Figure S1. Structures for chemical markers analyzed in Chinese date.

(A): The chemical structures of nucleotide and its derivatives. **(B):** The chemical structures of flavonoids. The number in the parenthesis is the notation from the LC analysis on Figure 3.

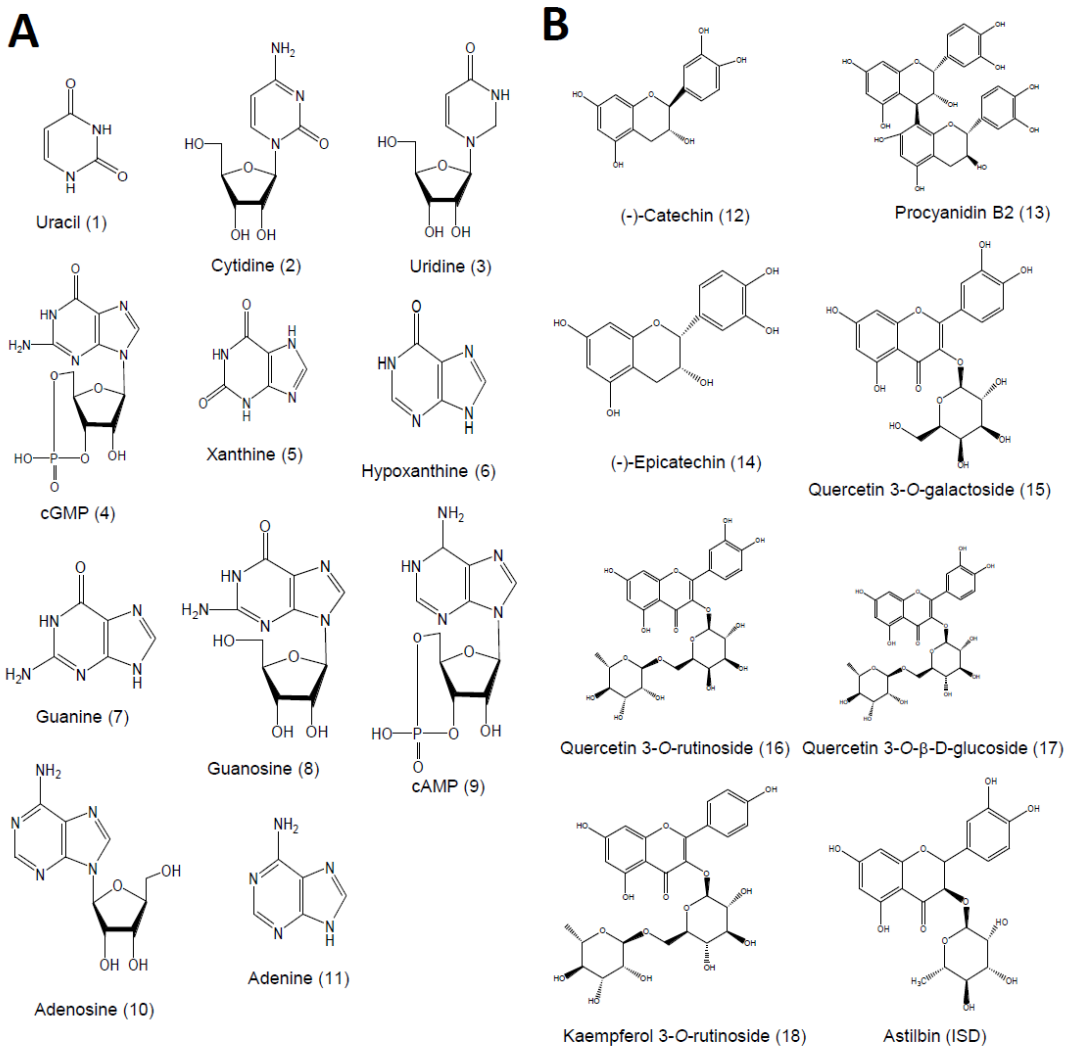


Figure S2. Optimization of extraction cycles in extraction.

