separations

MDPI

Article; Supplementary Data

Cost-effective Simultaneous Separation and Quantification of Phenolics in Green and Processed Tea using HPLC-UV-ESI Single Quadrupole MS Detector and Python Script

Chan-Su Rha [1,*], Yun-Mi Choi [2], Jong-Chan Kim [3] and Dae-Ok Kim [4,*]

[1] AMOREPACIFIC R&D Center, Yongin 17074, Republic of Korea
[2] Solution Centre, Waters Korea Ltd., Seoul 07241, Republic of Korea; Eunice_Choi@waters.com
[3] Innovalue Partners, Seoul 06724, Republic of Korea; 0001jmp@gmail.com
[4] Department of Food Science and Biotechnology, Kyung Hee University, Yongin 17104, Republic of Korea
* Correspondence: teaman@amorepacific.com (C.-S.R.); DOKIM05@khu.ac.kr (D.-O.K.); Tel.: +82-31-280-5981 (C.-S.R.); +82-31-201-3796 (D.-O.K.)

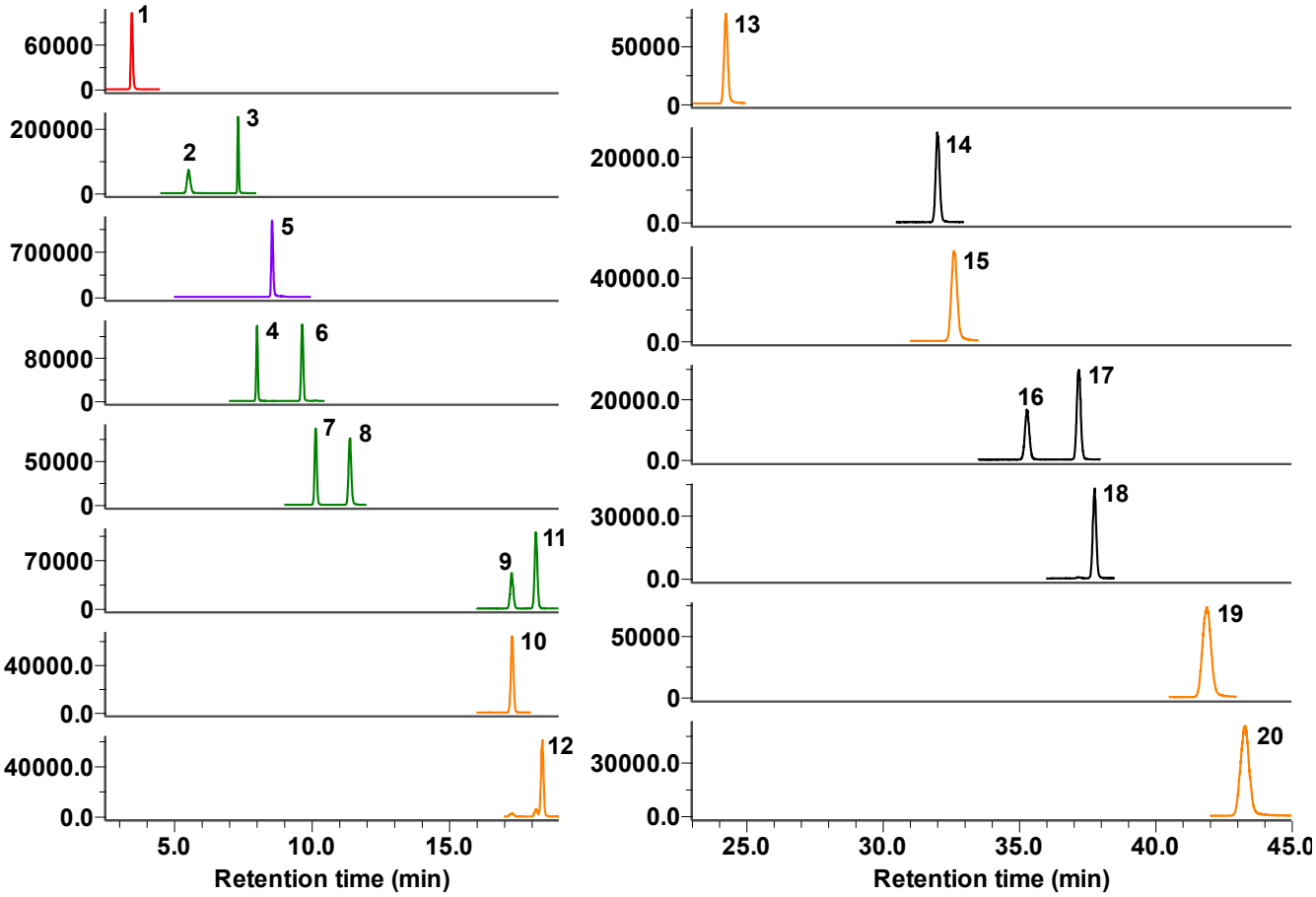

Figure S1. Mass spectrometry chromatograms of 19 phenolics and caffeine standards found in green tea and processed tea. Peaks are indicated listed as follows; 1, gallic acid; 2, (–)-gallocatechin; 3, (–)-epigallocatechin; 4, (–)-catechin; 5, caffeine; 6, (–)-epicatechin; 7, (–)-epigallocatechin gallate; 8, (–)-gallocatechin gallate; 9, (–)-epicatechin gallate; 10, rutin; 11, (–)-catechin gallate; 12, isoquercitrin; 13, myricetin; 14, theaflavin; 15, quercetin; 16, theaflavin 3-*O*-gallate; 17, theaflavin-3′-*O*-gallate; 18, theaflavin-3,3′-*O*-digallate; 19, apigenin; and 20, kaempferol. The concentration of standard aliquots was as follows: 10 μg/mL, peaks 5, 10, 12, 13, 15, 19, and 20; and 20 μg/mL, peaks 1–4, 6–9, 14, and 16–18.

