Journal of Chromatography B, 1029–1030 (2016) 10–15

Contents lists available at ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Short communication

Multivariate analyses of NP-TLC chromatographic retention data for grouping of structurally-related plant secondary metabolites

Eman Shawky

Department of Pharmacognosy, Faculty of Pharmacy, Alexandria University, Alkhartoom Square, Alexandria 21521, Egypt

ARTICLE INFO

Article history:
Received 8 May 2016
Received in revised form 8 June 2016
Accepted 21 June 2016
Available online 24 June 2016

Keywords:
Multivariate
NP-TLC
Retention data
Principal component analysis
Discriminant analysis
Hierarchical clustering heat map

ABSTRACT

The chromatographic behavior of 28 plant secondary metabolites belonging to four chemically similar classes (alkaloids, flavonoids, flavone glycosides and sesquiterpenes) was studied by normal-phase thin-layer chromatography (NP-TLC) under 5 different chromatographic systems commonly used in plant drug analysis with the aim to explore whether the retention properties of these metabolites can determine the chemical group they belong to. The use of R_M values as the retention parameter is implemented as a relatively new approach in plant analysis. Principal component analysis (PCA), hierarchical clustering heat maps and discriminant analysis (DA), were used for statistical evaluation of the chromatographic data and extraction of similarities between chemically related compounds. The twenty eight metabolites were classified into four groups by principal component analysis. The heat map of hierarchical clustering revealed that all metabolites were clustered into four groups, except for caffeine, while linear discriminant analysis showed that 96.4% of metabolites are predicted correctly as the groupings identified by chemical class in original and cross-validated data. The main advantage of the approach described in current paper is its simplicity which can assist with preliminary identification of metabolites in complex plant extracts.

1. Introduction

During recent decades, traditional herbal medicines and their natural products constituents were under special scientific research interests [1]. On the other side, herbal medicines have complex matrices and therefore, there are different problems in their chromatographic fingerprint analysis [2].

Obtaining precise information on the chemical composition of complex natural extracts (metabolomes) that are primarily obtained from plants is a challenging task that requires sophisticated analytical methods. In this respect, significant advances in hyphenated chromatographic techniques (LC–MS, GC–MS and LC-NMR), as well as data mining and processing methods, have occurred over the last decade [2]. Thin-layer chromatography (TLC), a type of planar chromatography, is another option of affordable equipment and is still frequently used in herbal analysis because of its simplicity, versatility, high velocity, specific sensitivity and simple sample preparation [3]. One of the fundamental issues in planar chromatography is the relationship between structure of the solute and its retention in a particular system. The quantitative structure–retention relationship (QSRR) chemometric investigations are widely established and often used in prediction of a retention for new solutes, finding the most informative structure descriptors for retention explaining and checking their compliance with the molecular theory of the separation [3,4]. To this direction, it was assumed that systematic information on the behavior of a series of metabolites in a number of chromatographic conditions can be exploited to predict the chemical class they belong to and that, in turn, may serve a role in metabolomics for the purposes of both peak annotation and de-replication in natural product research. Computerized methods of multivariate data processing, such as principal component analysis, cluster analysis, discriminant analysis, etc., allow the extraction of systematic information from large sets of retention data [6].

In the presented work, 28 metabolites belonging to four chemical classes were subjected to TLC analysis in different chromatographic systems aiming at clustering of metabolites, in accordance to their chemical structure classification. For this purpose, the generated retention data were subjected to three multivariate methods, principal component analysis (PCA) [7], heat map hierarchical cluster analysis (HCA) [8] and discriminant analysis (DA) [9,10].

E-mail address: shawkyeman@yahoo.com

http://dx.doi.org/10.1016/j.jchromb.2016.06.037
1570-0232/

Table 1
Retention data[a] obtained under five different chromatographic systems for 28 secondary metabolites.

