Profiling of illicit cocaine seized in China by ICP-MS analysis of inorganic elements

Cuimei Liu*, Zhendong Hua, Xin Meng
National Narcotics Laboratory, Drug Intelligence and Forensic Center of the Ministry of Public Security, Beijing, China

A R T I C L E   I N F O

Article history:
Received 9 January 2017
Received in revised form 19 April 2017
Accepted 22 April 2017
Available online 29 April 2017

Keywords:
Drug profiling
Cocaine
Inorganic element
Inductive coupled plasma-mass spectrometry
HCA

A B S T R A C T

For the first time in China, the inorganic element profiling of cocaine specimens was performed at the National Narcotics Laboratory. An inductive coupled plasma-mass spectrometry (ICP-MS) method was developed and validated for simultaneous analysis of 26 inorganic elements, including sodium (Na), magnesium (Mg), aluminum (Al), potassium (K), calcium (Ca), vanadium (V), chromium (Cr), manganese (Mn), iron (Fe), cobalt (Co), nickel (Ni), copper (Cu), zinc (Zn), arsenic (As), selenium (Se), molybdenum (Mo), silver (Ag), cadmium (Cd), antimony (Sb), barium (Ba), lanthanum (La), cerium (Ce), thallium (Tl), lead (Pb), thorium (Th), and uranium (U) in illicit cocaine samples. Ultrasonic assisted dissolution was applied for sample preparation. Minimum sample preparation and analysis time were required, which was suitable for routine analysis. After the analysis of 183 cocaine samples seized from 2011 to 2015, the element concentration ranges of cocaine sample were obtained. Based on the quantitative data set of 26 inorganic elements in 131 linked/un-linked cocaine samples, fifty combinations of pre-treatment methods and distance/correlation measurements were tested for their potential discrimination power for cocaine profiling, and normalization + standardization + logarithm (N + S + L) + cosine correlation exhibited the best result. After hierarchical cluster analysis (HCA) analysis of 183 cocaine samples, 21 groups of linked samples were found within and between provinces, which provide intelligence for case connection and revealing of the distribution networks.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Throughout the world, characterization and impurity profiling of seized drugs are increasingly being used to complement routine investigative work by law enforcement authorities. Chemical links between samples may be established, material from different seizures may be classified into groups of related samples and the origin of samples may be identified. That information may be used for evidential purposes or it may be used as a source of more general intelligence to identify drug trafficking patterns and distribution networks [1–3].

Several methods for cocaine profiling have been developed by determination of the major alkaloids [4–8], occluded solvents [9,10], and stable isotopes [11–13] in recent years. All these methods are based upon organic impurities, whereas another option for drug profiling is to study the inorganic impurities. The elemental impurities found in substances manufactured from constituents derived from natural plant origin might be expected to reflect the inorganic elements present in the original natural product overlaid by those added or subtracted during the manufacturing process [18]. Statistical evaluation of trace element composition of drug seizures therefore has potential both in providing a fingerprint by which different samples can be compared and in tracing geographical origin and/or method of manufacture.

Inorganic elemental analysis has been used to trace the origin or production process of many illicit drugs as heroin [14–18], ecstasy tablets [19,20], methamphetamine [21,22], and cannabis [23]. However, its application in the cocaine profiling is somewhat less emphasized. Violante et al. [24] proposed an analytical pilot study on trace elements in cocaine. From their conclusions, the quantification of metallic species can contribute to the characterization and traceability of additive drugs. Bermejo-Barrera et al. [25] reported the determination of 15 metallic species in 46 cocaine samples. Classification of these cocaine samples according to their geographic origin (Colombia and Venezuela) was achieved by application of pattern recognition techniques to the metallic content data. Other studies just focused on the optimization of the determination methods or the variation of elemental composition in cocaine samples [26–30].

http://dx.doi.org/10.1016/j.forsciint.2017.04.014
0379-0738 © 2017 Elsevier B.V. All rights reserved.
Atomic absorption spectrometry (AAS) [25–30], flameatomic emission spectrometry (F-AES) [25] and inductively coupled plasma atomic emission spectrometry (ICP-AES) [24] have been reported for determination of inorganic elements in cocaine samples. But none of related research employing inductive coupled plasma-mass spectrometry (ICP-MS) has been reported yet. ICP-MS is a reliable method for determining trace elements in illicit drug samples [14–23]. Rapid multielement capability, simplicity in spectra, high accuracy and precision and the ability to measure specific isotopes are some major advantages of ICP-MS over other analytical methods.