The fruit powder (sample A: Jinsixiaozao) (50 g) was boiled in 20 volume of water (v/w) for 1 hour. The extraction was performed 4 times as above. Total amount of tested nucleobase, nucleoside and cyclic nucleotide, including uracil, cytidine, uridine, cGMP, xanthine, hypoxanthine, guanine, guanosine, cAMP and adenine and the flavanoid markers, including (-)-catechin, procyanidin B2, (-)-epicatechin, quercetin 3-*O*-galactoside, quercetin 3-*O*-rutinoside, quercetin 3-*O*-β-D-glucoside and kaempferol 3-*O*-rutinoside were determined by HPLC analysis. The total extraction efficiency of three times is defined as 100%, and each of the extraction efficiency is compared with the total extraction efficiency. They are in Mean ± SD, where n=3. Statistical comparison was made with the lowest total amount of nucleotide or flavonoid in the sample; *p<0.05; **p<0.01. N.D.: not detected.

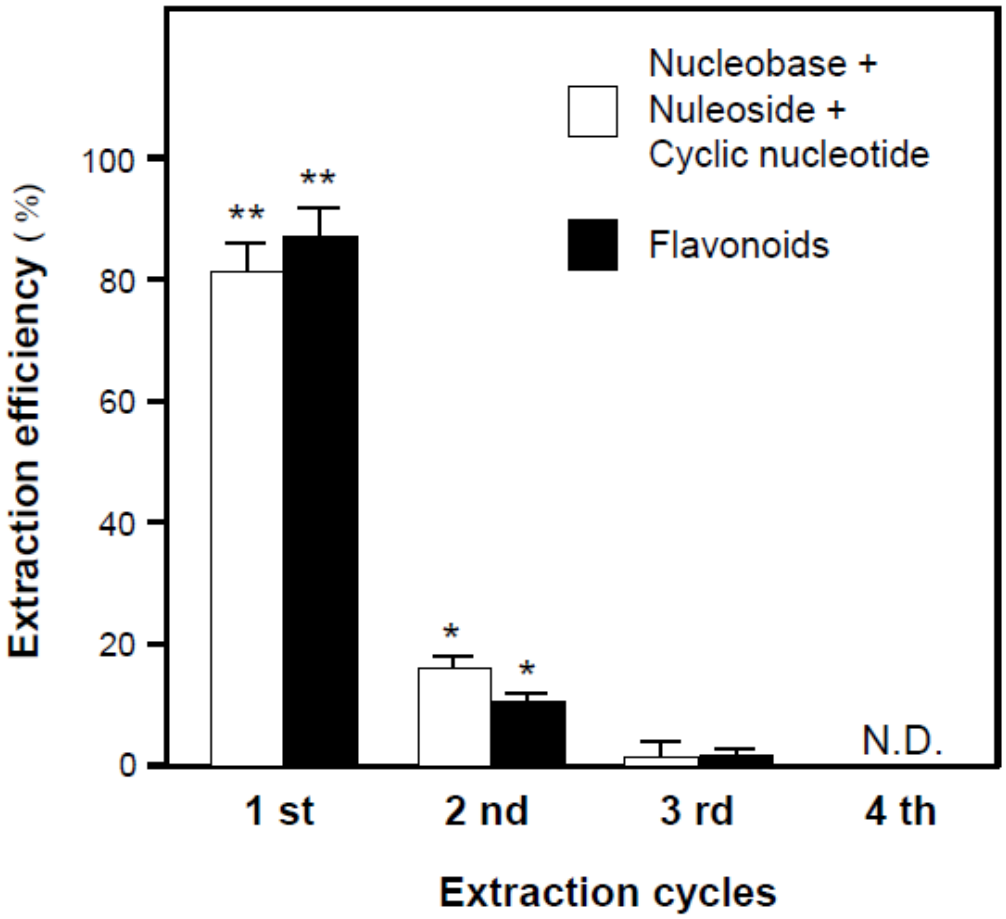
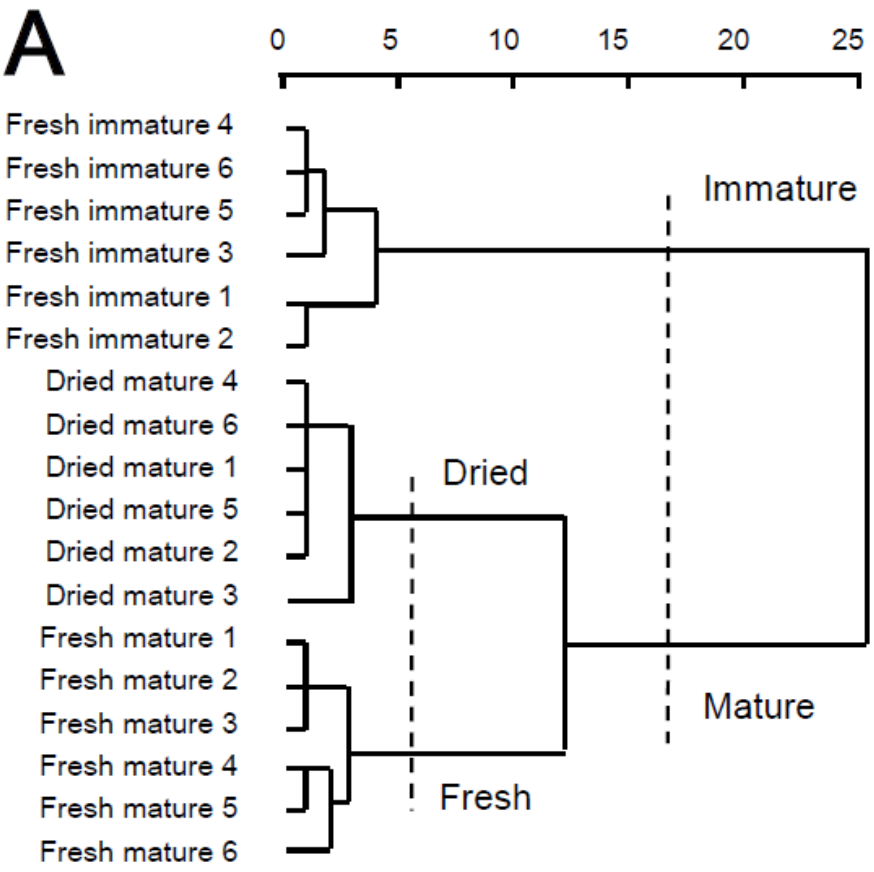

