



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

약학박사학위논문

크로마토그래프지문과  
상대 정량을 이용한  
천연물 품질 평가 연구

A Study for Quality Assessment of Herbal  
Materials Using Chromatographic Fingerprints  
and Relative Quantification

2017년 8월

서울대학교 대학원  
약학과 약품분석학전공  
이 광 진

# 크로마토그래피지문과 상대 정량을 이용한 천연물 품질 평가 연구

A Study for Quality Assessment of Herbal Materials  
Using Chromatographic Fingerprints and Relative  
Quantification

지도교수 박 정 일

이 논문을 약학박사 학위논문으로 제출함

2017년 8월

서울대학교 대학원

약학과 약품분석학전공

이 광 진

이광진의 박사학위논문을 인준함

2017년 8월

위 원 장	<u>                    </u>	권 성 원	(인)
부 위 원 장	<u>                    </u>	양 현 옥	(인)
위 원	<u>                    </u>	박 정 일	(인)
위 원	<u>                    </u>	임 요 한	(인)
위 원	<u>                    </u>	이 성 입	(인)



## 국문초록

천연물은 다양한 활성성분을 포함하고 있고 생물다양성으로 인해 원료마다 그 조성의 차이가 크므로 정확하게 원료의 품질을 평가하는 방법론의 개발이 중요시되고 있다. 현재 이를 위해 다성분을 종합적으로 검토하여 평가하는 크로마토그래프지문법과 소수 지표성분을 정량 평가하는 방법이 같이 사용되고 있다. 본 연구에서는 각 방식에서 보편적으로 사용되는 방법들을 개선하여 엄격하면서 효율적으로 품질을 평가할 수 있는 방법들을 제안한다. 먼저 크로마토그래프지문을 이용한 기존의 품질 평가법을 보완하기 위해 복합 유사성 지표 체계(CAM)를 개발하였다. 다양한 관점에서 지문 간의 유사성을 확인할 수 있었고 단일 유사성 지표를 활용하였을 때 나타나는 단점이 보완되었다. 위 체계는 백출 시료 40종에 적용하였고 품질 기준 설정에는 커널밀도추정을 활용하였다. 그 결과, 기준에 의해 품질이 보장된 원료 22종은 기존 방법에 의해 선택된 원료 22종에 비해 상정한 기준 시료와 더 화학적으로 동등했다. 둘째로, 표준물질 확보의 어려움을 극복하고 효율적으로 성분 함량을 평가하기 위해 증기화광산란검출기를 이용한 상대 정량법을 개발하였다. HPLC와 결합하였을 때 검출감도에 영향을 미치는 인자들을 조사하였고 이에 따른 이동상 조건을 최적화하여 다양한 성분의 검출감도를 평준화 할 수 있었다. 인삼 및 가공인삼제품의 주요활성성분인 Ginsenoside 9종 분석에 활용한 결과, 최적 조건에서 직접 정량법 대비 최대 차이는 8.26%, 전체 함량 차이는 0.91%에 불과했다.

**주요어** : 천연물 원료, 크로마토그래프지문, 상대 정량, 화학적 동등성, 품질 평가

**학 번** : 2012-21606

# 목차

국문초록 .....	i
목차 .....	ii
List of Tables .....	iv
List of Figures .....	v
List of Abbreviations .....	vii
I. 서론 .....	1
1. 크로마토그래프지문을 이용한 품질 평가법 .....	1
2. 상대정량을 이용한 품질 평가법 .....	5
II. 실험방법 .....	7
1. 크로마토그래프지문을 이용한 품질 평가법 .....	7
1.1. 시료, 시약 .....	7
1.2. 분석기기 및 조건 .....	7
1.3. 반응표면분석법을 이용한 추출법 최적화 .....	8
1.4. 크로마토그래프지문 획득 .....	8
1.5. 복합 유사성 지표(CAM) 계산 .....	9
1.6. 커널밀도추정을 이용한 이상치 검출 .....	9
2. 상대정량을 이용한 품질 평가법 .....	10
2.1. 시료, 시약 .....	10

2.2. 분석기기 및 조건 .....	11
2.3. 이동상 조성이 검출 감도에 미치는 영향 .....	11
2.4. Peak bandwidth 가 검출 감도에 미치는 영향 .....	12
2.5. 분석법 최적화 및 유효성 검증 .....	12
2.6. 직접 정량법과의 비교 평가 .....	13
III. 결과 및 고찰 .....	14
1. 크로마토그래프지문을 이용한 품질 평가법 .....	14
1.1. 크로마토그래프지문 획득 .....	14
1.2. 복합 유사성 지표(CAM) 계산 .....	23
1.3. 커널밀도추정을 이용한 이상치 검출 .....	27
1.4. 기존 품질 평가법과의 비교 .....	36
2. 상대정량을 이용한 품질 평가법 .....	39
2.1. 이동상 조성이 검출 감도에 미치는 영향 .....	39
2.2. Peak bandwidth 가 검출 감도에 미치는 영향 .....	43
2.3. 순차적 Gradient 변경을 통한 분석 조건 최적화 .....	46
2.4. 상대 정량법의 유효성 검증 .....	48
2.5. 직접 정량법과의 비교 .....	53
IV. 결론 .....	55
V. 참고문헌 .....	56

## List of Tables

- Table 1.** Design matrix and responses for the face-centered central composite design (FCCCD)
- Table 2.** Analysis of variance (ANOVA) table for the response surface quadratic model
- Table 3.** Validation results of the HPLC method
- Table 4.** The CAM values of the 40 samples
- Table 5.** Regression data, LOD, LOQ and recovery for the nine ginsenosides
- Table 6.** Intra- and inter-day precision of the direct and group quantification method for the nine ginsenosides (n=3)
- Table 7.** Individual and global uncertainty of the direct and group quantification method for the nine ginsenosides (calculated at quantification result)
- Table 8.** Comparison of contents for the nine ginsenosides between the direct and group quantification method (n=5)



## List of Figures

- Figure 1.** The limitations of the angular-based similarity indices revealed through virtual data vectors.
- Figure 2.** Response surface plots showing reciprocal interaction of two individual factors on the total response.
- Figure 3.** Representative fingerprints of the mixed standards, the *A. japonica* samples, and the *A. macrocephala* samples.
- Figure 4.** Selective elimination of the response of the unknown compound with the reference wavelength and 3D spectra in the corresponding area.
- Figure 5.** The elements of the three data vectors, KJ8, CJ10, and KJ11, having a high  $c$  value over 0.97 and their CAM similarity values.
- Figure 6.** The limitations of the congruence coefficient found in evaluating the similarity between the actual samples and the reference.
- Figure 7.** The distribution of the  $c$  (the congruence coefficient) and  $a$  (the average of peak area ratios) values for the 40 *Atractylodes* samples.
- Figure 8.** The distribution of the  $c$  (the congruence coefficient) and  $a$  (the average of peak area ratios) values for the 31 *A. Japonica* samples.
- Figure 9.** The distribution of the CAM values for the 31 *A. Japonica* samples and the three levels of the kernel density estimation with plug-in bandwidth matrix.
- Figure 10.** The distribution of the CAM values for the 31 *A. Japonica* samples and the two levels of the kernel density estimation with plug-in bandwidth matrix.

This figure shows the views from the lateral side of Figure 9 to understand the surface of the areas exactly.

**Figure 11.** The two-dimensional contour plots of the estimated kernel density at the level of 70%.

**Figure 12.** Comparative analysis of the samples qualified by the proposed approach and the conventional approach.

**Figure 13.** Structures and molecular weights of the selected ginsenosides.

**Figure 14.** ELSD response curves for the nine ginsenosides in flow injection analysis.

**Figure 15.** Evaluation of elution bandwidth effect on the ELSD response and stepwise changes of the gradient condition for the response normalization in the protopanaxadiol (PPD) group.

**Figure 16.** Evaluation of elution bandwidth effect on the ELSD response and stepwise changes of the gradient condition for the response normalization in the less-polar (LP) group.

**Figure 17.** HPLC-ELSD chromatograms of mixed standards by the initial condition, mixed standards by the final condition, and processed ginseng concentrate by the final condition.

## List of Abbreviations

WHO: World Health Organization  
FDA: Food and Drug Administration  
CFDA: China Food and Drug Administration  
EMA: European Medicines Agency  
MFDS: Ministry of Food and Drug Safety  
IR: Infrared spectroscopy  
NMR: Nuclear Magnetic Resonance  
MS: Mass Spectrometry  
HPTLC: High Pressure Thin Layer Chromatography  
HPLC: High Pressure Liquid Chromatography  
GC: Gas Chromatography  
DAD: Diode Array Detector  
UPLC: Ultra Pressure Liquid Chromatography  
ELSD: Evaporative Light Scattering Detector  
FCCCD: Face-Centered Central Composite Design  
KJ: *A. japonica* samples from Korea  
CJ: *A. japonica* samples from China  
CM: *A. macrocephala* samples from China  
PPT: Protopanaxatriol ginsenosides  
PPD: Protopanaxadiol ginsenosides  
LP: Less-Polar ginsenosides

# I. 서론

## 1. 크로마토그래프지문을 이용한 품질 평가법

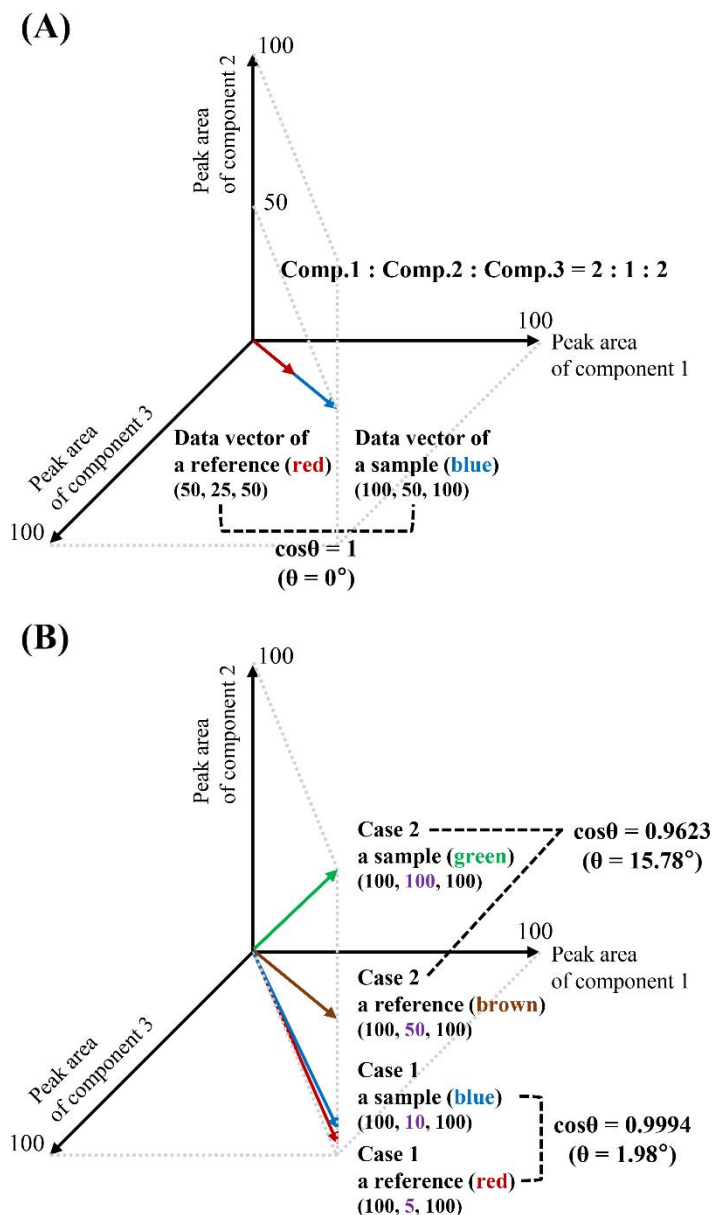
기원중, 재배 지역과 환경, 전처리 방법, 보관 방법 같은 변수들은 천연물 원료가 개체에 따라 구성 성분 함량 차이가 나도록 하는 원인들이고 결국 이 성분들의 복합적인 약리작용을 규명하여 우수한 신제품을 개발하는데 방해가 된다. 또한, 약효와 안전성이 입증되는 동일한 제품의 지속적인 생산을 어렵게 한다 [1]. 이러한 상황은 기준 원료와 화학적으로 동등한 원료를 선택할 수 있도록 하는 엄격한 방법의 필요성을 부각시킨다.