We reported here an ICP-MS method for the simultaneous determination of 26 inorganic elements of Na, Mg, Al, K, Ca, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Mo, Ag, Cd, Sb, Ba, La, Ce, Tl, Pb, Th, and U in cocaine sample. The ICP-MS system, with its lower detection limits and wider dynamic linear range for simultaneous analysis of almost all kinds of inorganic elements, was able to effectively and rapidly obtain inorganic profiles of cocaine. Ultrasonic assisted dissolution, which exhibited the feature of simple, fast, high efficiency [18], was selected as the sample preparation method. Multivariate statistics were applied on cocaine elemental data so as to evaluate the possibility to differentiate linked and unlinked samples. Different combinations of pre-treatments and distance/correlation calculated methods were evaluated to obtain the best differentiating results. The profiling results can be used as a piece of intelligence to support police investigation as well as to provide knowledge about the structure and organization of the drug trafficking, therefore contributing to forensic intelligence through operational and strategic perspectives.

2. Experimental

2.1. Reagents and standards

Nitric acid (65%, w/w; ultrapure reagent grade) was purchased from Merck (Germany). 18.2 MΩ/cm de-ionised water (Milli-Q, USA) was used throughout.

Environmental calibration standard of (1000 mg/L Ca, Fe, K, Mg, Na; 10 mg/L of Ag, As, Ba, Be, Cd, Co, Cr, Cu, Mn, Mo, Ni, Pb, Sb, Se, Th, Tl, U, V, and Zn) and multi-element calibration standard for lanthanides (10 mg/L of Ce and La) were purchased from Agilent (Santa Clara, CA, USA). 1000 mg/L of Re single-element standard was purchased from Central Iron & Steel Research Institute (Beijing, China).

10% nitric acid containing 10 μg/L Re was used to prepare the calibration curves solutions and dissolve samples.

All polyethylene plastic bottles were rinsed with de-ionised water before use.

2.2. Sampling

A total number of 183 cocaine specimens (cases) were collected by National Narcotics Laboratory of China from 2011 to 2015. 131 linked/unlinked samples were also prepared for the batch analysis method validation. The “linked” groups composed of nine cocaine cases, and each group consisted of a variable number of individual packaged samples. The picture of cocaine samples from five seized cases is shown in Fig. 1. Meanwhile, samples that came from two different groups were defined as unlinked ones.

A representative cocaine control sample (CT) that contains relatively high levels of all selected elements was prepared. Daily analysis of the prepared control sample solution served as a check of the instrumental system prior to sample analysis.

2.3. Sample preparation

Ultrasonic assisted dissolution was used to dissolve cocaine samples. Approximately 50 mg of each homogenized sample was weighed into a 50 mL polyethylene plastic bottle. 10 mL of 10% nitric acid solution containing 10 μg/L Re (Section 2.1) was added using a 10 mL plastic pipette, and the end volumes of the solution were determined gravimetrically. The solution was shaken vigorously and sonicated for 10 min. Each solution was checked for the presence of undissolved substance. If undissolved
substance could not be dissolved by another 10 min of sonication, centrifugation was performed then the supernatant was used for analysis. A reagent blank solution was prepared according to the same procedure applied to the sample.

2.4. Instrumentation

An Agilent 7700s inductively coupled plasma-mass spectrometry (Santa Clara, CA, USA) with octopole reaction system (ORS) collision/reaction cell was used. The instrument was operated in standard He mode to remove the possible polyatomic interferences. An ASX-500 Series ICP-MS autosampler (Agilent, Santa Clara, CA, USA) was used for auto sample. A peristaltic pump was used for sample uptake, and a glass concentric spray chamber and a micro-mist nebulizer were used for sample solution nebulization.

Typical instrumental operating conditions used were: 1510 W forward power, 15 L/min plasma flow, 1.0 L/min auxiliary gas flow, 0.771 L/min carrier gas flow, and 0.401 L/min makeup gas flow, respectively. A peak hopping scan mode was used with a dwell time of 1 s for As and Se, and 30 ms for other elements.

The following 26 elements were analyzed with ICP-MS in the present study (isotope used): Na (23), Mg (24), Al (27), K (39), Ca (44), V (51), Cr (52), Mn (55), Fe (56), Co (59), Ni (60), Cu (63), Zn (66), As (75), Se (78), Mo (95), Ag (107), Cd (112), Sb (121), Ba (137), La (139), Ce (140), Ti (205), Pb (208), Th (232), and U (238).