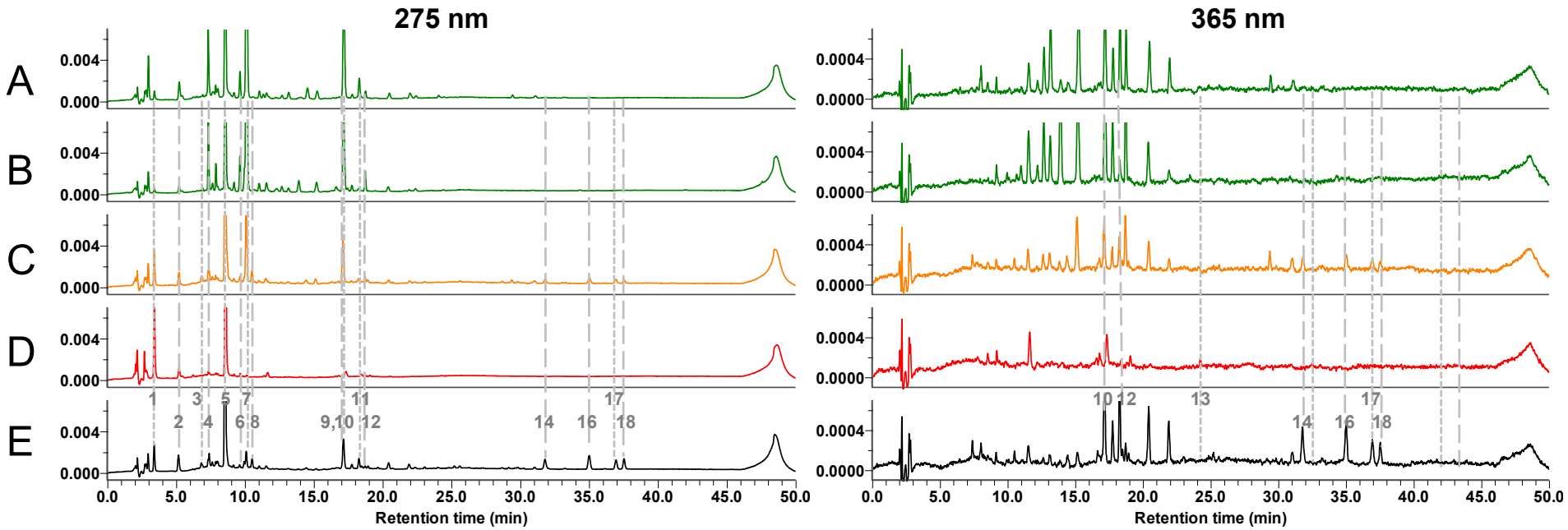

Figure S2. HPLC chromatograms at UV 275 nm (left panel) and UV 365 nm (right panel) of five commercial green teas and processed teas. Samples are listed as follows: (**A**), green tea of Taiwan; (**B**), green tea of Korea; (**C**), oolong tea of Taiwan; (**D**), post-fermented tea of Korea; and (**E**), black tea of England. All tea samples were made into a concentration of 500 μg/mL in 50% (*v*/*v*) aqueous methanol. Refer to Figure 1 for peaks and their corresponding chemical names. The peaks 15, 19, and 20 were not showed due to being below limit of quantification.