다성분의 표준물질을 구하기가 어렵고 정확하게 분석하는 것이 쉽지 않기 때문에 소수의 지표성분의 함량만을 평가함으로써 원료의 품질을 결정하는 것이 그 동안의 일반적인 품질 평가 방법이었다 [2]. 이와 달리, 크로마토그래프지문 비교법은 표준물질의 필요 없이 다성분의 화학적 동등성을 평가할 수 있기 때문에 새로운 품질 평가의 대안으로 제시되었다. 그런 이유로 WHO는 크로마토그래프지문 비교법을 천연물 원료의 품질을 평가하는 공식적인 방법으로 인정하였고 잇따라 미국 FDA, 중국 CFDA, 유럽 EMA, 국내 식약처(MFDS)에서 역시 이를 공식적인 방법으로 채택하였다 [3-5].

천연물 원료의 지문을 얻기 위해 그 동안 IR, NMR, MS, HPTLC, HPLC, GC같은 다양한 기기분석 방법들이 이용되어 왔다 [6-15]. 연구자들은 때때로 더 많은 정보를 함유하는 지문을 얻기 위해 DAD를 이용하여 3차원 지문을 얻거나 더 신속하게 지문을 얻기 위해 UPLC를 이용하기도 하였다 [16-17]. 얻은 지문은 다차원 데이터 벡터로 변환된다. 데이터 벡터의 각 요소로 특정 파수,

화학적 이동, 단위전하 당 질량 값, 머무름 시간에서의 피크의 높이나 면적을 사용한다. 두 개의 데이터 벡터를 비교함으로써 두 개의 천연물 원료 지문 사이의 유사성을 평가할 수 있고 따라서 데이터 처리와 비교 방법은 품질 평가의 정확도를 높이기 위해 중요하다 [18].

두 데이터 벡터 사이의 유사성 지표를 계산하여 두 지문의 유사성을 비교하는 방법은 원료의 품질을 결정할 때 직관적인 기준을 이용할 수 있고 (예; 지표 값이 0.9 이상인 것이 합격) 상대적으로 사용하기 쉽기 때문에 데이터 비교 방법 중 가장 선호된다 [19-20]. 몇 가지 유사성 지표들 중 Congruence coefficient와 Pearson's correlation coefficient가 가장 보편적으로 사용된다 [21-23]. 두 지표의 공통적인 특징은 두 데이터 벡터 사이의 각도 코사인 값을 유사성의 척도로 이용한다는 것이다. 이것은 두 데이터 벡터 사이의 각도가 작을수록 두 데이터 벡터가 유사하다는 것을 의미한다. 하지만, 이 두 지표의 한계점이 있다. 첫째로, 두 지표는 두 벡터의 방향이 같으면 그 크기에 상관없이 두 벡터를 동등한 것으로 평가한다는 것이다. 예를 들어, 두 지표는 크로마토그램에서 전반적인 농도 차이는 배제하고 패턴만 유사하면 동등한 것으로 판단한다는 것이다. 둘째로, 두 지표의 값은 주로 큰 값(큰 피크들의 면적)에 의해 결정된다는 것이다. 때때로 데이터 내에서 작은 값이 생리활성에 큰 영향을 미칠 수도 있음에도 불구하고 이들의 차이는 유사성 평가에 크게 반영되지 않는다 (Figure 1). 종합하면 가장 보편적으로 사용되는 이 두 가지 지표는 두 지문의 화학적 동등성을 정확히 표현할 수가 없고 이를 활용한 품질 평가는 엄격한 품질 평가를 보증하지 못한다는 것이다. 따라서 이 연구의 우선적인 목표는 위 지표들의 단점을 보완한 새로운 유사성 지표 평가 체계를 형성하는 것이었다.



**Figure 1.** The limitations of the angular-based similarity indices revealed through virtual data vectors. (A) The two vectors, whose composition ratios of the three components are two to one to two, are regarded as equivalents irrespective of their absolute values. (B) The two-fold difference at the concentration of the component 2 in Case 1 has little effect on the angle of the two vectors, but the other two-fold difference in Case 2 clearly influences.

품질 판정을 위한 기준 설정 연구 역시 수행하였다. 대부분의 연구에서 기준은 통계학적 접근 없이 연구자의 경험에 의해서 설정되었다 [24-26]. 일부의 경우에는 관리도(control chart)를 이용하였다 [20]. 이는 기준 원료와 기 합격 판정된 다수 원료의 유사성 지표 값의 정규분포에 따라 도출해 낸 경계하한선과 관리하한선을 이용하여 새로운 원료의 품질을 판정하는 것이다. 하지만 이는 이미 품질 평가에 대한 연구가 상당히 진행되어 선행 연구 데이터가 많은 경우에만 활용할 수 있는 한계가 있다. 실제로, 연구나 생산 초기부터 원료의 품질을 엄격하게 관리하는 것은 아주 바람직하나 이 상황에서는 선행 연구 결과가 없어 품질 기준을 설정하기가 어렵다. Goodzari 등은 유사성 평가의 중요한 이슈로 이 쟁점을 이미 지목한 바 있다 [19]. 따라서, 본 연구의 두번째 목표는 연구나 생산 초기와 같이 선행 품질 연구 결과가 없는 상황에서 적절한 통계학적 방법을 이용해 원료들 중 동등하지 않다고 판단되는 것을 분류할 수 있는 체계를 확립하는 것이었다.

본 연구의 대상 원료로 *Atractylodes*속의 두 가지 식물종인 *Atractylodes japonica*와 *Atractylodes macrocephala*를 이용하였다. 현재 두 천연물은 대한약전에서 백출로 정의되어 있고 소화불량, 식욕부진과 같은 소화기계 질환에 빈번하게 사용되고 있다 [27-28]. 이 두 천연물은 비슷한 외관과 치료 목적으로 동일하게 간주되고 있으나 화학적 조성(구성 성분)이 상이하여 상세한 논의가 필요하다. 따라서 본 연구를 통해 확립한 품질 평가 체계를 검증하기 위한 대상 원료로 선정하였다.

## 2. 상대정량을 이용한 품질 평가법

소비자들과 관련 규제 기관에서는 천연물을 포함하는 제품의 안전성과 유효성에 관련하여 제품 안에 들어 있는 다성분의 함량에 대한 더 많은 정보를 제공하기를 원하고 있고 중요 천연물 원료의 경우에는 이미 법적으로 규제되고 있다. 반면, 업체의 입장에서는 지속적인 정보 제공을 위해 성분의 정량 분석에 필요한 고비용의 개별 표준물질들을 확보해야 한다. 여기서 경제적이고 효율적인 분석을 위해 상대 정량의 개념이 대두된다. 상대 정량은 소수의 표준물질로 다양한 성분을 정량하는 것을 의미한다. 다양한 물질들의 다른 화학적 성질로 인해 같은 양을 분석했을 때 검출 감도(피크 면적)가 다양하게 나오는데 특수한 장치나 방법을 이용해 평준화 시킴으로써 상대 정량을 가능하게 할 수 있다.

많은 연구자들이 증기화광산란검출기(ELSD)가 상대 정량에 적합하다고 보고했다. 증기화광산란검출기는 물질의 화학적 특성이 아닌 물질의 양에 비례한 산란광의 정도에 따라 그 검출감도가 결정되므로 다양한 종류의 물질들을 분석했을 때 검출감도(피크 면적)가 크게 차이 나지 않았기 때문이다 [29-31]. 또한, 작동이 쉽고 저비용이며 다양한 HPLC 조건에 호환이 잘 되는 것이 그동안 증기화광산란검출기를 선호해 온 큰 이유였다 [32]. 그러나, 증기화광산란검출기의 감도가 HPLC 이동상 조건에 크게 영향을 받는다는 사실이 보고되면서 무분별하게 상대 정량을 진행해서는 안 된다는 것이 밝혀졌다 [33].

최근 상대 정량법에 대한 몇몇 연구들이 보고되었다. 먼저 3차원의 검량 표면 곡선(Three-dimensional calibration surface plot)을 이용한 방법이 개발되었다 [34]. 고정상과 상호작용을 하지 않는 분석체를 분석 시간 동안 주기적으로 주입하고 검출 감도를



관찰한다. 마찬가지로 농도를 변화시키며 같은 작업을 수행한다. 그 결과 3차원의 검량 표면 곡선을 얻을 수 있고 실제 원료를 분석할 때에 곡선으로부터 도출되는 검출 감도 보정 상수(Response factor)를 이용, 상대 정량을 가능케 하였다. 다른 방법으로는 이차 펌프(Secondary pump)를 설치함으로써 증기화광산란검출기로 들어가는 이동상 조성을 동일하게 유지시키는 이동상 조성 유지법(Mobile phase compensation method)이 개발되었다 [35]. 이 방법에서 이차 펌프는 고정상 뒤와 검출기 앞에 설치되며 기본 펌프가 흘려주는 이동상 조성의 역으로 이동상을 흘려주어 검출기 안에서의 이동상 조성이 동일하도록 만든다. 이를 통해 검출 감도 평준화 및 상대 정량을 가능케 하였다. 가장 최근에는 이동상 Gradient 대신에 온도 Gradient와 유속 Gradient를 결합하여 검출 감도 평준화를 시도한 방법도 개발되었다 [36]. 하지만, 이러한 방법들은 아직 일선에서 쉽게 사용하기에는 불편한 점들이 있다. 예를 들어, 많은 실험을 통해 얻어진 표면 곡선이나 HPLC 시스템 내에 큰 변화를 주어야 한다. 따라서 본 연구에서는 현장에서 좀 더 쉽게 활용할 수 있도록 기존 HPLC 분석에서 큰 변화를 주지 않은 채로 상대 정량을 가능케 하는 방법을 개발하고자 하였다.

본 연구의 대상 원료로 가공 인삼(Processed ginseng)을 이용하였다. 인삼 제품 시장은 천연물 제품 시장 중 가장 큰 규모를 가지며 그 만큼 제품 정보에 대한 소비자의 지식 수준 역시 높다. 특히 가공 인삼을 만들 때에 인삼 고유 Ginsenoside의 구조 변화가 일어나 새로운 Ginsenoside가 형성되며 이 물질들의 생리 활성은 다양하고 강한 것으로 알려져 있다 [37-40]. Ginsenoside는 구조 특성상 증기화광산란검출기로 분석하기 좋으며 가공 인삼은 관련 제품의 시장 규모를 고려하였을 때 상대 정량법의 개발이 필요하므로 본 연구의 대상 원료로 선정하였다.

## II. 실험방법

### 1. 크로마토그래프지문을 이용한 품질 평가법

#### 1.1. 시료, 시약

국내 시장에서 유통되는 백출 시료 1) 국산 *Atractylodes japonica* (KJ group) 20종; KJ1-20, 2) 중국산 *Atractylodes japonica* (CJ group) 11종; CJ1-11, 3) 중국산 *Atractylodes macrocephala* (CM group) 9종; CM1-9을 구입하였다. 분석법 밸리데이션을 위한 표준물질 atractylenolide I, II, III는 Aktin Chemical (Chengdu, China)에서 구입하였다. HPLC 등급 물, 메탄올, 아세토니트릴은 J.T. Baker (Phillipsburg, NJ, USA)에서 구입하였다.

