AminoacidDB: A liquid chromatography-tandem mass spectrometry-based toolkit for untargeted analysis of non-protein amino acids.

Pawanjit K Sandhu, Ryland T. Giebelhaus, Ryan Hayward, Tingting Zhao, Alix Tucker, Daniel Gaudet, Tao Huan, Susan J Murch

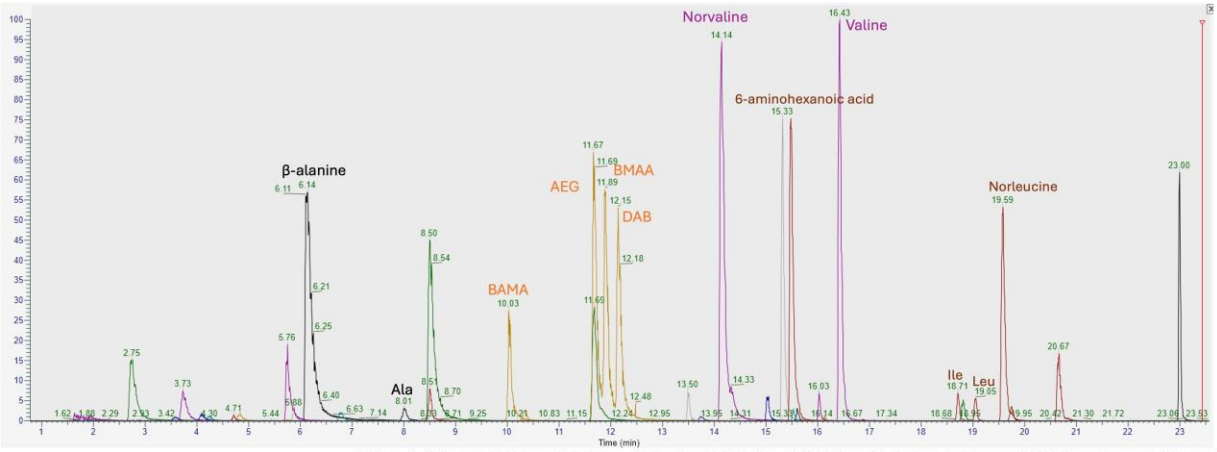
Supplementary Information

Corresponding Author

Susan Murch PhD, P.Ag.
Professor
Chemistry | Fipke 350
The University of British Columbia | Okanagan | Syilx Okanagan Nation Territory
3247 University Way | Kelowna British Columbia | V1V1V7 Canada
Phone 250 807 9566
e-mail: susan.murch@ubc.ca

The UBC Okanagan campus is situated on the traditional, ancestral, and contemporary territory of the Syilx Okanagan Nation.

Figure S1: (A) Extracted Ion Chromatogram (EIC) depicting elution profile of 39 AQC-derivatized amino acids from UHPLC-MS/MS analysis. The structural isomers are represented by same colored peaks and are highlighted by their names in the EIC.



DAB: L-2,4-Diaminobutyric acid; AEG: n-(2)-Aminoethylglycine; BMAA: β-methylamino-L-alanine; BAMA: β-aminomethyl-L-alanine

(B) Corresponding MS1 and MS2 spectra scans of BMAA separation from isomers.

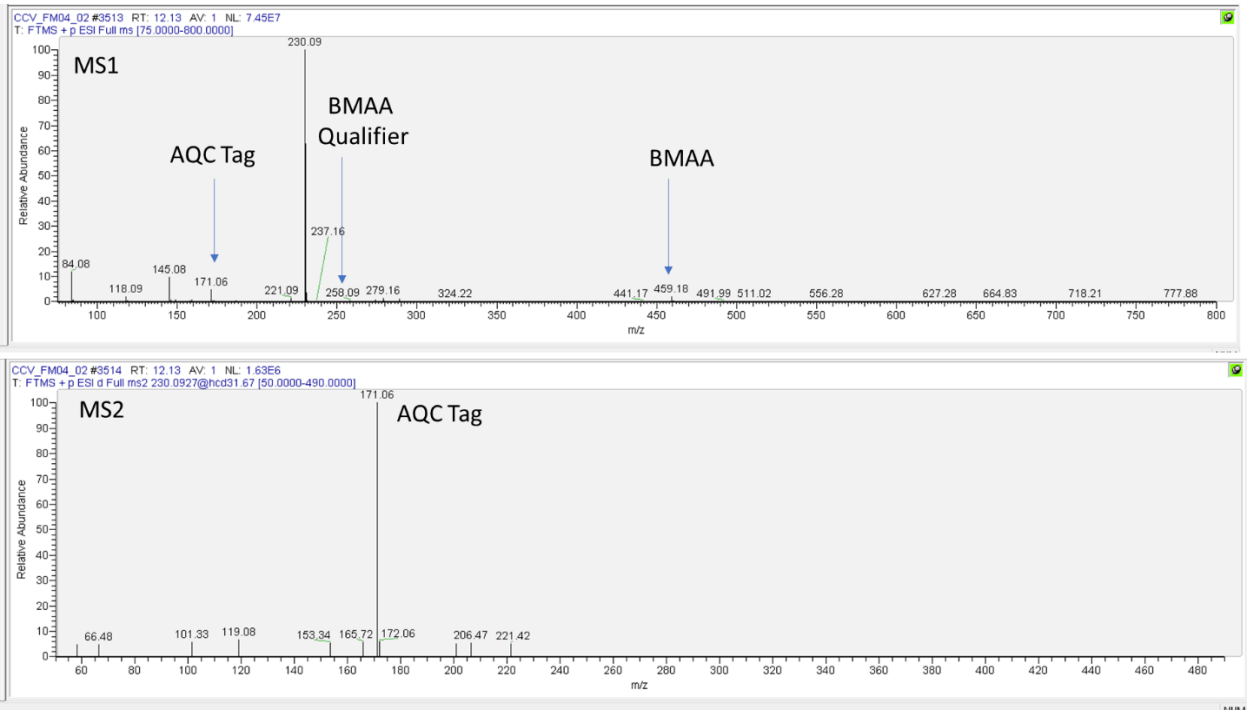


Figure S2: Monoisotopic mass distribution of 2,184,228 amino acids from PubChem (compounds containing (>1 C(=O)O + >1 (N[H])). The grey dashed lines indicate the region of 90%.

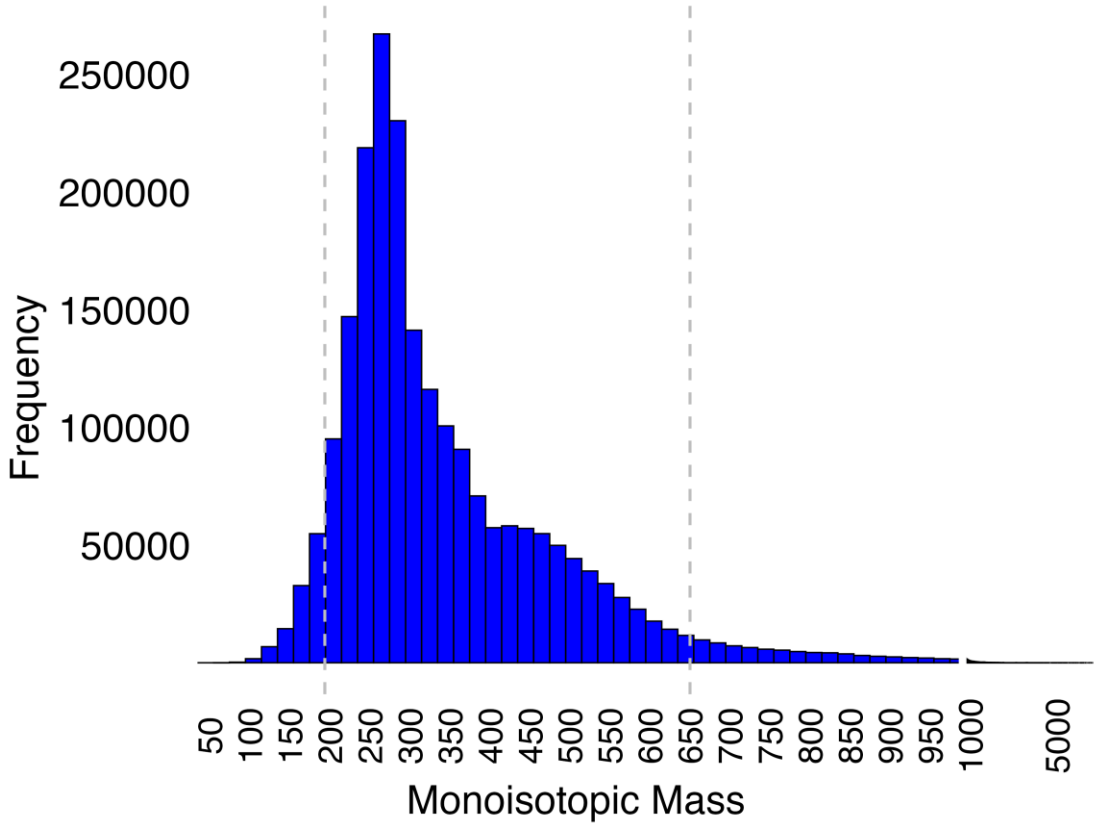


Figure S3: Number of isomers for each protein amino acid from the PubChem search. *Leucine or Isoleucine.

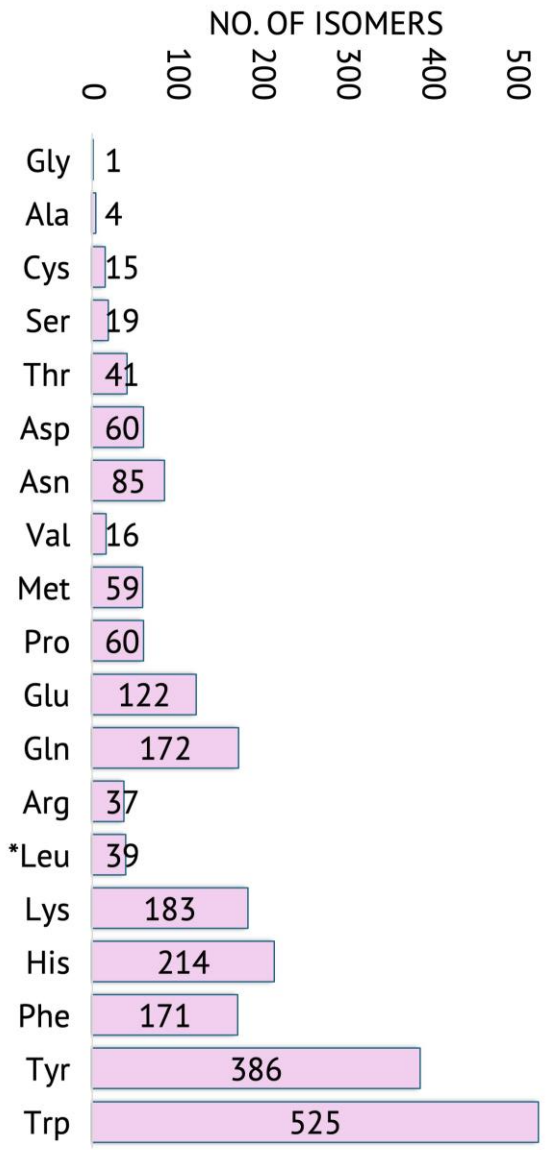

NO. OF ISOMERS
0
100
200
300
400
500
Gly
1
Ala
4
Cys
15
Ser
19
Thr
41
Asp
60
Asn
85
Val
16
Met
59
Pro
60
Glu
122
Gln
172
Arg
37
*Leu
39
Lys
183
His
214
Phe
171
Tyr
386
Trp
525

Table S1: Authentic amino acid standards used for method development and training of the machine learning algorithm.

S.No.	Name	MF	Monoisotopic mass	Amine	Source	Stock Conc. (mM)	Solvent
1	Glycine	C2H5NO2	75.032	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
2	Beta-alanine	C3H7NO2	89.04767	p	CAS No. 107-95-9; Fluka Analytical	500	0.1M aqueous HCl
3	L-Alanine	C3H7NO2	89.04767	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
4	gamma-aminobutyric acid	C4H9NO2	103.06332	p	CAS No. 56-12-2; Supelco 03835-250mg	500	0.1M aqueous HCl
5	L-Serine	C3H7NO3	105.0425	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
6	L-Proline	C5H9NO2	115.0633	s	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
7	DL-Norvaline	C5H11NO2	117.0789	p	CAS No. 760-78-1; SIGMA Kit	200	0.1M aqueous HCl
8	L-Valine	C5H11NO2	117.0789	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
9	β-aminomethyl-L-alanine	C4H10N2O2	118.07422	p	In-house custom synthesis	65	0.1M aqueous HCl
10	L-β-N-methylamino-L-alanine	C4H10N2O2	118.07422	p	CAS No. 16012-55-8, Sigma Aldrich	65	0.1M aqueous HCl
11	N-(2)-Aminoethylglycine	C4H10N2O2	118.07422	p	CAS No. 24123-14-6, Sigma Aldrich	65	0.1M aqueous HCl
12	L-2,4-Diaminobutyric acid	C4H10N2O2	118.07422	p	CAS No. 1883-09-06, TCI America	65	0.1M aqueous HCl

13	L-Threonine	C4H9NO3	119.0582	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
14	L-Cysteine	C3H7NO2S	121.0197	p	CAS No. 52-90-4; SIGMA; SIGMA	200	0.1M aqueous HCl
15	L-Isoleucine	C6H13NO2	131.0946	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
16	L-Leucine	C6H13NO2	131.0946	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
17	DL-Norleucine	C6H13NO2	131.0946	p	CAS No. N-6752; SIGMA Kit	140	0.1M aqueous HCl
18	6-aminohexanoic acid	C6H13NO2	131.09462	p	CAS No. 60-32-2; ALFA Aesar	500	0.1M aqueous HCl
19	L-Asparagine	C4H8N2O3	132.0534	p	CAS No. 70-47-3; SIGMA	250	E-pure H2O
20	DL-Ornithine	C5H12N2O2	132.0898	p	CAS No. O-2250; SIGMA Kit	200	0.1M aqueous HCl
21	L-Aspartic acid	C4H7NO4	133.0375	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
22	L-Glutamine	C5H10N2O3	146.0691	p	CAS No. 56-85-9; Fluka Analytical; Aldrich	100	0.1M aqueous HCl
23	L-Lysine	C6H14N2O2	146.1055	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
24	L-Glutamic acid	C5H9NO4	147.0531	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
25	L-Methionine	C5H11NO2S	149.051	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl

26	L-Histidine	C6H9N3O2	155.0694	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
27	L-Phenylalanine	C9H11NO2	165.0789	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
28	L-Arginine	C6H14N4O2	174.11167	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
29	DL-Citrulline	C6H13N3O3	175.09569	p	CAS No. C-6131; SIGMA Kit	200	0.1M aqueous HCl
30	L-Tyrosine	C9H11NO3	181.0738	p	Amino acid Standard H; Product #20088, Thermo Scientific	2.5	0.1M aqueous HCl
31	DL-β-3,4-dihydroxyphenylalanine	C9H11NO4	197.0688	p	CAS No. D-9503; SIGMA Kit	50	0.1M aqueous HCl
32	4-choloro-L-phenylalanine	C9H10ClNO2	199.04	p	CAS No. 14173-39-8; SIGMA	70	0.1M aqueous HCl
33	L-Tryptophan	C11H12N2O2	204.0898	p	CAS No. 73-22-3; SIGMA	80	0.1M aqueous HCl
34	5-hydroxy-L-tryptophan	C11H12N2O3	220.08479	p	CAS No. H-3753; sigma kit	100	0.1M aqueous HCl
35	L-Cystine	C6H12N2O4S2	240.02384	p	Amino acid Standard H; Product #20088, Thermo Scientific	1.25	0.1M aqueous HCl
36	DL-4-chlorophenylalanine methyl ester	C10H13Cl2NO2	249.03233	p	CAS No. 14173-40-1; aldrich	100	0.1M aqueous HCl
37	Kynurenic acid	C10H7NO3	189.04259	s	CAS No. 492-27-3; Millipore Sigma	52	100% Ethanol
38	1-Aminocyclopropane-1-carboxylic acid	C4H7NO2	101.04767	p	CAS No. 22059-21-8; SIGMA	100	E-pure H2O
39	Kynurenine	C10H12N2O3	208.08479	p	CAS No. 2922-83-0; Sigma-Aldrich	48	10% Methanol in Water

40	3-hydroxy-DL-kynurenine	C10H12N2O4	224.0797	p	CAS No. 484-78-6; Sigma-Aldrich	44	E-pure H2O
41	S-(5'-Adenosyl)-L-methionine	C15H23N6O5S+	399.14506	p	CAS No. 86867-01-8; Millipore Sigma	0.78	0.1M aqueous HCl

Table S2: Dilution series concentrations of authentic amino acid standards used for method development and training of aminoacidDB.