Nucleobase +
Nuleoside +
Cyclic nucleotide
Flavonoids
Extraction efficiency (%)
100
80
60
40
20
0
**
**
*
*
N.D.
1 st
2 nd
3 rd
4 th
Extraction cycles

Figure S3. HCA analysis of Chinese date (sample A: Jinsixiaozao) at different stages.

(A) Cluster plots for fresh immature, fresh mature and dried mature dates, using nucleotide components as input data.

(B) Cluster plots for fresh immature, fresh mature and dried mature dates, using flavonoid components as input data. The scale here is rescaled distance cluster combine.


A
0
5
10
15
20
25
Fresh immature 4
Fresh immature 6
Fresh immature 5
Fresh immature 3
Fresh immature 1
Fresh immature 2
Dried mature 4
Dried mature 6
Dried mature 1
Dried mature 5
Dried mature 2
Dried mature 3
Fresh mature 1
Fresh mature 2
Fresh mature 3
Fresh mature 4
Fresh mature 5
Fresh mature 6
Immature
Dried
Mature
Fresh

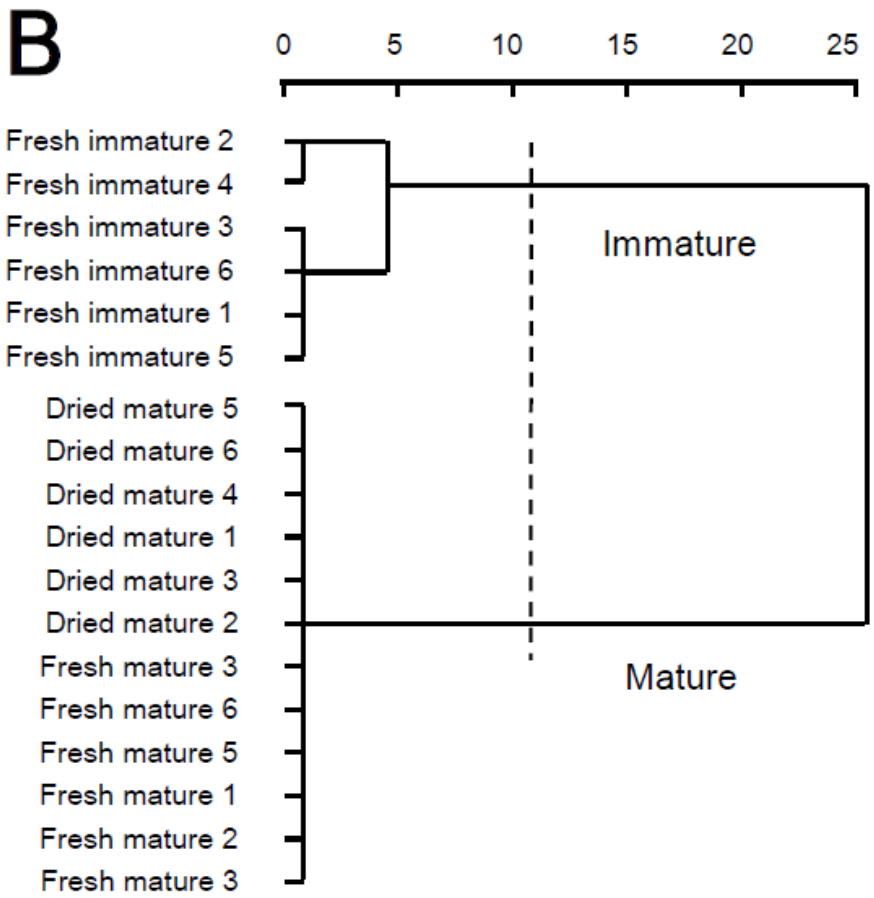

B
0
5
10
15
20
25
Fresh immature 2
Fresh immature 4
Fresh immature 3
Fresh immature 6
Fresh immature 1
Fresh immature 5
Dried mature 5
Dried mature 6
Dried mature 4
Dried mature 1
Dried mature 3
Dried mature 2
Fresh mature 3
Fresh mature 6
Fresh mature 5
Fresh mature 1
Fresh mature 2
Fresh mature 3
Immature
Mature

Table S1. Calibration curves, LOD and LOQ for 11 nucleotide markers

Chemical	Calibration curve [a]	Correlation coefficient	Linear range (μg/L)	LOD [b] (μg/L)	LOQ [c] (μg/L)
Uracil	Y=0.0620X+0.7498	0.9999	156.25-5000	1.60	5.00
Cytidine	Y=0.0240X+0.0488	0.9999	156.25-5000	1.60	5.00
Uridine	Y=0.0358X-0.2478	0.9997	156.25-5000	0.32	1.00
cGMP	Y=0.0226X-3.4292	0.9996	312.50-5000	0.06	0.20