#### 1.2. 분석기기 및 조건

Agilent (Santa Clara, CA, USA)사의 1260 Infinity HPLC system과 diode array detector를 시료 분석에 사용하였다. 고정상은 Agilent사의 Polaris 5 Amide-C<sub>18</sub> 컬럼 (250 × 4.6 mm, 5 μm)을 사용하였다. 컬럼 온도는 30°C로 유지되었으며 이동상 유속은 1 mL/min로 설정하였다. 이동상 A는 물, 이동상 B는 아세토니트릴을 사용했으며 조건은 다음과 같다. 0-5분 (30-50% B), 5-20분 (50-55.5% B), 20-35분 (55.5-78% B), 35-40분 (78-100% B), 40-45분 (100% B). 시료는 10 μL를 주입하여 분석하였다. 크로마토그래프지문은 파장 236 nm에서 획득하였다. 분석법 밸리데이션을 위한 atractylenolide I의 분석은 278 nm에서 참조 파장 345 nm를 추가하여 진행하였고 atractylenolide II, III의

분석은 220 nm에서 진행하였다.

### 1.3. 반응표면분석법을 이용한 추출법 최적화

Stat-Ease (MN, USA)사의 Design Expert 9 소프트웨어를 통해 추출 조건 최적화를 위한 반응표면분석을 수행하였다. 세 종류의 변인 각각에서 세 가지 수준의 조건이 설정된 FCCCD를 모델링에 사용하였다. 해당하는 독립변인은 다음과 같다. 온도( $X_1$ ; 30, 40, 50°C), 시간( $X_2$ ; 30, 60, 90분), % 메탄올( $X_3$ ; 60, 80, 100% 메탄올). 크로마토그램으로부터 17개의 유의한 피크를 선정하고 이들 피크 면적의 합을 종속변인으로 하였다. 반응표면분석으로 얻은 모델에서 독립변인과 종속변인의 관계를 다항식으로 확인하였고 모델의 통계적 유의성을 ANOVA 분석을 통해 확인하였다. 얻어진 최적 추출 조건의 정확성을 검증하기 위하여 그 조건에서 실제 실험을 3번 진행하여 모델이 예측했던 값과 일치하는지 확인하였다.

### 1.4. 크로마토그래프지문 획득

500 mg의 백출 시료에 10 mL의 81% 메탄올을 넣고 50°C 조건에서 90분 동안 초음파 추출하였다. 원심분리를 통해 상정액을 분리한 후 동일용매 10 mL를 다시 넣고 같은 추출을 진행하였다. 총 3번 추출하여 시료당 30 mL의 상정액을 얻었고 농축기를 이용해 용매를 제거한 후 다시 10 mL의 81% 메탄올에 녹여 분석 시료로 활용하였다. HPLC 분석 전에 0.45  $\mu\text{m}$  membrane filter (ADVANTEC, Tokyo, Japan)로 방해물을 제거하였다.

## 1.5. 복합 유사성 지표(CAM) 계산

A. *Japonica* 시료들의 평균 크로마토그래프지문을 연구 진행을 위한 임시 기준 시료로 상정하였다. 기준 시료와 평가 대상 시료의 유사성을 확인하기 위한 유사성 지표  $c$ ,  $a$ ,  $m$ 은 다음과 같이 계산한다.

$$c = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n (x_i)^2 \sum_{i=1}^n (y_i)^2}} \quad (1)$$

$$a = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{x_i} \quad (2)$$

$$m = \max\left(\max\left(\frac{y_i}{x_i}\right), \frac{1}{\min\left(\frac{y_i}{x_i}\right)}\right) \quad (i = 1, 2, \dots, n) \quad (3)$$

여기서  $x_i$ 와  $y_i$ 는 각각 기준 시료와 평가 대상 시료의 크로마토그래프지문 내  $i$ 번째 피크 면적을 의미하고  $n$ 은 총 피크의 개수를 의미한다. 지표  $c$ 는 0 ~ 1의 값을 가지며 1일 때 두 크로마토그래프지문 내 성분 피크들의 조성비가 동일하다. 지표  $a$ 는 각 성분에서의 면적 상대비(평가 대상 시료/기준 시료)의 평균으로써 값이 1에 가까울수록 두 시료의 전반적인 농도차이가 없어 유사하다. 지표  $m$ 은 성분 농도차이의 최대값(면적 상대비의 최대값과 최소값의 역수 중 큰 값)으로 이를 통해 두 크로마토그래프지문 사이에서 가장 큰 농도차이를 확인할 수 있고 그 값이 1일 때 두 시료는 화학적으로 동등하다.

## 1.6. 커널밀도추정을 이용한 이상치 검출

각 시료의 CAM값( $c$ ,  $a$ ,  $m$ )은 품질 기준 설정과 시료의 품질 판정을 위한 통계 분석에 활용되었다. 아래의 삼차원 커널 밀도

추정을 통해 기준 시료와 동등하다고 판단할 수 있는 시료들을 분류하였다.

$$\hat{f}_H(x) = \frac{1}{n} \sum_{i=1}^n K_H(x - x_i) \quad (4)$$

$x$ 는 CAM값의 벡터( $x = (c, a, m)^T$ ),  $K_H(x)$ 는  $H$ 를 가질 때의 커널 함수를 의미한다.  $n$ 은 시료의 숫자를 의미한다. 이 연구에서 커널 함수  $K$ 는 아래와 같은 일반적인 커널 함수(standard multivariate normal kernel function)를 사용하였다 [41].

$$K_H(x) = (2\pi)^{-3/2} |H|^{-1/2} e^{-\frac{1}{2}x^T H^{-1}x} \quad (5)$$

여기서  $H$ 는 데이터의 공분산 행렬( $\hat{\Sigma}^{-1/2}$ )에 비례하게 결정되었다. R program의 ks package를 통해 이 분석을 수행하였고 평가 대상 시료  $z$ 는 시료의 추정 확률 밀도  $\hat{f}(z)$ 가 선정한 역치  $\eta$ 보다 클 때 합격인 것으로 하였다.

## 2. 상대정량을 이용한 품질 평가법

### 2.1. 시료, 시약

Ginsenoside Rg1, Re, Rb1, Rc, Rb2, Rd, 20(S)-Rg3, Rk1, Rg5는 인삼과 가공인삼에서 분리하였고 그 순도가 98% 이상임을 HPLC-UV-ELSD로 확인하였다 [40]. HPLC 등급의 물, 메탄올, 아세트니트릴은 J.T. Baker (Phillipsburg, NJ, USA)에서 구입하였다. 분석법을 적용할 다양한 주요 Ginsenoside들을 포함하고 있는 인삼 제품 원료 농축물은 Ginseng Science Inc. (Seoul, Korea)에서 제공받았다. 각 Ginsenoside는 4 mg/mL의

농도로 메탄올에 녹여 저장용액으로 보관하였고 매일 목적에 따라 용액끼리 혼합 혹은 메탄올로 희석하여 사용하였다.

## 2.2. 분석기기 및 조건

Waters (Milford, MA, USA)사의 2695 Alliance HPLC system과 Sedere (Virty-sur-Seine, France)사의 Sedex 75 Evaporative Light Scattering Detector (ELSD, 증기화광산란검출기)를 시료 분석에 사용하였다. 고정상은 Phenomenex (Torrance, CA, USA)사의 Gemini C<sub>18</sub> 컬럼 (250 × 4.6 mm, 5 μm)을 사용하였다. 컬럼 온도는 40°C로 유지되었으며 이동상 유속은 1 mL/min로 설정하였다. 이동상 A는 물, 이동상 B는 아세토니트릴을 사용했으며 최종 분석 조건은 다음과 같다. 0-20분 (22% B), 20-25분 (22-32% B), 25-30분 (32% B), 30-45분 (32-40% B), 45-55분 (40-55% B), 55-57분 (55% B), 57-67분 (55-100% B), 67-70분 (100% B). 일관적으로 시료를 주입하기 위하여 Injector에 10 μL 루프를 설치하고 full-loop injection하였다. ELSD 조건은 다음과 같다. Nebulizer gas pressure; 2.5 bar, Drift tube temperature; 50°C, Detector gain 8.

## 2.3. 이동상 조성이 검출 감도에 미치는 영향

Ginsenoside 종류와 이동상 조성에 따라 검출기의 감도 변화를 확인하기 위하여 분석 컬럼 대신 Union을 연결하여 Flow injection을 실시하였다. 이동상 유속은 1 mL/min로 고정하였으며 Ginsenoside 용액은 25, 50, 100, 200, 400 μg/mL의 농도로

준비하였다. 이동상 조성은 4가지 조건(20, 40, 60, 80% 아세토니트릴)에서 확인하였다.

## 2.4. Peak bandwidth가 검출 감도에 미치는 영향

9개의 Ginsenoside는 화학적 구조와 설정한 분석 환경에서의 용리 순서를 고려하여 세 가지 그룹으로 분류되었다; (1) Protopanaxatriol (PPT) 계에 속하는 Rg1, Re, (2) Protopanaxadiol (PPD) 계에 속하는 Rb1, Rc, Rb2, Rd, (3) Less-polar (LP) 계에 속하는 Rg3(S), Rk1, Rg5. Peak bandwidth가 검출 감도에 영향을 주는 정도를 확인하기 위하여 각 그룹 내 Ginsenoside를 22% (PPT), 35% (PPD), 55% (LP) 아세토니트릴로 isocratic 분석하고 검출 감도를 확인하였다.

## 2.5. 분석법 최적화 및 유효성 검증

검출 감도 평준화(Response normalization)을 위해 중요 요소인 이동상 조성과 Peak bandwidth의 영향력을 고려하여 분석 조건을 순차적으로 최적화하였다. 최적 조건은 Section 2.2에 명시되어 있다. Ginsenoside 표준물질 9종 혼합물과 인삼 제품 원료 농축물은 최적 조건에서 분석하였다.

분석법의 유효성을 검증하기 위해 분석법의 직선성(Linearity), 검출 한계(Limit of Detection), 정량 한계(Limit of Quantification), 정밀성(Inter-, Intra-day precision), 정확성(Accuracy), 불확정도(Uncertainty)를 측정하였다. 직선성은 다섯 농도 조건에서 Ginsenoside를 검출하고 농도와 그 때의 면적 값을

로그(Log) 값으로 변형한 것을 바탕으로 검량선을 작성하여 확인하였다. 검출 한계와 정량 한계는 signal-to-noise비가 각각 3, 10이 되는 분석 농도를 확인하여 도출하였다. 정밀성은 하루에 3번, 3일 연속 같은 실험을 반복하여 결과의 편차를 확인함으로써 평가하였다. 정확도는 원료 농축물에 Ginsenoside 표준물질 적당량을 가하여 동일한 방법으로 추출, 분석을 진행한 뒤 추가로 가한 양이 분석 결과에 정확히 반영되는지를 확인함으로써 평가하였다. 불확정도( $U$ )는 EURACHEM/CITAC Guide CG 4에 고시된 방법대로 평가하였으며 5개의 항목(검량선;  $U_1$ , Inter-day 정밀성;  $U_2$ , Intra-day 정밀성;  $U_3$ , 정확도;  $U_4$ , 다른 분석자가 측정한 정확도;  $U_5$ )에서 발생하는 개별 불확정도를 모두 고려하여 계산하였다 [42].

$$U = (U_1^2 + U_2^2 + U_3^2 + U_4^2 + U_5^2)^{1/2} \quad (6)$$

## 2.6. 직접 정량법과의 비교 평가

400 mg의 건조된 원료 농축물에 10 mL의 메탄올을 넣고 40°C 조건에서 60분 동안 초음파 추출하였다. 원심분리를 통해 상정액을 분리한 후 동일용매 10 mL를 다시 넣고 같은 추출을 진행하였다. 총 3번 추출하여 시료당 30 mL의 상정액을 얻었고 농축기를 이용해 용매를 제거한 후 다시 10 mL의 메탄올에 녹여 분석 시료로 활용하였다. HPLC 분석 전에 0.45  $\mu$ m membrane filter로 방해물을 제거하였다. 상대 정량에서는 Ginsenoside Rg1(PPT 그룹), Rc(PPD 그룹), Rg3(S)(LP 그룹)가 각 그룹의 대표물질로 선택되었으며 상대 정량을 통한 농축물의 정량 결과와 직접 정량을 통한 농축물의 정량 결과를 비교하였다.