S.No.	Amino acid	Conc. (pg/uL) post-derivatization with AQC									
		Dilution 1	**Dilution 2**	**Dilution 3**	**Dilution 4**	**Dilution 5**	**Dilution 6**	**Dilution 7**	**Dilution 8**	**Dilution 9**	**Dilution 10**
1	Glycine	93.84	46.92	23.46	11.73	5.86	2.93	1.47	0.73	0.37	0.18
2	Beta-alanine	6186.81	3093.40	1546.70	773.35	386.68	193.34	96.67	48.33	24.17	-
3	L-Alanine	111.36	55.68	27.84	13.92	6.96	3.48	1.74	0.87	0.44	0.22
4	gamma-aminobutyric acid	7161.11	3580.56	1790.28	895.14	447.57	223.78	111.89	55.95	27.97	-
5	L-Serine	131.36	65.68	32.84	16.42	8.21	4.11	2.05	1.03	0.51	0.26
6	L-Proline	143.91	71.96	35.98	17.99	8.99	4.50	2.25	1.12	0.56	0.28
7	DL-Norvaline	3254.17	1627.08	813.54	406.77	203.39	101.69	50.85	25.42	12.71	-
8	L-Valine	146.44	73.22	36.61	18.30	9.15	4.58	2.29	1.14	0.57	0.29
9	β-aminomethyl-L-alanine	1066.45	533.23	266.61	133.31	66.65	33.33	16.66	8.33	4.17	-
10	L-β-N-methylamino-L-alanine	1066.45	533.23	266.61	133.31	66.65	33.33	16.66	8.33	4.17	-
11	N-(2)-Aminoethylglycine	1066.45	533.23	266.61	133.31	66.65	33.33	16.66	8.33	4.17	-
12	L-2,4-Diaminobutyric acid	1066.45	533.23	266.61	133.31	66.65	33.33	16.66	8.33	4.17	-
13	L-Threonine	148.90	74.45	37.23	18.61	9.31	4.65	2.33	1.16	0.58	0.29
14	L-Cysteine	3365.56	1682.78	841.39	420.69	210.35	105.17	52.59	26.29	13.15	-

15	L-Isoleucine	163.96	81.98	40.99	20.50	10.25	5.12	2.56	1.28	0.64	0.32
16	L-Leucine	163.96	81.98	40.99	20.50	10.25	5.12	2.56	1.28	0.64	0.32
17	DL-Norleucine	2368.35	1184.17	592.09	296.04	148.02	74.01	37.01	18.50	9.25	-
18	6-aminohexanoic acid	9109.24	4554.62	2277.31	1138.65	569.33	284.66	142.33	71.17	35.58	-
19	L-Asparagine	4587.50	2293.75	1146.88	573.44	286.72	143.36	71.68	35.84	17.92	-
20	DL-Ornithine	2880.51	1440.26	720.13	360.06	180.03	90.02	45.01	22.50	11.25	-
21	L-Aspartic acid	166.38	83.19	41.59	20.80	10.40	5.20	2.60	1.30	0.65	0.32
22	L-Glutamine	2029.72	1014.86	507.43	253.72	126.86	63.43	31.71	15.86	7.93	-
23	L-Lysine	182.74	91.37	45.68	22.84	11.42	5.71	2.86	1.43	0.71	0.36
24	L-Glutamic acid	183.91	91.96	45.98	22.99	11.49	5.75	2.87	1.44	0.72	0.36
25	L-Methionine	186.51	93.26	46.63	23.31	11.66	5.83	2.91	1.46	0.73	0.36
26	L-Histidine	193.94	96.97	48.48	24.24	12.12	6.06	3.03	1.52	0.76	0.38
27	L-Phenylalanine	206.49	103.24	51.62	25.81	12.91	6.45	3.23	1.61	0.81	0.40
28	L-Arginine	217.75	108.88	54.44	27.22	13.61	6.80	3.40	1.70	0.85	0.43
29	DL-Citrulline	4866.39	2433.19	1216.60	608.30	304.15	152.07	76.04	38.02	19.01	-
30	L-Tyrosine	226.49	113.24	56.62	28.31	14.16	7.08	3.54	1.77	0.88	0.44
31	DL-β-3,4-dihydroxyphenylalanine	1369.38	684.69	342.34	171.17	85.59	42.79	21.40	10.70	5.35	-
32	4-chloro-L-phenylalanine	1848.43	924.21	462.11	231.05	115.53	57.76	28.88	14.44	7.22	-
33	L-Tryptophan	2269.11	1134.56	567.28	283.64	141.82	70.91	35.45	17.73	8.86	-
34	5-hydroxy-L-tryptophan	3058.75	1529.38	764.69	382.34	191.17	95.59	47.79	23.90	11.95	-
35	L-Cystine	150.18	75.09	37.55	18.77	9.39	4.69	2.35	1.17	0.59	0.29
36	1-Aminocyclopropane-1-carboxylic acid	351.04	175.52	87.76	43.88	21.94	10.97	5.49	2.74	1.37	-
37	L-Kynurenine	1388.65	694.32	347.16	173.58	86.79	43.40	21.70	10.85	5.42	-

38	3-hydroxy-DL-kynurenine	1388.86	694.43	347.21	173.61	86.80	43.40	21.70	10.85	5.43	-
39	S-(5'-Adenosyl)-L-methionine	5557.20	2778.60	1389.30	694.65	347.32	173.66	86.83	43.42	21.71	-

Table S3: Optimized UHPLC elution gradient for separation of non-protein amino acids.

Time (min)	%A	%B	Curve
0	98	2	
1	98	2	5
15	87	13	7
20	85	15	5
25	40	60	5
27	5	95	5
28	5	95	5
28.1	98	2	5
30	98	2	

Table S4: Method performance characteristics of amino acid standards used for method development and training of aminoacidDB. LOD and LOQ refers to limits of detection and limits of quantification in picograms respectively. Comparison to standard curves provides MS1 level data in the plant sample datasets. %RSD is percent residual standard deviation for retention time (RT) and Intra-Day and Inter-Day MS detector response for each of the amino acids.

S.No.	Name	pg on column		Calibration curve parameters			% RSD		
		LOD	LOQ	Slope	Intercept	R^2	RT	Intra-day MS detector Response	Inter-day MS detector Response
1	Glycine	0.5	1.8	2209105.3	1521990.5	0.996	2.2	13.4	25.3
2	Beta-alanine	73.2	241.7	24787063.3	1150464.9	0.996	1.5	8.5	8.5
3	L-Alanine	5.3	17.4	1177311.0	999694.8	0.997	1.4	7.9	14.8
4	gamma-aminobutyric acid	84.8	279.7	24965023.3	1002223.2	0.996	1.5	8.3	7.0
5	L-Serine	6.2	20.5	1460811.5	385088.4	0.996	3.1	23.9	30.3
6	L-Proline	0.8	2.8	877477.2	652924.1	0.999	1.2	8.4	12.0
7	DL-Norvaline	38.5	127.1	-9675131.2	1044633.4	0.994	0.6	2.3	4.3
8	L-Valine	0.9	2.9	146001.4	846432.7	0.997	0.3	4.4	8.1
9	β-aminomethyl-L-alanine	12.6	41.7	-5691710.4	1779419.3	0.995	0.8	2.3	5.4
10	L-β-N-methylamino-L-alanine	12.6	41.7	-326887.2	92702.3	0.999	0.6	8.6	12.2
11	N-(2)-Aminoethylglycine	12.6	41.7	-9222132.4	4938907.7	1.000	0.5	3.1	4.2
12	L-2,4-Diaminobutyric acid	12.6	41.7	-10036367.3	3825571.8	0.999	0.7	4.9	4.7
13	L-Threonine	7.1	23.3	776628.3	566669.1	1.000	1.5	7.1	13.3
14	L-Cysteine	39.8	131.5	-2544367.8	144953.8	0.971	0.8	73.7	48.7
15	L-Isoleucine	1.0	3.2	16517.1	758364.3	0.998	0.3	5.2	6.5
16	L-Leucine	1.0	3.2	1355.1	950911.2	0.997	0.3	4.8	6.5
17	DL-Norleucine	28.0	92.5	-3701615.9	1356414.6	0.997	0.3	4.9	6.2

18	6-aminohexanoic acid	107.8	355.8	6714550.7	1198310.1	0.994	0.3	5.3	5.6
19	L-Asparagine	54.3	179.2	-730180.1	420706.3	0.997	3.8	13.2	14.3
20	DL-Ornithine	34.1	112.5	344994077.6	1260188.1	1.000	0.7	5.1	4.3
21	L-Aspartic acid	2.0	6.5	742623.4	1583624.4	0.948	2.0	10.6	20.1
22	L-Glutamine	24.0	79.3	-152050.5	409198.5	0.999	2.8	15.5	18.6
23	L-Lysine	8.7	28.6	1481893.6	1914840.3	0.997	0.3	4.4	11.5
24	L-Glutamic acid	8.7	28.7	131877.1	253481.4	1.000	1.7	14.5	21.5
25	L-Methionine	1.1	3.6	-147623.5	347857.7	0.999	0.2	7.8	11.6
26	L-Phenylalanine	1.2	4	13284.0	444891.2	1.000	0.2	4.8	7.8
27	L-Arginine	10.3	34	-65841.0	59687.0	0.996	3.6	28.1	35.2
28	DL-Citrulline	57.6	190.1	-3209533.1	266649.6	1.000	1.7	3.6	6.6
29	L-Tyrosine	1.3	4.4	-91754.8	280900.7	0.999	0.2	6.9	7.1
30	DL-β-3,4-dihydroxyphenylalanine	16.2	53.5	-341633.6	244913.1	0.996	0.6	7.0	7.2
31	4-chloro-L-phenylalanine	21.9	72.2	-2011432.4	1263527.9	0.999	0.1	10.8	9.1
32	L-Tryptophan	26.8	88.6	-3853486.8	393327.2	0.997	0.2	5.0	7.0
33	5-hydroxy-L-tryptophan	36.2	119.5	-4811209.1	109692.6	0.991	0.7	5.5	9.2
34	L-Cystine	0.9	2.9	-160554.2	593094.8	0.999	0.2	4.1	4.0
35	1-Aminocyclopropane-1-carboxylic acid	4.2	13.7	-540194.8	1813561.4	0.997	1.2	5.3	4.7
36	L-Kynurenine	16.4	54.2	-679829.6	189970.8	0.997	0.2	3.3	5.8
37	3-hydroxy-DL-kynurenine	16.5	54.3	-174727.6	36810.3	0.999	0.6	14.9	15.3
38	S-(5’-Adenosyl)-L-methionine	65.8	217.1	-409613.8	30351.0	0.997	4.7	16.5	14.6

Table S5: List of standards used to build retention time prediction model using Retip 2.0.

S.No.	Name	Monoisotopic Mass	M+H	RT (minutes)	SMILES	InChiKey
1	β-aminomethyl-L-alanine (BAMA)	118.07422	119.0815	1.03	NCC(NC)C(O)=O	LYESEIVAYITDEH-UHFFFAOYSA-N
2	β-aminomethyl-L-alanine_AQC	458.17022	459.1775	9.27	O=C(NCC(N€C(NC1=CC=C2N=CC=CC2=C1)=O)C(O)=O)NC3=CC=C4N=CC=CC4=C3	YVMXLIZLVIKBJX-UHFFFAOYSA-N
3	(-)-Caryophyllene oxide	220.182715	221.189995	27.67	C[C@@]12CC[C@@H]3[C@H](CC3€C)C(=C)CC[C@H]1O2	NVEQFIOZRFFVFW-RGCMKSIDSA-N
4	(-)-Epinephrine	183.08954	184.09682	1.17	CNC[C@@H](C1=CC(=C(C=C1)O)O)O	UCTWMZQNUQWSLP-VIFPVBQESA-N
5	(-)-Guaiol	222.198365	223.205645	27.71	C[C@H]1CC[C@H](CC2=C1CC[C@@H]2C)C€€O	TWVJWDMOZJXUID-SDDRHHMPSA-N
6	(-)-Norepinephrine	169.07389	170.08117	1.2	C1=CC(=C(C=C1[C@H](CN)O)O)	SFLSHLFXELFNJZ-QMMMGPOBSA-N
7	(-)-Sabinene	136.1252	137.13248	25.36	CC€[C@@]12CCC(=C)[C@@H]1C2	NDVASEGYNIMXJL-UWVGGRQHSA-N
8	(-)-α-Bisabolol	222.198365	223.205645	27.71	CC1=CC[C@H](CC1)[C@]€(CCC=C€C)O	RGZSQWQPBWRIAQ-CABCVRRESA-N
9	(+)-Alpha-Damascone	192.151415	193.158695	26.9	C/C=C/C(=O)C1C(=CCCC1€C)C	CRIGTVCBMUKRSL-FNORWQNLSA-N
10	(+)-Nootkatone	218.167065	219.174345	26.76	C[C@@H]1CC(=O)C=C2[C@]1(C[C@@H](CC2)C(=C)C)C	WTOYNNBCKUYIKC-JMSVASOKSA-N
11	(+/-)-Abscisic acid	264.136159	265.143439	22.88	CC1=CC(=O)CC(C1(/C=C/C(=C\C(=O)O)/C)O)€C	JLIDBLDQVAYHNE-LXGGSRJLSA-N
12	(Aminomethyl) Phosphonic Acid	111.00853	112.01581	1.3	C(N)P(=O)(O)O	MGRVRXRGTBOSHW-UHFFFAOYSA-N
13	€-(+)-Limonene	136.1252	137.13248	25.55	CC1=CC[C@@H](CC1)C(=C)C	XMGQYMWWDOXHJM-JTQLQIEISA-N
14	1-Aminocyclopropane-1-carboxylic acid	101.04767	102.05495	1.03	C1CC1(C(=O)O)N	PAJPWUMXBYXFCZ-UHFFFAOYSA-N