2.5. Statistical analysis

2.5.1. Statistical software

The numerical data obtained were treated statistically using SPSS® 12.0 software (SPSS Inc., Chicago, Delaware) and Matlab R. 2016b (Math Works Inc.).

2.5.2. Pre-treatment and distance/correlation measurement methods

Systematic or uncontrolled variations may hinder the observation of interesting variations in a data set. A pretreatment step is a very important part of the chemometric analyses used to reduce instrumental influence on the results to compare [31,32]. Therefore, ten types of pre-treatment methods showed in Table 1 were tested on the data. Various distance/correlation measure- ment methods are available for the characterization of links or proximity between samples. They were proved reliable for drug profiling in previous works [7,31,32]. Five distance/correlation measurements showed in Table 1 were tested in order to highlight the similarities between samples. On the whole, fifty combinations of pretreatment and distance/correlation measurement were tested to determine their potential for cocaine profiling.

2.5.3. Estimation of the discrimination

An estimation of the overall discrimination was done for the selecting of the best pre-treatment and distance/correlation measurements combinations that appeared promising in handing our research data. Averages and standard deviations of distance/correlation coefficient values between linked and non-linked samples were calculated. Thus, a discriminating power was calculated according to Eq. (1) [31,32]. It showed how well the linked and non-linked samples were separated. The higher the quotient the better the separation.

\[
\text{Estimation of the discrimination} = \frac{\mu_{\text{linked}} - \sigma_{\text{linked}}}{\mu_{\text{non-linked}}} - \frac{\sigma_{\text{non-linked}}}{\mu_{\text{linked}}} + \sigma_{\text{linked}}
\]

where \( \mu \) is the average, \( \sigma \) is the standard deviation, and \( 3\sigma \) is the 99% of the confidence interval.

2.5.4. Receiver operating characteristic (ROC) curve

The overall efficiency of different statistical combinations was also evaluated using receiver operating characteristic (ROC) curve [33]. Such curve graphically emphasizes the best connection between linked samples and the best separation between non-linked ones. A ROC curve is a plot of the sensitivity as a function of 1-specificity for different distances/correlations obtained. Sensitivity is defined as the probability to get a positive test if the cocaine samples were truly linked (true positive fraction). Specificity is defined as the probability to get a negative test if the cocaine samples were truly non-linked (true negative fraction). Thus, 1-specificity threshold variations can then be observed according to the sensitivity and specificity selected of desired.

2.5.5. Outlier

There are various methods for outlier detection, such as Chauvenet's criterion, Grubbs' test for outliers, Dixon's Q test, In this study, Chauvenet's criterion Eq. (2) [34] was used as the method for the detection of outlier.

\[
\text{ABS}(x_i - \mu) > D_{\text{max}} \sigma
\]

where \( D_{\text{max}} \) is the maximum allowable deviation, ABS is the absolute value, \( x_i \) is the value of suspected outlier, \( \mu \) is the average, \( \sigma \) is the standard deviation.

In this study, the number of samples was 183, which was less than 200 and more than 100, so the value of \( D_{\text{max}} \) was set at 2.81. Therefore, if the value of \( \text{ABS}(x_i - \mu) \) is greater than 2.81\( \sigma \), it is a possible outlier.

2.5.6. Hierarchical cluster analysis (HCA)

Hierarchical cluster analysis (HCA) was performed on 183 cocaine samples collected from 2010 to 2015, which was proved to be an excellent technique to differentiate samples from different batches and to highlight operational links between samples.

3. Results and discussion

3.1. Selection of elements

Elements were selected on the basis of accurate quantification and timesaving. Lanthanides and noble metals were previously reported to have significant effect for geography classification of heroin [14]. The lanthanide elements of dysprosium (Dy), erbium (Er), europium (Eu), gadolinium (Gd), holmium (Ho), lutetium (Lu), neodymium (Nd), praseodymium (Pr), samarium (Sm), terbium (Tb), thorium (Th), thulium (Tm), yttrium (Y), and ytterbium (Yb); as well as the noble metals of gold (Au), hafnium (Hf), iridium (Ir), palladium (Pd), platinum (Pt), rhodium (Rh), and ruthenium (Ru) were found to be rather low in concentration in a preliminary test, hence they were excluded. Some elements like Si and Ti required special treatment in sample preparation and were also not analyzed to simplify the method. Finally, 26 elements were selected as target elements for ICP-MS analysis.

