

# Qualification of coumarin and some phenylpropanoids in cinnamon by HPLC - PDA and preliminary discrimination of cinnamon by growing area

Bui Thi Lan Phuong<sup>1</sup>, Ngo Quang Trung<sup>1</sup>, Trinh Tuan Nam<sup>1</sup>,  
Nguyen Thi Ha Ly<sup>2</sup>, Nguyen Thi Kieu Anh<sup>1\*</sup>

<sup>1</sup> Hanoi University of Pharmacy

<sup>2</sup> National Institute of Medicinal Materials

\* Corresponding author: (+84)915 158 462, email: anhntk@hup.edu.vn

## ABSTRACT

**Background:** Cinnamon growing area is an important factor, deciding its quality. Classification by origin by multivariate statistics is applied in many studies in the world. The objective of this study was to use assays of the main components in the cinnamon cortex to classify cinnamon by growing area.

**Method:** Contents of coumarin and 4 phenylpropanoids in 130 cinnamon-bark samples in Quang Ninh, Quang Nam, Yen Bai, and Thanh Hoa were determined by HPLC-PDA. It was combined with chemometrics to classify cinnamon by growing area

Chromatographic conditions were an Inertsil ODS column (25 cm × 4.6 mm, 5 μm), the gradient mobile phase program consisting of acetonitrile and 0.05 % phosphoric acid in water. Flow rate: 1.0 mL/minute, UV detection: 265 nm, injection volume: 10 μL, retention time: 35 minutes. The discrimination method of cinnamon by the growing source was developed using 5 input variables, which were the content of 5 quantifiable substances in 130 research samples.

**Results:** The study successfully applied a multivariate statistical mathematical method to discriminate cinnamon by growing source with the correct classification rate is 85 %, 95 %, and 100 % respectively for the partial least squares discriminant analysis (PLS-DA) model, principal component analysis discriminant analysis (PCA-DA) model and back-propagation neural network (BNN) model. In which, the BNN model had given the best classification result with an accuracy of 100 %.

**Conclusions:** These results indicated that the application of chemometrics combining chemical data is feasible and can be developed into a set of tools to check the origin of cinnamon.

*Key word:* Cinnamon, chemometrics, principal component analysis PCA, back-propagation neural network.



# Nghiên cứu định lượng coumarin và một số phenylpropanoid trong quế bằng HPLC- PDA và bước đầu phục vụ phân loại quế theo vùng trồng

Bùi Thị Lan Phương<sup>1</sup>, Ngô Quang Trung<sup>1</sup>, Trịnh Tuấn Nam<sup>1</sup>,  
Nguyễn Thị Hà Ly<sup>2</sup>, Nguyễn Thị Kiều Anh<sup>1\*</sup>

<sup>1</sup>Trường Đại học Dược Hà Nội

<sup>2</sup>Viện Dược liệu

\* Tác giả liên hệ: (+84) 915 158 462, anhntk@hup.edu.vn

(Ngày gửi đăng: 21/11/2022 – Ngày duyệt đăng: 25/12/2022)

## TÓM TẮT

**Dẫn nhập và mục tiêu nghiên cứu:** Chất lượng quế bị ảnh hưởng rất nhiều bởi yếu tố vùng trồng, do đó để kiểm soát quế thì việc phân loại theo nguồn gốc là rất cần thiết. Nghiên cứu này được tiến hành với mục đích phân loại quế theo vùng trồng dựa vào kết quả định lượng các thành phần chính.

**Phương pháp nghiên cứu:** Sử dụng phương pháp HPLC-PDA để định lượng các thành phần chính có trong 130 mẫu quế có nguồn gốc từ Quảng Ninh, Quảng Nam, Yên Bái và Thanh Hóa với các thông số: cột Inertsil ODA (25 cm × 4.6 mm, 5 μm), pha động gồm acetonitril và dung dịch acid phosphoric 0,05 %, theo chương trình gradien, tốc độ dòng 1,0 mL/ phút, thể tích tiêm mẫu 10 μL, bước sóng phát hiện 265 nm. Sử dụng phương pháp thống kê đa biến với dữ liệu đầu vào là hàm lượng thành các phần chính của 130 mẫu để phân loại.

**Kết quả nghiên cứu:** Phương pháp phân tích đa biến với các mô hình PLS-DA, PCA-DA và BNN, dữ liệu đầu vào là hàm lượng 5 thành phần chính trong 130 mẫu quế đã phân loại nguồn gốc mẫu quế với tỷ lệ phân loại đúng lần lượt là 85 %, 95 % và 100 %. Trong đó mô hình BNN phân loại đúng 100 % mẫu với độ chính xác 100 %.

**Kết luận:** Đây là cơ sở khoa học giúp nhận biết đúng nguồn gốc địa lý và kiểm soát chất lượng sản phẩm quế.

**Từ khóa:** Quế, chemometris, phân tích thành phần chính PCA, mạng nơ-ron nhân tạo lan truyền ngược, phân loại theo nguồn gốc.