### III. 결과 및 고찰

#### 1. 크로마토그래프지문을 이용한 품질 평가법

##### 1.1. 크로마토그래프지문 획득

크로마토그래프지문 내의 정보량을 극대화하기 위해서 최대한 많은 수의 성분들이 검출한계 이상으로 크로마토그래프지문 내에 나타나야 했다. 따라서, FCCCD를 이용하여 40개의 시료에서 유의한 17개의 성분들이 최대한으로 추출될 수 있는 조건을 설정하고자 했다. 최적 조건을 확립하기 위한 모델 내 실험 조건과 해당하는 조건에서의 피크 면적 합은 Table 1에 나타내었다. 실험결과로부터 독립 변인들과 종속 변인의 관계를 다중 회귀 분석 (multiple regression analysis)을 통해 다항식으로 도출하였다.

$$Y = 21640.98 + 525.60X_1 + 174.00X_2 + 280.80X_3 - 54.00X_1X_2 - 81.00X_1X_3 - 87.00X_2X_3 - 119.45X_1^2 + 90.55X_2^2 - 887.45X_3^2 \quad (7)$$

ANOVA 분석 결과(Table 2),  $p$ -value가 0.05보다 작은 값을 갖는 요소인  $X_1$ ,  $X_2$ ,  $X_3$ ,  $X_3^2$ 이 추출 효율을 결정하는데 중요한 요소로 밝혀졌다. 다시 말하면, 온도는 1차 함수의 형태로 추출 효율에 큰 영향을 미치며, % 메탄올(용매 극성)은 2차 함수의 형태로 추출 효율에 영향을 미쳤다. 또한, 두 가지 요소 사이의 상호작용으로 인한 추출 효율의 변화는 미미한 것으로 밝혀졌다. 다항식의 정확성은 결정계수;  $R^2$ , adjusted- $R^2$ , adequate precision를 통해 확인하였다.  $R^2$ 은 결과 값이 모델 내에서 평균 값으로부터 편차가 얼마나 발생했는지를 나타내는 지표이며 0.9741로 1에 가까웠다. adjusted- $R^2$ 은 요소의 개수를 고려하여 계산한 지표이며 역시 0.9507로 높은 값을 나타내어 다항식이

유의함을 나타냈다. Adequate precision은 계획한 포인트에서 예측 값의 범위를 평균 예측 오차로 나눈 값이며 (signal-to-noise ratio) 그 값이 4보다 클 때 바람직하다. 본 모델의 경우 24.3의 값을 나타냈다.

Figure 2는 한 가지 변인이 중간 값으로 고정되어 있을 때 나머지 두 가지 독립 변인 사이의 상관관계를 시각화 한 반응 표면 곡선이다. 위에 서술한 변인들의 영향력을 시각적으로 확인할 수 있다. 최종적으로 다항식으로부터 desirability 값을 고려하여 최적 추출 조건을 도출한 것은 50°C, 90분, 81% 메탄올의 조건이었다. 이 예측의 정확성은 이 조건에서의 3번의 추가 추출 실험 결과가 예측 값이랑 부합하는지 확인함으로써 검증하였다. 3번의 실험 결과, 실험 값이 예측 값과 거의 일치하는 것을 확인하였다. 분석은 지문 정보가 가장 많이 포함된 236 nm의 파장에서 진행되었다.

한편, 다수의 시료 수집과 측정은 대부분의 경우에 단기간에 이루어지지 않기 때문에 측정 사이에 편차가 발생할 가능성이 충분히 있다. 이것은 데이터의 왜곡, 잘못된 시료 DB 구축, 품질 평가에 오류를 발생 시킬 수 있기 때문에 정확하게 실험되었음을 증명하고자 분석법의 유효성을 검증하였다. 유효성 검증에 사용된 표준물질은 백출의 중요한 성분으로 잘 알려진 atractylenolide I, II, III였다. 일부 *A. macrocephala* 시료에서 atractylenolide의 양이 적어 검출이 잘 되지 않았기 때문에 (236 nm) 표준물질의  $\lambda_{max}$ 인 278 nm (attractylenolide I), 220 nm (attractylenolide II, III)에서 진행하였다. Figure 3는 표준물질과 *A. japonica*, *A. macrocephala* 시료의 대표 크로마토그램을 보여준다.

하지만 Figure 3C에 나타난 것처럼 식별되지 않은 물질이 atractylenolide I과 겹쳐서 용리 되었고 이동상 조건의 변화로도

분리가 되지 않아 validation이 어려운 상황이었다. Figure 4는 불순물의 피크만 선택적으로 제거할 수 있는 Peak suppression method를 설명하는 그림이다. 불순물의 UV 흡수가 278 nm와 345 nm에서 동일한 반면 atractylenolide I은 345 nm에서 UV 흡수가 없다. 따라서, 345 nm를 참조 파장으로 처리함으로써 345 nm에서 발생하는 UV 흡수를 제거해서 atractylenolide I의 피크 변화 없이 선택적으로 불순물 피크를 제거할 수 있었다. 최종적으로 atractylenolide I의 분석은 345 nm의 참조파장 아래 278 nm에서 분석하였고 atractylenolide II, III의 경우는 220 nm에서 분석하였다.

Validation 항목으로 분석법의 직선성(Linearity), 검출 한계(Limit of Detection, LOD), 정량 한계(Limit of Quantification, LOQ), 정밀성(Inter-, Intra-day precision), 정확성(Accuracy), 재현성(Repeatability)를 평가하였다. 직선성은 다섯 농도 조건에서 atractylenolide를 검출하고 농도와 그 때의 면적 값을 바탕으로 검량선(Calibration curve)을 작성하여 확인하였다. 검출 한계와 정량 한계는 signal-to-noise비가 각각 3, 10이 되는 분석 농도를 확인하여 도출하였다. 정밀성은 하루에 5번, 5일 연속 같은 실험을 반복하여 결과의 편차를 확인함으로써 평가되었다. 정확도는 원료 농축물에 atractylenolide 표준물질 적당량을 가하여 동일한 방법으로 추출, 분석을 진행한 뒤 추가로 가한 양이 분석 결과에 정확히 반영되는지를 확인함으로써 평가되었다. 재현성은 전체 실험을 총 5번 반복했을 때 나타나는 상대표준편차로 확인하였다. Table 3는 Validation 결과를 보여준다. 이 결과는 확립한 분석법이 *Atractylodes*속 식물의 크로마토그래프지문을 재현성 있게 얻기에 적절하다는 것을 입증한다.

**Table 1**

Design matrix and responses for the face-centered central composite design (FCCCD)

Set	Temperature ( $X_1$ )	Time ( $X_2$ )	% MeOH ( $X_3$ )	Total Peak Area
1	30	30	60	19604
2	50	30	60	20872
3	30	90	60	20068
4	50	90	60	21344
5	30	30	100	20336
6	50	30	100	21504
7	30	90	100	20676
8	50	90	100	21404
9	30	60	80	21108
10	50	60	80	21924
11	40	30	80	21444
12	40	90	80	22008
13	40	60	60	20360
14	40	60	100	21136
15	40	60	80	21476
16	40	60	80	21624
17	40	60	80	21776
18	40	60	80	21536
19	40	60	80	21612
20	40	60	80	21844

**Table 2**

Analysis of variance (ANOVA) table for the response surface quadratic model

Source	Sum of squares <sup>a</sup>	d.f. <sup>b</sup>	Mean square <sup>c</sup>	<i>F</i> -value <sup>d</sup>	<i>p</i> -value, Prob> <i>F</i> <sup>e</sup>	Significance
Model	$8.13 \times 10^6$	9	$9.03 \times 10^5$	41.71	<0.0001	significant
$X_1$	$2.76 \times 10^6$	1	$2.76 \times 10^6$	127.56	<0.0001	significant
$X_2$	$3.03 \times 10^5$	1	$3.03 \times 10^5$	13.98	0.0039	significant
$X_3$	$7.89 \times 10^5$	1	$7.89 \times 10^5$	36.41	0.0001	significant
$X_1X_2$	$2.33 \times 10^4$	1	$2.33 \times 10^4$	1.08	0.3238	not significant
$X_1X_3$	$5.25 \times 10^4$	1	$5.25 \times 10^4$	2.42	0.1506	not significant
$X_2X_3$	$6.06 \times 10^4$	1	$6.06 \times 10^4$	2.80	0.1254	not significant
$X_1^2$	$3.92 \times 10^4$	1	$3.92 \times 10^4$	1.81	0.2080	not significant
$X_2^2$	$2.25 \times 10^4$	1	$2.25 \times 10^4$	1.04	0.3316	not significant
$X_3^2$	$2.17 \times 10^6$	1	$2.17 \times 10^6$	100.01	<0.0001	significant
Residual <sup>f</sup>	$2.17 \times 10^5$	10	$2.17 \times 10^4$			
Lack of fit <sup>g</sup>	$1.18 \times 10^5$	5	$2.36 \times 10^4$	1.19	0.4254	not significant
Pure error <sup>h</sup>	$9.87 \times 10^4$	5	$1.97 \times 10^4$			
CorTotal <sup>i</sup>	$8.35 \times 10^6$	19				

<sup>a</sup> Sum of the squared differences between the average values and the overall mean.

<sup>b</sup> Degrees of freedom.

<sup>c</sup> Sum of squares divided by d.f.

<sup>d</sup> Test for comparing term variance with residual (error) variance.

<sup>e</sup> Probability of seeing the observed *F*-value if the null hypothesis is true.

<sup>f</sup> Consists of terms used to estimate experimental error.

<sup>g</sup> Variation of the data around the fitted model.

<sup>h</sup> Variation in the response in replicated design points.

<sup>i</sup> Totals of all information corrected for the mean.