3.2. Method validation

The limit of quantification (LOQ) was set at 9 times the standard deviation of the concentration of each element in the control blank. The linearity ranges were also estimated within the ranges listed in Table 2, the counts were found to be linearly correlated to concentration with correlation coefficient from 0.9979 to 1.000. The control sample that contains relatively high levels of all selected elements was selected for method validation. The
Table 1

Descriptions of pre-treatment methods and distance/correlation measurements (xi: element i concentration; xj: element j concentration; S.D.; element i standard deviation).

<table>
<thead>
<tr>
<th>Pre-treatment/treatment</th>
<th>Abbreviation</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalization</td>
<td>N</td>
<td>$\frac{x_i}{\sum_{k=1}^{n} x_k} \times 100$</td>
</tr>
<tr>
<td>Standardization</td>
<td>S</td>
<td>$\frac{x_i}{SD_i}$</td>
</tr>
<tr>
<td>Logarithm</td>
<td>L</td>
<td>$\log x_i$</td>
</tr>
<tr>
<td>Fourth square root</td>
<td>4R</td>
<td>$\sqrt{x_i}$</td>
</tr>
<tr>
<td>Normalization + standardization</td>
<td>N*S</td>
<td>$\frac{N_i}{\sum_{k=1}^{n} x_k} \times 100$</td>
</tr>
<tr>
<td>Standardization + normalization</td>
<td>S*N</td>
<td>$\sqrt{x_i}$</td>
</tr>
<tr>
<td>Normalization + logarithm</td>
<td>N*L</td>
<td>$\log(N_i)$</td>
</tr>
<tr>
<td>Normalization + fourth square root</td>
<td>N*4R</td>
<td>$\sqrt{N_i}$</td>
</tr>
<tr>
<td>Normalization + standardization + logarithm</td>
<td>N<em>S</em>L</td>
<td>$\log(N \times S_i)$</td>
</tr>
<tr>
<td>Normalization + standardization + fourth square root</td>
<td>N*S+4R</td>
<td>$\sqrt{N \times S_i}$</td>
</tr>
<tr>
<td>Pearson correlation</td>
<td>Pear $r_0$</td>
<td>$\frac{\sum_{k=1}^{n} (x_k - \bar{x})(x_0 - \bar{x})}{\sqrt{\sum_{k=1}^{n} (x_k - \bar{x})^2 \sum_{k=1}^{n} (x_0 - \bar{x})^2}}$</td>
</tr>
<tr>
<td>Cosine correlation</td>
<td>Cos $r_0$</td>
<td>$\frac{\sum_{k=1}^{n} x_k x_0}{\sqrt{\sum_{k=1}^{n} x_k^2 \sum_{k=1}^{n} x_0^2}}$</td>
</tr>
<tr>
<td>Euclidean distance</td>
<td>Euc $D_0$</td>
<td>$\sqrt{\sum_{k=1}^{n} (x_k - x_0)^2}$</td>
</tr>
<tr>
<td>Manhattan distance</td>
<td>Manh $D_0$</td>
<td>$\sum_{k=1}^{n}</td>
</tr>
<tr>
<td>Chebychev</td>
<td>Cheb $D_0$</td>
<td>$\max_{k=1}^{n}</td>
</tr>
</tbody>
</table>

Table 2

The LOQ, linear range, standard curve, repeatability, and reproducibility of the method.