## Đặt vấn đề

Quế được sử dụng rộng rãi trong đời sống hàng ngày như một gia vị trong thực phẩm, phụ gia cho các ngành công nghiệp. Đồng thời, quế là vị thuốc đem lại nhiều lợi ích cho cơ thể, đã được sử dụng từ lâu trong y học cổ truyền. Trong quế chứa nhiều thành phần hóa

học khác nhau như các dẫn xuất của terpenoid, phenylpropanoid (PNP), glycosid, lignan, lacton...[4], [13]... Trong đó, hàm lượng các PNP có trong quế được coi là chất quyết định mùi hương và vị của quế, cũng như quyết định tác dụng dược lý, đặc biệt là giá trị kinh tế của loại cây này. Bên cạnh yếu



tổ giống, chế độ canh tác, kĩ thuật chế biến, ... thì điều kiện khí hậu, thổ nhưỡng có ảnh hưởng rất lớn đến các thành phần hóa học trong quế. Ở nước ta quế thành phẩm được bảo hộ chỉ dẫn địa lý ở 04 vùng gồm quế Thường Xuân, quế Văn Yên, quế Trà Bồng, quế Trà My. Để kiểm soát chất lượng và tránh hiện tượng trộn lẫn quế của mỗi vùng trồng thì việc phân loại quế theo vùng trồng là rất cần thiết.

Sự kết hợp phân tích thành phần hóa học trong mẫu bằng các phương pháp hóa lý với toán hóa (chemometrics) để phân loại nguồn gốc vùng trồng đã mang lại hiệu quả rất khả quan. Ở nước ngoài, sử dụng phương pháp thống kê đa biến dựa trên tập số liệu về kết quả phân tích các thành phần hóa học có hoạt tính [5], [6], [9] hoặc các kim loại [12] trong mẫu để phân biệt nguồn gốc đã được áp dụng từ những năm 2010. Trong những năm gần đây ở Việt Nam đã có một số công bố phát triển phương pháp này trên đối tượng nông sản [2], dược liệu [3] ...

Trong nghiên cứu này, tiến hành áp dụng phương pháp HPLC-UV để xác định hàm lượng các thành phần chính là coumarin (CM), cinnamyl alcohol (CALC), cinnamaldehyd (CAL), acid cinnamic (CA), 2-methoxycinnamaldehyd (MCAL) và acid 2-methoxy cinnamic (MCA) trong 130 mẫu quế thu hái từ 4 vùng trồng khác nhau tại Việt Nam. Đồng thời, kết hợp phương pháp phân tích đa biến để phân loại nguồn gốc mẫu theo vùng trồng làm cơ sở khoa học giúp nhận biết đúng nguồn gốc địa lý và kiểm soát chất lượng sản phẩm quế.

**Đối tượng, nguyên liệu và thiết bị nghiên cứu**

**Đối tượng nghiên cứu:** Vỏ thân cây quế thu hái ở Quảng Nam, Yên Bái, Quảng Ninh, Thanh Hóa thu hái từ tháng 9 -10 năm 2021. Mỗi mẫu khoảng 100 g đóng vào túi PE kín. Bảo quản ở chỗ mát tránh ánh sáng

**Chất đối chiếu:** acid trans-cinnamic (Sigma-Aldrich), Code C80857, độ tinh khiết 99 %; Cinnamaldehyd (Sigma-Aldrich), Code C80687, độ tinh khiết 99 %; Dược liệu đối chiếu *Cinnamomum Cassia* nguồn gốc từ Việt Nam được PGS.TS Trần Văn Ôn (Trường Đại

học Dược Hà Nội) cung cấp sau khi đã xác định đúng tên khoa học.

**Hoá chất dung môi:** acetonitril (ACN) (HPLC, Merck, Đức), methanol (MeOH) (HPLC, Fisher, Mỹ), acid phosphoric (HPLC, Merck, Đức), methanol, xylen (tinh khiết hoá học).

**Thiết bị:** Hệ thống HPLC – DAD Shimadzu 20A(Nhật); Cân phân tích Mettler – Toledo XP 105, độ chính xác 0,0001 g; Máy siêu âm DAIHAN WUC – A22H; Máy li tâm lạnh MIRKO 220H HETTICH.

**Phương pháp nghiên cứu**

**Phương pháp định lượng hàm lượng coumarin và các phenylpropanoid trong quế**

Sau khi tham khảo các dược điển Mỹ [11], dược điển Trung Quốc [8], dược điển Hồng Kong [7] và đồng thời căn cứ vào mục đích phân loại quế theo vùng trồng từ phân tích thành phần chính, nghiên cứu đã chọn định lượng các thành phần trong quế được thực hiện theo phương pháp trình bày trong chuyên luận *Cinnamomum cassia* của USP 2022 có thay đổi về dược liệu đối chiếu và bổ sung dung dịch chuẩn hỗn hợp CA và CAL trong đánh giá độ phù hợp hệ thống. Cụ thể như sau:

**Điều kiện sắc ký:** Hệ thống HPLC – DAD Shimadzu 20A(Nhật); cột tách Inertsil ODS (25 cm x 4,6 mm, 5 µm), bước sóng phát hiện tại 265 nm, thể tích tiêm mẫu: 10 µl, tốc độ dòng: 1,0 mL/phút, chế độ rửa giải gradien gồm acetonitril (ACN) và acid phosphoric 0,05 % trong nước như sau:

Bảng 1. Chương trình dung môi chạy HPLC

Thời gian (phút)	Dung dịch acid phosphoric 0,05% (%)	Acetonitril (%)
0	75	25
1	75	25
21	62	38
30	60	40
35	60	40

**Chuẩn bị các dung dịch**

- Dung dịch chuẩn CA: Pha dung dịch chuẩn CA nồng độ 0,05 mg/ml trong MeOH từ chất chuẩn.