**Table 3**

## Validation results of the HPLC method

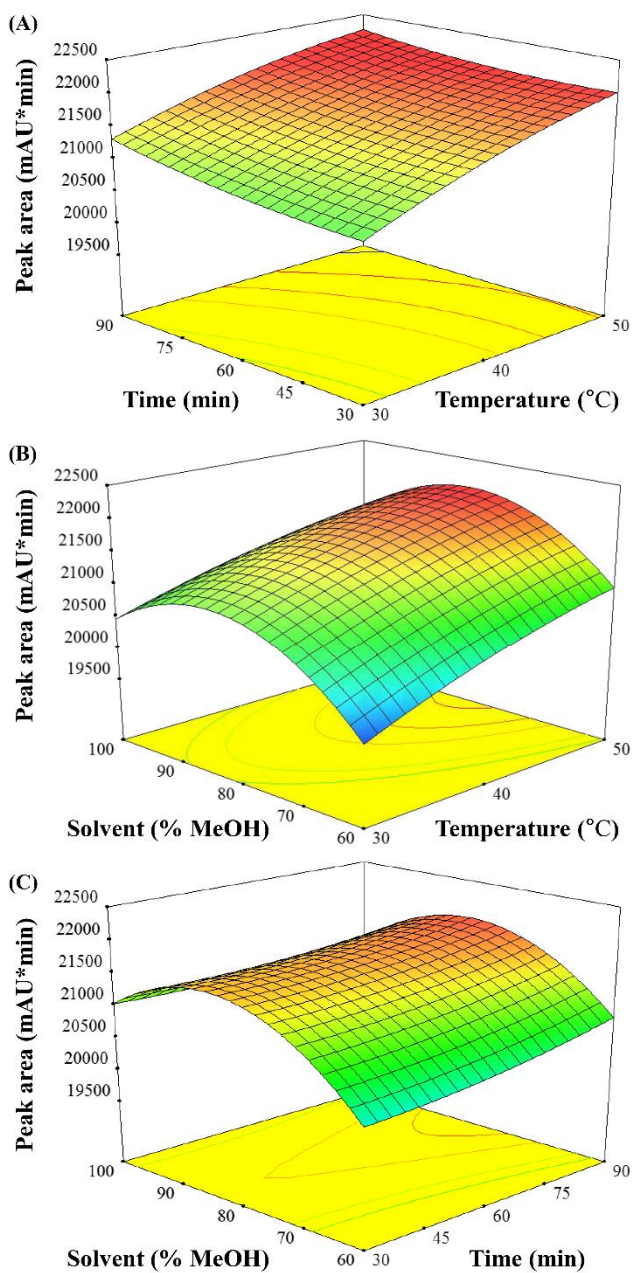
	Calibration curve <sup>a</sup>	$R^2$	LOD (ng)	LOQ (ng)	Precision ( $n=5$ )		Recovery <sup>b</sup> ( $n=3$ )		Repeatability
					Intra-day	Inter-day	Mean (%)	RSD (%)	RSD (%)
					RSD <sup>c</sup> (%)	RSD (%)			
AT-1	$y = 51.655x + 44.476$	1.0000	0.26	0.86	0.37	1.98	92.96	2.80	3.93
AT-2	$y = 39.555x + 24.414$	1.0000	0.37	1.22	0.37	1.90	93.67	1.80	3.83
AT-3	$y = 25.357x + 22.989$	0.9999	0.50	1.65	0.75	1.00	99.79	1.68	5.35

<sup>a</sup>  $y$ , value of peak area;  $x$ , value of concentration.

<sup>b</sup> Recovery (%) = (observed amount – original amount)/spiked amount  
× 100%.

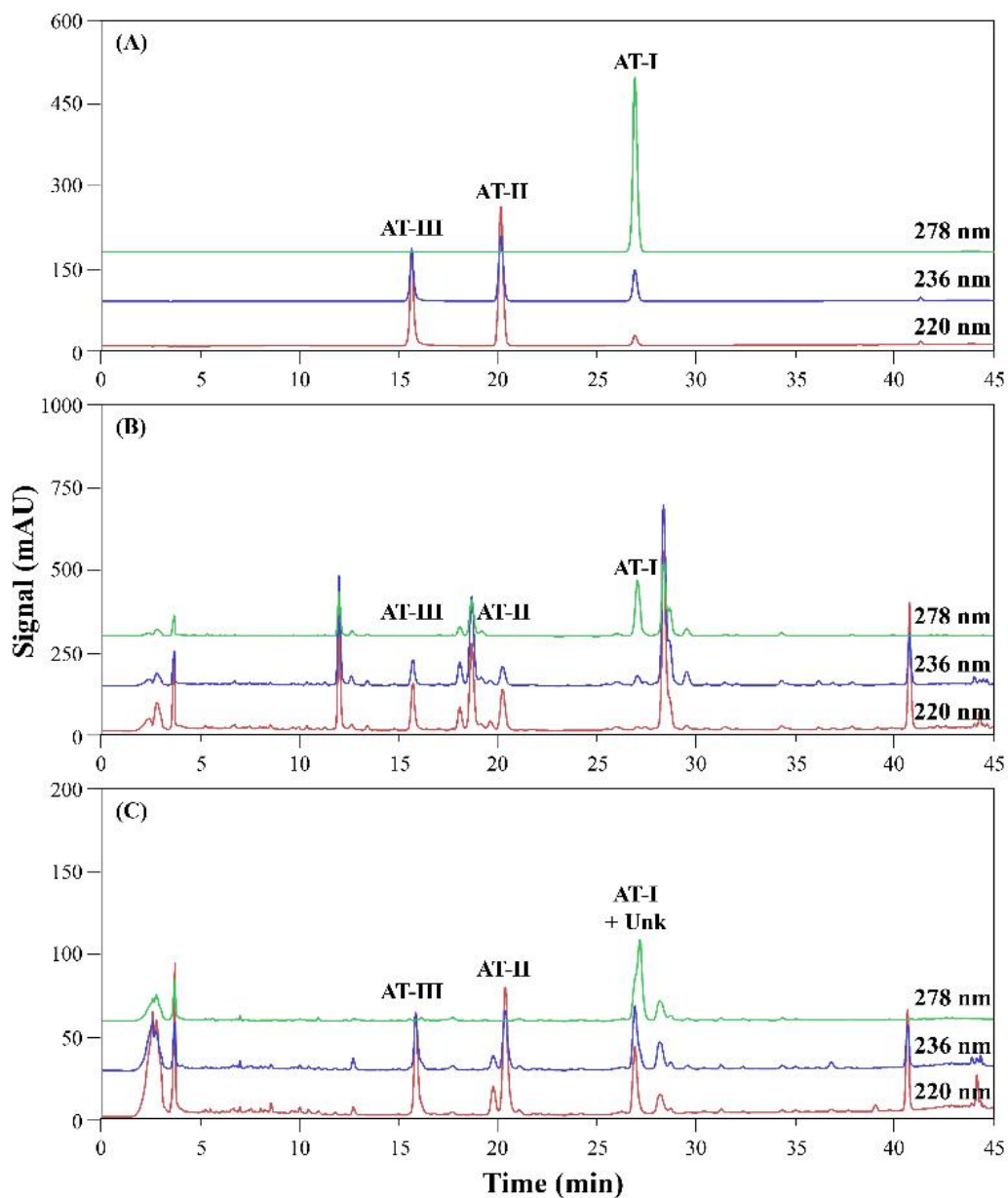
<sup>c</sup> RSD (%) = standard deviation/mean × 100%.

<sup>d</sup> AT-1; atractylenolide I, AT-2; atractylenolide II, AT-3; atractylenolide  
III

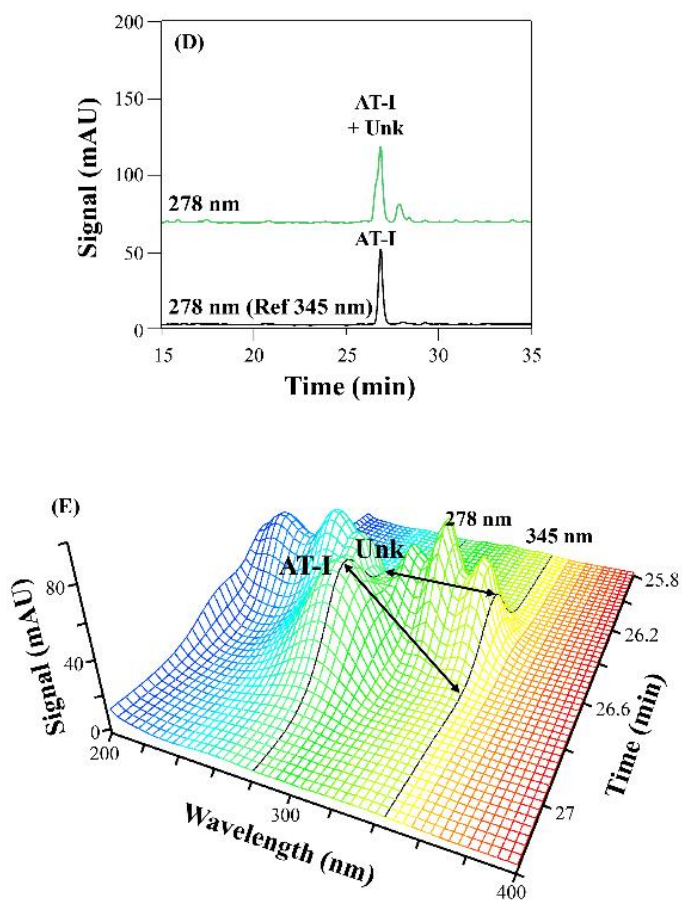


**Figure 2.** Response surface plots showing reciprocal interaction of two individual factors on the total response. (A) Temperature ( $X_1$ ) – Time ( $X_2$ ), (B) Temperature ( $X_1$ ) – Percentage of methanol ( $X_3$ ), and (C) Time ( $X_2$ ) – Percentage of methanol ( $X_3$ ).





**Figure 3.** Representative fingerprints of (A) the mixed standards, (B) the *A. japonica* samples, and (C) the *A. macrocephala* samples. AT-1; atractylenolide I, AT-2; atractylenolide II, AT-3; atractylenolide III, Unk; Unknown compound.



**Figure 4.** Selective elimination of the response of the unknown compound with the reference wavelength and 3D spectra in the corresponding area.

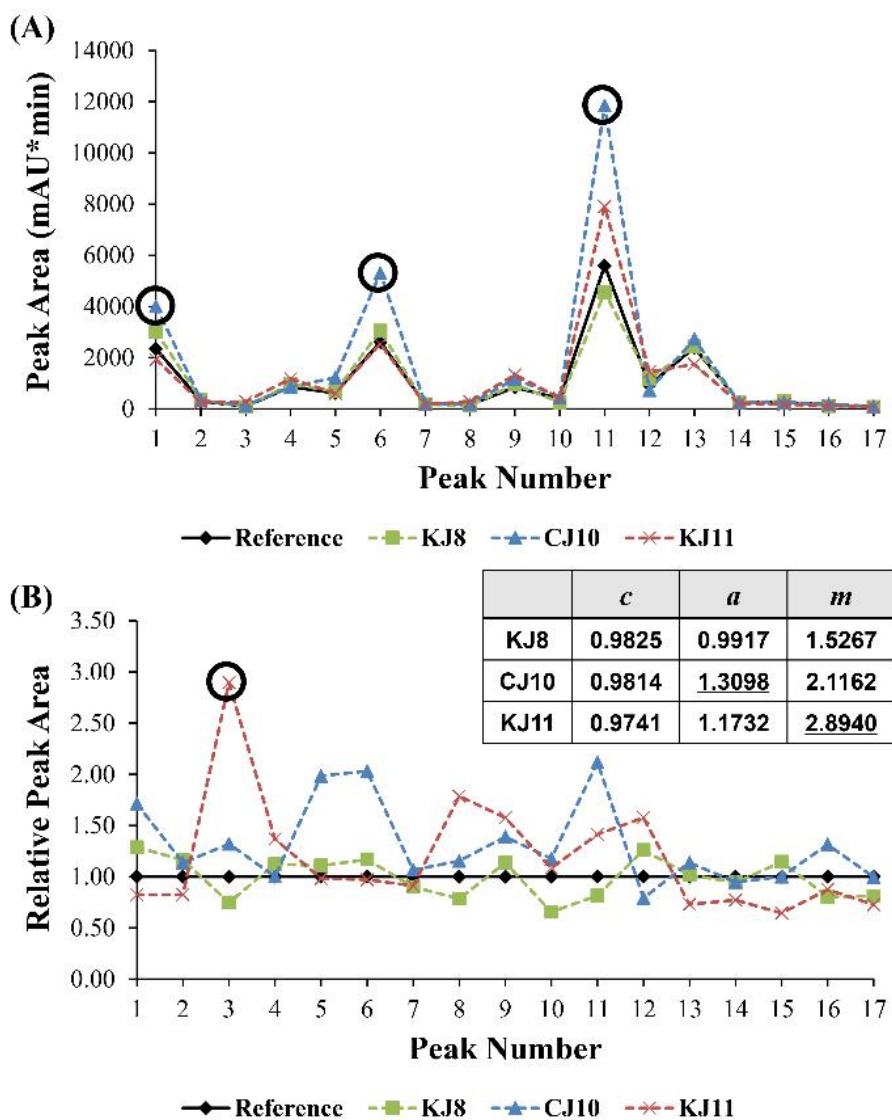
## 1.2. 복합 유사성 지표(CAM) 계산

40개의 크로마토그래프지문들 내 피크들의 배열은 피크들의 UV 스펙트럼을 확인하여 정밀하게 수행했고 최종적으로 17개의 피크가 데이터 벡터를 형성하는 요소가 되었다. 기준 시료는 가장 약효가 높은 것으로 선정되는 것이 바람직하나 질환 타겟에 따라 기준 시료가 달라질 수 있어 방법론을 개발하는 본 연구에서는 *A. japonica* 시료들의 평균을 연구 진행을 위한 임의 기준 시료로 상정하였다. 기준 시료의 지문과 비교하여 40개 시료들의 유사성 지표 값( $c, a, m$ )을 계산하였다.