<table>
<thead>
<tr>
<th>Element</th>
<th>LOQ (mg/kg)</th>
<th>Linear range (mg/kg)</th>
<th>Standard curve calibration slope</th>
<th>Correlation coefficient ($R^2$)</th>
<th>Concentration of control sample (mg/kg)</th>
<th>Repeatability (%)</th>
<th>Reproducibility (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Na</td>
<td>3.96</td>
<td>3.96-50,000</td>
<td>y = 2020 x + 79228</td>
<td>1.000</td>
<td>14.999</td>
<td>1.2</td>
<td>4.7</td>
</tr>
<tr>
<td>Mg</td>
<td>0.174</td>
<td>0.174-50,000</td>
<td>y = 927 x + 9563</td>
<td>1.000</td>
<td>40.5</td>
<td>6.5</td>
<td>9.5</td>
</tr>
<tr>
<td>Al</td>
<td>0.121</td>
<td>0.121-500</td>
<td>y = 376 x + 261</td>
<td>0.9982</td>
<td>1.59</td>
<td>7.8</td>
<td>11.8</td>
</tr>
<tr>
<td>K</td>
<td>1.62</td>
<td>1.62-5000</td>
<td>y = 847 x + 18286</td>
<td>0.9999</td>
<td>5.13</td>
<td>2.1</td>
<td>18.6</td>
</tr>
<tr>
<td>Ca</td>
<td>2.18</td>
<td>2.18-50,000</td>
<td>y = 46.7 x + 2286</td>
<td>0.9999</td>
<td>74.2</td>
<td>3.6</td>
<td>10.9</td>
</tr>
<tr>
<td>Cr</td>
<td>0.0088</td>
<td>0.0088-100</td>
<td>y = 1700 x + 108</td>
<td>0.9999</td>
<td>0.323</td>
<td>0.4</td>
<td>14.8</td>
</tr>
<tr>
<td>Mn</td>
<td>0.029</td>
<td>0.029-100</td>
<td>y = 8060 x + 507</td>
<td>0.9999</td>
<td>0.531</td>
<td>2.1</td>
<td>14.1</td>
</tr>
<tr>
<td>Fe</td>
<td>0.0238</td>
<td>0.0238-5000</td>
<td>y = 1990 x + 41580</td>
<td>0.9999</td>
<td>6.39</td>
<td>1.7</td>
<td>13.2</td>
</tr>
<tr>
<td>Co</td>
<td>0.00076</td>
<td>0.00076-10</td>
<td>y = 2260 x + 658</td>
<td>1.000</td>
<td>0.071</td>
<td>2.8</td>
<td>14.9</td>
</tr>
<tr>
<td>Ni</td>
<td>0.00185</td>
<td>0.00185-10</td>
<td>y = 5570 x + 662</td>
<td>0.9979</td>
<td>0.172</td>
<td>2.5</td>
<td>12.9</td>
</tr>
<tr>
<td>Cu</td>
<td>0.0178</td>
<td>0.0178-100</td>
<td>y = 16340 x + 1891</td>
<td>0.9999</td>
<td>0.132</td>
<td>1.5</td>
<td>14.6</td>
</tr>
<tr>
<td>Zn</td>
<td>0.018</td>
<td>0.018-500</td>
<td>y = 2440 x + 383</td>
<td>0.9995</td>
<td>5.42</td>
<td>0.9</td>
<td>15.2</td>
</tr>
<tr>
<td>As</td>
<td>0.0352</td>
<td>0.0352-10</td>
<td>y = 1370 x + 28.5</td>
<td>0.9999</td>
<td>0.0303</td>
<td>8.0</td>
<td>15.1</td>
</tr>
<tr>
<td>Se</td>
<td>0.0146</td>
<td>0.0146-10</td>
<td>y = 120 x + 503</td>
<td>0.9989</td>
<td>0.141</td>
<td>9.6</td>
<td>12.1</td>
</tr>
<tr>
<td>Mo</td>
<td>0.002</td>
<td>0.002-10</td>
<td>y = 8550 x + 114</td>
<td>1.000</td>
<td>0.0033</td>
<td>2.3</td>
<td>7.8</td>
</tr>
<tr>
<td>Ag</td>
<td>0.003</td>
<td>0.003-10</td>
<td>y = 27850 x + 24.2</td>
<td>0.9999</td>
<td>0.0018</td>
<td>5.3</td>
<td>15.1</td>
</tr>
<tr>
<td>Cd</td>
<td>0.00298</td>
<td>0.00298-10</td>
<td>y = 43900 x + 57.5</td>
<td>1.000</td>
<td>0.0075</td>
<td>4.5</td>
<td>13.1</td>
</tr>
<tr>
<td>Sb</td>
<td>0.00164</td>
<td>0.00164-10</td>
<td>y = 1710 x + 27938</td>
<td>0.9984</td>
<td>0.766</td>
<td>1.5</td>
<td>12.7</td>
</tr>
<tr>
<td>Ba</td>
<td>0.0078</td>
<td>0.0078-100</td>
<td>y = 43300 x + 674</td>
<td>1.000</td>
<td>0.563</td>
<td>3.8</td>
<td>13.6</td>
</tr>
<tr>
<td>La</td>
<td>0.00016</td>
<td>0.00016-5</td>
<td>y = 57904 x + 52.7</td>
<td>1.000</td>
<td>0.00456</td>
<td>5.3</td>
<td>11.4</td>
</tr>
<tr>
<td>Ce</td>
<td>0.00006</td>
<td>0.00006-5</td>
<td>y = 69229 x + 85.5</td>
<td>1.000</td>
<td>0.000605</td>
<td>7.4</td>
<td>8</td>
</tr>
<tr>
<td>Ti</td>
<td>0.00036</td>
<td>0.00036-10</td>
<td>y = 72420 x + 613</td>
<td>1.000</td>
<td>0.000402</td>
<td>11.4</td>
<td>17.5</td>
</tr>
<tr>
<td>Pb</td>
<td>0.0046</td>
<td>0.0046-500</td>
<td>y = 49850 x + 4838</td>
<td>0.9996</td>
<td>0.394</td>
<td>0.6</td>
<td>11.9</td>
</tr>
<tr>
<td>Th</td>
<td>0.00786</td>
<td>0.00786-5</td>
<td>y = 85460 x + 4191</td>
<td>0.9986</td>
<td>0.00786</td>
<td>7.4</td>
<td>12.5</td>
</tr>
<tr>
<td>U</td>
<td>0.00094</td>
<td>0.00094-5</td>
<td>y = 102220 x + 759</td>
<td>1.000</td>
<td>0.00140</td>
<td>4.4</td>
<td>18.9</td>
</tr>
</tbody>
</table>
prepared control samples were analyzed six times on the same day to
determine the repeatability of the method (prepared from the
weighting, dissolution, and analyze) and six times on six different
days to determine the reproducibility of the method (one
weighting, dissolution, and analyze per day with different
operators). Satisfied repeatability with RSDs less than 15% and
reproducibility with RSDs less than 20% were observed for all
elements (Table 2). Not surprisingly, the elements with low
concentration in the control sample, for example Ti, showed
relatively higher relative standard deviations (RSDs) for repeatability
and reproducibility.