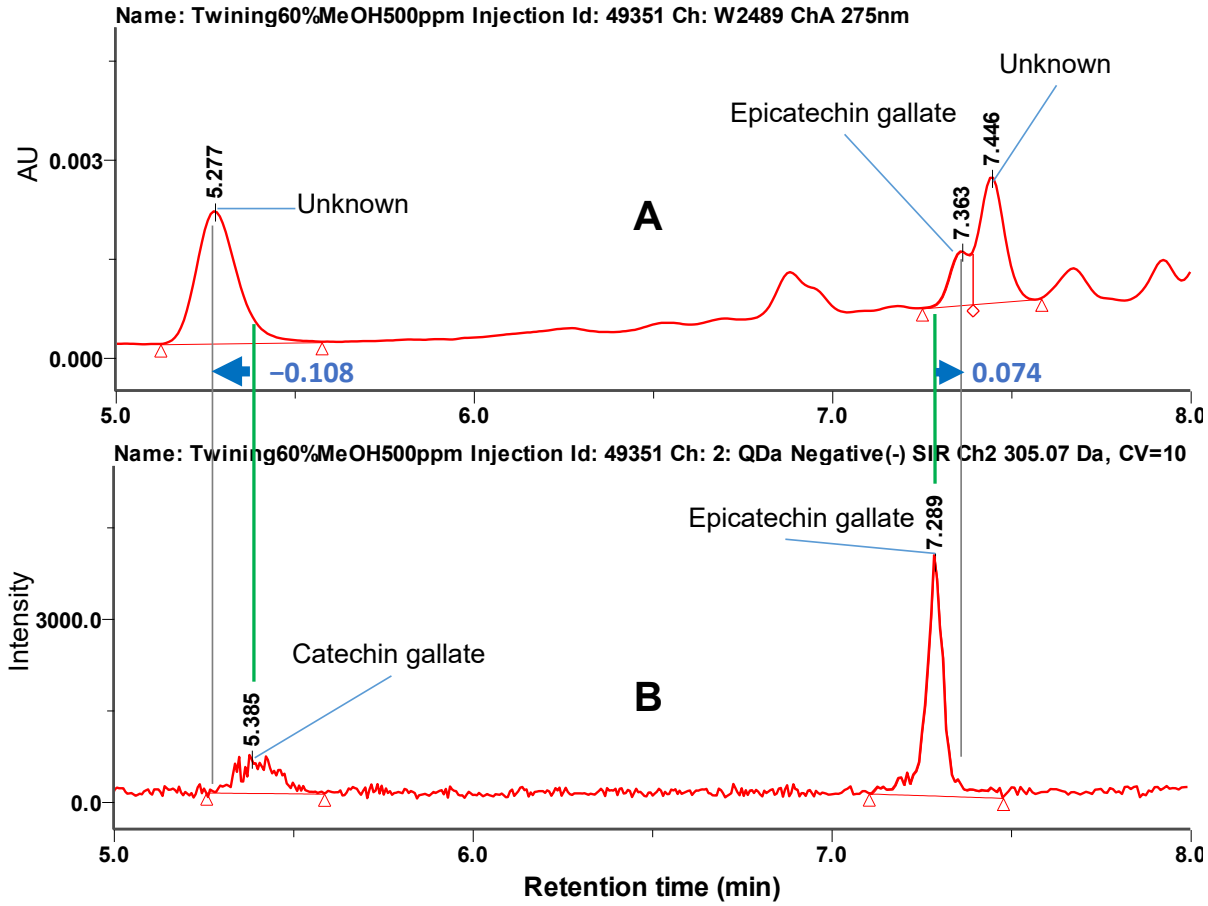

Figure S3. Retention time difference between UV detection and MS detection. Chromatograms of the same analyte (black tea by full-fermentation: 500 mg/L of 60% (*v*/*v*) aqueous methanol, same injection ID) are depicted in (**A**) UV 275 nm and (**B**) single-ion recording of MS (*m*/*z*: 305.07). The vertical green lines indicate the retention time of compounds in the MS corresponding to that of UV. Blue numbers are the gap of t_R (min) calculated as follows: t_R gap = t_R (UV) – t_R (MS^1).

Table S1. Concentrations (mg/g) of phenolics and caffeine in commercial teas determined by an HPLC system equipped with a UV detector (manually processed).

Compound [a]	**Tea extract** [b]									
	GTT MeOH	GTT Water	GTK MeOH	GTK Water	OTT MeOH	OTT Water	PTK MeOH	PTK Water	BTE MeOH	BTE Water
Gallic acid	0.48 ± 0.07d[c]	0.51 ± 0.01d	0.28 ± 0.04d	0.28 ± 0.02d	1.95 ± 0.04c	1.91 ± 0.10c	5.00 ± 0.32a	4.50 ± 0.09b	1.72 ± 0.17c	1.84 ± 0.06c
GC	n.d.[d]	3.43 ± 0.25a	n.d.	1.76 ± 0.08b	n.d.	n.d.	n.d.	1.91 ± 0.15b	n.d.	n.d.
EGC	26.20 ± 0.43ab	23.19 ± 0.04b	28.67 ± 1.90a	22.06 ± 1.01b	5.04 ± 2.62c	2.79 ± 0.38c	4.77 ± 0.55c	5.69 ± 1.19c	4.72 ± 5.30c	3.00 ± 0.20c
C	n.d.	n.d.	n.d.	n.d.	n.d.	<LOQ[e]	2.64 ± 0.52c	n.d.	11.94 ± 1.15a	8.67 ± 0.60b
Caffeine	24.86 ± 0.93e	19.72 ± 0.55f	27.28 ± 1.02d	20.83 ± 0.56f	31.92 ± 0.72b	24.16 ± 0.54e	33.91 ± 0.88a	29.47 ± 0.49c	28.71 ± 0.27cd	23.48 ± 0.29e
EC	7.26 ± 0.63bc	6.54 ± 0.10c	9.81 ± 0.49a	7.86 ± 0.19b	3.71 ± 0.78e	1.64 ± 0.60f	1.60 ± 0.30f	0.98 ± 0.37f	5.11 ± 0.49d	1.79 ± 0.30f
EGCG	60.54 ± 1.34b	21.24 ± 0.50c	65.30 ± 1.07a	12.47 ± 1.20d	12.17 ± 0.52d	1.95 ± 0.17e	n.d.	n.d.	2.52 ± 0.72e	n.d.
GCG	0.67 ± 0.02b	<LOQ	<LOQ	n.d.	n.d.	n.d.	n.d.	<LOQ	n.d.	n.d.
ECG	15.49 ± 0.35a	3.23 ± 0.47e	14.10 ± 0.81b	2.55 ± 0.05ef	5.96 ± 0.28c	<LOQ	<LOQ	n.d.	4.87 ± 0.16d	1.40 ± 0.03fg
Rutin	1.84 ± 0.07ab	1.52 ± 0.07b	n.d.	n.d.	0.76 ± 0.22c	0.68 ± 0.06cd	<LOQ	n.d.	1.93 ± 0.38a	1.71 ± 0.15ab
Isoquercitrin	0.95 ± 0.15ab	0.71 ± 0.04bc	0.56 ± 0.29cd	0.19 ± 0.04e	0.41 ± 0.11de	0.24 ± 0.06e	n.d.	n.d.	1.18 ± 0.22a	0.94 ± 0.10ab
TF	n.d.	n.d.	n.d.	n.d.	0.74 ± 0.05b	n.d.	n.d.	n.d.	2.64 ± 0.39a	0.33 ± 0.17b
TF3G	n.d.	n.d.	n.d.	n.d.	1.71 ± 0.09b	<LOQ	n.d.	n.d.	5.93 ± 0.16a	n.d.
TF3′G	n.d.	n.d.	n.d.	n.d.	0.66 ± 0.11b	n.d.	n.d.	n.d.	2.14 ± 0.14a	n.d.
TF3,3′DG	n.d.	n.d.	n.d.	n.d.	1.40 ± 0.25b	n.d.	n.d.	n.d.	3.22 ± 0.19a	n.d.
Apigenin	<LOQ	n.d.	n.d.	n.d.	<LOQ	n.d.	n.d.	n.d.	n.d.	n.d.