유사성 지표  $c$ 는 두 데이터 벡터 사이의 각도를 기준으로 (코사인 값) 두 시료의 유사성을 측정하며 따라서 성분들의 조성비(시너지 효과)를 강조하는 천연물의 품질평가에 종종 활용된다. 하지만, 기존 방법과 같이 지표  $c$ 만 단독으로 사용하여 품질을 평가하는 것은 이미 언급한 것처럼 다양한 크기의 데이터를 포함하고 있는 크로마토그래프지문에 적용했을 때 품질평가의 오류를 야기한다. 본 연구에서도 확인할 수 있었고 그 실제 예시가 Figure 5, 6에 설명되어 있다.

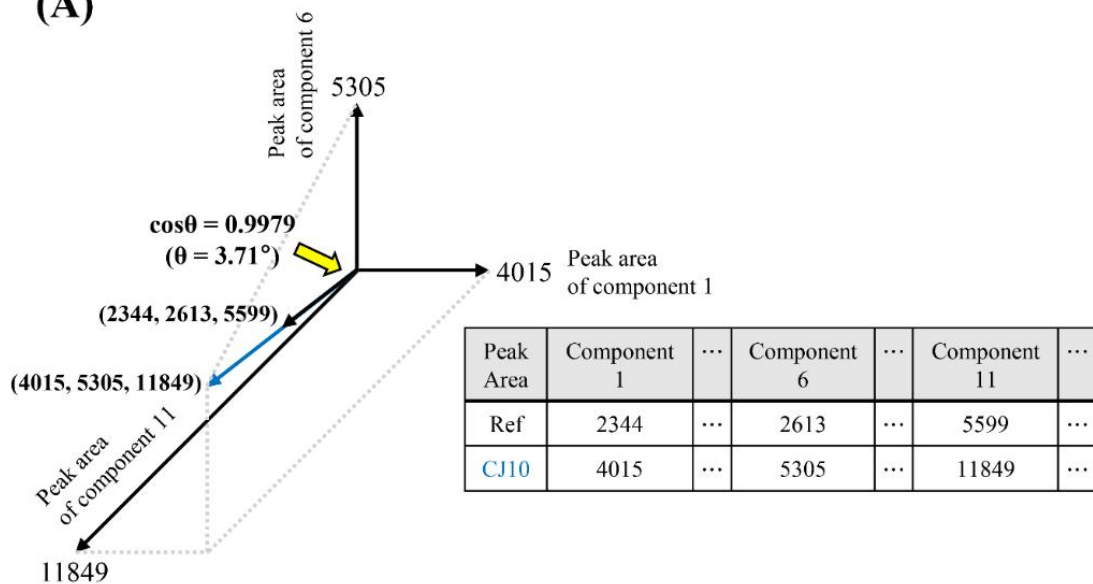
Figure 5는 지표  $c$ 값이 0.97 이상으로 높은 세 가지 예시 시료 KJ8, CJ10, KJ11의 크로마토그래프지문 내 절대 피크 면적과 기준 시료 대비 상대 피크 면적을 보여준다. KJ8 시료는 실제로 기준 시료와 상당히 화학적으로 동등했다. 반면, CJ10, KJ11 시료는 KJ8처럼 높은  $c$ 값을 갖고 있음에도 불구하고 일부분에서 기준 시료와 꽤 상이한 것을 확인할 수 있었다. 먼저, CJ10 시료는 성분 1, 6, 11번의 면적 값이 기준 시료에 비해 약 2배 가량 높았다. 이 성분들은 크로마토그래프지문에서 큰 비중을 차지하고 있기 때문에 이 성분들의 조성비율이 기준 시료의 것과 동등하다면  $c$ 값이 1에

가깝게 나타날 수 있을 것이라 예상했고 확인 결과 이 세 개의 성분들이 주도적으로 CJ10 시료의 높은  $c$ 값을 결정했다는 것이 밝혀졌다 (Figure 6A). CJ10과 같은 동등하지 않은 시료를 KJ8과 같이 동등한 시료와 구분하기 위하여 지표  $a$ 를 도입하였다. 어떤 시료의 성분들이 전반적으로(혹은 부분적으로) 기준 시료에 비해 많거나 적은 농도로 시료에 함유되어 있다면 이 시료의  $a$ 값은 1보다 크거나 작다. 실제로 KJ8의  $a$ 값은 0.9917인 반면 CJ10의  $a$ 값은 1.3098로 1보다 꽤 큰 값이었다. 지표  $a$ 를 통해 시료 CJ10과 KJ8을 구분할 수 있음을 확인하였다. KJ11 시료의 경우 전반적으로 기준 시료와 동등한데 성분 3번에서는 유난히 차이가 많이 나타났다. 기준 시료 보다 약 3배 가량 많았다. 하지만 이 큰 차이는 지표  $c$ 와  $a$ 를 통해서는 드러나지 않았다. 그 이유는 이 3번 성분은 전체 크로마토그래피문에서 매우 작은 부분을 차지하며 17개의 성분 중 단지 하나에 불과하기 때문이었다 (Figure 6B). 지표  $m$ 은 KJ11처럼 지표  $c$ 와  $a$ 를 통해서도 드러나지 않는 품질에 이상이 있는 시료를 검출하기 위해 도입되었다. 지표  $m$ 은 시료와 기준 시료 사이의 가장 큰 농도 차이를 검출하는 지표이며 KJ8의  $m$ 값은 10번 성분의 상대 면적비의 역수인 1.5267인 반면, KJ10의  $m$ 값은 3번 성분의 농도 차이에 해당하는 2.8940으로 나타났고 지표  $m$ 을 통해 두 시료를 구분할 수 있음을 확인하였다. 요약하면, 지표  $c$ 는 원료의 시너지 효과와 관련된 성분 조성비를 검증하는 지표이며 지표  $a$ 는 전반적 혹은 부분적 농도 차이를 검증하고 지표  $m$ 은 미세한 부분에서 발생하는 큰 농도 차이를 검증하는 지표다. CAM 체계를 이용함으로써 생물다양성으로부터 오는 원료들의 성분 차이, 가공 과정에서 발생하는 원료들의 성분 차이 등을 다양한 측면에서 엄격하게 파악할 수 있었다. 이하 동등함의 기준을 선정하는 과정, 원료를 분류하는 과정에 대하여 논의하고자 한다.

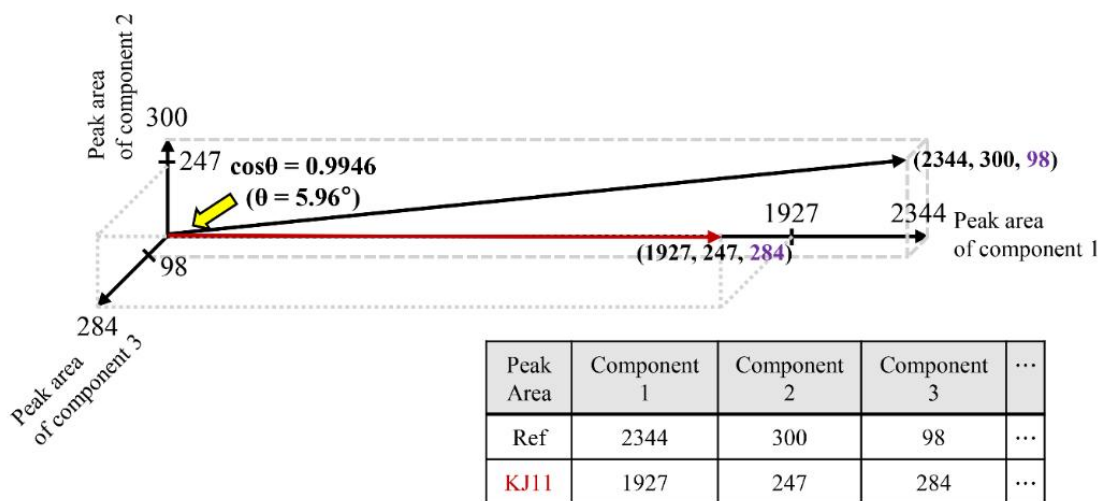


**Figure 5.** The elements of the three data vectors, KJ8, CJ10, and KJ11, having a high  $c$  value over 0.97 and their CAM similarity values. (A) Peak areas of the fingerprints and (B) Relative peak areas of the fingerprints to the reference.

(A)



(B)



**Figure 6.** The limitations of the congruence coefficient found in evaluating the similarity between the actual samples and the reference. (A) The case of CJ10 and (B) The case of KJ11.

### 1.3. 커널밀도추정을 이용한 이상치 검출

Figure 7은 40개 시료의  $c$ ,  $a$ 값으로 구성된 2차원 품질 평가 플롯(two-dimensional quality evaluation plot)을 나타낸다. 이 플롯은 40개 시료의 기준 시료 대비 화학적 동등성을 한 눈에 파악하는데 매우 용이하다. 플롯에서 *A. macrocephala* (CM group) 시료들은 왼쪽 아래에 위치해 있고 *A. japonica* (KJ, CJ group) 시료들과 구분되어 있다. 비슷한 외관 때문에 이들 시료들은 현장에서 상호 혼동되지만 두 개의 지표 값을 통해 쉽게 종 간 차이(Inter-species difference)를 구분할 수 있었다. CM group의  $c$ 값이 낮은 것은 크로마토그래프지문의 패턴이 기준 시료의 것과 상당히 다르다는 것을 의미하며  $a$ 값이 낮은 것은 성분들의 농도가 전반적으로 기준 시료의 것들과 비교했을 때 작다는 것을 의미한다. 결국, 이 시료들은 기준 시료와 완전히 이질적인 것으로 판단 되어 통계 분석 전에 이상치(anomaly or outlier)로 선 제거되었다. 반면에, *A. japonica* 시료들은 1에 가까운  $c$ ,  $a$ 값을 가지면서 군집되어 있다 (Figure 8). 이것은 대다수의 *A. japonica* 시료들이 기준 시료와 화학적으로 상당히 유사한 것을 의미하며 통계 분석을 통해 이 시료들 중 기준과 동등하다고 할 수 있는 것들을 판별하고자 하였다.