- Dung dịch đối chiếu 1: Pha dung dịch chuẩn hỗn hợp CAL nồng độ 0,005 mg/ml và CA nồng độ 0,05 mg/ml trong MeOH.

- Dung dịch đối chiếu 2: Pha dung dịch dược liệu đối chiếu (*Cinnamomum Cassia*) tương tự mẫu thử

- Dung dịch thử: Cân chính xác khoảng 0,25 g bột mẫu thử vào bình định mức 25 ml. Thêm 1 thể tích hỗn hợp dung môi chiết (MeOH – nước 7:3). Lắc nhẹ 1 phút. Siêu âm 30 phút. Để nguội. Thêm dung môi chiết đến vạch. Lắc đều. Lấy khoảng 5 ml dung dịch vào ống ly tâm và ly tâm 3000 vòng/ phút x 5 phút. Lọc lớp dịch trong phía trên qua màng lọc 0,45 µm được dung dịch tiêm sắc ký.

- Độ ẩm của mẫu thử được xác định độ ẩm bằng phương pháp cất với dung môi (Phụ lục 12.13- ĐĐVN V) [1].

*Định danh các chất nghiên cứu:* dựa vào thời gian lưu tương đối so với CA trong sắc ký đồ dung dịch đối chiếu 2, theo bảng 2

*Bảng 2. Thời gian lưu tương đối so với acid cinnamic và hệ số đáp ứng của coumarin và các phenylpropanoid*

Chất phân tích	Thời gian lưu tương đối	Hệ số đáp ứng
Coumarin (CM)	0,80	2,10
Cinnamyl alcohol (CALC)	0,90	2,12
Acid cinnamic (CA)	1,00	1,00
Acid 2-methoxyl cinnamicmic (MCA)	1,10	1,82
Cinnamaldehy (CAL)	1,30	1,47
2-Methoxy cinnamaldehyd (MCAL)	1,50	2,69

*Yêu cầu:* Dung dịch đối chiếu 1: Thời gian lưu tương đối của pic CAL và pic CA so với pic acid cinamic khoảng 1,3 và 1,0.

Dung dịch đối chiếu 2: Độ phân giải (Rs) giữa pic CALC và CA; giữa pic CM và CALC không dưới 1,5.

Dung dịch chuẩn CA: Hệ số kéo đuôi (Tailing factor) của pic CA không quá 2,0. Độ lệch chuẩn tương đối (RSD) của pic CA trong

6 lần tiêm lặp lại không vượt quá 2,0%.

Hàm lượng phần trăm chất phân tích trong mẫu thử khô kiệt được tính theo công thức sau:

$$(r_u / r_s) \times C_s \times (V/W / (1-H/100)) \times F \times 100$$

Trong đó:

$r_u$  : diện tích pic của chất tương ứng trên sắc ký đồ dung dịch thử

$r_s$  : diện tích pic của CA trong sắc ký đồ của dung dịch chuẩn

$C_s$  : Nồng độ CA của dung dịch chuẩn (mg/ml)

V: Là thể tích pha mẫu thử (ml);

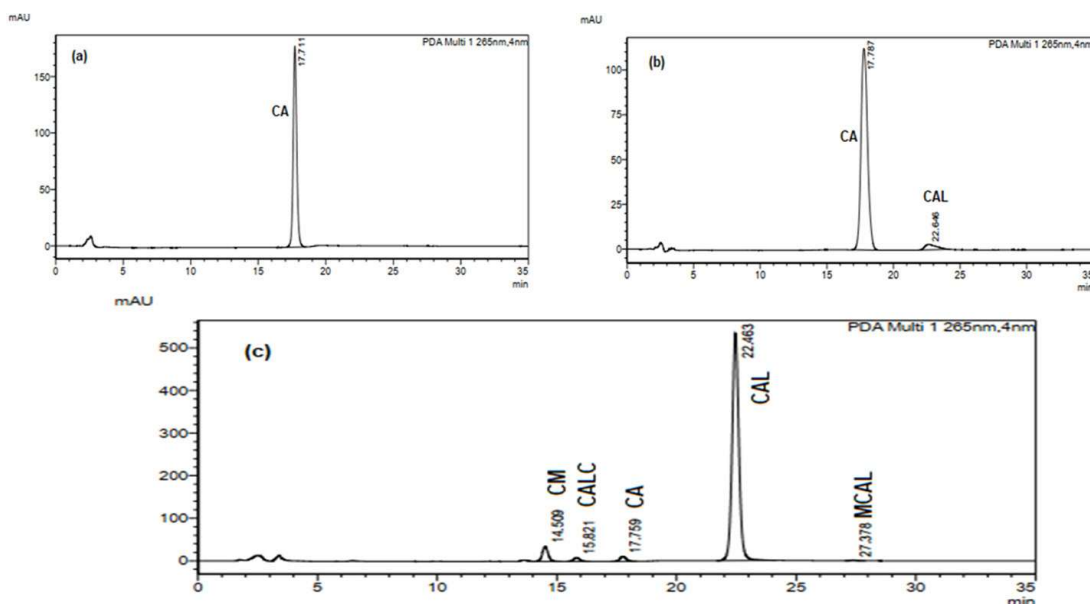
W: Khối lượng mẫu thử (mg)