3.3. Evaluation of data pre-treatment methods

3.3.1. Discriminating power estimation

Different combinations of statistical treatments were tested to
give the best separation result of the 131 linked/unlinked cocaine
samples. Each sample was analyzed twice. Also included were data
from regular re-analysis of the control sample over a prolonged
period, which was used to test the ruggedness of the process.
Each combination of pre-treatment and treatment (Table 1) was applied
on all samples. Finally, a total number of 3798 linked and
24,068 unlinked distance/correlation coefficient values were
observed. An average and standard deviation were thus obtained
for each combination. The discriminating power was evaluated for
all statistical combinations and shown in Fig. 2. According to the
results, Cosine and Pearson correlations allowed better discrimi-
nation between linked and non-linked samples than other three
distance methods.

3.3.2. ROC curves

After the statistical pre-treatment of the concentration values
for every target element, the distance/correlation values for linked/
unlinked cocaine samples were measured. ROC curves were
calculated for each statistical combination, and some representa-
tive ROC curves are shown in Fig. 3. Table 3 represents the AUC
(area under the ROC curve) for the four most discriminative
combinations. All these four combinations gave good discrimina-
tion results (bigger than 0.99); however, N+S+L/Cosine and N+L/
Cosine demonstrated much better result with an AUC of
0.9981 and 0.9975, respectively. In this study, N+S+L/Cosine
was the final choice of the statistical treatment for cocaine
profiling.

3.3.3. Decision threshold value

Overlap was often seen for the distributions of the linked and
unlinked samples [31,32]. For the statistical combination of
N+S+L/Cosine, the correlation coefficient values for linked
samples were dispersed on a short range of high values from
0.7460 to 0.9999, while the ones for unlinked one were dispersed
on a large range of values from −0.2657 to 0.9167 (Fig. 4).
Therefore, a correlation coefficient/distance value above or below
which two samples can be considered as linked or not should be
chosen; also, the true positive, true negative, false positive and
false negative rates (TP, TN, FP, and FN, respectively) obtained
according to the specific threshold values need to be evaluated.
Table 4 lists the lower and upper threshold values of the four most
discriminative statistical combinations at confidence interval of
95%. It is important to highlight that, according to different
practical applications when using this information, the kind of
error rates to minimize is different and may be adjusted. For
example, FP rate should be minimized if the profiling result is used
in court, while FN rate should be minimized if the profiling result is
used to support police investigation.

3.4. Analysis of seized cocaine samples

3.4.1. Concentration ranges of different elements in cocaine samples

A total number of 183 cocaine specimens (cases) were collected from
2011 to 2015. Each sample was analyzed twice following the
procedure listed in Section 3.3, and the average concentration was
used for later calculation. The concentration ranges of the studied
elements were very different (Fig. 5). The minimum, maximum,
mean value and RSD of all investigated elements are given in
Table 4. The most abundant element found in cocaine samples was
Na, 4.89–298.328 mg/kg. The high concentration of Na in some
samples could be ascribed to the use of NaCO3 in the