15	1-Aminocyclopropane-1-carboxylic acid_AQC	271.09567	272.10295	8.03	O=C(N([H])C1(C(O)=O)CC1)NC2=CC=C3N=CC=CC3=C2	NAUFGQKGGJGWSM-UHFFFAOYSA-N
16	1-Ephedrine	165.11536	166.12264	4.5	CC(C(C1=CC=CC=C1)O)NC	KWGRBVOPPLSCSI-UHFFFAOYSA-N
17	1-Napthaleneacetic acid	186.068	187.07582	27.4	C1=CC=C2C(=C1)C=CC=C2CC(=O)O	PRPINYUDVPFIRX-UHFFFAOYSA-N
18	2-Iodomelatonin	358.01783	359.02511	23.35	CC(=O)NCCC1=C(NC2=C1C=C(C=C2)OC)I	FJDDSMSDZHURBJ-UHFFFAOYSA-N
19	2-Isopentenyl adenine	203.11709	204.12491	12.62	CC(=CCNC1=NC=NC2=C1NC=N2)C	HYVABZIGRDEKCD-UHFFFAOYSA-N
20	3-hydroxy-DL-kynurenine	224.0797	225.08698	1.64	C1=CC(=C(C(=C1)O)N)C(=O)CC(C(=O)O)N	VCKPUUFAIGNJHC-UHFFFAOYSA-N
21	3-hydroxy-DL-kynurenine_AQC	394.1277	395.13498	15.52	O=C(NC(CC(C1=C(C(O)=CC=C1)N)=O)C(O)=O)NC2=CC=C3N=CC=CC3=C2	BALHTGIPAWNDAS-UHFFFAOYSA-N
22	3-Hydroxytyramine	153.07897	154.08625	1.2	C1=CC(=C(C=C1CCN)O)O	VYFYYTLLBUKUHU-UHFFFAOYSA-
23	3-Indoleacetic acid	175.063328	176.070608	17.13	C1=CC=C2C(=C1)C(=CN2)CC(=O)O	SEOVTRFCIGRIMH-UHFFFAOYSA-N
24	3’,5’-Dimethoxy-4’-hydroxyacetophenone	196.073558	197.080838	16.32	CC(=O)C1=CC(=C(C(=C1)OC)O)OC	OJOBTAOGJIWAGB-UHFFFAOYSA-N
25	4-chloro-L-phenylalanine	199.04	200.04728	6.6	C1=CC(=CC=C1C[C@@H](C(=O)O)N)Cl	NIGWMJHCCYYCSF-QMMMGPOBSA-N
26	4-choloro-L-phenylalanine _AQC	369.088	370.09528	22.63	O=C(NC(CC1=CC=C(Cl)C=C1)C(O)=O)NC2=CC=C3N=CC=CC3=C2	ZSNRDIAJCJIEDD-UHFFFAOYSA-N
27	4-Oxoisophorone – 98%	152.08372	153.091	18.1	CC1=CC(=O)CC(C1=O)€C	AYJXHIDNNLJQDT-UHFFFAOYSA-N
28	4-P-PDOT	279.162314	280.169594	25.72	CCC(=O)NC1CC(C2=CC=CC=C2C1)C3=CC=CC=C3	RCYLUNPFECYGDW-UHFFFAOYSA-N
29	5-Chlorotryptamine	194.06107	195.06835	13.7	C1=CC2=C(C=C1Cl)C(=CN2)CCN	FVQKQPVVCKOWLM-UHFFFAOYSA-N

30	5-Deoxy-5'-(methylthio)adenosine	297.08956	298.09684	6.21	CSC[C@@H]1[C@H]([C@H]([C@@H](O1)N2C=NC3=C(N=CN=C32)N)O)O	WUUGFSXJNOTRMR-IOSLPCCCSA-N
31	5-hydroxy-L-tryptophan	220.08479	221.09207	1.8	C1=CC2=C(C=C1O)C(=CN2)CC(C(=O)O)N	LDCYZAJDBXYCGN-UHFFFAOYSA-N
32	5-hydroxy-L-tryptophan_AQC	390.13279	391.14007	14.18	O=C(NC(CC1=CNC2=C1C=C(O)C=C2)C(O)=O)NC3=CC=C4N=CC=CC4=C3	MGARJOPHAJUMPH-UHFFFAOYSA-N
33	5-hydroxy-L-tryptophan_AQC	560.18079	561.18807	17.8	O=C(NC(CC1=CN(C(NC2=CC=C3N=CC=CC3=C2)=O)C4=C1C=C(O)C=C4)C(O)=O)NC5=CC=C6N=CC=CC6=C5	GHPHVJOASJNJAI-UHFFFAOYSA-N
34	6-(Gamma,gamma-dimethylallylamino)Purine	203.117095	204.124375	13.16	CC(=CCNC1=NC=NC2=C1NC=N2)C	HYVABZIGRDEKCD-UHFFFAOYSA-N
35	6-aminohexanoic acid	131.0946	132.10188	1.2	C(CCC(=O)O)CCN	SLXKOJJOQWFEFD-UHFFFAOYSA-N
36	6-aminohexanoic acid_AQC	301.14262	302.1499	14.78	OC(CCCCCNC(NC1=CC=C2N=CC=CC2=C1)=O)=O	NTMMYJWQCRNXHW-UHFFFAOYSA-N
37	6-Benzylaminopurine	225.1014	226.10868	14.1	C1=CC=C(C=C1)CNC2=NC=NC3=C2NC=N3	NWBJYWHLCVSVIJ-UHFFFAOYSA-N
38	7-Methoxycoumarin-4-acetic acid	234.05282	235.0601	18.8	COC1=CC2=C(C=C1)C(=CC(=O)O2)CC(=O)O	ZEKAXIFHLIITGV-UHFFFAOYSA-N
39	7,8-Dihydro-α-ionone	194.16706	195.17434	27.24	CC1=CCCC(C1CCC(=O)C)€C	JHJCHCSUEGPIGE-UHFFFAOYSA-N
40	Acequinolyl_hydrolyzed	342.21949	343.2271	28.75	CCCCCCCCCCCCC1=C(O)C(C2=CC=CC=C2C1=O)=O	KUUFMNYPHOKBFD-UHFFFAOYSA-N
41	Acetaminophen	151.06332	152.0706	3.2	CC(=O)NC1=CC=C(C=C1)O	RZVAJINKPMORJF-UHFFFAOYSA-N
42	Acetovanillone	166.06299	167.07027	14.24	CC(=O)C1=CC(=C(C=C1)O)OC	DFYRUELUNQRZTB-UHFFFAOYSA-N
43	Adenine	135.05449	136.06177	1.2	C1=NC2=NC=NC(=C2N1)N	GFFGJBXGBJISGV-UHFFFAOYSA-N

44	Aflatoxin B1 (from Aspergillus flavus)	312.06338	313.0712	23.62	COC1=C2C3=C(C(=O)CC3)C(=O)OC2=C4[C@@H]5C=CO[C@@H]5OC4=C1	OQIQSTLJSLGHID-WNWIJWBNSA-N
45	Aflatoxin B2	314.07903	315.08685	23.3	COC1=C2C3=C(C(=O)CC3)C(=O)OC2=C4[C@@H]5CCO[C@@H]5OC4=C1	WWSYXEZEXMQWHT-WNWIJWBNSA-N
46	Aflatoxin G1 (from Aspergillus flavus)	328.0583	329.06612	23.12	COC1=C2C3=C(C(=O)OCC3)C(=O)OC2=C4C5C=COC5OC4=C1	XWIYFDMXXLINPU-UHFFFAOYSA-N
47	Aflatoxin G2	330.07395	331.08177	22.45	COC1=C2C3=C(C(=O)OCC3)C(=O)OC2=C4[C@@H]5CCO[C@@H]5OC4=C1	WPCVRWVBBXIRMA-WNWIJWBNSA-N
48	Alpha-Ionone	192.151415	193.158695	26.9	CC1=CCCC(C1/C=C/C(=O)C)€C	UZFLPKAIBPNNCA-BQYQJAHWSA-N
49	Alpha-Naphthoflavone	272.083729	273.091009	26.52	C1=CC=C(C=C1)C2=CC(=O)C3=C(O2)C4=CC=CC=C4C=C3	VFMMPHCGEFXGIP-UHFFFAOYSA-N
50	alpha-Terpinene	136.1252	137.13248	25.37	CC1=CC=C(CC1)C€C	YHQGMYUVUMAZJR-UHFFFAOYSA-N
51	Ancymidol	256.121177	257.128457	23.34	COC1=CC=C(C=C1)C(C2CC2)(C3=CN=CN=C3)O	HUTDUHSNJYTCAR-UHFFFAOYSA-N
52	Apigenin	270.05282	271.0601	23.62	C1=CC(=CC=C1C2=CC(=O)C3=C(C=C(C=C3O2)O)O)O	KZNIFHPLKGYRTM-UHFFFAOYSA-N
53	Artemisinin	282.146723	283.154003	25.83	C[C@@H]1CC[C@H]2[C@H](C(=O)O[C@H]3[C@@]24[C@H]1CC[C@](O3)(OO4)C)C	BLUAFEHZUWYNDE-NNWCWBAJSA-N
54	b-alanine	89.04767	90.05495	1.17	C(CN)C(=O)O	UCMIRNVEIXFBKS-UHFFFAOYSA-N
55	B-Eudesmol	222.1984	223.20568	26.97	C[C@]12CCCC(=C)[C@@H]1C[C@@H](CC2)C€€O	BOPIMTNSYWYZOC-VNHYZAJKSA-N
56	Baicalein	270.05282	271.0601	24.14	C1=CC=C(C=C1)C2=CC(=O)C3=C(O2)C=C(C(=C3O)O)O	FXNFHKRTJBSTCS-UHFFFAOYSA-N
57	Baicalin	446.084911	447.092191	22.55	C1=CC=C(C=C1)C2=CC(=O)C3=C(C(=C(C=C3O2)O[C@H]4[C	IKIIZLYTISPENI-ZFORQUDYSA-N

					@@H]([C@H]([C@@H]([C@H](O4)C(=O)O)O)O)O)O)O	
58	Beta-alanine_AQC	259.09567	260.10295	5.81	O=C(NCCC(O)=O)NC1=CC=C2N=CC=CC2=C1	CKQDPPRGGAJCBA-UHFFFAOYSA-N
59	Beta-Ionone 96%	192.151415	193.158695	26.9	CC1=C(C(CCC1)€C)/C=C/C(=O)C	PSQYTAPXSHCGMF-BQYQJAHWSA-N
60	Brassinoiide	480.345089	481.352369	25.44	CC€C€C(C(C€C1CCC2C1(CCC3C2COC(=O)C4C3(CC(C(C4)O)O)C)C)O)O	IXVMHGVQKLDRKH-UHFFFAOYSA-N
61	Caffeine	194.08037	195.08765	9.35	CN1C=NC2=C1C(=O)N(C(=O)N2C)C	RYYVLZVUVIJVGH-UHFFFAOYSA-N
62	Cannabicycolic acid	358.2144	359.22222	28.64	CCCCCC1=CC2=C(C3C4C(C3€C)CCC4(O2)C)C(=C1C(=O)O)O	JVOHLEIRDMVLHS-UHFFFAOYSA-N
63	Cannabidiol (CBD)	314.22458	315.2324	27.69	CCCCCC1=CC(=C(C(=C1)O)[C@@H]2C=C(CC[C@H]2C(=C)C)C)O	QHMBSVQNZZTUGM-ZWKOTPCHSA-N
64	Cannabigerol (CBG)	316.24023	317.24805	27.61	CCCCCC1=CC(=C(C(=C1)O)C/C=C(\C)/CCC=C€C)O	QXACEHWTBCFNSA-SFQUDFHCSA-N
65	Cannabigerolic Acid (CBGA)	360.230059	361.237879	27.52	CCCCCC1=CC(=C(C(=C1C(=O)O)O)C/C=C(\C)/CCC=C€C)O	SEEZIOZEUUMJME-FOWTUZBSSA-N
66	Cannabinol (CBN)	310.19328	311.2011	28.04	CCCCCC1=CC(=C2C(=C1)OC(C3=C2C=C(C=C3)C)€C)O	VBGLYOIFKLUMQG-UHFFFAOYSA-N
67	Cannabinolic Acid (CBNA)	354.183109	355.190929	28.24	CCCCCC1=CC2=C(C3=C(C=CC(=C3)C)C(O2)€C)C(=C1C(=O)O)O	KXKOBIRSQLNUPS-UHFFFAOYSA-N
68	CBDV	286.19328	287.2011	27.11	CCCC1=CC(=C(C(=C1)O)[C@@H]2C=C(CC[C@H]2C(=C)C)C)O	REOZWEGFPHTFEI-JKSUJKDBSA-N
69	Cedrol	222.198365	223.205645	27.71	C[C@@H]1CC[C@@H]2[C@]13CC[C@@]([C@H](C3)C2€C)€O	SVURIXNDRWRAFU-OGMFBOKVSA-N
70	Chlorogenic acid	354.09508	355.10236	8.1	C1[C@H]([C@H]([C@@H](C[C@@]1(C(=O)O)O)OC(=O)/C=C/C2=CC(=C(C=C2)O)O)O)O	CWVRJTMFETXNAD-JUHZACGLSA-N

71	Citric acid	192.027	193.03428	1.5	C(C(=O)O)C(CC(=O)O)(C(=O)O)O	KRKNYBCHXYNGOX-UHFFFAOYSA-N
72	Citronellol	156.151415	157.158695	25.85	CC(CCC=C€C)CCO	QMVPMAAFGQKVCJ-UHFFFAOYSA-N
73	Corticosterone, >92%	346.214409	347.221689	24.26	C[C@]12CCC(=O)C=C1CC[C@@H]3[C@@H]2[C@H](C[C@]4([C@H]3CC[C@@H]4C(=O)CO)C)O	OMFXVFTZEKFJBZ-HJTSIMOOSA-N
74	Cyanidin-3-Arabinoside Standard	419.09782	420.1051	11.9	C1[C@@H]([C@@H]([C@H]([C@@H](O1)OC2=CC3=C(C=C(C=C3[O+]=C2C4=CC(=C(C=C4)O)O)O)O)O)O)O	KUCVMQMKRICXJC-DEYWJSPQSA-O
75	Cyanidin-3-Galactoside Standard	449.10838	450.11566	1.58	C1=CC(=C(C=C1C2=[O+]C3=CC(=CC(=C3C=C2O[C@H]4[C@@H]([C@H]([C@H]([C@H](O4)CO)O)O)O)O)O)O)O	RKWHWFONKJEUEF-WVXKDWSHSA-O
76	Cytosine	111.04326	112.05054	1.12	C1=C(NC(=O)N=C1)N	OPTASPLRGRRNAP-UHFFFAOYSA-N
77	d-Valerolactam; 2-piperidone	99.068413	100.075693	2.75	C1CCNC(=O)C1	XUWHAWMETYGRKB-UHFFFAOYSA-N
78	d8-Tetrahydrocannabinol (d8-THC)	314.22458	315.2324	28.41	CCCCCC1=CC(=C2[C@@H]3CC(=CC[C@H]3C(OC2=C1)€C)C)O	HCAWPGARWVBULJ-IAGOWNOFSA-N
79	d9-Tetrahydrocannabinol (d9-THC)	314.22458	315.2324	28.25	CCCCCC1=CC(=C2[C@@H]3C=C(CC[C@H]3C(OC2=C1)€C)C)O	CYQFCXCEBYINGO-IAGOWNOFSA-N
80	DL-4-chlorophenylalanine methyl ester	213.05565	214.06293	13.58	COC(=O)C(CC1=CC=C(C=C1)Cl)N	FBHPVOLRIXZZMD-UHFFFAOYSA-N
81	DL-beta-3,4-DOPA	197.0688	198.07608	1.5	C1=CC(=C(C=C1CC(C(=O)O)N)O)O	WTDRDQBEARUVNC-UHFFFAOYSA-N
82	DL-beta-3,4-DOPA_AQC	367.1168	368.12408	12.64	O=C(NC(CC1=CC(O)=C(C=C1)O)C(O)=O)NC2=CC=C3N=CC=CC3=C2	BFXBRTAYUCZMFZ-UHFFFAOYSA-N