H: độ ẩm (%)

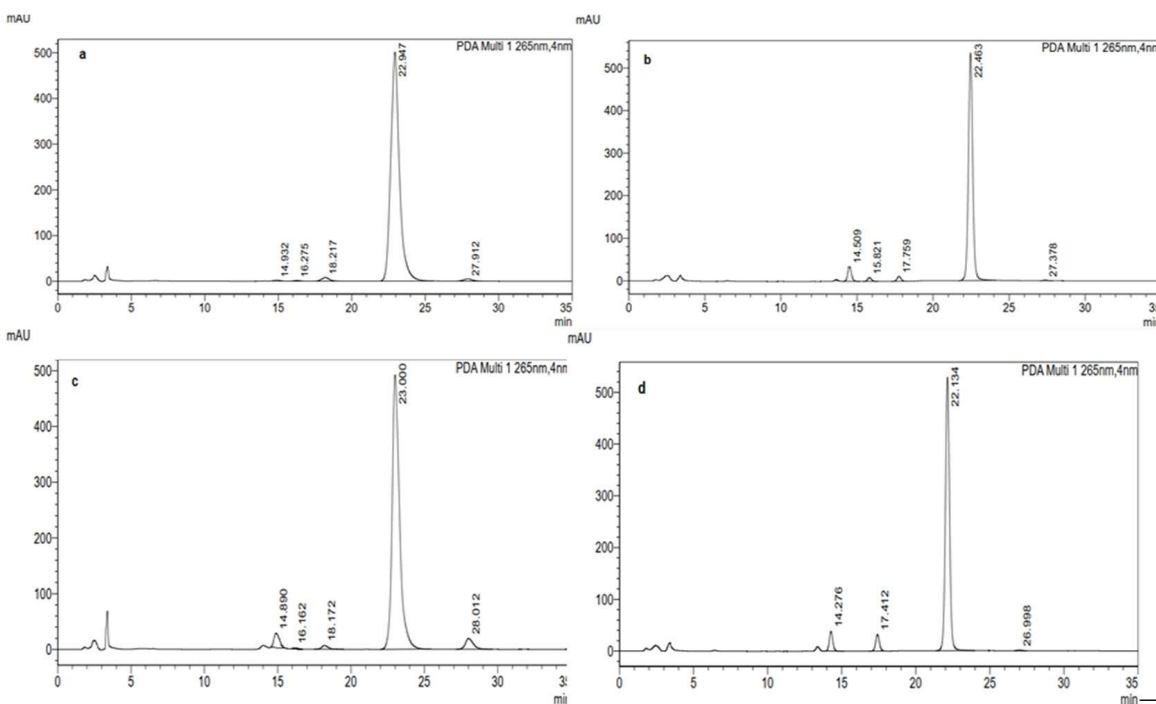
F: Hệ số đáp ứng của chất phân tích theo bảng 2.

**Phương pháp thống kê đa biến để phân loại quế theo vùng trồng:** Sử dụng một số mô hình thống kê đa biến chạy trên phần mềm lập trình tính toán ma trận Matlab R2018a cụ thể gồm: Phân tích thành phần chính kết hợp phân tích phân biệt (principal components analysis discriminant analysis, PCA-DA), Phân tích biệt số bình phương tối thiểu từng phần (partial least squares discriminant analysis, PLS-DA) và mạng nơron nhân tạo lan truyền ngược (back-propagation neural network, BNN) là những mô hình có nhiều nghiên cứu lựa chọn khi phân loại nguồn gốc dựa vào thành phần chính cho kết quả tốt.

Sử dụng 5 biến đầu vào là hàm lượng 5 chất định danh được trong 130 mẫu nghiên cứu (ma trận 130 dòng và 5 cột). Trong đó, chọn ngẫu nhiên khoảng 15 % số mẫu để làm bộ kiểm tra (testing set) (tương ứng 5 mẫu /tỉnh x 4 tỉnh, tổng cộng có 20 mẫu kiểm tra), còn lại 110 mẫu được sử dụng để xây dựng bộ luyện (training set) cho mô hình phân loại. Bộ kiểm tra (testing set) được sử dụng để kiểm tra độ chính xác của bộ luyện (training set). Chấp nhận mô hình phân loại đúng từ 95 % và có đồ thị phân bố rõ ràng.



Hình 1. Sắc ký đồ sử dụng để định danh các chất nghiên cứu (a) dung dịch chuẩn CA; (b) Dung dịch đối chiếu 1; (c) Dung dịch đối chiếu 2



Hình 2. Sắc ký đồ mẫu quế thu hái ở Quảng Ninh (a), Quảng Nam (b), Yên Bái (c), Thanh Hóa (d)

**Kết quả nghiên cứu**

**Định lượng coumarin và một số phenylpropanoid trong quế bằng phương pháp HPLC**

Định danh các chất nghiên cứu và đánh giá độ phù hợp của hệ thống

Chuẩn bị các dung dịch chuẩn CA, dung dịch đối chiếu 1, dung dịch đối chiếu 2 như



Bảng 3. Kết quả trung bình hàm lượng của các PNP và % hàm lượng CAL trong tổng PNP