[a] Compounds are listed in ascending order of retention time. Refer to the abbreviations in Section 2.1. Some compounds (CG, quercetin, and kaempferol) were not detected by UV.

[b] Abbreviations: GTT, green tea of Taiwan; GTK, green tea of Korea; OTT, oolong tea of Taiwan; PTK, post-fermented tea of Korea; BTE, black tea of England. MeOH: 60% (*v*/*v*) aqueous methanol, Water: 80°C hot water.

[c] Results are shown as means ± standard deviations (n = 3–6). Different letters in the same row indicate significant differences according to the Tukey–Kramer honestly significant difference test ($p < 0.05$).

[d] n.d.: not detected.

[e] <LOQ: below limit of quantification.

Post-processing of UV data based on MS data

The Python programming language was utilized for the automated calculation of UV sample data according to the MS^1 sample data. Modules and functions of ANACONDA software (www.anaconda.com; Austin, TX, USA) were used to establish the virtual environment of Python. Coded scripts were executed using Jupyter Notebook v.5 software (www.jupyter.org) embedded with the Pandas, NumPy, and Seaborn library of Python (Script S1). Valid UV sample data were collected by matching the injection identification (ID) of each compound based on the MS^1 sample data, of which gap of retention time (t_R gap = t_R (UV)−t_R (MS^1)) were within +0.06 to +0.09. Quantification of each peak was recalculated using the corresponded equation of quantification.