83	DL-Citrulline	175.09569	176.10297	1.2	C(CC(C(=O)O)N)CNC(=O)N	RHGKLRLOHDJJDR-UHFFFAOYSA-N
84	DL-Citrulline_AQC	345.14369	346.15097	5.35	O=C(NC(C(O)=O)CCCNC(N)=O)NC1=CC=C2N=CC=CC2=C1	QWLPTLROMFJBHD-UHFFFAOYSA-N
85	DL-Norleucine	131.0946	132.10188	2	CCCCC(C(=O)O)N	LRQKBLKVPFOOQJ-UHFFFAOYSA-N
86	DL-Norleucine_AQC	301.1426	302.14988	18.4	OC(C(CCCC)NC(NC1=CC=C2N=CC=CC2=C1)=O)=O	WBEASBBDJFMOND-UHFFFAOYSA-N
87	DL-Norvaline	117.0789	118.08618	1.46	CCCC(C(=O)O)N	SNDPXSYFESPGGJ-UHFFFAOYSA-N
88	DL-Norvaline_AQC	287.1269	288.13418	15.97	OC(C(CCC)NC(NC1=CC=C2N=CC=CC2=C1)=O)=O	YJSIWLITUTWKFD-UHFFFAOYSA-N
89	DL-Ornithine	132.0898	133.09708	1.03	C(CC(C(=O)O)N)CN	AHLPHDHHMVZTML-UHFFFAOYSA-N
90	DL-Ornithine hydrochloride_AQC	302.1378	303.14508	12.77	O=C(NCCCC(N)C(O)=O)NC1=CC=C2N=CC=CC2=C1	GXBHPVKIPQORGA-UHFFFAOYSA-N
91	Farnesol	222.198365	223.205645	27.71	CC(=CCC/C(=C/CC/C(=C/CO)/C)/C)C	CRDAMVZIKSXKFV-YFVJMOTDSA-N
92	Ferulic Acid	194.057908	195.065188	14.94	COC1=C(C=CC(=C1)/C=C/C(=O)O)O	KSEBMYQBYZTDHS-HWKANZROSA-N
93	gamma-aminobutyric acid	103.06332	104.0706	1.18	C(CC(=O)O)CN	BTCSSZJGUNDROE-UHFFFAOYSA-N
94	gamma-Terpinene	136.1252	137.13248	25.1	CC1=CCC(=CC1)C€C	YKFLAYDHMOASIY-UHFFFAOYSA-N
95	Gibberellic Acid, GA3	346.141638	347.148918	16.5	C[C@@]12[C@H](C=C[C@@]3([C@@H]1[C@@H]([C@]45[C@H]3CC[C@](C4)(C(=C)C5)O)C(=O)O)OC2=O)O	IXORZMNAPKEEDV-OBDJNFEBSA-N
96	Glycine_AQC	245.08	246.08728	3.94	O=C(NCC(O)=O)NC1=CC=C2N=CC=CC2=C1	DXEKOJIDPPOBDC-UHFFFAOYSA-N
97	Guanine	151.049409	152.056689	1.2	C1=NC2=C(N1)C(=O)NC(=N2)N	UYTPUPDQBNUYGX-UHFFFAOYSA-N
98	Homovanillic acid	182.057908	183.065188	18.54	COC1=C(C=CC(=C1)CC(=O)O)O	QRMZSPFSDQBLIX-UHFFFAOYSA-N

99	Indole	117.0578	118.06562	1.6	C1=CC=C2C(=C1)C=CN2	SIKJAQJRHWYJAI-UHFFFAOYSA-N
100	Indole-3-butyric acid	203.09462	204.1019	23.57	C1=CC=C2C(=C1)C(=CN2)CCCC(=O)O	JTEDVYBZBROSJT-UHFFFAOYSA-N
101	Isophorone	138.104465	139.111745	23.24	CC1=CC(=O)CC(C1)€C	HJOVHMDZYOCNQW-UHFFFAOYSA-N
102	Jasmonic acid	210.12559	211.13341	23.35	CCC=CCC1C(CCC1=O)CC(=O)O	ZNJFBWYDHIGLCU-UHFFFAOYSA-N
103	Karrikin 1	150.03169	151.03897	12.4	CC1=C2C=COC=C2OC1=O	JUTMAMXOAOYKHT-UHFFFAOYSA-N
104	Karrikin 2	136.016043	137.023323	5.35	C1=COC=C2C1=CC(=O)O2	OBHRCUWVUFZNMG-UHFFFAOYSA-N
105	Karrikin 3	164.04734	165.05462	19.74	CC1=CC2=C(C(=O)OC2=CO1)C	LVIAGZUAWKSKRF-UHFFFAOYSA-N
106	Karrikin 4	164.04734	165.05462	18.67	CC1=C2C=COC(=C2OC1=O)C	TYHXSNBTCKZPMK-UHFFFAOYSA-N
107	Kinetin	215.0807	216.08798	8	C1=COC(=C1)CNC2=NC=NC3=C2NC=N3	QANMHLXAZMSUEX-UHFFFAOYSA-N
108	Kynurenic acid	189.04259	190.04987	6.21	C1=CC=C2C(=C1)C(=O)C=C(N2)C(=O)O	HCZHHEIFKROPDY-UHFFFAOYSA-N
109	L-2,4-Diaminobutyric acid (DAB)	118.07422	119.0815	1.06	C(CN)C(C(=O)O)N	OGNSCSPNOLGXSM-UHFFFAOYSA-N
110	L-2,4-Diaminobutyric acid_AQC	458.17022	459.1775	11.26	O=C(NCCC(NC(NC1=CC=C2N=CC=CC2=C1)=O)C(O)=O)NC3=CC=C4N=CC=CC4=C3	MLGLHUQXOCUKOX-UHFFFAOYSA-N
111	L-Alanine_AQC	259.09567	260.10295	7.51	O=C(NC€C(O)=O)NC1=CC=C2N=CC=CC2=C1	FVKVIQZOFGVLST-UHFFFAOYSA-N
112	L-Arginine	174.11167	175.11895	1.04	C(C[C@@H](C(=O)O)N)CN=C(N)N	ODKSFYDXXFIFQN-BYPYZUCNSA-N
113	L-Arginine_AQC	344.15967	345.16695	2.98	O=C(NC(CCCNC(N)=N)C(O)=O)NC1=CC=C2N=CC=CC2=C1	HALOHUBWKUUSLS-UHFFFAOYSA-N
114	L-Asparagine	132.0534	133.06068	1.04	C([C@@H](C(=O)O)N)C(=O)N	DCXYFEDJOCDNAF-REOHCLBHSA-N

115	L-Asparagine_AQC	302.1014	303.10868	2.66	OC(C(CC(N)=O)NC(NC1=CC=C2N=CC=CC2=C1)=O)=O	BIHJMSHZULSXMU-UHFFFAOYSA-N
116	L-Aspartic acid	133.0375	134.04478	1.07	C([C@@H](C(=O)O)N)C(=O)O	CKLJMWTZIZZHCS-REOHCLBHSA-N
117	L-Aspartic acid_AQC	303.0855	304.09278	4.53	OC(C(CC(O)=O)NC(NC1=CC=C2N=CC=CC2=C1)=O)=O	PGRPCKNZRCZILV-UHFFFAOYSA-N
118	L-Cysteine	121.0197	122.02698	1.2	C([C@@H](C(=O)O)N)S	XUJNEKJLAYXESH-REOHCLBHSA-N
119	L-Cysteine_AQC	291.0677	292.07498	9.4	O=C(NC(CS)C(O)=O)NC1=CC=C2N=CC=CC2=C1	BEHPXUDEHCQWLR-UHFFFAOYSA-N
120	L-Cystine	240.02384	241.03112	1.06	C([C@@H](C(=O)O)N)SSC[C@@H](C(=O)O)N	LEVWYRKDKASIDU-IMJSIDKUSA-N
121	L-Cystine_AQC	580.11984	581.12712	14.61	O=C(NC(C(O)=O)CSSCC(NC(NC1=CC=C2N=CC=CC2=C1)=O)C(O)=O)NC3=CC=C4N=CC=CC4=C3	JQSPRGNKNHSZDG-UHFFFAOYSA-N
122	L-Glutamic acid	147.0531	148.06038	1.16	C(CC(=O)O)[C@@H](C(=O)O)N	WHUUTDBJXJRKMK-VKHMYHEASA-N
123	L-Glutamic acid_AQC	317.1011	318.10838	5.53	OC(C(CCC(O)=O)NC(NC1=CC=C2N=CC=CC2=C1)=O)=O	JVJPRJRNTNXLRL-UHFFFAOYSA-N
124	L-Glutamine	146.0691	147.07638	1.13	C(CC(=O)N)[C@@H](C(=O)O)N	ZDXPYRJPNDTMRX-VKHMYHEASA-N
125	L-Glutamine_AQC	316.1171	317.12438	3.5	OC(C(CCC(N)=O)NC(NC1=CC=C2N=CC=CC2=C1)=O)=O	XCYFNCHKDHNFDR-UHFFFAOYSA-N
126	L-Histidine	155.0694	156.07668	1.04	C1=C(NC=N1)CC(C(=O)O)N	HNDVDQJCIGZPNO-YFKPBYRVSA-N
127	L-Isoleucine	131.0946	132.10188	1.71	CC[C@H]€[C@@H](C(=O)O)N	AGPKZVBTJJNPAG-WHFBIAKZSA-N
128	L-Isoleucine_AQC	301.1426	302.14988	17.84	OC(C(C€CC)NC(NC1=CC=C2N=CC=CC2=C1)=O)=O	UOFNOPJOTMYQKN-UHFFFAOYSA-N
129	L-Kynurenine	208.08479	209.09207	2.5	C1=CC=C(C(=C1)C(=O)C[C@@H](C(=O)O)N)N	YGPSJZOEDVAXAB-QMMMGPOBSA-N

130	L-Kynurenine_AQC	378.13279	379.14007	17.81	O=C(NC(CC(C1=C(C=CC=C1)N)=O)C(O)=O)NC2=CC=C3N=CC=CC3=C2	ZDWXDIWXIQYSTO-UHFFFAOYSA-N
131	L-Leucine	131.0946	132.10188	1.71	CC€C[C@@H](C(=O)O)N	ROHFNLRQFUQHCH-YFKPBYRVSA-N
132	L-Leucine_AQC	301.1426	302.14988	18.11	OC(C(CC€C)NC(NC1=CC=C2N=CC=CC2=C1)=O)=O	HBEOIKITMLTQSK-UHFFFAOYSA-N
133	L-Lysine	146.1055	147.11278	1.04	C(CCN)C[C@@H](C(=O)O)N	KDXKERNSBIXSRK-YFKPBYRVSA-N
134	L-Lysine_AQC	486.2015	487.20878	14.53	O=C(NC(C(O)=O)CCCCNC(NC1=CC=C2N=CC=CC2=C1)=O)NC3=CC=C4N=CC=CC4=C3	XKOGDNDUCKFMPN-UHFFFAOYSA-N
135	L-Methionine	149.051	150.05828	1.47	CSCC[C@@H](C(=O)O)N	FFEARJCKVFRZRR-BYPYZUCNSA-N
136	L-Methionine_AQC	319.099	320.10628	15.03	O=C(NC(CCSC)C(O)=O)NC1=CC=C2N=CC=CC2=C1	QIRVFWHLEKKETR-UHFFFAOYSA-N
137	L-Phenylalanine	165.0789	166.08618	2.63	C1=CC=C(C=C1)C[C@@H](C(=O)O)N	COLNVLDHVKWLRT-QMMMGPOBSA-N
138	L-Phenylalanine_AQC	335.1269	336.13418	18.58	O=C(NC(CC1=CC=CC=C1)C(O)=O)NC2=CC=C3N=CC=CC3=C2	SVNHSVRVKWMTHW-UHFFFAOYSA-N
139	L-Proline	115.0633	116.07058	1.2	C1C[C@H](NC1)C(=O)O	ONIBWKKTOPOVIA-BYPYZUCNSA-N
140	L-Proline_AQC	285.1113	286.11858	9.71	OC(C1CCCN1C(NC2=CC=C3N=CC=CC3=C2)=O)=O	ICOJMKASMOSKSV-UHFFFAOYSA-N
141	L-Serine	105.0425	106.04978	1.07	C([C@@H](C(=O)O)N)O	MTCFGRXMJLQNBG-REOHCLBHSA-N
142	L-Serine_AQC	275.0905	276.09778	3.44	O=C(NC(CO)C(O)=O)NC1=CC=C2N=CC=CC2=C1	DINGYVQDFDKUIH-UHFFFAOYSA-N
143	L-Theanine	174.10044	175.10772	1.43	CCNC(=O)CC[C@@H](C(=O)O)N	DATAGRPVKZEWHA-YFKPBYRVSA-N
144	L-Threonine	119.0582	120.06548	1.27	C[C@H]([C@@H](C(=O)O)N)O	AYFVYJQAPQTCCC-GBXIJSLDSA-N
145	L-Threonine_AQC	289.1062	290.11348	6.36	O=C(NC(C€O)C(O)=O)NC1=CC=C2N=CC=CC2=C1	AMXKFUHOPRBCRP-UHFFFAOYSA-N

146	L-Tryptophan	204.0898	205.09708	4.37	C1=CC=C2C(=C1)C(=CN2)C[C@@H](C(=O)O)N	QIVBCDIJIAJPQS-VIFPVBQESA-N
147	L-Tryptophan_AQC	374.1378	375.14508	19.08	O=C(NC(CC1=CNC2=C1C=CC=C2)C(O)=O)NC3=CC=C4N=CC=CC4=C3	SSVZSVSMGWVJIC-UHFFFAOYSA-N
148	L-Tyrosine	181.0738	182.08108	1.65	C1=CC(=CC=C1C[C@@H](C(=O)O)N)O	OUYCCCASQSFEME-QMMMGPOBSA-N
149	L-Tyrosine_AQC	351.1218	352.12908	14.85	O=C(NC(CC1=CC=C(O)C=C1)C(O)=O)NC2=CC=C3N=CC=CC3=C2	JLRDKGHKBMKYNU-UHFFFAOYSA-N
150	L-Valine	117.0789	118.08618	1.39	CC€C(C(=O)O)N	KZSNJWFQEVHDMF-UHFFFAOYSA-N
151	L-Valine_AQC	287.1269	288.13418	15.52	OC(C(C€C)NC(NC1=CC=C2N=CC=CC2=C1)=O)=O	FOJGXQXTMBRUTJ-UHFFFAOYSA-N
152	Lidocaine hydrochloride monohydrate	234.173213	235.180493	11.3	CCN(CC(NC1=C€C=CC=C1C)=O)CC	NNJVILVZKWQKPM-UHFFFAOYSA-N
153	Linoleic acid	280.24023	281.24751	28.37	CCCCC/C=C\C/C=C\CCCCCCCC(=O)O	OYHQOLUKZRVURQ-HZJYTTRNSA-N
154	Linolenic acid	278.22458	279.23186	28.03	CC/C=C\C/C=C\C/C=C\CCCCCCCC(=O)O	DTOSIQBPPRVQHS-PDBXOOCHSA-N
155	Maleic Acid	116.01095	117.01823	1.3	C(=C\C(=O)O)\C(=O)O	VZCYOOQTPOCHFL-UPHRSURJSA-N
156	Melamine	126.06539	127.07267	1.14	C1(=NC(=NC(=N1)N)N)N	JDSHMPZPIAZGSV-UHFFFAOYSA-N
157	Melatonin	232.121177	233.128457	18.27	CC(=O)NCCC1=CNC2=C1C=C(C=C2)OC	DRLFMBDRBRZALE-UHFFFAOYSA-N
158	Menthol	156.151415	157.158695	25.85	CC1CCC(C(C1)O)C€C	NOOLISFMXDJSKH-UHFFFAOYSA-N
159	Meta-Topolin	241.09635	242.10363	8.04	C1=CC(=CC(=C1)O)CNC2=NC=NC3=C2NC=N3	BUDWTFCZGZYQHZ-UHFFFAOYSA-N
160	Methyl salicylate	152.04734	153.05462	24.46	COC(=O)C1=CC=CC=C1O	OSWPMRLSEDHDFF-UHFFFAOYSA-N