	Độ ẩm	CM (%)	CALC (%)	CA (%)	MCA (%)	CAL (%)	MCAL (%)	Tổng hàm lượng 5 PNP (%) (không tính CM)	% hàm lượng CAL trong tổng PNP
<b>Quảng Ninh (n=42)</b>									
<b>Giá trị thấp nhất</b>	10,786	0,007	0,002	0,017	-	2,290	0,032	2,681	85,409
<b>Giá trị lớn nhất</b>	14,733	0,240	0,052	0,068	-	8,670	0,328	9,199	98,092
<b>Trung bình</b>	12,827	0,060	0,017	0,041	-	4,484	0,167	4,709	94,077
<b>RSD (%)</b>	0,822	0,053	0,010	0,011	-	1,097	0,132	1,158	2,565
<b>Quảng Nam (n=30)</b>									
<b>Giá trị thấp nhất</b>	14,034	0,053	0,025	0,033	-	2,440	0,005	2,542	90,062
<b>Giá trị lớn nhất</b>	17,742	0,770	0,350	0,085	-	8,205	0,219	8,605	98,662
<b>Trung bình</b>	15,321	0,400	0,084	0,051	-	4,390	0,023	4,548	96,432
<b>RSD (%)</b>	5,477	45,479	84,695	27,037	-	29,491	181,649	29,286	1,643
<b>Yên Bái (n=48)</b>									
<b>Giá trị thấp nhất</b>	9,120	0,043	0,006	0,013	-	2,220	0,013	2,367	87,282
<b>Giá trị lớn nhất</b>	14,667	0,930	0,120	0,068	-	7,423	0,481	7,626	98,466
<b>Trung bình</b>	10,397	0,387	0,034	0,035	-	3,843	0,171	4,468	86,434
<b>RSD (%)</b>	8,424	68,529	82,730	44,931	-	35,080	79,935	36,400	6,185
<b>Thanh Hóa (n=5)</b>									
<b>Giá trị thấp nhất</b>	14,067	0,120	-	0,059	-	1,410	0,007	1,473	90,874
<b>Giá trị lớn nhất</b>	29,983	0,310	-	0,105	-	2,920	0,120	3,041	96,145
<b>Trung bình</b>	22,072	0,203	-	0,086	-	2,156	0,033	2,478	86,942
<b>RSD (%)</b>	26,283	37,785	-	23,806	-	32,237	148,578	31,005	4,472

mục phương pháp nghiên cứu. Tiến hành sắc ký. Kết quả được trình bày ở hình 1 cho thấy: Các chất trên sắc ký đồ của dung dịch dược liệu chuẩn (hình 1c) đều đã được định danh, các pic tách nhau tốt, cân xứng,  $R_s$  giữa pic CM và pic CALC là 3,487; giữa pic CALC và pic CA là 4,732 (đều lớn hơn 1,5). Hệ số đuôi (Tailing factor) của pic CA nhỏ hơn 2,0 (hình 1a) và RSD diện tích pic CA khi tiêm lặp lại 6 lần dung dịch CA chuẩn là 1,84 % (< 2%). Thời gian lưu tương đối của pic CAL và CA trong dung dịch đối chiếu 1 so với pic tương ứng với CA là 1,3 và 1,0 (hình 1b).

Hình 1. Kết quả quá trình lựa chọn các nghiên cứu theo PRISMA

Trên sắc ký đồ 1c, xuất hiện các pic có RRT so với CA lần lượt là 0,8 – 0,9 – 1,0 – 1,3 – 1,5 so sánh với các giá trị bảng 2 xác nhận tương ứng với các chất coumarin (CM), cinnamylalcohol (CALC), acid cinnamic (CA), cinnamadehyd (CAL) và 2-methoxy cinnamaldehyd (MCAL).