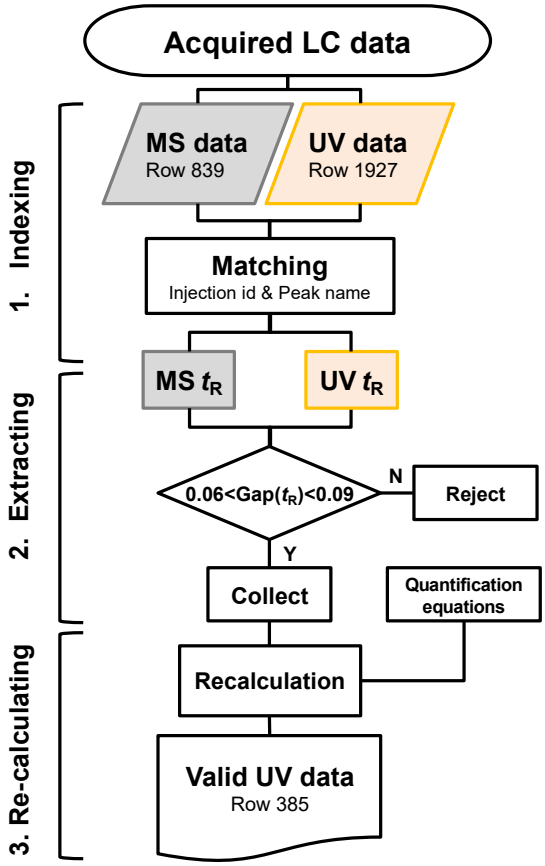

Figure S4. Post-processing scheme for obtaining valid UV data

Script S1. Python input code for processing valid UV sample data

Library and option settings

```
import pandas as pd
import numpy as np
```

UV sample data indexing according to MS^{s} sample data

```
data1 = pd.read_excel('MS_sample_data.xlsx')
df2 = pd.read_excel('UV_sample_data_275.xlsx')
df3 = pd.read_excel('UV_sample_data_365.xlsx')
data1['Retention_Min_range']= data1['Retention Time (min)'] + 0.06
data1['Retention_Max_range']= data1['Retention Time (min)'] + 0.09
data1['unique']= data1['Injection Id'].astype(str) + '_' + data1['Peak Name']
ms_table1 = data1[['Injection Id','Peak Name','unique','Retention Time (min)','Retention_Min_range', 'Retention_Max_range']]
ms_table1.index = pd.IntervalIndex.from_arrays(ms_table1['Retention_Min_range'],ms_table1['Retention_Max_range'],closed='both')
```

Extracting valid UV data

```
# main Code for 'UV_sample_data_275.xlsx'
for j in range(0,len(ms_table1['Injection Id'].unique())):
    try:
        # extracting same 'Injection ID' in MS- and UV-sample
        num1 = ms_table1['Injection Id'].unique()[j]
        cond1 = (df2['Injection Id']==num1)
        cond2 = (ms_table1['Injection Id']==num1)
        id_sample_uv = df2.loc[cond1].reset_index()
        id_sample_ms = ms_table1.loc[cond2]
            # indexing UV data in the tRrange of MS
            for i in range(0,len(id_sample_uv['Retention Time (min)'])) :
                try:
                    x = id_sample_uv['Retention Time (min)'][i]
                    result = id_sample_ms.iloc[id_sample_ms.index.get_loc(x)]
                    # merge the data table for obtaining quantification data
                    dfn = pd.DataFrame(result00).T
                    dfuv = pd.DataFrame(id_sample_uv.iloc[i]).T
                    merge_table = pd.merge(dfuv,dfn, on='Injection Id')
                    # combine MS and UV data within the correct tR
                    df_concat = pd.concat([df_index,merge_table])
                    df_total = pd.concat([df_total,df_concat])
                except Exception as e:
                    pass
                finally:
                    pass
        except Exception as e :
            print('Matching Error')
        finally :
            pass
```

```
# main Code for 'UV_sample_data_365.xlsx'
for j in range(0,len(ms_table1['Injection Id'].unique())):
    try:
        # extracting same 'Injection ID' in MS- and UV-sample
        num1 = ms_table1['Injection Id'].unique()[j]
        cond1 = (df3['Injection Id']==num1)
        cond2 = (ms_table1['Injection Id']==num1)
        id_sample_uv = df3.loc[cond1].reset_index()
        id_sample_ms = ms_table1.loc[cond2]
            # indexing UV data in the tRrange of MS
            for i in range(0,len(id_sample_uv['Retention Time (min)'])) :
```

```
                        try:
                            x = id_sample_uv['Retention Time (min)'][i]
                            result = id_sample_ms.iloc[id_sample_ms.index.get_loc(x)]
                        # merge the data table for obtaining quantification data
                        dfn = pd.DataFrame(result00).T
                        dfuv = pd.DataFrame(id_sample_uv.iloc[i]).T