이 문제는 이상치에 대한 사전 경험 혹은 데이터가 없는 상황에서 시료를 구분해야 하기 때문에 일반적인 two-classes 분류문제보다는 좀 더 까다로운 one-class 분류문제에 해당한다. 단일 종류 분류에 있어 가장 바람직한 방법은 데이터로부터 확률밀도를 추정하고 여기에 역치(기준)를 정하는 것으로 알려져 있다 [43-44]. 본 연구에서는 커널 밀도 함수를 추정하여 비모수적 방법으로 분석을 진행하기로 하였다. 커널 함수는

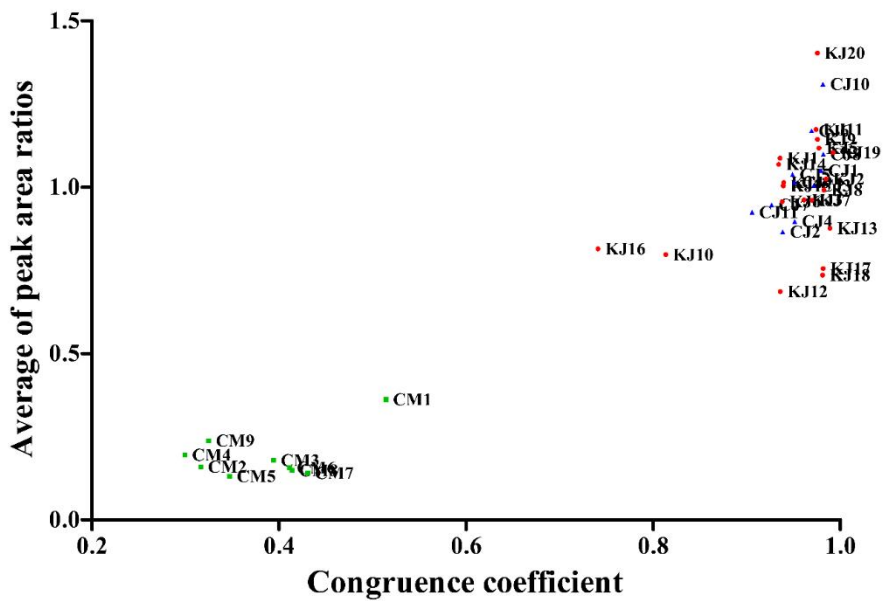
데이터의 값에 중심을 두고 대칭이며 적분 값이 1인 non-negative 함수이다. Gaussian, Epanechnikov, uniform 함수가 대표적인 커널 함수로 알려져 있다. 커널 함수를 결합한 커널 밀도 추정(kernel density estimation)은 기존 히스토그램과 같은 고전적인 방법들과 달리 모델링에 유연함(flexibility), 부드러움(smoothness), 연속성(continuity)을 제공하기 때문에 본 연구에 사용되었다 [45].

Figure 9는 31개 시료의  $c$ ,  $a$ ,  $m$  값으로 구성된 3차원 품질 평가 플롯(three-dimensional quality evaluation plot)과 이 값들로부터 얻은 커널 밀도에 3 종류의 경계를 설정한 그림이다. 경계를 설정했을 때 높은 밀도에 위치한 시료들만 동등한 것으로 판정된다. 예를 들어 Figure 9A에서 타원과 유사한 노란색의 경계는 시료들 중 90%를 포함할 것으로 예상되는 경계다. 여기서 KJ7, KJ10, KJ12, KJ16은 이상치로 판정되었다. 이 시료들은 31개 시료 중 가장 높은  $m$  값을 가지며 KJ7을 제외하고는  $c$ ,  $a$  값 역시 1보다 한참 작은 값이었기 때문에 우선적으로 이상치로 선정되었다 (Table 4). Figure 9B와 10A는 시료들 중 80%를 포함할 것으로 예상되는 경계다 (Figure 10는 경계의 정확한 이해를 위해 Figure 9을 측면에서 바라본 그림이다). 이 수준에서 KJ17, KJ18, KJ20이 추가로 이상치로 선정되었다. 이 시료들의  $c$ 와  $m$  값 보다는  $a$  값이 이상치 결정에 크게 기여한 것으로 보인다. KJ17, KJ18의  $a$  값은 이미 제거된 KJ12를 제외하면 가장 작은 2개였고 KJ20은 가장 큰  $a$  값을 가졌다. Figure 9C와 10B는 70% 수준을 나타내고 이 수준에서는 CJ10과 KJ11이 추가로 이상치로 선정되었다. CJ10은 이미 제거된 KJ20 다음으로 큰  $a$  값을 가지며 KJ11의  $a$  값과  $m$  값은 남은 동등하다고 판단된 시료들의 값들과 상당히 동떨어져 있었다. 이 시료들은 특히  $c$  값을 기준으로 봤을 때는 (기존 방법) 상위 50% 이내에 들 정도로 좋은 시료들에 속하지만 실제 유사성은 앞서

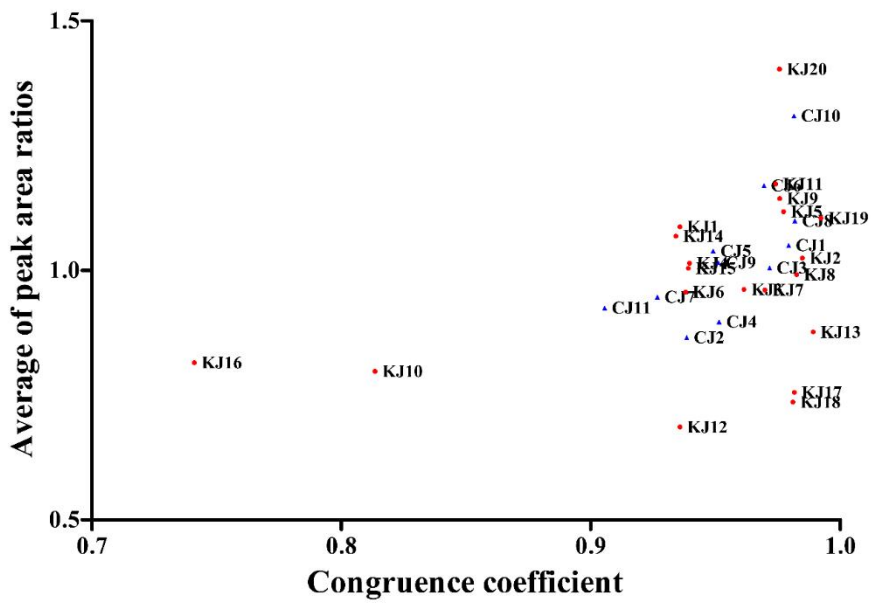


Figure 5에 언급한 것처럼 그렇게 높지 않았다. 이 시료들의 CAM값을 고려하였을 때 70% 수준에서 제거되는 것은 마땅한 결과였다. 위 결과를 바탕으로, 삼차원 커널 밀도 추정(three-dimensional kernel density estimation)을 이용하면 데이터 의존적인 방법이면서 정확한 순서대로 이상치를 검출해낼 수 있음을 알 수 있다.

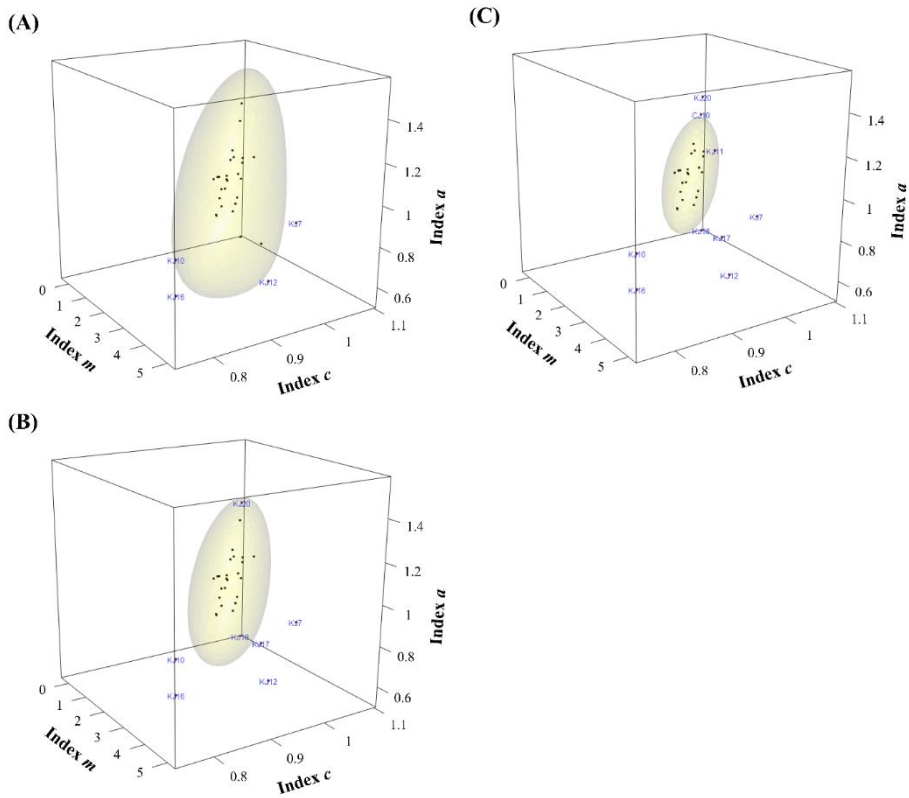
Figure 11은 Figure 9C(70% 수준)에서  $m$ 값이 주어졌을 때의 2D-contour plot을 보여준다. 하나의 점(시료)이 밀도가 높은 붉은 색 영역에 위치할 수록 정상치(동등한 것)일 가능성이 높은 것을 의미한다.  $m$ 값의 관점에서 가장 높은 밀도의 영역은  $m$ 값이 2.1과 2.4 사이였다. 이를 통해  $m$ 값이 1에 가까운 아주 동등한 시료도 과도한 이상치 제거 시에는 데이터에 의존하는 모델링의 특성상 역시 제거 될 수 있음을 알 수 있었다. 경계를 설정할 때에 이런 오류들이 발생하지 않도록 주의를 기울였다. 실제로 사전 품질 결정에 대한 지식이 없는 경우에 시료의 분류를 하고자 하면 연구자들은 주관적인 결정을 할 수 밖에 없는 상황에 놓인다. 직관에 의해 결정하는 것과 다른 연구에서 설정된 기준을 단순히 사용하는 것과 달리 커널 밀도 추정을 이용하면 우수한(동등한) 시료들이 몰려 있는 영역을 쉽게 확인할 수 있고 이를 통해 경계 설정을 상대적으로 수월하게 진행할 수 있다. 마찬가지로 본 연구의 경우에는 시료의 70% 수준을 동등함의 기준으로 설정하였다.



**Figure 7.** The distribution of the  $c$  (the congruence coefficient) and  $a$  (the average of peak area ratios) values for the 40 *Atractylodes* samples; the KJ group (circle, red), the CJ group (triangle, blue), and the CM group (square, green).



**Figure 8.** The distribution of the  $c$  (the congruence coefficient) and  $a$  (the average of peak area ratios) values for the 31 *A. Japonica* samples; the KJ group (circle, red) and the CJ group (triangle, blue).



**Figure 9.** The distribution of the CAM values for the 31 *A. Japonica* samples and the three levels of the kernel density estimation with plug-in bandwidth matrix. The yellow colored regions are expected to contain (A) 90, (B) 80, and (C) 70% of the points. The anomalies written in blue color: (A) KJ7, KJ10, KJ12, and KJ16; (B) KJ17, KJ18, and KJ20 added on to those in (A); (C) CJ10 and KJ11 added on to those in (B). (The probability density area with the *c* value greater than 1 is formed due to the feature of the kernel density estimation, but is not the area where actual data exists and is not relevant to the procedure of the anomaly detection.)