161	Mitragynine	398.2205	399.22832	22.85	CC[C@@H]1CN2CCC3=C([C@@H]2C[C@@H]1/C(=C\OC)/C(=O)OC)NC4=C3C(=CC=C4)OC	LELBFTMXCIIKKX-QVRQZEMUSA-N
162	Mitraphylline	368.173607	369.180887	18.6	C[C@H]1[C@H]2CN3CC[C@]4([C@@H]3C[C@@H]2C(=CO1)C(=O)OC)C5=CC=CC=C5NC4=O	JMIAZDVHNCCPDM-DAFCLMLCSA-N
163	Myristicin	192.0786	193.08588	25.87	COC1=CC(=CC2=C1OCO2)CC=C	BNWJOHGLIBDBOB-UHFFFAOYSA-N
164	n-(2)-Aminoethylglycine _AQC	458.17022	459.1775	10.86	O=C(N(CC(O)=O)CCNC(NC1=CC=C2N=CC=CC2=C1)=O)NC3=CC=C4N=CC=CC4=C3	DABZHFFHOVWQGQ-UHFFFAOYSA-N
165	n-(2)-Aminoethylglycine (AEG)	118.07422	119.0815	1.11	C(CNCC(=O)O)N	PIINGYXNCHTJTF-UHFFFAOYSA-N
166	N-Acetyl-5-Hydroxytryptamine	218.105527	219.112807	7.85	CC(=O)NCCC1=CNC2=C1C=C(C=C2)O	MVAWJSIDNICKHF-UHFFFAOYSA-N
167	N-Bromoacetyltryptamine	280.02113	281.02841	23.35	C1=CC=C2C(=C1)C(=CN2)CCNC(=O)CBr	TZRUZONGYSVOGE-UHFFFAOYSA-N
168	N-Feruloyl Serotonin	352.142307	353.149587	22.25	COC1=C(C=CC(=C1)/C=C/C(=O)NCCC2=CNC3=C2C=C(C=C3)O)O	WGHKJYWENWLOMY-XVNBXDOJSA-N
169	Nerolidol	222.198365	223.205645	27.71	CC(=CCC/C(=C/CCC€(C=C)O)/C)C	FQTLCLSUCSAZDY-SDNWHVSQSA-N
170	Nerylacetone	194.16706	195.17434	27.25	CC(=CCC/C(=C\CCC(=O)C)/C)C	HNZUNIKWNYHEJJ-XFXZXTDPSA-N
171	Ochratoxin A (from Petromyces albertensis)	403.08226	404.09008	25.42	C[C@@H]1CC2=C(C=C(C(=C2C(=O)O1)O)C(=O)N[C@@H](CC3=CC=CC=C3)C(=O)O)Cl	RWQKHEORZBHNRI-BMIGLBTASA-N
172	p-Coumaric acid	164.04734	165.05462	11.1	C1=CC(=CC=C1/C=C/C(=O)O)O	NGSWKAQJJWESNS-ZZXKWVIFSA-N
173	Peonidin-3-Galactoside Standard	463.12404	464.13132	28.04	COC1=C(C=CC(=C1)C2=[O+]C3=CC(=CC(=C3C=C2O[C@H]4[	ZZWPMFROUHHAKY-VRRLNDPFSA-O

					C@@H]([C@H]([C@H]([C@H](O4)CO)O)O)O)O)O	
174	Picloram	239.92601	240.93329	9.04	C1(=C(C(=NC(=C1Cl)Cl)C(=O)O)Cl)N	NQQVFXUMIDALNH-UHFFFAOYSA-N
175	Piperonal	150.03169	151.03951	11.23	C1OC2=C(O1)C=C(C=C2)C=O	SATCULPHIDQDRE-UHFFFAOYSA-N
176	Progesterone	314.22458	315.23186	26.46	CC(=O)[C@H]1CC[C@@H]2[C@@]1(CC[C@H]3[C@H]2CCC4=CC(=O)CC[C@]34C)C	RJKFOVLPORLFTN-LEKSSAKUSA-N
177	Quercetin	302.04265	303.04993	22.85	C1=CC(=C(C=C1C2=C(C(=O)C3=C(C=C(C=C3O2)O)O)O)O)O	REFJWTPEDVJJIY-UHFFFAOYSA-N
178	Quinine anhydrous	324.18378	325.19106	14.29	COC1=CC2=C(C=CN=C2C=C1)[C@H]([C@@H]3C[C@@H]4CCN3C[C@@H]4C=C)O	LOUPRKONTZGTKE-WZBLMQSHSA-N
179	Rutin Hydrate	610.15338	611.16066	17.38	C[C@H]1[C@H](O)[C@@H](O)[C@@H](O)[C@H](OC[C@@H]2[C@@H](O)[C@H](O)[C@@H](O)[C@H](OC3=C(C4=CC(O)=C(O)C=C4)OC5=CC(O)=CC(O)=C5C3=O)O2)O1	IKGXIBQEEMLURG-NVPNHPEKSA-N
180	S-adenosyl-L-methionine	399.14506	399.14506	1.05	C[S+](CC[C@@H](C(=O)[O-])N)C[C@@H]1[C@H]([C@H]([C@@H](O1)N2C=NC3=C(N=CN=C32)N)O)O	MEFKEPWMEQBLKI-AIRLBKTGSA-N
181	S-adenosyl-L-methionine_AQC	284.09289	285.10017	1.7	C[S+](CC1C(O)C(O)C(N2C=NC3=C(N)N=CN=C32)O1)CCC(NC(NC4=CC=C5N=CC=CC5=C4)=O)C(O)=O	JLZMCLIMGHFIBV-UHFFFAOYSA-O
182	Scutellarin	462.07982	463.0871	17.5	C1=CC(=CC=C1C2=CC(=O)C3=C(C(=C(C=C3O2)O[C@H]4[C@@H]([C@H]([C@@H]([C@H](O4)C(=O)O)O)O)O)O)O)O	DJSISFGPUUYILV-ZFORQUDYSA-N
183	Serotonin	176.09496	177.10224	1.4	C1=CC2=C(C=C1O)C(=CN2)CCN	QZAYGJVTTNCVMB-UHFFFAOYSA-N

184	Strychnine	334.1681279	335.1754079	12.29	C1CN2CC3=CCO[C@H]4CC(=O)N5[C@H]6[C@H]4[C@H]3C[C@H]2[C@@]61C7=CC=CC=C75	QMGVPVSNSZLJIA-FVWCLLPLSA-N
185	Sulfaguanidine	214.05244	215.05972	1.43	C1=CC(=CC=C1N)S(=O)(=O)N=C(N)N	BRBKOPJOKNSWSG-UHFFFAOYSA-N
186	Tandosporine	383.232125	384.239405	19.34	C1C[C@H]2C[C@@H]1[C@H]3[C@@H]2C(=O)N(C3=O)CCCCN4CCN(CC4)C5=NC=CC=N5	CEIJFEGBUDEYSX-FZDBZEDMSA-N
187	Tasimelteon	245.1415788	246.1488588	24.04	CCC(=O)NC[C@@H]1C[C@H]1C2=C3CCOC3=CC=C2	PTOIAAWZLUQTIO-GXFFZTMASA-N
188	tetrahydrocannabinolic acid	358.2144	359.22222	28.47	CCCCCC1=CC2=C([C@@H]3C=C(CC[C@H]3C(O2)€C)C)C(=C1C(=O)O)O	UCONUSSAWGCZMV-HZPDHXFCSA-N
189	Tetrahydrocannabivarin (THCV)	286.19328	287.2011	27.68	CCCC1=CC(=C2[C@@H]3C=C(CC[C@H]3C(OC2=C1)€C)C)O	ZROLHBHDLIHEMS-HUUCEWRRSA-N
190	Thidiazuron	220.04188	221.04916	22.74	C1=CC=C(C=C1)NC(=O)NC2=CN=NS2	HFCYZXMHUIHAQI-UHFFFAOYSA-N
191	Tryptamine	160.10005	161.10733	5.2	C1=CC=C2C(=C1)C(=CN2)CCN	APJYDQYYACXCRM-UHFFFAOYSA-N
192	Uracil	112.027277	113.034557	1.2	C1=CNC(=O)NC1=O	ISAKRJDGNUQOIC-UHFFFAOYSA-N
193	Val-Tyr-Val	379.21072	380.218	11.85	CC€C(C(=O)NC(CC1=CC=C(C=C1)O)C(=O)NC(C€C)C(=O)O)N	ZNGPROMGGGFOAA-UHFFFAOYSA-N
194	Vanillin	152.04734	153.05462	10.73	COC1=C(C=CC(=C1)C=O)O	MWOOGOJBHIARFG-UHFFFAOYSA-N
195	Wogonin	284.06847	285.07575	24.95	COC1=C(C=C(C2=C1OC(=CC2=O)C3=CC=CC=C3)O)O	XLTFNNCXVBYBSX-UHFFFAOYSA-N
196	Zeatin (Trans)	219.11201	220.11929	4.1	C/C(=C\CNC1=NC=NC2=C1NC=N2)/CO	UZKQTCBAMSWPJD-FARCUNLSSA-N
197	β-methylamino-L-alanine (BMAA)	118.07422	119.0815	1.12	CNC[C@@H](C(=O)O)N	UJVHVMNGOZXSOZ-VKHMYHEASA-N

198	β-methylamino-L-alanine_AQC	458.17022	459.1775	11.04	O=C(N€CC(NC(NC1=CC=C2N=CC=CC2=C1)=O)C(O)=O)NC3=CC=C4N=CC=CC4=C3	DVJBZWLSAFWOJA-UHFFFAOYSA-N

Table S6: Putatively annotated metabolites in *Arabidopsis thaliana* and *Cannabis sativa* leaf tissues using aminoacidDB.

Sno.	**Metabolite Name**	**MSI Level**	**Adduct**	**Monoisotopic Mass**	**Mass error (ppm)**	**RT (min)**	**Predicted RT (min)**	**ΔRT (min)**	**RT match (%)**
1	Glycine	1	[M+H]+	75.032	2.25	3.98	4.1	0.12	97.0
2	beta-Alanine	1	[M+H]+	89.04767	2.22	6.07	6.16	0.09	98.5
3	L-Alanine	1	[M+H]+	89.04767	2.32	7.94	7.94	0.00	100.0
4	1-aminocyclopropane-1-carboxylic acid	1	[M+H]+	101.04767	2.54	8.46	8.47	0.01	99.9
5	Gamma-aminobutyric acid	1	[M+H]+	103.0633	2.37	8.53	8.48	0.04	99.5
6	L-Serine	1	[M+H]+	105.0425	1.75	3.41	3.6	0.19	94.4
7	L-Proline	1	[M+H]+	115.0633	2.11	10.30	10.28	0.02	99.8
8	L-Valine	1	[M+H]+	117.0789	1.99	16.08	16.03	0.05	99.7
9	beta-N-methylamino-L-alanine	1	[M+H]+	118.07422	1.17	11.90	11.89	0.01	99.9
10	L-2,4-diaminobutyric acid	1	[M+H]+	118.07422	2.04	12.13	12.17	0.04	99.7
11	n-(2)-aminoethylglycine	1	[M+H]+	118.07422	2.20	11.69	11.74	0.05	99.6
12	L-Threonine	1	[M+H]+	119.0582	2.12	6.68	6.74	0.06	99.1
13	L-Cysteine	1	[M+H]+	121.0197	1.99	9.90	9.94	0.04	99.6
14	L-Isoleucine	1	[M+H]+	131.0946	1.81	18.75	18.7	0.05	99.7
15	L-Leucine	1	[M+H]+	131.0946	1.87	19.08	19.24	0.16	99.2
16	6-aminohexanoic acid	1	[M+H]+	131.09462	1.40	15.50	15.51	0.01	99.9
17	L-Asparagine	1	[M+H]+	132.0534	1.92	2.61	2.87	0.26	90.0
18	DL-Ornithine	1	[M+H]+	132.0898	2.28	13.75	13.76	0.01	99.9
19	DL-Norleucine	1	[M+H]+	132.10242	1.18	19.62	19.64	0.02	99.9
20	L-Aspartic acid	1	[M+H]+	133.0375	2.09	4.66	4.79	0.13	97.2

21	L-Glutamine	1	[M+H]+	146.0691	1.80	3.59	3.72	0.13	96.4
22	L-Lysine	1	[M+H]+	146.1055	2.12	15.29	15.23	0.06	99.6
23	L-Glutamic acid	1	[M+H]+	147.0531	1.69	5.76	5.82	0.06	99.0
24	L-Methionine	1	[M+H]+	149.051	1.58	15.61	15.54	0.07	99.6
25	L-Phenylalanine	1	[M+H]+	165.0789	1.77	19.79	19.75	0.04	99.8
26	L-Arginine	1	[M+H]+	174.11167	1.96	2.97	3.19	0.22	92.6
27	DL-Citrulline	1	[M+H]+	175.09569	1.77	5.65	5.71	0.06	98.9
28	L-Tyrosine	1	[M+H]+	181.0738	1.51	15.54	15.52	0.02	99.9
29	DL-beta-3,4-dihydroxyphenylalanine	1	[M+H]+	197.0688	1.47	13.45	13.51	0.06	99.6
30	4-Chloro-L-phenylalanine	1	[M+H]+	199.04	1.21	23.01	23	0.01	100.0
31	L-Tryptophan	1	[M+H]+	204.0898	1.31	20.69	20.68	0.01	100.0
32	L-Kynurenine	1	[M+H]+	208.08479	1.65	18.85	18.84	0.01	99.9
33	5-hydroxy-L-tryptophan	1	[M+H]+	220.08479	-1.56	18.73	18.82	0.09	99.5
34	L-Cystine	1	[M+H]+	240.02384	1.98	15.33	15.33	0.00	100.0
35	Alpha-Aminobutyric Acid	2	[M+H]+	103.0633	1.16	11.73	10.99	0.74	93.7
36	N-Methylalanine	2	[M+H]+	103.0633	0.10	12.30	12.98	0.68	94.4
37	Pyroglutamic acid	2	[M+H]+	129.04259	0.23	11.04	9.72	1.32	88.1
38	4-Oxo-L-proline	2	[M+H]+	129.04259	-0.39	7.24	8.48	1.24	82.9
39	O-Acetylserine	2	[M+H]+	147.05315	0.27	9.58	7.95	1.63	83.0
40	4-hydroxyornithine	2	[2M+H]+	148.08479	-0.23	9.79	11.71	1.92	80.4
41	N(6)-methyllysine	2	[M+H]+	160.121	1.11	16.26	12.62	3.64	77.6
42	N-acetyl-L-ornithine	2	[M+H]+	174.10044	0.57	8.66	8.58	0.08	99.1
43	Homo-L-arginine	2	[M+H]+	188.12732	-3.83	4.38	4.22	0.16	96.3
44	Rhizobitoxine	2	[M+H]+	190.09498	1.95	4.99	5.55	0.56	88.8
45	Valylvaline	2	[M+H]+	216.14719	0.94	18.22	16	2.22	88.4
46	Leucyl-Histidine	2	[M+H]+	268.15366	0.45	15.53	12.02	3.51	77.4
47	Cyclo(-his-phe)	2	[M+H]+	284.12729	-0.12	16.01	18.22	2.21	86.2
48	Glycylprolylhydroxyproline	2	[M+H]+	285.13189	-2.03	9.46	7.76	1.70	82.0

49	N-(1-Deoxy-1-fructosyl)tyrosine	2	[M+H]+	343.12692	0.61	12.57	9.64	2.93	76.7
50	Terazosin	2	[M+H]+	387.19033	0.84	14.17	14.67	0.50	96.4
51	Oxiglutathione	2	[M+H]+	612.1512	3.20	7.68	6.74	0.94	87.7
52	2-aminoacrylic acid	3	[M+H]+	87.0318	2.53	11.10	7.49	3.61	67.5
53	DL-homoserine	3	[M+H]+	119.05822	0.19	4.32	5.88	1.56	63.9
54	Iminodiacetic acid	3	[M+H]+	133.03752	-0.09	4.08	6.32	2.24	45.0
55	L-2-Amino-5-hydroxypentanoic acid; (2s,4s)-4-Hydroxynorvaline; L-Pentahomoserine; 3-hydroxynorvaline	3	[M+H]+	133.07389	-0.08	7.26	Multiple isomers match	NA	NA
56	L-2-Amino-5-hydroxypentanoic acid; (2s,4s)-4-Hydroxynorvaline; L-Pentahomoserine; 3-hydroxynorvaline	3	[M+H]+	133.07389	0.00	6.61	Multiple isomers match	NA	NA
57	5-Methylnorleucine, N-Methyl-L-isoleucine, (2S)-2-amino-4-methylhexanoic acid	3	[M+H]+	145.11018	0.68	19.27	Multiple isomers match	NA	NA
58	Beta-Ureidoisobutyric acid	3	[M+H]+	146.06844	4.81	17.36	11.68	5.68	67.3
59	2-Oxo-4-hydroxy-5-aminovalerate, Methyl 2-acetamidooxyacetate, 5-Hydroxy-4-oxonorvaline, 3-Aminopentanedioic acid, L-4-Hydroxyglutamate semialdehyde, N-Acetylserine, threo-3-methyl-L-aspartic acid, 3-(Carboxymethylamino)propanoic acid, N-Methyl-D-aspartic acid	3	[M+H]+	147.05313	-0.18	5.86	Multiple isomers match	NA	NA
60	(1R,6S)-6-amino-5-oxocyclohex-2-ene-1-carboxylic acid, (5R,6R)-6-amino-5-hydroxycyclohexa-1,3-diene-1-carboxylic acid	3	[M+H]+	155.05798	1.70	13.96	Multiple isomers match	NA	NA
61	gamma-Acetyldiaminobutyric acid; 4-Acetamido-2-aminobutanoic acid	3	[M+H]+	160.08479	-2.00	7.87	Multiple isomers match	NA	NA