                        merge_table = pd.merge(dfuv,dfn, on='Injection Id')
                        # combine MS and UV data within the correct tR
                        df_concat = pd.concat([df_index,merge_table])
                        df_total = pd.concat([df_total,df_concat])
                        except Exception as e:
                            pass
                        finally:
                            pass
    except Exception as e :
        print('Matching Error')
    finally :
        pass
```

```
# calculation formula for quantification
def cal_Amount_func(row):
    peak_name = row['Peak Name_y']
    processing_method = row['Processing Method']
    area = row['Area']
    if (peak_name =='Apigenin') & (processing_method =='Flavonol365nm'):
        return (area - 1.22e+003)/(1.68e+004)
    elif (peak_name =='C') & (processing_method =='CateTF275nm'):
        return (area - 1.46e+002)/(3.10e+003)
    elif (peak_name =='Caffeine') & (processing_method =='CateTF275nm'):
        return (area - 7.44e+002)/2.28e+004
    elif (peak_name =='CG') & (processing_method =='CateTF275nm'):
        return (area - 1.93e+002)/1.23e+004
    elif (peak_name =='EC') & (processing_method =='CateTF275nm'):
        return (area - 2.37e+002)/4.44e+003
    elif (peak_name =='ECG') & (processing_method =='CateTF275nm'):
        return (area - 1.24e+002)/1.29e+004
    elif (peak_name =='EGC') & (processing_method =='CateTF275nm'):
        return (area - 4.96e+001)/3.29e+003
    elif (peak_name =='EGCG') & (processing_method =='CateTF275nm'):
        return (area + 9.26e+001)/1.06e+004
    elif (peak_name =='Gallic acid') & (processing_method =='CateTF275nm'):
        return (area - 2.74e+002)/(2.11e+004)
    elif (peak_name =='GC') & (processing_method =='CateTF275nm'):
        return (area + 1.05e+002)/1.84e+003
    elif (peak_name =='GCG') & (processing_method =='CateTF275nm'):
        return (area + 1.33e+002)/9.97e+003
    elif (peak_name =='Isoquercitrin') & (processing_method =='Flavonol365nm'):
        return (area - 6.92e+002)/1.31e+004
    elif (peak_name =='Kaempferol') & (processing_method =='Flavonol365nm'):
        return (area - 1.34e+003)/2.30e+004
    elif (peak_name =='Myricetin') & (processing_method =='Flavonol365nm'):
        return (area + 7.21e+002)/3.06e+004
    elif (peak_name =='Quercetin') & (processing_method =='Flavonol365nm'):
        return (area + 9.54e+002)/2.99e+004
    elif (peak_name =='Rutin') & (processing_method =='Flavonol365nm'):
        return (area - 4.74e+002)/1.03e+004
    elif (peak_name =='TF1') & (processing_method =='CateTF275nm'):
        return (area - 1.39e+002)/1.14e+004
    elif (peak_name =='TF2a') & (processing_method =='CateTF275nm'):
        return (area + 1.22e+002)/8.55e+003
    elif (peak_name =='TF2b') & (processing_method =='CateTF275nm'):
```

```
            return (area + 1.53e+002)/1.49e+004
        elif (peak_name =='TF3') & (processing_method =='CateTF275nm'):
            return (area + 5.57e+002)/1.66e+004
        else:
            pass
```

```
df_total1['Amount_Recalculation'] = df_total1.apply(cal_Amount_func,axis=1)
df_total2['Amount_Recalculation'] = df_total2.apply(cal_Amount_func,axis=1)
```