*Xác định hàm lượng các chất nghiên cứu*

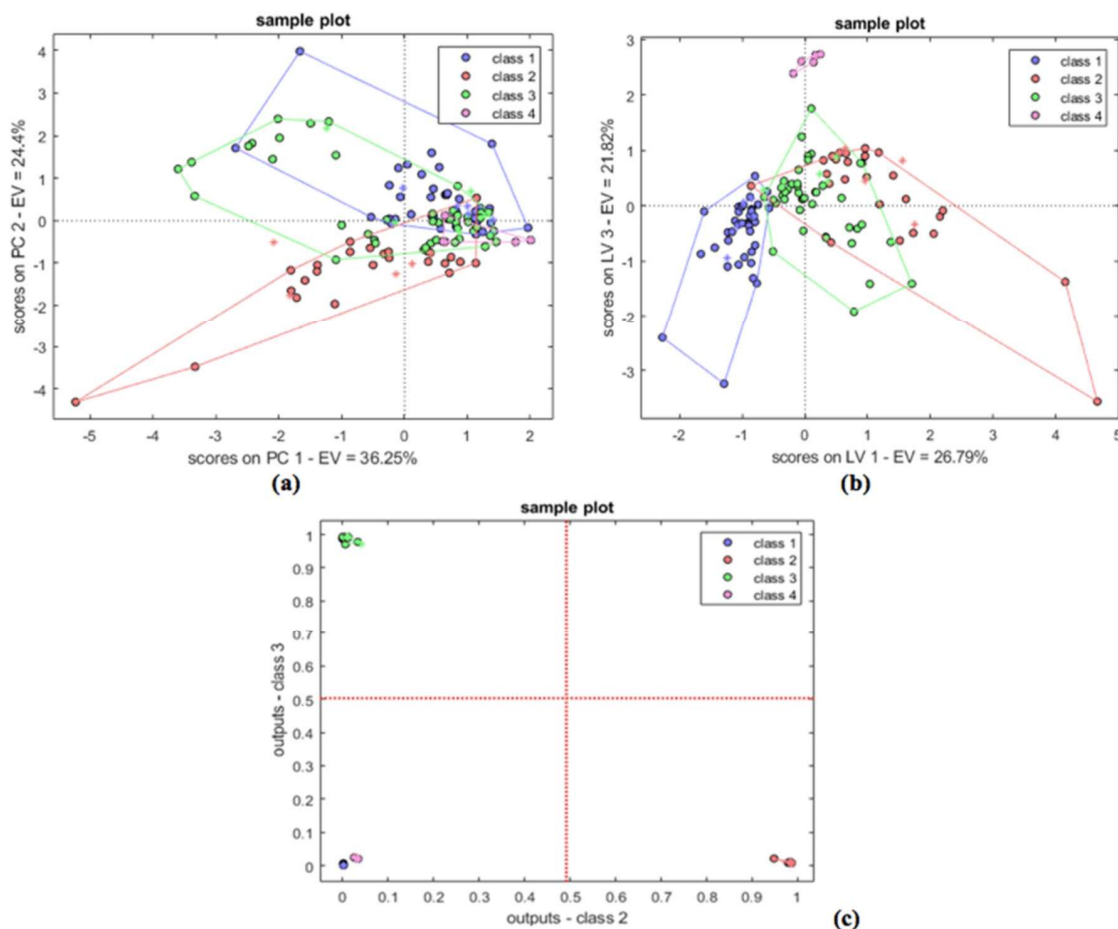
Các mẫu quế thu thập ở 4 vùng Quảng Ninh, Quảng Nam, Yên Bái, Thanh Hóa. Phân tích mẫu theo phương pháp phân tích đã trình bày. Kết quả định lượng các mẫu quế được trình bày trong bảng 3, sắc ký đồ mẫu quế đại diện các vùng trình bày ở hình 2.





Bảng 4. Kết quả phân loại mẫu kiểm tra theo 3 mô hình

Mô hình	Nhóm dự đoán	Quảng Ninh	Quảng Nam	Yên Bái	Thanh Hóa	Không xác định	Số mẫu phân loại sai	Tỷ lệ mẫu được phân loại đúng (%)
PCA - DA	Quảng Ninh	5	0	0	0	0	1	95
	Quảng Nam	0	5	0	0	0		
	Yên Bái	0	1(false)	4	0	0		
	Thanh Hóa	0	0	0	5	0		
PLS - DA	Quảng Ninh	5	0	0	0	0	3	85
	Quảng Nam	0	5	0	0	0		
	Yên Bái	0	0	3	0	2(false)		
	Thanh Hóa	0	0	0	4	1(false)		
BNN	Quảng Ninh	5	0	0	0	0	0	100
	Quảng Nam	0	5	0	0	0		
	Yên Bái	0	0	5	0	0		
	Thanh Hóa	0	0	0	5	0		



Hình 3. Biểu đồ phân loại theo 3 mô hình

(a). PCA-DA (b). PLS-DA (c). BNN

Chú thích: Class 1: Quảng Ninh - Class 2: Quảng Nam - Class 3: Yên Bái - Class 4: Thanh hoá

• Tập luyện \* Tập kiểm tra



### **Bước đầu phân loại quế theo vùng trồng bằng phương pháp thống kê đa biến.**

Sử dụng hàm lượng 5 chất được định danh của 130 mẫu khảo sát làm biến đầu vào cho mô hình PCA-DA, PLS – DA và BNN để phân loại quế theo vùng trồng. Kết quả thể hiện trong bảng 4 và hình 3.

#### *Phân loại bằng mô hình PCA-DA*

Kết quả từ bảng 4 cho thấy trong tập kiểm tra, toàn bộ 5 mẫu quế của tỉnh Quảng Ninh, Quảng Nam và Thanh Hóa đều được phân loại vào đúng nhóm. Tuy nhiên có 1 mẫu của tỉnh Yên Bái được phân loại vào tỉnh Quảng Nam. Như vậy có 95 % số mẫu được phân vào đúng nhóm và 5 % mẫu phân loại sai với độ chính xác 95 % Trong hình 3a biểu đồ biểu diễn sự phân bố của 110 mẫu tập luyện và 20 mẫu tập kiểm tra cho thấy, các mẫu có độ chụm không cao và khó phân biệt.

#### *Phân loại bằng mô hình PLS – DA*

Kết quả từ bảng 4 cho thấy toàn bộ 5 mẫu quế của tỉnh Quảng Ninh và Quảng Nam được phân loại vào đúng nhóm. Tuy nhiên có 2 mẫu của tỉnh Yên Bái và 1 mẫu của tỉnh Thanh Hóa không được phân loại. Như vậy có 85 % số mẫu được phân vào đúng nhóm và 15 % mẫu phân loại sai với độ chính xác 100 %. Trong hình 3b biểu đồ biểu diễn sự phân bố của 110 mẫu tập luyện và 20 mẫu tập kiểm tra, trừ các mẫu Thanh Hóa được phân loại rõ ràng, mẫu quế của 3 tỉnh còn lại khó phân biệt.