Xanthine	Y=0.0566X-4.6124	0.9994	156.25-5000	0.32	1.00
Hypoxanthine	Y=0.0648X-1.9602	0.9998	156.25-5000	0.32	1.00
Guanine	Y=0.0505X+0.3990	0.9998	156.25-5000	0.06	0.20
Guanosine	Y=0.0403X+2.7010	0.9994	156.25-5000	0.02	0.10
cAMP	Y=0.0201X-0.1294	0.9999	156.25-5000	0.02	0.10
Adenosine	Y=0.0296X+8.3042	0.9984	156.25-2500	0.02	0.10
Adenine	Y=0.0538X-2.4960	0.9996	156.25-5000	0.02	0.10

[a] The calibration curves were constructed by plotting the peak areas versus the concentration of each analyte. Each calibration curve included six data points.

[b] LOD refers to the limit of detection and is determined at an S/N of 3.

[c] LOQ refers to the limit of quantification and is determined at an S/N of 10.

Table S2. Precision, repeatability and recovery of 11 nucleotide makers

Chemicals	Precision		Repeatability		Recoverv [a] (*n*=3)	
	Intra-day [b]	Inter-day [c]	Mean	RSD	Mean	RSD
	RSD(%)	RSD(%)	(μg/g)	(%)	(%)	(%)
Uracil	1.5	2.0	5.3	2.5	97.7	2.4
Cytidine	2.9	2.1	24.8	3.6	94.5	3.8
Uridine	1.8	1.7	34.1	2.2	96.8	2.5
cGMP	2.5	2.4	124.0	2.0	97.3	2.3
Xanthine	1.6	2.4	39.8	3.1	98.6	2.6
Hypoxanthine	2.4	2.6	10.7	1.8	97.2	2.3
Guanine	3.1	2.9	15.5	2.7	95.8	2.6
Guanosine	2.3	2.7	16.3	3.2	96.4	3.4
cAMP	1.6	1.8	186.8	1.5	96.8	2.4
Adenosine	2.4	3.3	6.5	3.3	9423	2.6
Adenine	1.6	1.2	29.0	1.7	98.9	1.7

[a] Recovery (%)=100×(amount found − original amount)/amount spiked. The data were presented as average of three independent determinations.

[b] The intra-day analysis refers to the sample examined for six replicates within one day.

[c] The inter-day analysis refers to the sample examined in duplicates over three consecutive days.