62	3-(N-methylamino)glutaric acid, 4-Amino-5-methoxy-5-oxopentanoic acid; (±)-2,2'-Iminobispropanoic acid	3	[M+H]+	161.06869	0.68	12.84	Multiple isomers match	NA	NA
63	4-methylglutamate, (2S,4S,5S)-4,5-dihydroxypiperidine-2-carboxylic acid, O-acetyl-L-homoserine, 4-Amino-5-methoxy-5-oxopentanoic acid, Glutamic acid gamma-methyl ester, 2-Aminoadipic Acid, (±)-2,2'-Iminobispropanoic acid	3	[M+H]+	161.0688	0.81	9.74	Multiple isomers match	NA	NA
64	Serylglycine, L-4-Hydroxyglutamine	3	[M+H]+	162.06405	0.04	4.64	Multiple isomers match	NA	NA
65	4-Hydroxyphenylglycine	3	[M+H]+	167.05774	3.01	11.20	14.24	3.04	72.8
66	2,5-Dihydrophenylalanine, l-Dihydrophenylalanine, (2S)-2-amino-3-cyclohexa-2,4-dien-1-ylpropanoic acid, Norpandamarilactonine A	3	[M+H]+	167.09459	0.23	17.39	Multiple isomers match	NA	NA
67	Cyclo(deltaAla-L-Val)	3	[M+H]+	168.08965	1.35	8.78	14.2	5.43	38.2
68	N-Carboxyethyl-g-aminobutyric acid	3	[M+H]+	175.0844	0.00	16.23	11.85	4.38	73.0
69	Threonylglycine, Alanylserine	3	[M+H]+	176.07972	-0.07	5.38	Multiple isomers match	NA	NA
70	(3S,4R)-3-Hydroxy-4-methyl-L-glutamic acid, (2s,4r)-4-Hydroxy-4-methyl-glutamic acid, (2S,3R,4R,5S)-3,4,5-Trihydroxypiperidine-2-carboxylic acid	3	[M+H]+	177.06349	1.31	4.62	Multiple isomers match	NA	NA
71	Alanylvaline, Valylalanine, Isoleucyl-Glycine, Leucyl-Glycine, H-Gln(isopropyl)-OH, 2-[(2-Amino-3-methylbutanoyl)amino]propanoic acid	3	[M+H]+	188.11605	0.23	16.87	Multiple isomers match	NA	NA

72	N(alpha)-acetyl-L-lysine methyl ester, Alanylisoleucine, Isoleucyl-Alanine, Leucylalanine, Alanylleucine	3	[M+H]+	202.13174	-0.74	17.42	Multiple isomers match	NA	NA
73	Asymmetric dimethylarginine, Symmetric Dimethylarginine	3	[M+H]+	202.14301	-0.20	5.57	Multiple isomers match	NA	NA
74	4-carboxyphenylalanine, 3-Formyl-L-tyrosine, (S)-3-(2-Amino-2-carboxyethyl)benzoic acid, N-Benzyloxycarbonylglycine, 3-Carbamoyl-2-phenylpropionic acid, Hydroxyphenylacetylglycine	3	[M+H]+	209.06881	-0.97	17.02	Multiple isomers match	NA	NA
75	Valylproline, Dethiobiotin, Monascustin	3	[M+H]+	214.13187	-0.60	18.78	Multiple isomers match	NA	NA
76	gamma-Glutamylalanine, Serylhydroxyproline, Hydroxyprolyl-Serine, Glutamylalanine, Alanylglutamic acid	3	[M+H]+	218.08961	3.03	8.42	Multiple isomers match	NA	NA
77	gamma-Glutamylalanine, Serylhydroxyproline, Hydroxyprolyl-Serine, Glutamylalanine, Alanylglutamic acid	3	[M+H]+	218.08988	1.80	7.26	Multiple isomers match	NA	NA
78	Threonylvaline, Valylthreonine, Meprobamate, Pantothenamide	3	[M+H]+	218.12666	-0.01	16.81	Multiple isomers match	NA	NA
79	DL-3-Carboxytyrosine, 4-Amino-4-deoxychorismic acid, Madurastatin B2, Vanilloylglycine	3	[M+H]+	225.06376	-0.17	16.74	Multiple isomers match	NA	NA
80	Isoleucylproline, Leucylproline, 5S-hydroxynorvaline-S-Ile, Pro- leu	3	[M+H]+	228.14766	-1.17	22.33	Multiple isomers match	NA	NA
81	Spermic acid 2	3	[M+H]+	232.14241	-0.47	17.26	12.74	4.52	73.8
82	gamma-glutamyl-beta-cyanoalanine	3	[M+H]+	243.08548	0.16	5.73	8.82	3.09	46.1

83	Lysylvaline	3	[M+H]+	245.17409	-0.61	7.38	9.33	1.96	73.5
84	gamma-Glutamylvaline, Leucyl-Aspartate, L-beta-aspartyl-L-leucine, Glutamylvaline, Valylglutamic acid, Aspartyl-Leucine, Aspartyl-Isoleucine, Isoleucyl-Aspartate	3	[M+H]+	246.12059	3.99	14.90	Multiple isomers match	NA	NA
85	Alanyltyrosine, Tyr-Ala, Ser-Phe, Phe-Ser (2S,4S)-2-amino-5-benzamido-4-hydroxypentanoic acid, (2S)-2-amino-5-[(4-hydroxyphenyl)methylamino]-5-oxopentanoic acid	3	[M+H]+	252.11069	-1.26	13.93	Multiple isomers match	NA	NA
86	Alanyltyrosine, Tyr-Ala, Ser-Phe, Phe-Ser (2S,4S)-2-amino-5-benzamido-4-hydroxypentanoic acid, (2S)-2-amino-5-[(4-hydroxyphenyl)methylamino]-5-oxopentanoic acid	3	[M+H]+	252.11083	-0.70	12.71	Multiple isomers match	NA	NA
87	N5-(4-methoxybenzyl)glutamine, Threonylphenylalanine, Phenylalanylthreonine, gamma-Glutamyltyramine	3	[M+H]+	266.12665	1.05	17.36	Multiple isomers match	NA	NA
88	Ser-Tyr, Tyrosyl-Serine	3	[M+H]+	268.10561	-1.16	11.76	Multiple isomers match	NA	NA
89	Prolyl-Arginine, Arginylproline	3	[M+H]+	271.16456	-0.44	10.04	Multiple isomers match	NA	NA
90	Arginyl-Valine, Val-Arg	3	[M+H]+	273.18018	-0.33	10.55	Multiple isomers match	NA	NA
91	Norophthalmic acid, Glutaminylglutamic acid, Glutamyl-Gamma-glutamate, Glutamylglutamine, N-gamma-Glutamylglutamine	3	[M+H]+	275.11172	0.06	4.41	Multiple isomers match	NA	NA

92	Arg-Ile, L-isoleucyl-L-arginine, Leucylarginine, Arginylleucine	3	[M+H]+	287.19587	0.49	14.45	Multiple isomers match	NA	NA
93	Arg-Ile, L-isoleucyl-L-arginine, Leucylarginine, Arginylleucine	3	[M+H]+	287.19589	0.56	14.96	Multiple isomers match	NA	NA
94	Arg-Ile, L-isoleucyl-L-arginine, Leucylarginine, Arginylleucine	3	[M+H]+	287.19593	0.70	15.50	Multiple isomers match	NA	NA
95	Argininosuccinate, N2-(3-Hydroxysuccinoyl)arginine, N2-(3-Carboxy-2-hydroxy-1-oxopropyl)arginine	3	[M+H]+	290.12273	-0.33	4.42	Multiple isomers match	NA	NA
96	Histidylphenylalanine, Phenylalanylhistidine	3	[M+H]+	302.1381	0.70	15.14	Multiple isomers match	NA	NA
97	gamma-Glutamylarginine, Glutamylarginine, Arginylglutamic acid	3	[M+H]+	303.15414	0.42	3.54	Multiple isomers match	NA	NA
98	Asparaginyl-Tryptophan, Tryptophyl-Asparagine	3	[M+H]+	318.1329	0.30	9.79	Multiple isomers match	NA	NA
99	Lysyltryptophan, Tryptophyl-Lysine	3	[M+H]+	332.1848	-0.12	15.79	Multiple isomers match	NA	NA
100	2-Amino-5-((1-((carboxymethyl)amino)-3-((1-hydroxyhexan-3-yl)thio)-1-oxopropan-2-yl)amino)-5-oxopentanoic acid	3	[M+H]+	407.17248	0.35	17.30	6.62	10.68	38.3
101	Halocyamine B	3	[M+H]+	653.15764	4.56	14.74	8.56	6.18	58.1
102	3-methyl-2-aziridinecarboxylic acid; C4H7NO2	4	[M+H]+	101.04735	3.17	4.28	Multiple isomers match	NA	NA

103	€-3-(methylamino)-2-propenoic acid; C4H7NO2	4	[M+H]+	101.04764	0.30	1.96	Multiple isomers match	NA	NA
104	2,3-dihydroisoxazole-5-carboxylic acid; C4H5NO3	4	[M+H]+	115.02661	2.89	4.08	Multiple isomers match	NA	NA
105	2,3-dihydroisoxazole-3-carboxylic acid; C4H5NO3	4	[M+H]+	115.02663	2.72	4.50	Multiple isomers match	NA	NA
106	2-pyrrolidinecarboxylic acid; C5H9NO2	4	[M+H]+	115.06325	0.68	2.12	Multiple isomers match	NA	NA
107	3-methyl-2,3-dihydro-1H-pyrrole-5-carboxylic acid; C6H9NO2	4	[M+H]+	127.06335	-0.24	7.77	Multiple isomers match	NA	NA
108	2,3-dihydro-1H-pyrazine-4-carboxylic acid; C5H8N2O2	4	[M+H]+	128.05854	0.29	4.51	Multiple isomers match	NA	NA
109	3-aminopropyl(methyl)carbamic acid; C5H12N2O2	4	[M+H]+	132.08993	-0.40	2.75	Multiple isomers match	NA	NA
110	2,6-dimethyl-1,4-dihydropyridine-3-carboxylic acid; C8H11NO2	4	[M+H]+	153.07893	0.32	12.11	Multiple isomers match	NA	NA
111	3-(2-aminoethyl)-4-imidazolecarboxylic acid; C6H9N3O2	4	[M+H]+	155.06944	0.19	2.63	Multiple isomers match	NA	NA
112	2-(4,5-dihydro-1H-imidazol-2-yl)ethylcarbamic acid; C6H11N3O2	4	[M+H]+	157.08483	1.89	6.38	Multiple isomers match	NA	NA
113	2-[amino(propan-2-yl)amino]butanoic acid; C7H16N2O2	4	[M+H]+	160.12098	1.24	3.05	Multiple isomers match	NA	NA

114	4-amino-5-mercapto-2-methylpentanoic acid; C6H13NO2S	4	[M+H]+	163.06649	1.29	2.05	Multiple isomers match	NA	NA
115	(2S)-2-amino-3-ethylsulfinylpropanoic acid; C5H11NO3S	4	[M+H]+	165.04589	0.45	4.70	Multiple isomers match	NA	NA
116	2-(3-furanylmethylamino)propanoic acid; C8H11NO3	4	[M+H]+	169.07377	0.73	9.81	Multiple isomers match	NA	NA
117	2-(5-ethyl-1H-1,2,4-triazol-3-yl)propanoic acid; C7H11N3O2	4	[M+H]+	169.0846	3.08	2.83	Multiple isomers match	NA	NA
118	4-(3-methylbutan-2-ylamino)butanoic acid; C9H19NO2	4	[M+H]+	173.14145	0.74	22.47	Multiple isomers match	NA	NA
119	5-amino-5-(diaminomethylideneamino)pentanoic acid; C6H14N4O2	4	[M+H]+	174.1119	-1.29	5.35	Multiple isomers match	NA	NA
120	7-methyl-1H-pyrrolo[3,2-c]pyridine-3-carboxylic acid; C9H8N2O2	4	[M+H]+	176.05883	-1.43	4.62	Multiple isomers match	NA	NA
121	thiazolidine-2,5-dicarboxylic acid; C5H7NO4S	4	[M+H]+	177.0094	1.01	8.09	Multiple isomers match	NA	NA
122	2-amino-6-(dimethylamino)-2-methylhexanoic acid; C9H20N2O2	4	[M+H]+	188.15251	-0.17	3.14	Multiple isomers match	NA	NA
123	4-(3-amino-3-oxopropyl)-2-morpholinecarboxylic acid; C8H14N2O4	4	[M+H]+	202.09532	0.18	2.13	Multiple isomers match	NA	NA
124	3-[[[(1-methyl-4-piperidinyl)methylamino]-oxomethyl]amino]propanoic acid; C11H21N3O3	4	[M+H]+	243.158	1.20	4.26	Multiple isomers match	NA	NA

125	L-Ala-3-[[€-4-Amino-1,4-dioxo-2-butenyl]amino]-L-Ala-OH; C10H16N4O5	4	[M+H]+	272.11197	0.37	3.41	6.35	2.94	13.8
126	Saccharopine; C11H20N2O6	4	[M+H]+	276.13216	-0.11	4.49	8.66	4.17	7.0
127	(2R)-2-amino-3-[[4,6-bis(dimethylamino)-1,3,5-triazin-2-yl]thio]propanoic acid; C10H18N6O2S	4	[M+H]+	286.12216	-3.37	13.04	Multiple isomers match	NA	NA
128	N-{5-[(3,3-diaminoprop-2-en-1-ylidene)carbamoyl]-1h-pyrrol-3-yl}-5-iminopyrrolidine-2-carboximidic acid; C13H17N7O2	4	[M+H]+	303.14289	0.46	3.53	Multiple isomers match	NA	NA
129	2-acetamido-3-[[4-amino-6-(dimethylamino)-1,3,5-triazin-2-yl]methylthio]propanoic acid; C11H18N6O3S	4	[M+H]+	314.11696	-2.71	21.00	Multiple isomers match	NA	NA

Table S7: *Arabidopsis thaliana* studies selected from The Metabolomics Workbench used to validate the aminoacidDB workflow.