#### *Phân loại bằng mô hình BNN*

Kết quả từ bảng 4 cho thấy toàn bộ 5 mẫu của cả 4 tỉnh đã được phân loại vào đúng nhóm. Mô hình đã phân loại được đúng 100 % số mẫu quế của tập kiểm tra với độ chính xác 100 %. Trong hình 3c, các mẫu của cả tập luyện và tập kiểm tra đều được phân loại rất rõ ràng, không có đan xen mẫu các vùng vào nhau.

### **Bàn luận**

Các mẫu quế của 4 địa phương đã định

đanh được 05 trong tổng số 06 chất theo dược điển Mỹ, không có mẫu nào xác định được hàm lượng của MCA. Ở các mẫu đều cho kết quả hàm lượng tổng PNP chiếm khoảng 3 % so với dược liệu khô kiệt, trong đó hàm lượng CAL cao nhất chiếm hơn 80 % tổng lượng PNP. Kết quả này tương tự kết quả nghiên cứu về phân biệt các bộ phận của *Cinnamomum cassia* sử dụng LC-qTOF-MS với quế có nguồn gốc từ Trung Quốc và Đài Loan, CAL là chất có hàm lượng cao nhất với khoảng 2,6 % khối lượng quế, chiếm hơn 90 % tổng hàm lượng PNP [6]. Trong một nghiên cứu khác của nhóm nghiên cứu Su Yang Jeong [10] với mẫu vỏ quế và quế cành có nguồn gốc từ Trung Quốc, Việt Nam, Indonesia và Sri Lanka, CAL là chất có hàm lượng cao nhất phân tích được trong cả mẫu vỏ quế và quế cành. Đối với mẫu vỏ quế, hàm lượng CAL khoảng 3,21 % còn trong mẫu quế cành là 1,75 % so với dược liệu khô kiệt.

Mô hình thống kê đa biến đã được sử dụng để phân loại trong nhiều nghiên cứu cả trong nước và nước ngoài. Yan Dinh sử dụng để phân loại mẫu quế ở 3 nước gồm Trung Quốc, Việt Nam, Indonesia [5], nhóm của Chao Li sử dụng để phân loại Thông quan đằng [9], Y. Yücel, P. Sultanoglu phân loại mật ong [12]. Trần Thị Huế đã sử dụng mô hình PCA-DA, PLS-DA và ANN để phân loại chè xanh thu hái ở 6 tỉnh Thái Nguyên, Lâm Đồng, Hà Giang, Yên Bái, Tuyên Quang, Hòa Bình [2]. Tương tự, nhóm tác giả Nguyễn Thị Hà Ly phân loại Hà thủ ô đỏ [3]. Qua những nghiên cứu này có thể nhận thấy với nhóm mẫu có đặc tính rõ rệt thì tương đối dễ phân loại. Nhưng với những mẫu không thực sự rõ ràng thì cần cỡ mẫu lớn hoặc đưa thêm các chỉ tiêu khác để tăng độ chính xác cho mô hình như trong nghiên cứu phân loại trà xanh, nhóm nghiên cứu đã phải đưa thêm dữ liệu phổ IR, dữ liệu phổ UV-Vis toàn dải cùng khi dữ liệu phân tích HPLC phân loại không chính xác với tất cả các





mô hình [2]. Ngoài ra, lựa chọn mô hình phân loại cũng ảnh hưởng tới kết quả. Trong các mô hình thường được lựa chọn thì mô hình PLS-DA thường cho độ chính xác thấp nhất, mô hình mạng neuron nhân tạo cho kết quả phân loại chính xác nhất.

Để phân loại quế theo vùng trồng, nghiên cứu này đã sử dụng 05 biến đầu vào là hàm lượng 05 thành phần trong vỏ quế. Với lượng mẫu khá lớn (130 mẫu), nhưng thuộc tính đặc trưng lại không nhiều. Hàm lượng 05 chất định lượng được không có sự khác biệt rõ ràng ở các vùng, không có sự khu trú nên việc phân loại sẽ khó khăn. Do đó, nghiên cứu này sử dụng mô hình thống kê PCA-DA, PLS-DA và BNN là những mô hình đã có nhiều nghiên cứu lựa chọn với khả năng phân loại tốt trong phân tích thành phần chính. Kết quả khi sử dụng mô hình PCA-DA, kết quả khảo sát trên 20 mẫu tập kiểm tra phân loại đúng 95 % với độ chính xác là 95 %, nhưng biểu đồ phân loại 110 mẫu quế tập luyện cùng với 20 mẫu kiểm tra lại không khả quan, các mẫu có độ chụm không cao, không thể phân loại được mẫu của 4 tỉnh. Mô hình PLS-DA cho kết quả trên cả tập kiểm tra và tất cả các mẫu đều không tốt, phân loại đúng 85 % mẫu kiểm tra, biểu đồ phân bố không rõ ràng dù độ tin cậy đạt 100 %. Mô hình BNN cho kết quả phân loại chính xác nhất trên tất cả các mẫu cũng như trên tập kiểm tra, với độ tin cậy 100 %. Biểu đồ biểu diễn 110 mẫu tập luyện được phân loại rõ ràng thành 4 vùng, các mẫu kiểm tra cũng

được xác định chỉ khu trú trong 4 vùng đó, không có đan xen mẫu các vùng vào nhau. Mô hình phi tuyến tính BNN với nhiều thuật toán kết hợp cho kết quả phân loại vùng trồng được liệu chính xác hơn các mô hình tuyến tính PCA-DA và PLS-DA.

