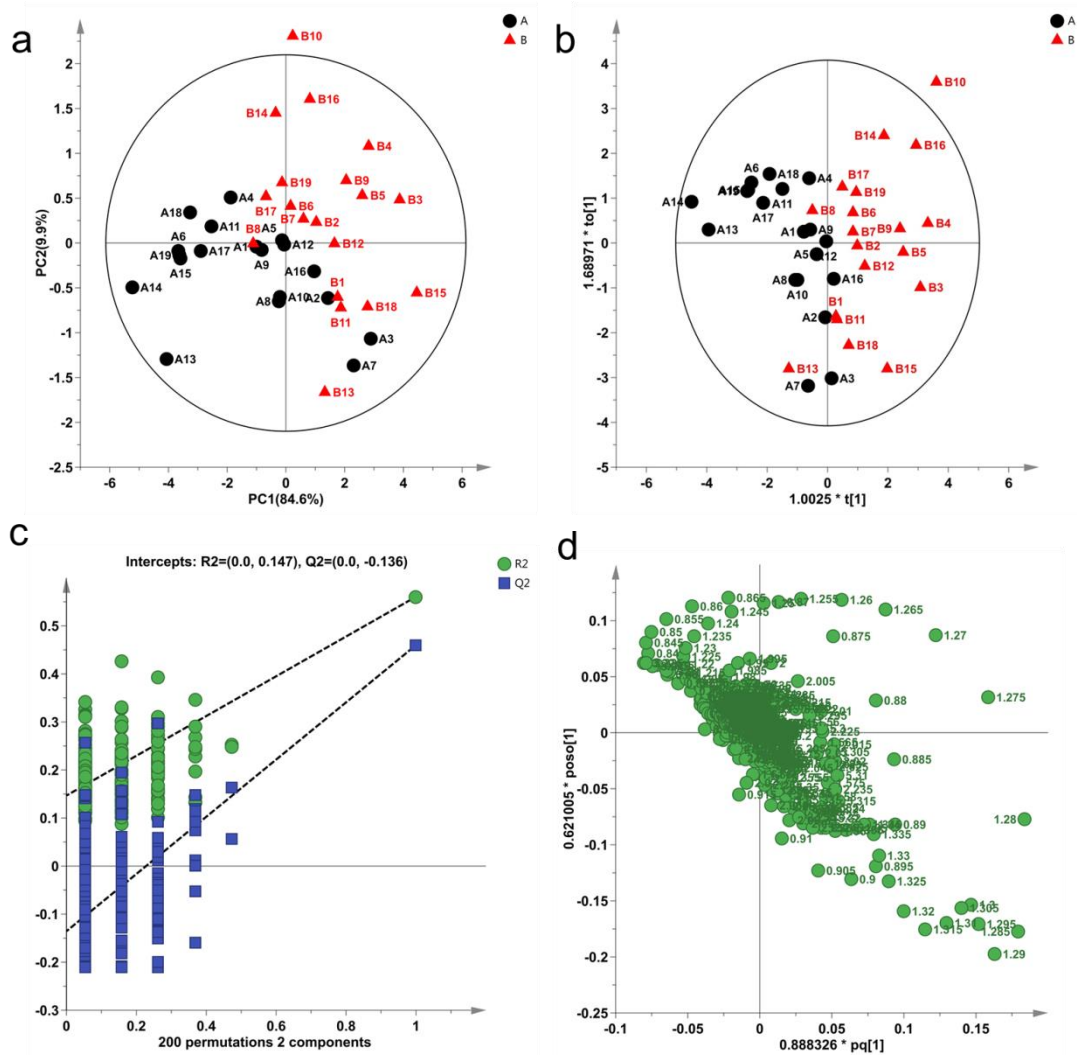in CPMG

Figure S4 Multivariate analyses of BPP-LED ^{1}H NMR spectra of healthy controls and ALL patients. (a) The scores of the PCA for healthy controls (A, black circle) and ALL patients (B, red triangle), (b) The scores of the OPLS-DA for healthy controls (A, black circle) and ALL patients (B, red triangle), (c) The permutation test result for the OPLS-DA model (R2 = (0.0, 0.147), Q2 = (0.0, -0.136)). (d) The corresponding loading plot of the OPLS-DA model.