Project DOI	Study Title	Source	Reverse phase column	Chromato-graphy Instrument	MS instrument	Solvents	LC Gradient
https://doi.org/10.21228/M8XM40	Data resource for fully 13C labelled and non-labelled plant tissues (part-I)	Plant-shoot	Waters Acquity BEH C18 (100 x 2.1mm,1.7um)	Waters Acquity	QToF: Waters Synpat G2	100% water; 0.1% formic acid; 100% acetonitrile; 0.1% formic acid	99.5% solvent A/0.5% solvent B at 0 min, 99.5%A/0.5%B at 0.1 min, 20%A/80%B at 10 min, 0.5%A/99.5%B at 10.1 min, 0.5%A/99.5%B at 12.0 min, 99.5%A/0.5%B at 12.1 min and 99.5%A/0.5%B at 15.0 min
https://doi.org/10.21228/M8XM40	Arabidopsis thaliana 25 accessions	Plant-shoot and root	Waters Acquity BEH C18 (100 x 2.1mm,1.7um)	Waters Acquity	QToF: Waters Synpat G2	100% water; 0.1% formic acid; 100% acetonitrile; 0.1% formic acid	99.5% solvent A/0.5% solvent B at 0 min, 99.5%A/0.5%B at 0.1 min, 20%A/80%B at 10 min, 0.5%A/99.5%B at 10.1 min, 0.5%A/99.5%B at 12.0 min, 99.5%A/0.5%B at 12.1 min and 99.5%A/0.5%B at 15.0 min
https://doi.org/10.21228/M8570V	Pollen metabolomics using Arabidopsis thaliana: Comparison of pollen at mature, hydration and germination stage	Plant-pollen	Higgins Analytica Targa C18 (150 x 0.3mm,3.5um)	Eksigent Ekspert microLC 200	Orbitrap: Thermo Q Exactive Orbitrap	100% water; 0.1% formic acid; 100% acetonitrile; 0.1% formic acid	Not provided

Notes on Validation and Use of the Database

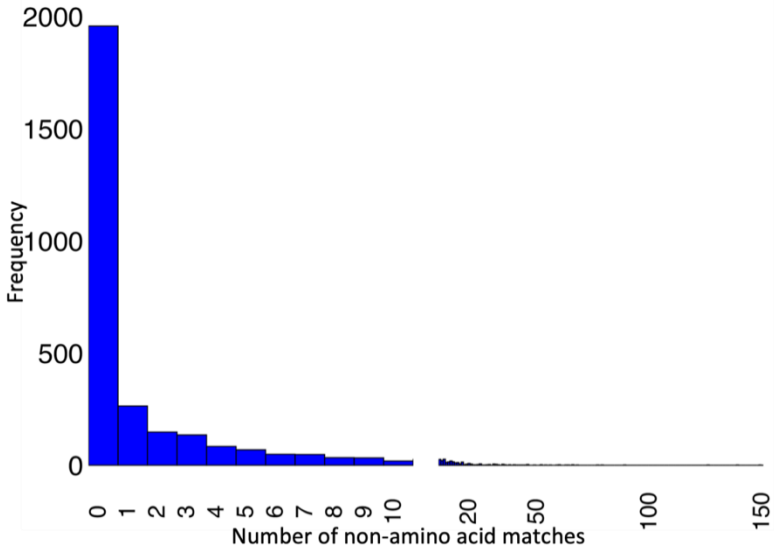
We performed an experiment matching unique masses in our dataset to determine whether non-amino acid matches are a significant issue. The AminoacidDB Lotus/HMDB database contains 3109 unique monoisotopic masses pertaining to 6485 structures. The experiment could not be performed for PubChem dataset in AminoacidDB due to the size of the database and computational constraints.

The unique masses in Lotus/HMDB datasets were searched against the entire Lotus (216302 compounds) and HMDB databases (217920 compounds) to find non-amino acid matches. The detailed results of the experiment are provided in the Table 1. Overall, there were 11103 non-amino acid matches for the 3109 unique monoisotopic masses. About 48% of the compounds do not result in any non-amino acid matches while 90% of the compounds have 10 or less non-amino acid matches (Figure 1) highlighting the limited error rate in AminoacidDB database. Some monoisotopic masses tend to produce a larger diversity of non-amino acid structures. For example, $C_{21}H_{24}N_2O_3$ and $C_{19}H_{21}NO_4$ matched to 154 and 144 non-amino acid compounds respectively, the highest in the dataset. This agrees with Reviewer's point of matching to non-amino acid compounds when using monoisotopic mass/molecular formula as a sole criterion but does not negate our original recommendations for the use of the databases. For revisions addressing this specific issue, please see lines 450-461.

This is the nature of all databases and is not unique to aminoacidDB. Orthogonal techniques such as Retention time (RT)[1–4], MS/MS[5–7], ion-mobility (IM)[8,9], two-dimensional separation[10–12] reduce the list of candidates and increase the rate of identification in metabolomics studies[7,13]. For example, using ReTip algorithm in MSFinder reduced the list of candidate structures by 68% in mouse blood plasma[2]. In our method validation dataset, AminoacidDB the monoisotopic mass 174.100442 matched to 9 amino acids, which was reduced to one annotation, N-acetyl-L-ornithine, with 99.1% RT match. Using CSI:FingerID, one of the best software for metabolite annotation, results in correct compound annotation for 74.8% of the compounds in the top 10 (34% for the top hit) in positive ionization and 60% (28% for the top hit) for negative ionization modes[5]. In the latest CASMI (Critical Assessment of Small Molecule Identification) event, CASMI 2022, the 2D structure annotation was achieved for only 58-64 of the >200 LC-MS/MS spectra inspected through a combined use of SIRIUS 5.7 with Ion Identity Networking in MZMine3, Feature-Based Molecular Networking in GNPS, spectral library search, and Mass Spectrometry Search Tool (MASST) [14]. Manual annotation achieved best results, with 38% of the compounds correctly annotated, however could only inspect 55 spectra compared to >200 spectra processed by most software[14]. The gold standard for unambiguous identification in metabolomics is the use of authentic standards, often isotope-labelled internal standards. This would translate to custom synthesis of >6000 compounds for aminoacidDB Lotus/HMDB dataset which is not practical at the current time. Further MS/MS spectral libraries are only available for less than 1% of the known compounds[14], due to limitation in availability of authentic standards. Use of in-silico MS/MS fragmentation methods, Molecular networking and RT prediction libraries can help increase coverage of

Figure 1: Frequency of occurrence of non-amino acid matches for unique monoisotopic masses in AminoacidDB Lotus/HMDB dataset.

metabolites in untargeted metabolomics studies by employing the latest advancements in machine learning algorithms[7,13,15,16].



The goal of AminoacidDB (and many other metabolomics databases) is to fill the gap to decipher the identity of "dark matter" in untargeted metabolomics. The databases always need complimentary knowledge from expert biologists and biochemists, referred to as "biology-driven metabolomics" to predict which chemical classes and metabolites could be expected in a given experiment[7,15]. There are many specific databases such as SteroidXtract[17], HormonomicsDB[18], FlavanoidSearch[19], Phenol-Explorer[20], targeting a specific class of metabolites that could be expected in a particular organism depending upon the objectives and methodology of the experiment. The reviews[7,13,15,16] highlighted here provide a great detail on the use of complimentary biological information, statistical analysis, pathway mapping in leading the user to putative annotation of metabolites in their untargeted metabolomics datasets. Databases are a good first step toward hypothesis generation but do not provide complete and final results.

Previous authors have demonstrated that RT prediction can be appropriately applied to real-life datasets for metabolite annotation[1–4]. Using a "Retention Projection" approach, which used compound's isocratic retention factor (k) vs solvent composition (Φ) relationship to project a retention time, Abate-Pella et al., tested the effect of unintentional (using same LC conditions) vs intentional (different LC conditions) changes on use of RT projection scores across eight labs, and found acceptable results in consistency of projected retention times[1]. The quantitative structure-retention relationship (QSRR) models, similar to the one used in development of AminoacidDB[2] are even more resilient to interlaboratory changes[2–4]. The RT matching algorithm could also be used to reject candidates with very low match scores, thus reducing the list of candidate targets[7].

In aminoacidDB, the RT matching algorithm seems quite rugged as shown in the meta-analysis of *Arabidopsis thaliana* datasets, if the same solvents and chromatography (C_{18}) are used with different column or gradient elution parameters [Table S7 in original manuscript]. The aminoacidDB was developed using a column and solvent system common to untargeted metabolomics studies. We advise

users to use a similar system to use RT matching criterion in AminoacidDB. We recommend users to run a set of standards, ideally covering a large space on the gradient elution, and use that as a guide to incorporate RT matching into their criterion.

PubChem is a chemical database with over 120 million compounds from >1000 data sources including many chemical databases such as ChEMBL (>4 million annotations), Catalogue of Life (>5 million annotations) in addition to metabolomics databases such as MoNA (1.8 million annotations). The original data downloaded from PubChem (>100 million compounds) was not filtered for element type, number of carbons, molecular weight, etc. resulting in dataset of >2.5 million compounds containing both carboxylic and amine (primary or secondary) groups. This data is used in Figure S2 to provide a general overview of the original PubChem data. The PubChem dataset in AminoacidDB was limited to uncharged chemicals containing CHNOS with number of carbons ranging from 2 to 11 and follows the distribution pattern shown in Figure D8 below.

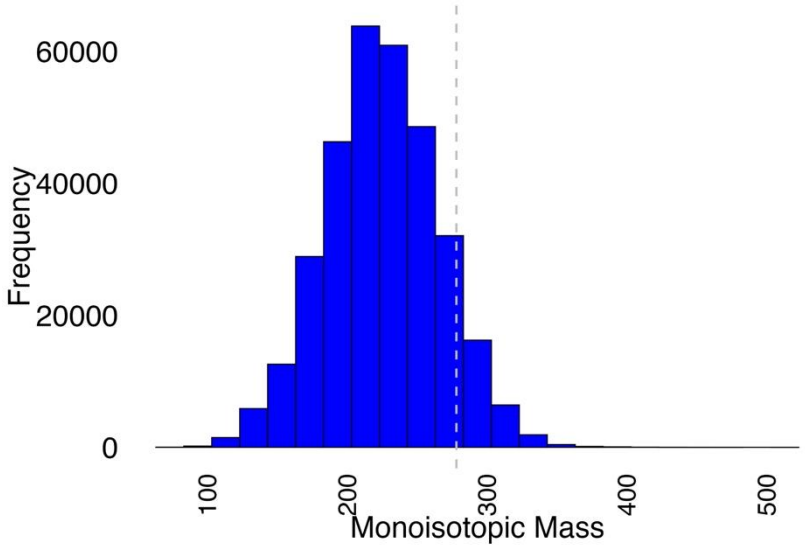


Figure S8: Monoisotopic mass distribution of 325843 compounds in AminoacidDB PubChem dataset. The grey dashed lines indicate 90th percentile of data.

Table S8: Amino acids annotated from three public datasets of *Arabidopsis thaliana* accessed from The Metabolomics Workbench using aminoacidDB and compared with our current data.

SNo.	Name	MSI Level	Adduct	Study	Monoisotopic mass	Retention time (RT; min)	ΔRT (min)	% RT match	Mass Error (ppm)	Found in our *A. thaliana* dataset?	Classification (Lotus; HMDB)
1	L-Arginine	2	[M+H]+	Study 2	174.1117	0.85	0.19	77.6	-4.9	Yes	Amino acids
2	L-Asparagine	2	[M+H]+	Study 2	132.0534	0.91	0.13	85.7	-5.5	Yes	Amino acids
3	L-Glutamine	2	[M+H]+	Study 2	146.0691	0.94	0.19	79.8	-6.3	Yes	Amino acids
4	L-Histidine	2	[M+H]+	Study 2	155.0694	0.85	0.19	77.6	-7.2	No	Amino acids
5	L-Isoleucine	2	[M+H]+	Study 1, 2	131.0946	1.95	0.24	87.7	5.5	Yes	Amino acids
6	L-Leucine	2	[M+H]+	Study 1, 2	131.0946	1.87	0.16	91.4	-0.9	Yes	Amino acids
7	L-Methionine	2	[M+H]+	Study 2	149.051	1.44	0.03	97.9	-6.2	Yes	Amino acids
8	L-Phenylalanine	2	[M+H]+	Study 1, 2, 3	165.0789	2.36	0.27	88.6	4.4	Yes	Amino acids
9	L-Proline	2	[M+H]+	Study 2	115.0633	0.97	0.23	76.3	-7.1	Yes	Amino acids
10	L-Tryptophan	2	[M+H]+	Study 1, 2, 3	204.0898	2.88	1.49	54.6	-1.0	Yes	Amino acids
11	L-Tyrosine	2	[M+H]+	Study 1, 2	181.0739	1.79	0.14	92.2	-0.1	Yes	Amino acids

12	L-Valine	2	[M+H]+	Study 2	117.0789	1.26	0.13	89.7	-7.0	Yes	Amino acids
13	Threonine	2	[M+H]+	Study 2	119.0582	0.89	0.38	57.3	-10.2	Yes	Amino acids
14	Homocysteinesulfinic acid	2	[M+H]+	Study 3	167.025	2.47	0.7	71.7	1.6	No	Alpha amino acids
15	Levetiracetam	2	[M+H]+	Study 3	170.1057	3.36	0.68	79.8	-1.0	No	Alpha amino acids and derivatives
16	(3xi,6xi)-cyclo(alanylvalyl)	2	[M+H]+	Study 3	170.1057	6.99	0.69	90.1	-1.1	No	Alpha amino acids and derivatives
17	N-allylglycine	2	[M+H]+	Study 2	115.0642	1.93	0.06	96.9	-7.7	Yes	Amino acids
18	Norvaline	2	[M+H]+	Study 3	117.079	2.46	0.31	87.4	-1.9	No	Amino acids
19	4-oxo-L-proline	2	[M+H]+	Study 2	129.0431	1.62	0.08	95.1	-4.1	Yes	Amino acids
20	DL-Norleucine	2	[M+H]+	Study 3	131.0946	3.21	0.15	95.4	-1.8	Yes	Amino acids
21	N-hydroxy-L-threonine	2	[M+H]+	Study 2	135.0537	1.87	0.05	97.3	-3.9	No	Amino acids
22	(2S,3R)-3-hydroxybaikiain	2	[M+H]+	Study 1, 2	143.0576	1.87	0.18	90.4	4.6	No	Amino acids
23	(2S,4R)-4-hydroxy-1-methylpyrrolidine-2-carboxylic acid	2	[M+H]+	Study 1, 2	145.0739	2.42	0	100.0	-1.6	No	Amino acids
24	5-hydroxy-4-oxonorvaline	2	[M+H]+	Study 2	147.0545	0.89	0.26	70.8	-9.3	Yes	Amino acids
25	(2R)-2-Amino-3-(methylsulfinyl)propanoic acid	2	[M+H]+	Study 3	151.0303	2.55	0.57	77.6	-0.9	No	Amino acids

26	(2S)-2-amino-2-[(1S,2S)-2-hydroxycyclopent-3-en-1-yl]acetic acid	2	[M+H]+	Study 2	157.0739	2.41	0.9	62.7	-5.3	No	Amino acids
27	4-aminohydratropic acid	2	[M+H]+	Study 2	165.0776	4.15	1.42	65.8	8.5	No	Amino acids
28	O-succinyl-L-homoserine	2	[M+H]+	Study 1, 2	219.0761	6.41	2.84	55.7	-8.3	No	Amino acids
29	Deoxyeritadenine	2	[M+H]+	Study 1, 2	237.0867	2.6	1.02	64.9	-3.1	No	Amino acids
30	Muramic acid	2	[M+H]+	Study 3	251.1015	8.47	0.77	90.9	-3.9	No	Amino acids
31	CID 25113628	2	[M+H]+	Study 1, 2	591.1597	3.67	0.6	83.7	-8.4	No	Amino acids
32	6-amino-5[N-methylformylamino]-1-methyluracil	2	[M+H]+	Study 3	198.0741	2.2	0.48	78.2	5.9	No	Amino acids and derivatives
33	Rocuronium	2	[M+H]+	Study 3	529.398	14.92	4	73.2	4.7	No	Amino acids and derivatives
34	Deferoxamine	2	[M+H]+	Study 3	560.3568	14.33	6.04	57.9	-6.0	No	Amino acids and derivatives
35	Pyroglutamic acid	2	[M+H]+	Study 1, 2	129.0426	1.73	0.4	76.9	5.5	Yes	Amino acids; Alpha amino acids and derivatives
36	L-4-hydroxyglutamate semialdehyde	2	[M+H]+	Study 2	147.0539	0.92	0.05	94.6	-5.2	Yes	Amino acids; Alpha amino acids and derivatives