Metabolite	Rf syst1	Rf sys2	Rf sys 3	Rf sys 4	Rf sys 5	Rm syst1	Rm sys2	Rm sys 3	Rm sys 4	Rm sys 5
Strychnine	0.206	0.634	0.269	0.873	0.285	0.585	−0.240	0.432	−0.837	0.397
Cinchonidine	0.174	0.412	0.222	0.555	0.238	0.674	0.153	0.544	−0.096	0.505
Quinine	0.238	0.761	0.269	0.936	0.301	0.505	−0.505	0.432	−1.168	0.364
Physostigmine	0.222	0.730	0.253	0.904	0.285	0.544	−0.432	0.467	−0.977	0.397
Ephedrine	0.1583	0.365	0.222	0.476	0.253	0.724	0.240	0.544	0.041	0.467
Caffeine	0.3492	0.809	0.714	0.952	0.873	0.270	−0.628	−0.397	−1.301	−0.837
Lycorine	0.1904	0.380	0.256	0.619	0.269	0.628	0.210	0.467	−0.210	0.432
Apigenin	0.253	0.253	0.841	0.285	0.952	0.467	0.467	−0.724	0.397	−1.301
Rutin	0.0476	0.063	0.317	0.079	0.412	1.301	1.168	0.332	1.064	0.153
Vitexin	0.126	0.126	0.396	0.142	0.555	0.837	0.837	0.181	0.778	−0.096
Chrysin	0.555	0.555	0.873	0.571	0.968	−0.096	−0.096	−0.837	−0.124	−1.484
Tricin	0.222	0.222	0.761	0.349	0.873	0.544	0.544	−0.505	0.270	−0.837
Kaempferol	0.269	0.285	0.873	0.396	0.936	0.432	0.397	−0.837	0.181	−1.168
Astragalin	0.111	0.126	0.365	0.142	0.507	0.903	0.837	0.240	0.778	−0.013
Isorhamnetin neo	0.063	0.079	0.285	0.111	0.460	1.168	1.064	0.397	0.903	0.0690
Quercetin-3-*O*-gluco	0.079	0.079	0.333	0.126	0.428	1.064	1.064	0.301	0.837	0.124
Apigenin 7-*O*-digluco	0.095	0.111	0.396	0.174	0.476	0.977	0.903	0.181	0.674	0.0413
Aguerin B	0.317	0.587	0.603	0.666	0.698	0.332	−0.153	−0.181	−0.301	−0.364
Melampolide	0.349	0.539	0.523	0.714	0.746	0.270	−0.069	−0.041	−0.397	−0.467
Santonin	0.412	0.619	0.650	0.746	0.777	0.153	−0.210	−0.270	−0.467	−0.544
Cynaropicrin	0.333	0.555	0.571	0.666	0.714	0.301	−0.096	−0.124	−0.301	−0.397
Galanthamine	0.253	0.603	0.380	0.825	0.412	0.467	−0.181	0.210	−0.674	0.153
Homatropine	0.126	0.444	0.174	0.539	0.206	0.837	0.096	0.674	−0.069	0.585
Atropine	0.095	0.380	0.158	0.492	0.190	0.977	0.210	0.724	0.013	0.628
Quercetin	0.190	0.190	0.650	0.206	0.793	0.628	0.628	−0.270	0.585	−0.585
Kaempferol-3-*O*-rhamonoglucoside	0.063	0.079	0.349	0.095	0.460	1.168	1.064	0.270	0.977	0.069
Deacylcyanropicrin	0.174	0.444	0.492	0.619	0.650	0.674	0.096	0.013	−0.210	−0.270
Luteolin	0.222	0.238	0.793	0.253	0.920	0.544	0.505	−0.585	0.467	−1.064

[a] Each chromatographic value is the average of three experiments.

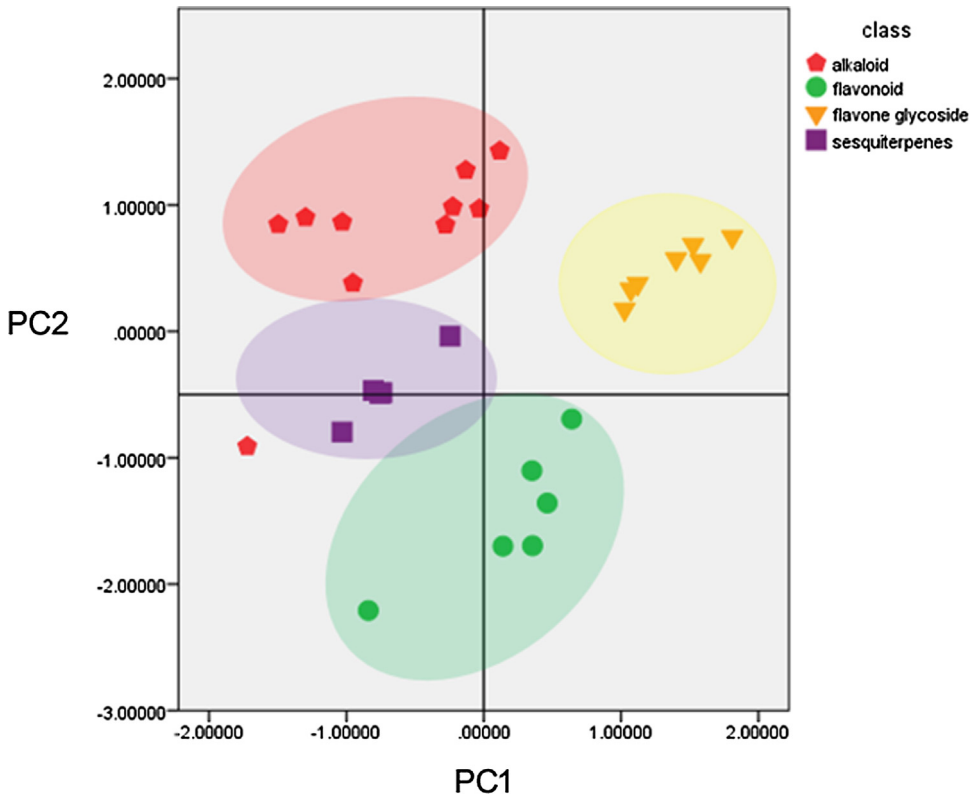

Fig. 1. PCA score scatter plot of all tested analytes belonging to four chemically similar metabolite classes – (a) alkaloids, (b) flavonoids, (c) flavone glycosides and (d) sesquiterpenes – onto the space of the first two principal components PC1 and PC2 extracted from their retention data obtained under all chromatographic conditions tested.

The author is unaware of any previous similar report attempting classification of chemically related plant secondary metabolites by statistical evaluation of their NP-TLC chromatographic retention data which may serve as a preliminary step for improvement of metabolites identification in complex plant extracts.