### Kết luận

Nghiên cứu đã định danh được các thành phần chính có trong mẫu vỏ quế thu hái ở 4 địa phương gồm: coumarin, cinnamyl alcohol, acid cinnamic, cinnamaldehyd và 2-methoxycinnamaldehyd với độ tích hợp hệ thống và độ phân giải đạt yêu cầu. Đồng thời, định lượng được các thành phần đã định danh trong 130 mẫu vỏ quế thu hái ở 4 tỉnh Quảng Ninh, Quảng Nam, Yên Bái và Thanh Hóa bằng HPLC. Với dữ liệu hàm lượng các thành phần chính trong mẫu, nghiên cứu bước đầu đã phân loại được quế theo vùng trồng sử dụng phương pháp toán hoá thống kê đa biến. Trong đó mô hình BNN cho kết quả phân loại đúng tất cả các mẫu ở bộ kiểm tra và có độ chính xác 100 %, biểu đồ phân bố rõ ràng không có sự đan xen mẫu giữa vào các vùng khác nhau. Kết quả của nghiên cứu này có thể là cơ sở khoa học góp phần xây dựng tiêu chuẩn cơ sở cũng như phân loại quế theo vùng trồng. Bên cạnh đó cần tăng quy mô lấy mẫu ở các vùng khác đồng thời đưa thêm dữ liệu phổ UV, IR làm biến đầu vào nhằm nhằm xây dựng cơ sở dữ liệu chặt chẽ để phát triển bộ công cụ kiểm soát chất lượng và nguồn gốc quế ở Việt Nam.

### TÀI LIỆU THAM KHẢO

1. Bộ Y tế (2017), *Dược điển Việt Nam V*, 1296.
2. Trần Thị Huế, Bùi Đức Thọ, Nguyễn Văn Ri, Tạ Thị Thảo, (2019), "Phân loại nguồn gốc địa lí và phân tích đồng thời các methylxanthine trong chè Việt Nam sử dụng phổ hồng ngoại phản xạ kết hợp với thống kê đa biến", *Tạp chí Khoa học và công nghệ*. 6, tr. 33-39.
3. Nguyen Thi Ha Ly, Ta Thi Thao, Phuong Thien Thuong (2019), "Quality Evaluation of *Fallopia multiflora* in Vietnam Based on HPLC/FL and Chemometrics", *Tạp chí Dược liệu*. 3, tr. 131-138.



4. Chunling Zhang et al. (2019), "*Cinnamomum cassia* Presl: A Review of Its Traditional Uses, Phytochemistry, Pharmacology and Toxicology", *Molecules* 24(19), pp. 3473.
5. Ding Yan, Wu En Qi et al. (2011), "Discrimination of cinnamon bark and cinnamon twig samples sourced from various countries using HPLC-based fingerprint analysis", *Food chemistry*. 127(2), pp. 755 - 760.
6. Pei-Yi Chen, Jhe-Wei Yu, Fen-Ling Lu, Mei-Chih Lin, Hwei-Fang Cheng (2016), "Differentiating parts of *Cinnamomum cassia* using LC-qTOF-MS in conjunction with principal component analysis", *Biomedical Chromatography* 30(9), pp. 1449 - 1457.
7. The Government of the Hong Kong Special Administrative Region of the People's Republic of China (2009), *Hong Kong Chinese Materia Medica Standards*, Vol. 5, 6.
8. Chinese Pharmacopoeia Commission (2015), *Pharmacopoeia of the people's republic of china*, Vol. IA.
9. Chao Li, S. C. Yang, Q. S. Guo, K. Y. Zheng, Y. F. Shi, X. F. Xiao, G. Q. Long (2014), "Determining the geographical origin of the medicinal plant *Marsdenia tenacissima* with multi-element analysis and data mining techniques", *Chemometrics and Intelligent Laboratory Systems*. 136, pp. 115 - 120.
10. Dong Cheul Moon Su Yang Jeong, Jong Seong Kang, Je Hyun Lee, Byung Sun Min, Jong Keun Son, Mi Hee Woo, (2013), "Quantitative Analysis of Bioactive Marker Compounds from *Cinnamomi Ramulus* and *Cinnamomi Cortex* by HPLC-UV", *Natural Product Science*. 19(1).
11. USP (2022), Current DocID: GUID-8940E36B-6037-4043 826C7251B82ACCFB \_1\_en-US; GUID-2D357928-6650-4556-8400-13676D7B7446 \_1\_en-US .
12. Pinar Sultanog lu Yasin Yucel (2013), "Characterization of Hatay honeys according to their multi-element analysis using ICP-OES combined with chemometrics", *Food Chemistry*. 140.
13. Shi-Lei Geng, Zhao-Xue Cui, Bin Shu, Sheng Zhao, Xin-Hua Yu, (2012), "Histochemistry and Cell Wall Specialization of Oil Cells related to the Essential Oil Accumulation in the Bark of *Cinnamomum cassia* Presl. (Lauraceae)", *Plant Production Science*. 15(1), pp. 1 - 09.