37	tiglylglycine	2	[M+H]+	Study 2	157.0746	2.95	0.85	71.2	-4.6	No	Amino acids; N-acyl-alpha amino acids
38	m-Aminobenzoic acid	2	[M+H]+	Study 2, 3	137.0477	2.33	0.86	63.1	-1.2	No	Aminobenzoic acids
39	1-(3-(5-Isobutyl-3,6-dioxopiperazin-2-yl)propyl)guanidine	2	[M+H]+	Study 3	269.184	8.26	1.39	83.2	4.6	No	Dipeptides
40	Resorcinomycin B	2	[M+H]+	Study 1, 2	310.1259	2.29	0.04	98.3	5.9	No	Dipeptides
41	Resorcinomycin A	2	[M+H]+	Study 1, 2	324.1414	2.86	0.87	69.6	6.1	No	Dipeptides
42	Eponemycin	2	[M+H]+	Study 3	398.2434	14.14	3.15	77.7	-4.2	No	Dipeptides
43	2-acetamido-3-[[5-[[(2-aminoacetyl)amino]-carboxymethyl]-3-[(2E)-2-hydrazinylideneacetyl]oxy-1-hydroxypyrrolidin-2-yl]methylsulfanyl]propanoic acid	2	[M+H]+	Study 1	478.147	4.07	0.89	78.1	2.5	No	Dipeptides
44	((4-(4-amidinophenoxy)butanoyl)aspartyl)valine	2	[M+H]+	Study 3	436.2001	9.82	2.69	72.6	-9.8	No	Dipeptides, Tripeptides
45	hexanoylglycine	2	[M+H]+	Study 2	173.1045	3.76	1.41	62.5	4.1	No	Dipeptides; N-acyl-alpha amino acids, Alpha amino acid amides
46	N-acetyl-L-methionine	2	[M+H]+	Study 2	191.0604	5.18	1.07	79.3	6.2	No	Dipeptides; N-acyl-

											alpha amino acids, Alpha amino acid amides
47	sulforaphane-N-acetylcysteine	2	[M+H]+	Study 3	340.0564	10.87	1.77	83.7	6.3	No	N-acyl-alpha amino acids
48	Endomorphin-2	2	[M+H]+	Study 3	571.2809	11.4	1.48	87.0	-2.5	No	N-acyl-alpha amino acids and derivatives
49	glycyl-cysteine	2	[M+H]+	Study 2	178.0425	1.53	0.17	88.9	-7.3	No	N-acyl-alpha amino acids, Alpha amino acid amides
50	Isoleucyl-serine	2	[M+H]+	Study 2	218.1268	2.43	0.53	90.9	-0.7	No	N-acyl-alpha amino acids, Alpha amino acid amides
51	gamma-Glutamyltryptophan	2	[M+H]+	Study 3	333.1354	10.8	3.27	69.7	-8.6	No	N-acyl-alpha amino acids, Alpha

											amino acid amides
52	Gluten exorphin A5	2	[M+H]+	Study 3	599.265	10.04	0.82	91.8	-9.8	No	N-acyl-alpha amino acids, Alpha amino acid amides
53	Rimosamide D	2	[M+H]+	Study 2	532.2878	8.54	1.6	81.3	3.6	No	Tripeptides
54	Aeruginozamide AEG603	2	[M+H]+	Study 3	603.255	10.13	4.45	56.1	-5.7	No	Tripeptides
55	Fulicin	2	[M+H]+	Study 3	653.316	12.25	5.49	55.2	2.0	No	Tripeptides
56	(E)-2-amino-4-(3,4-dihydroxyphenyl)-3-butenoic acid; C10H11NO4	4	[M+H]+	Study 1, 2	209.0679	2.25	0.25	88.9	4.4	Yes	-
57	4,6-dihydro-1H-furo[3,4-b]pyrrole-3-carboxylic acid; C7H7NO3	4	[M+H]+	Study 3	153.0427	9.75	0.64	93.4	-1.0	No	-
58	(E)-5-(3-aminophenyl)pent-2-en-4-ynoic acid; C11H9NO2	4	[M+H]+	Study 3	187.0635	8.01	1.36	83.0	-0.8	No	-
59	2-amino-2-(3-benzofuranyl)acetic acid, C10H9NO3	4	[M+H]+	Study 1	191.0577	5.97	0.02	99.7	2.7	No	-
60	(2R)-2-amino-3-(3,3-dihydroxypropylthio)propanoic acid; C6H13NO4S	4	[M+H]+	Study 3	195.0566	2.3	0.05	97.8	-0.5	No	-
61	2-amino-2-(4-hydroxy-2,3-dihydro-1H-inden-5-yl)acetic acid; C11H13NO3	4	[M+H]+	Study 2	207.0887	4.15	0.11	97.3	4.2	No	-
62	2-diazenyl-3-methyl-2-methylsulfonylbutanoic acid; C6H12N2O4S	4	[M+H]+	Study 3	208.0526	12.38	0.2	98.4	-3.7	No	-
63	2-carbamothioyl-5-methoxybenzoic acid; C9H9NO3S	4	[M+H]+	Study 3	211.0297	10.66	0.36	96.6	2.9	No	-
64	2-amino-2-methyl-3-(3-sulfanylidene-1-cyclohexenyl)propanoic acid; C10H15NO2S	4	[M+H]+	Study 3	213.0825	11.86	0.06	99.5	-0.8	No	-

65	4-[(3-aminopropylamino)methyl]-1-cyclohexanecarboxylic acid; C11H22N2O2	4	[M+H]+	Study 3	214.1682	9.4	0.11	98.8	-0.5	No	-
66	2-[[3,4-dihydroxy-5-(hydroxymethyl)cyclohexyl]amino]acetic acid; C9H17NO5	4	[M+H]+	Study 2	219.1099	2.47	0.13	94.7	3.6	No	-
67	5-(2,4-dimethyl-5-thiazolyl)-2H-triazole-4-carboxylic acid; C8H8N4O2S	4	[M+H]+	Study 3	224.036	11.64	0.72	93.8	3.6	No	-
68	(2S)-6-amino-2-(sulfoamino)hexanoic acid; C6H14N2O5S	4	[M+H]+	Study 3	226.0628	2.47	0.03	98.8	-2.0	No	-
69	4-amino-4-oxo-2-[[oxo(1H-pyrazol-4-yl)methyl]amino]butanoic acid; C8H10N4O4	4	[M+H]+	Study 3	226.0691	2.19	0.01	99.5	4.8	No	-
70	4-[(2,3-dihydroxypropylamino)methyl]-1-cyclohexanecarboxylic acid; C11H21NO4	4	[M+H]+	Study 3	231.1472	10.41	0.33	96.8	-0.4	No	-
71	3-[(1-methyl-4-pyrazolyl)methylamino]-2-pyrazinecarboxylic acid; C10H11N5O2	4	[M+H]+	Study 1, 2	233.0922	7.17	0.1	98.6	-4.1	No	-
72	(2R)-2-[(2,4-dihydroxy-3,3-dimethyl-1-oxobutoxy)amino]propanoic acid, C9H17NO6	4	[M+H]+	Study 3	235.1057	9.95	0.75	92.5	-0.4	No	-
73	2-[2-(diethylamino)ethylamino]-4-pyrimidinecarboxylic acid; C11H18N4O2	4	[M+H]+	Study 3	238.1418	7.77	0.02	99.7	4.9	No	-
74	(2R)-1-[(1S)-3-amino-1-carboxy-3-oxopropyl]-5-oxo-2-pyrrolidinecarboxylic acid; C9H12N2O6	4	[M+H]+	Study 3	244.0697	3.27	0	100.0	-0.7	No	-
75	(2R)-2-amino-3-[(2-amino-1-thiophen-2-ylethyl)thio]propanoic acid; C9H14N2O2S2	4	[M+H]+	Study 3	246.0506	2.46	0	100.0	-3.7	No	-
76	(2S)-2-amino-3-(3-hydroxy-3-methylpentyl)pentanedioic acid; C11H21NO5	4	[M+H]+	Study 3	247.1421	11.17	0.38	96.6	-0.3	No	-
77	1-[2-[4-(dimethylamino)butylamino]ethyl]-4-triazolecarboxylic acid; C11H21N5O2	4	[M+H]+	Study 3	255.1683	7.77	0.23	97.0	4.6	No	-
78	5-[[[2-(carbamoylamino)ethylamino]-oxomethyl]amino]-3-pyridinecarboxylic acid; C10H13N5O4	4	[M+H]+	Study 2, 3	267.0969	3.34	0.07	97.9	-0.5	No	-
79	4-[[(2-methyl-5-tetrazolyl)amino]-oxomethyl]-3-thiomorpholinecarboxylic acid; C8H12N6O3S	4	[M+H]+	Study 3	272.0686	10.86	2.7	75.1	2.1	No	-

80	1-[1-(2-amino-3-carboxy-1-oxopropyl)-3-azetidinyl]-4-triazolecarboxylic acid; C10H13N5O5	4	[M+H]+	Study 2, 3	283.0921	3.35	0.15	97.3	-0.3	No	-
81	2-[4-[(1H-imidazol-5-ylsulfonylamino)methyl]-1-triazolyl]acetic acid; C8H10N6O4S	4	[M+H]+	Study 2	286.0491	4.42	1.08	75.6	-2.4	No	-
82	2-[[2-(carboxymethyl)-3-hydroxy-1-aziridinyl]methoxymethyl]aziridine-2,3-dicarboxylic acid, C10H14N2O8	4	[M+H]+	Study 3	290.0756	3.26	0.01	99.7	-2.0	No	-
83	(2R,3aR,6R,7R,7aR)-2-[(2S)-2-amino-2-carboxyethyl]-6,7-dihydroxy-3,3a,5,6,7,7a-hexahydrofuro[3,2-b]pyran-2-carboxylic acid (neodysiherbaine A); C11H17NO8	4	[M+H]+	Study 1, 2	291.0948	1.74	0.17	92.7	3.2	No	-
84	8-(2-methyl-2-thiazolidinyl)-2,6-dioxo-3H-purine-4-carboxylic acid; C10H11N5O4S	4	[M+H]+	Study 3	297.0546	2.27	0.21	90.7	-4.7	No	-
85	2-amino-2-[3-[[(diaminomethylidenehydrazinylidene)methylthio]methyl]-4-hydroxyphenyl]acetic acid; C11H15N5O3S	4	[M+H]+	Study 2	297.0905	2.7	0.33	87.8	-0.5	No	-
86	1-[[(4,6-diamino-1,3,5-triazin-2-yl)amino]methyl]butane-1,2,4-tricarboxylic acid; C11H16N6O6	4	[M+H]+	Study 1	328.1133	2.99	1.34	55.2	-0.5	No	-

References:

(1) Abate-Pella, D.; Freund, D. M.; Ma, Y.; Simón-Manso, Y.; Hollender, J.; Broeckling, C. D.; Huhman, D. V.; Krokhin, O. V.; Stoll, D. R.; Hegeman, A. D.; Kind, T.; Fiehn, O.; Schymanski, E. L.; Prenni, J. E.; Sumner, L. W.; Boswell, P. G. Retention Projection Enables Accurate Calculation of Liquid Chromatographic Retention Times across Labs and Methods. J. Chromatogr. A 2015, 1412, 43–51. https://doi.org/10.1016/j.chroma.2015.07.108.

(2) Bonini, P.; Kind, T.; Tsugawa, H.; Barupal, D. K.; Fiehn, O. Retip: Retention Time Prediction for Compound Annotation in Untargeted Metabolomics. Anal. Chem. 2020, 92 (11), 7515–7522. https://doi.org/10.1021/acs.analchem.9b05765.

(3) Aalizadeh, R.; Nika, M.-C.; Thomaidis, N. S. Development and Application of Retention Time Prediction Models in the Suspect and Non-Target Screening of Emerging Contaminants. J. Hazard. Mater. 2019, 363, 277–285. https://doi.org/10.1016/j.jhazmat.2018.09.047.

(4) Falchi, F.; Bertozzi, S. M.; Ottonello, G.; Ruda, G. F.; Colombano, G.; Fiorelli, C.; Martucci, C.; Bertorelli, R.; Scarpelli, R.; Cavalli, A.; Bandiera, T.; Armirotti, A. Kernel-Based, Partial Least Squares Quantitative Structure-Retention Relationship Model for UPLC Retention Time Prediction: A Useful Tool for Metabolite Identification. Anal. Chem. 2016, 88 (19), 9510–9517. https://doi.org/10.1021/acs.analchem.6b02075.

(5) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. Nat. Methods 2019, 16 (4), 299–302. https://doi.org/10.1038/s41592-019-0344-8.

(6) Blaženović, I.; Kind, T.; Ji, J.; Fiehn, O. Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics. Metabolites 2018, 8 (2), 31. https://doi.org/10.3390/metabo8020031.

(7) Collins, S. L.; Koo, I.; Peters, J. M.; Smith, P. B.; Patterson, A. D. Current Challenges and Recent Developments in Mass Spectrometry–Based Metabolomics. Annu. Rev. Anal. Chem. 2021, 14 (Volume 14, 2021), 467–487. https://doi.org/10.1146/annurev-anchem-091620-015205.

(8) Wu, Q.; Wang, J.-Y.; Han, D.-Q.; Yao, Z.-P. Recent Advances in Differentiation of Isomers by Ion Mobility Mass Spectrometry. TrAC Trends Anal. Chem. 2020, 124, 115801. https://doi.org/10.1016/j.trac.2019.115801.

(9) Joshi, R.; Sharma, S.; Kumar, D. Advances of Ion Mobility Platform for Plant Metabolomics. Crit. Rev. Anal. Chem. 2024, 54 (1), 175–191. https://doi.org/10.1080/10408347.2022.2070000.

(10) Pérez-Cova, M.; Tauler, R.; Jaumot, J. Two-Dimensional Liquid Chromatography in Metabolomics and Lipidomics. In Metabolomics; Wood, P. L., Ed.; Springer US: New York, NY, 2021; pp 25–47. https://doi.org/10.1007/978-1-0716-0864-7_3.

(11) Grübner, M.; Dunkel, A.; Steiner, F.; Hofmann, T. Systematic Evaluation of Liquid Chromatography (LC) Column Combinations for Application in Two-Dimensional LC Metabolomic Studies. Anal. Chem. 2021, 93 (37), 12565–12573. https://doi.org/10.1021/acs.analchem.1c01857.

(12) Keppler, E. A. H.; Jenkins, C. L.; Davis, T. J.; Bean, H. D. Advances in the Application of Comprehensive Two-Dimensional Gas Chromatography in Metabolomics. Trends Anal. Chem. TRAC 2018, 109, 275–286. https://doi.org/10.1016/j.trac.2018.10.015.

(13) Misra, B. B. New Software Tools, Databases, and Resources in Metabolomics: Updates from 2020. Metabolomics 2021, 17 (5), 49. https://doi.org/10.1007/s11306-021-01796-1.

(14) Fiehn Lab - CASMI 2022 - Results. https://fiehnlab.ucdavis.edu/casmi/casmi-2022-results (accessed 2025-11-26).

(15) Guo, J.; Yu, H.; Xing, S.; Huan, T. Addressing Big Data Challenges in Mass Spectrometry-Based Metabolomics. 2022. https://doi.org/10.1039/D2CC03598G.

(16) Krassowski, M.; Das, V.; Sahu, S. K.; Misra, B. B. State of the Field in Multi-Omics Research: From Computational Needs to Data Mining and Sharing. Front. Genet. 2020, 11. https://doi.org/10.3389/fgene.2020.610798.

(17) Xing, S.; Jiao, Y.; Salehzadeh, M.; Soma, K. K.; Huan, T. SteroidXtract: Deep Learning-Based Pattern Recognition Enables Comprehensive and Rapid Extraction of Steroid-Like Metabolic Features for Automated Biology-Driven Metabolomics. Anal. Chem. 2021, 93 (14), 5735–5743. https://doi.org/10.1021/acs.analchem.0c04834.

(18) Giebelhaus, R. T.; Erland, L. A. E.; Murch, S. J. HormonomicsDB: A Novel Workflow for the Untargeted Analysis of Plant Growth Regulators and Hormones. F1000Research October 18, 2022. https://doi.org/10.12688/f1000research.124194.1.

(19) Akimoto, N.; Ara, T.; Nakajima, D.; Suda, K.; Ikeda, C.; Takahashi, S.; Muneto, R.; Yamada, M.; Suzuki, H.; Shibata, D.; Sakurai, N. FlavonoidSearch: A System for Comprehensive Flavonoid Annotation by Mass Spectrometry. Sci. Rep. 2017, 7 (1), 1243. https://doi.org/10.1038/s41598-017-01390-3.

(20) Rothwell, J. A.; Perez-Jimenez, J.; Neveu, V.; Medina-Remón, A.; M'Hiri, N.; García-Lobato, P.; Manach, C.; Knox, C.; Eisner, R.; Wishart, D. S.; Scalbert, A. Phenol-Explorer 3.0: A Major Update of the Phenol-Explorer Database to Incorporate Data on the Effects of Food Processing on Polyphenol Content. Database 2013, 2013, bat070. https://doi.org/10.1093/database/bat070.