Recalculating result by the quantification formulas of standard compounds

```
df_total1['Conc(ppm)'] = ((df_total1['Amount_Recalculation']*1000) / df_total1['ppm'].astype(int)).astype(float)
df_total2['Conc(ppm)'] = ((df_total2['Amount_Recalculation']*1000) / df_total2['ppm'].astype(int)).astype(float)
cond1 = (df_total1['Amount']>20)
df_total0275 = df_total1.loc[~(cond1)]
cond2 = (df_total2['Amount']>20)
df_total0365 = df_total2.loc[~(cond2)]
```

Pivoting modified values

```
df_total1['Peak Name'] = df_total1['Peak Name_x'].replace(np.nan, '(blank)')
p_result_RA2 = pd.pivot_table(data= df_total1, index='Peak Name',columns='Group',
                    values='Conc(ppm)',aggfunc=[np.average,'std'], fill_value='n.d')
p_result_RA2
p_result_RA2.to_excel('Pivot_Result_Conc_275.xlsx')
df_total2['Peak Name'] = df_total2['Peak Name_x'].replace(np.nan, '(blank)')
p_result_RA2 = pd.pivot_table(data= df_total2, index='Peak Name',columns='Group',
                    values='Conc(ppm)',aggfunc=[np.average,'std'], fill_value='n.d')
p_result_RA2
p_result_RA2.to_excel('Pivot_Result_Conc_365.xlsx')
```

Table S2. Operation cost comparison of phenolic separation using HPLC-UV-MS and UPLC-MS/MS.