| Table 3 |
|-----------------|-----------------|------------------|
| Statistical combinations | AUC | Confidence interval at 95% |
| Upper threshold | Lower threshold |
|-----------------|-----------------|------------------|
| N+S+L/Cosine | 0.9975 | 0.9984 | 0.9968 |
| N+S+L/Cosine | 0.9981 | 0.9342 | 0.8507 |
| N+S+4R/Cosine | 0.9968 | 0.9866 | 0.9714 |
| N+S+4R/Cosine | 0.9974 | 0.8769 | 0.7521 |
manufacturing process [35]. The second highest element was Ca, 40.3–201,743 mg/kg, the high concentration of Ca encountered in some samples could be explained by using calcium salts (bicarbonate, chalk, talcum powder) as diluting agents [36]. Al (0.242–571 mg/kg), Fe (0.385–186.2 mg/kg), Zn (0.251–475 mg/kg), and Pb (0.0122–203.7 mg/kg) were found at moderately high levels, possibly because of metal pots in process and for storage. The order of element concentrations of Al, Ba, Cd, Cr, Cu, Fe, Mg, Mn, Ni, Pb, and Zn in this work was the same as that determined by Violante et al. [25]. The Ag and Cd of this study also showed similarities to literature data [30]. Ti (n.d.–0.01275 mg/kg), Th (n. d.–0.203 mg/kg), and U (n.d.–0.0589 mg/kg) were at the lowest measured levels.

3.4.2. Detection of outlier data

It was noted that some samples contained element(s) in extremely high abundance, which is visually displayed in way of boxplot in Fig. 5. Take element Na as an example, the max concentration of it was 298,328 mg/kg, which was almost 26 fold of the average concentration (11,138 mg/kg). This sample was absolutely an outlier. In order to get a more objective concept of average concentrations and concentration ranges for different elements, outlier samples were found and deleted from the sample data. Chauvenet’s criterion (Section 2.5.5) was used to distinguish outlier element, and the sample contained one or more outlier elements was taken as outlier. Finally, a total number of 45 outlier samples were found. After excluding these samples, the concentration ranges and average concentrations of the 137 samples left were newly calculated and showed in Table 4. Elements Mg and Mo changed the most, the average concentration of these two elements decreased to 15% and 18%, respectively.

Deletion of outlier data is a controversial practice frowned on by many scientists and science instructors; while Chauvenet’s criterion provides an objective and quantitative method for data rejection, it does not make the practice more scientifically or methodologically sound, especially in small sets or where a normal distribution cannot be assumed. Therefore, whether to delete the outlier data or not need more consideration in practice work.

3.4.3. Handling of the concentration value below the LOQ

It has been noticed that the elemental concentration in some samples are below the LOQ. To facilitate the statistics calculation, several approaches have been used, for example, attributing a value of zero or assign an arbitrary value to replace that value. Because the value of zero cannot perform the logarithm calculation, the replacement values of $1 \times 10^{-6}$, 1/2 LOQ, and LOQ were

![Fig. 4. (A) Histogram of the Cosine correlation values of unlinked cocaine samples. (B) Histogram of the Cosine correlation values of linked cocaine samples.](image-url)

![Fig. 5. Box plot of the concentration ranges for different elements in cocaine samples.](image-url)
Table 4
Inorganic element concentration levels found in cocaine samples.