Table S3. Mass spectra properties of marker chemicals in Chinese date

Chemical	Formula	Calculated mass [M]	Precursor ion $[M-H]^-$ [a]	Fragmentor energy [b]	Collison energy [c]	Product ion [d]	Retention time (min) [e]
(-)-Catechin	$C_{15}H_{14}O_6$	290.1	289.1	154	9 29	245.1 109.0	2.14
Procyanidin B2	$C_{30}H_{26}O_{12}$	578.1	577.1	154	21 9	289.1 425.1	2.96
(-)-Epicatechin	$C_{15}H_{14}O_6$	290.1	289.1	154	9 29	245.1 109.0	3.82
Quercetin 3-*O*-galactoside	$C_{12}H_{20}O_{12}$	464.1	463.1	202	25 45	300.0 271.0	11.89
Quercetin 3-*O*-rutinoside	$C_{27}H_{30}O_{16}$	610.1	609.1	250	37 65	300.0 255.0	13.00
Quercetin 3-*O*-β-D-glucoside	$C_{12}H_{20}O_{12}$	464.1	463.1	202	25 45	300.0 271.0	13.73
Kaempferol 3-*O*-rutinoside	$C_{27}H_{30}O_{15}$	594.1	593.1	202	29 61	110.0 285.0	17.60
Astilbin	$C_{21}H_{22}O_{11}$	450.1	449.1	202	13 17	285.0 151.0	12.40

[a] The detected chemicals had the greatest responses under the negative mode: the $[M-H]^-$ was used as the precursor ion.

[b] The fragmentor energy was optimized to have the greatest ionize efficiency.

[c] The collision energy was optimized to have the greatest product ion intensity, which was the key factor in the MRM mode.

[d] Two product ions were used for the MRM analysis. The upper one was used for quantitative analysis and the lower one was for qualitative analyses, which guarantee the precision of chemicals.

[e] The retention time was determined by 3 different individual analyses (n=3).

Table S4. Calibration curves, LOD and LOQ for seven chemicals

Chemical	Calibration curve [a]	Correlation coefficient (r^2)	Linear range (μg/L)	LOD [b] (μg/L)	LOQ [c] (μg/L)
(-)-Catechin	Y=0.682453X+0.014939	0.9954	28.32-7250	0.50	2.00
Procyanidin B2	Y=0.920743X-0.007297	0.9938	56.44-28900	1.00	5.00
(-)-Epicatechin	Y=0.710261X+0.023583	0.9989	28.32-7250	0.50	2.00
Quercetin 3-*O*-galactoside	Y=8.731864X+0.013852	0.9993	0.045-145	0.01	0.02
Quercetin 3-O-rutinoside	Y=4.672127X-0.097741	0.9993	59.57-15250	0.06	0.40
Quercetin 3-O-β-D-glucoside	Y=12.041716X+0.017686	0.9989	4.53-1450	0.30	1.50
Kaempferol 3-O-rutinoside	Y=0.0538X-2.4960	0.9999	5.80-928	0.05	0.20

[a] The calibration curves were constructed by plotting the peak areas versus the concentration of each analyte. Each calibration curve included six data points.

[b] LOD refers to the limit of detection and is determined at an S/N of 3.

[c] LOQ refers to the limit of quantification and is determined at an S/N of 10.

Table S5. Precision, repeatability and recovery of seven flavonoid markers

Chemicals	Precision		Repeatability (*n*=5)		Recovery [a] (*n*=3)	
	Intra-day [b] (*n*=5)	Inter-day [c] (*n*=5)	Mean (mg/g)	RSD (%)	Mean (%)	RSD (%)
	RSD (%)	RSD (%)				
(-)-Catechin	2.3	1.2	7.15	4.8	99.2	5.0
Procyanidin B2	3.5	1.5	36.24	4.6	100.5	4.8
(-)-Epicatechin	0.9	1.6	22.94	5.2	98.1	4.5
Quercetin 3-*O*-galactoside	3.6	4.4	0.40	4.4	95.5	4.8
Quercetin 3-O-rutinoside	2.8	1.4	6.10	4.1	97.5	2.5
Quercetin 3-O-β-D-glucoside	4.6	5.1	0.19	2.3	98.7	3.2
Kaempferol 3-O-rutinoside	0.8	2.2	0.76	2.4	97.4	3.1

[a] Recovery (%)=100×(amount found − original amount)/amount spiked. The data were presented as average of three independent determinations.

[b] The intra-day analysis refers to the sample examined for six replicates within one day.

[c] The inter-day analysis refers to the sample examined in duplicates over three consecutive days.