Category	No.	Content [a]	(A) HPLC-UV-MS1 [b]	(B) UPLC-MS/MS [c]	Source of price [d]
Analytical conditions	1	Flowrate (mL/min)	1.0	0.6	
	2	Run time per injection (min)	50	10	
	3	Injections per one separation column	2,000	2,000	
	4	Injections per year	5,000	5,000	
	5	Cumulative time of analysis per year (h)	4,167	833	
	6	Volume of mobile phase solvent per year (L)	250	30	
Analytical consumables (US dollar)	7	Price of column (1 ea)	910	957	waters.com
	8	Price of vial (100 ea)	57	89	waters.com
	9	Price of syringe filter (300 ea)	724	724	waters.com
	10	Price of mobile phase solvent per 4×4 liter (average ratio; Water: Acetonitrile = 80 : 20)	668	967	fishersci.com
	11	Discipline cost (per year per person)	5,165	9,125	waters.com
Operational cost (US dollar per year)	12	Solvent	10,443	1,813	
	13	Column	2,275	2,393	
	14	Vial	2,850	4,450	
	15	Filter	12,060	12,060	
	16	Discipline	5,165	9,125	
	17	Maintenance (per year)	2,875	4,379	waters.com
	18	Sub-Sum (No. 12–17)	35,668	34,220	
Initial cost (US dollar)	19	Price of instruments	80,000	300,000	
Total cost (US dollar)	20	Operation for 5 years (No. 18 × 5 y + No. 19)	258,342	471,098	

[a] Conditions and specifications of the analytical materials and tools and price of discipline: No. 3 and 4 – We set the same limit of injection in the two systems per column and per year. No. **7** – (**A**) CORTECS C18, 90 Å, 2.7 µm, 4.6 × 150 mm; (**B**) Acquity UPLC HSS T3, 100 Å, 1.8 µm, 2.1 × 50 mm. No. **8** – (**A**) Polypropylene 12 × 32 mm Screw Neck Vial with Cap and Preslit PTFE/Silicone Septum, 300 µL; (**B**) TruView LCMS Certified Clear Glass 12 × 32 mm Screw Neck Vial with Cap and Preslit PTFE/Silicone Septum, 2 mL. No. **8** – (**A**) Acrodisc, Minispike Syringe Filter, GHP, 13 mm, 0.45 µm; (**B**) Acrodisc, Minispike Syringe Filter, GHP, 13 mm, 0.2 µm. No. 9 – (**A**) Water for HPLC (Cat: 6003013, 4×4 L, USD 387.47); Acetonitrile for HPLC (Cat: 02002180, 4×4 L, USD 1,792); (**B**) Water for MS (Optima TM, 4×4 L, USD 627), Acetonitrile for MS (Optima TM, 4×4L, USD 2,327). No. 11 – (**A**) HPLC eLearning Track (USD 750), Empower 3 workflow eLearning Track (USD 500), Acquity QDa eLearning Track (USD 300), Method development and validation for HPLC with Lab (USD 3,615); (**B**) UPLC eLearning Track (USD 1,100), Xevo TQ-XS/Xevo TQ-S micro (USD 4,010), MassLynx eLearning Track (USD 400), Method development and validation for HPLC with Lab (USD 3,615). No. **17** – (**A**) 2695 Performance Maintenance kit +Acquity QDa Standard (KAD) Performance Maintenance Kit; (**B**) Xevo TQ-S (Rotary) Performance maintenance Kit with Chemical Kit. No. 19 – The prices of instruments are assumed according to the private experience without the authentic information from the manufacturer. No. 20 – We assumed the total cost for 5 years without any depreciation of the instruments in terms of accounting aspects.

[b] HPLC-UV-MS[1]: high-performance liquid chromatography UV system coupled with an electrospray ionization (ESI) single quadrupole mass spectrometer, Alliance HPLC (Waters Corp., Milford, MA, USA) equipped with a quaternary pump, Waters 2996 UV detector, and ACQUITY QDa™ mass detector

[c] UPLC-MS/MS: ultra-high-performance liquid chromatography UV system coupled with an ESI tandem (triple) quadrupole mass spectrometer, Xevo TQ-S

[d] The price of each analytical materials and tools was accessed in Sep 6, 2020 via authentic web sites, however, variances can be expected according to the local circumstance such as rate of currency, promotions, and market size of the country.