<table>
<thead>
<tr>
<th>Element</th>
<th>All 183 samples</th>
<th>138 samples after deletion of 45 outline samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range (mg/kg)</td>
<td>Average concentration (mg/kg)</td>
</tr>
<tr>
<td>Na</td>
<td>4.89–298.328</td>
<td>11.38</td>
</tr>
<tr>
<td>Mg</td>
<td>0.306–17.517</td>
<td>1.41</td>
</tr>
<tr>
<td>Al</td>
<td>0.242–271</td>
<td>13.4</td>
</tr>
<tr>
<td>K</td>
<td>2.67–2870</td>
<td>34.5</td>
</tr>
<tr>
<td>Ca</td>
<td>40.3–201,743</td>
<td>15.9</td>
</tr>
<tr>
<td>V</td>
<td>n.d.–0.27</td>
<td>0.036</td>
</tr>
<tr>
<td>Cr</td>
<td>0.0209–25.8</td>
<td>0.564</td>
</tr>
<tr>
<td>Mn</td>
<td>n.d.–36.4</td>
<td>1.33</td>
</tr>
<tr>
<td>Fe</td>
<td>0.385–186.2</td>
<td>11.4</td>
</tr>
<tr>
<td>Co</td>
<td>n.d.–0.05</td>
<td>0.040</td>
</tr>
<tr>
<td>Ni</td>
<td>n.d.–2.44</td>
<td>0.201</td>
</tr>
<tr>
<td>Cu</td>
<td>0.0476–13.6</td>
<td>0.816</td>
</tr>
<tr>
<td>Zn</td>
<td>0.251–475</td>
<td>16.5</td>
</tr>
<tr>
<td>As</td>
<td>n.d.–4.21</td>
<td>0.114</td>
</tr>
<tr>
<td>Se</td>
<td>0.0501–0.724</td>
<td>0.132</td>
</tr>
<tr>
<td>Mo</td>
<td>0.00837–0.257</td>
<td>0.0992</td>
</tr>
<tr>
<td>Ag</td>
<td>n.d.–0.405</td>
<td>0.0104</td>
</tr>
<tr>
<td>Cd</td>
<td>n.d.–0.529</td>
<td>0.018</td>
</tr>
<tr>
<td>Sb</td>
<td>0.330–2.82</td>
<td>0.801</td>
</tr>
<tr>
<td>Ba</td>
<td>0.0444–10.1</td>
<td>0.488</td>
</tr>
<tr>
<td>La</td>
<td>0.00488–0.4726</td>
<td>0.0140</td>
</tr>
<tr>
<td>Ce</td>
<td>0.00902–0.690</td>
<td>0.0215</td>
</tr>
<tr>
<td>Ti</td>
<td>n.d.–0.01275</td>
<td>0.000481</td>
</tr>
<tr>
<td>Pb</td>
<td>0.0122–203.7</td>
<td>2.67</td>
</tr>
<tr>
<td>Th</td>
<td>n.d.–0.203</td>
<td>0.00390</td>
</tr>
<tr>
<td>U</td>
<td>n.d.–0.0589</td>
<td>0.00180</td>
</tr>
</tbody>
</table>

Fig. 6. Dendrogram resulting from hierarchical cluster analysis of 183 cocaine samples seized from 2011 to 2015.

selected to test the best discriminate power. The discriminate power and AUC calculated based on the data set of 131 linked/unlinked samples were used as the criterion for evaluation. As shown in Table 5, three replacement plans were equivalent for their discriminate effect; while for the AUC result, 1/2 LOQ and LOQ replaced value showed greater accuracy than value replaced with $1 \times 10^{-10}$. Finally, LOQ was chosen to replace the concentration value that below the LOQ.

3.5. Hierarchical cluster analysis (HCA) of cocaine samples collected from 2011 to 2015

The application of multidimensional statistical methods should be a very good approach for group classification of numerous samples in large data sets. Compared with other data reduction methods, dendograms from hierarchical cluster analysis gives the most understandable and digestible presentations of data with similarities and differences between seizures being shown. The elements concentration in 183 cocaine samples was first calibrated by internal standard, and then followed by N+S+L/Cosine treatment. The obtained data set was used to generate a hierarchical clustering scheme based on the "similarity" between
samples. The resulting dendrogram shown in Fig. 6 was useful to get a pre-selection of profiles of high similarity. When Cosine correlation value of 0.9342 (Table 5) was set as the thresholds for linked samples (with a false positive rate of 5%), 21 groups of linked samples were found, and several links were highlighted within and between provinces. The case numbers of all the cases in linked groups were fed back to the narcotics control departments of different provinces/cities concerned for further case information analysis.

4. Conclusions

An ultrasonic assisted dissolution-ICP-MS method was validated for simultaneous analysis of 26 inorganic elements (Na, Mg, Al, K, Ca, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Mo, Ag, Cd, Sb, Ba, La, Ce, Tl, Pb, Th, and U) in cigarette samples. The reported method is fast, convenient and environmentally friendly, which makes it an ideal solution for high throughput screening of large samples. After the analysis of 183 cigarette samples seized from 2011 to 2015, element concentration ranges in cigarette samples were obtained. Moreover, based on the data set of major, minor and trace inorganic elements in 131 linked/unlinked cigarette samples, fifty combinations (ten pre-treatment methods plus five distance/correlation measurements) were tested for their potential discrimination power for cigarette profiling, and N + 5 + L/Cosine exhibited the best result.

In addition to comparative (sample to sample) analyzes, the determination of element content in illicit cigarette samples may be helpful in obtaining information about their geographical origin. The analysis of illicit cigarette sample of known origin will be required for accurate determination of geographical origins and/or trafficking routes.

Acknowledgements

We are gratefully acknowledging the financial support of the Program of National Science & Technology Pillar of China (2011BAK04B06) and the National Key Research and Development Program of China (2016YFC0800903).

References