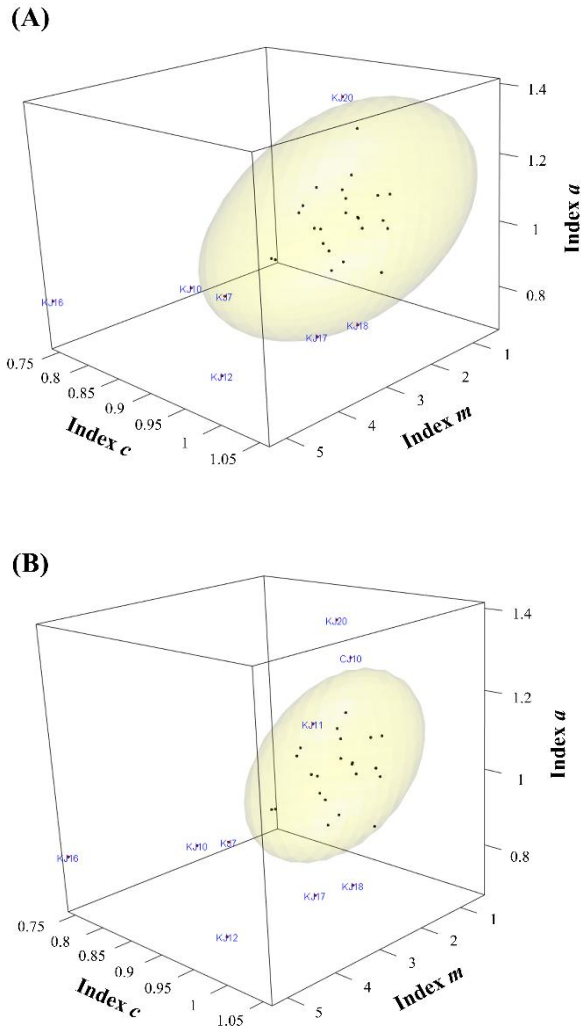


**Table 4**

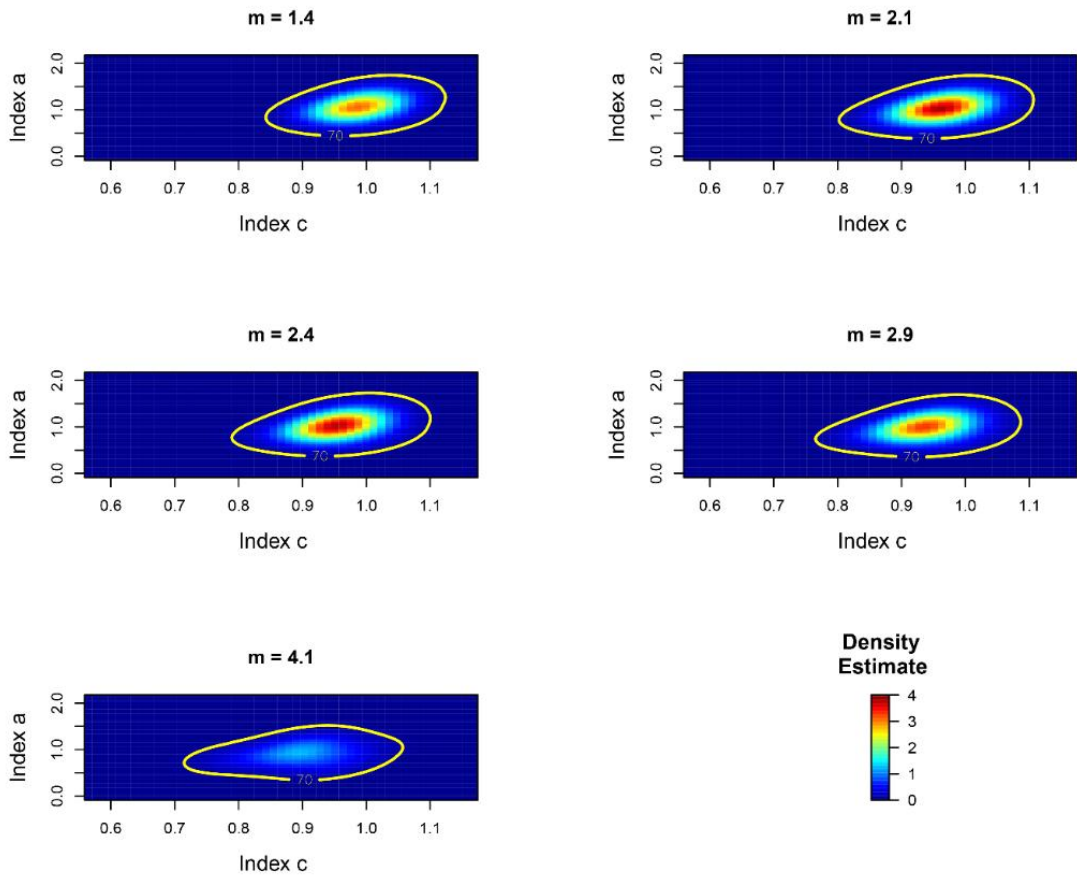
The CAM values of the 40 samples

Sample	The <i>c</i> value	The <i>a</i> value	The <i>m</i> value
KJ1	0.9357	1.0872	2.6082
KJ2	0.9848	1.0245	2.1336
KJ3	0.9614	0.9618	2.4933
KJ4	0.9396	1.0143	2.4500
KJ5	0.9773	1.1178	2.3431
KJ6	0.9380	0.9567	2.2507
KJ7	0.9698	0.9607	4.7017
KJ8	0.9825	0.9917	1.5267
KJ9	0.9757	1.1441	2.3852
KJ10	0.8135	0.7977	3.3592
KJ11	0.9741	1.1732	2.8940
KJ12	0.9358	0.6860	4.4168
KJ13	0.9891	0.8766	1.8052
KJ14	0.9341	1.0684	2.6762
KJ15	0.9391	1.0042	2.3066
KJ16	0.7411	0.8148	5.2372
KJ17	0.9816	0.7554	3.1258
KJ18	0.9810	0.7358	2.2400
KJ19	0.9923	1.1046	1.6241
KJ20	0.9756	1.4031	2.3045
CJ1	0.9793	1.0501	2.1209
CJ2	0.9385	0.8655	2.0964
CJ3	0.9717	1.0051	1.4399
CJ4	0.9514	0.8969	2.0318
CJ5	0.9491	1.0386	1.9009
CJ6	0.9694	1.1699	2.0768
CJ7	0.9267	0.9465	3.1003
CJ8	0.9818	1.0993	1.7144
CJ9	0.9511	1.0150	1.6712
CJ10	0.9814	1.3098	2.1162
CJ11	0.9056	0.9246	2.8825
CM1	0.5144	0.3615	N.C. <sup>a</sup>
CM2	0.3165	0.1585	N.C.
CM3	0.3942	0.1788	N.C.
CM4	0.2995	0.1950	N.C.
CM5	0.3473	0.1307	N.C.
CM6	0.4114	0.1565	N.C.
CM7	0.4309	0.1405	N.C.
CM8	0.4140	0.1489	N.C.
CM9	0.3249	0.2373	N.C.

<sup>a</sup> Not calculated because the samples of the CM group were ruled out from the statistical analysis.



**Figure 10.** The distribution of the CAM values for the 31 *A. Japonica* samples and the two levels of the kernel density estimation with plug-in bandwidth matrix. The yellow colored regions are expected to contain (A) 80 and (B) 70% of the points. The anomalies written in blue color: (A) KJ7, KJ10, KJ12, KJ16, KJ17, KJ18, and KJ20; (B) CJ10 and KJ11 added on to those in (A). This figure shows the views from the lateral side of Figures 9B and 9C to understand the surface of the areas exactly. (The probability density area with the  $c$  value greater than 1 is formed due to the feature of the kernel density estimation, but is not the area where actual data exists and is not relevant to the procedure of the anomaly detection.)



**Figure 11.** The two-dimensional contour plots of the estimated kernel density at the level of 70% (Figure 9C). If the point is situated far from the red region of the high density area, the likelihood of it being anomalous is high. (The probability density area with the  $c$  value greater than 1 is formed due to the feature of the kernel density estimation, but is not the area where actual data exists and is not relevant to the procedure of the anomaly detection.)

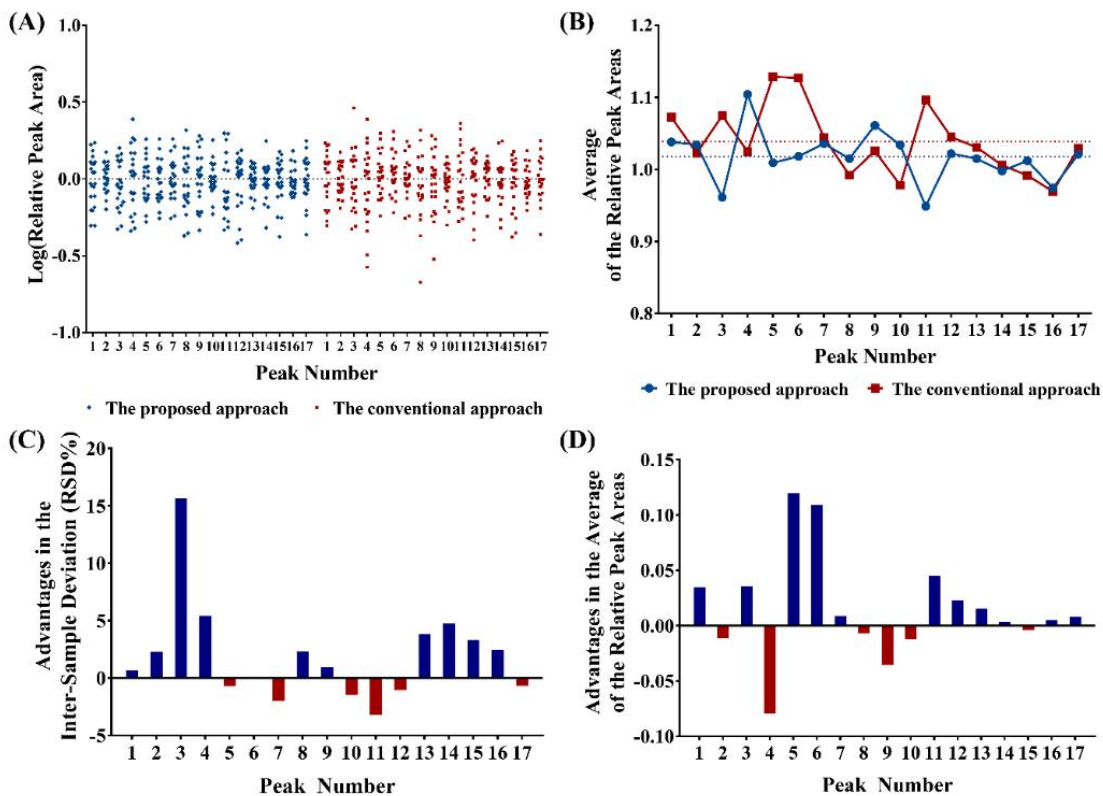


#### 1.4. 기존 품질 평가법과의 비교

본 연구에서 제안한 품질 평가 체계의 이점을 보여주기 위해 유사성 지표  $c$ 만 고려하는 일반적인 방법과 비교하였다. 최종적으로 동등하다고 판정된 22개의 시료를 40개의 시료 중에서  $c$ 값이 높은 상위 22개의 시료와 비교하였다 (Figure 12). Figure 12A는 각 방법에서 선정된 22개 합격 시료 군집들의 기준 시료 대비 피크 면적(시료의 피크 면적/기준 시료의 피크 면적)을 보여주는 산점도(scatter plot)이다. 제안된 방법으로 선정된 시료 성분들의 상대면적비가 더 1에 가깝게 분포한 것을 확인할 수 있다 (큰 outlier가 관찰되지 않음). 시료들의 상대면적비의 평균은 Figure 12B에 요약 되어 있다. 여기서 각 성분에서의 평균값을 연결한 선(Figure 12B의 실선)이 일직선의 형태와 가깝다는 것은 구성 성분의 조성비가 기준 시료의 것과 동등하다는 것을 의미하고, 즉 해당하는 방법이 시너지 효과를 잘 유지하는 시료들만 선정한다는 것을 의미한다. 각 방법에서 선정된 시료들이 얼마나 기준 시료의 시너지 효과가 유지되는지를 확인하기 위해 편차의 제곱합(Sum of squares)을 계산하였다. Figure 12B의 점선은 각 평균값들의 전반적인 평균이며 이 전반적인 평균값으로부터 개별 값들이 얼마나 차이가 나는지를 확인하고 이 차이들을 합하여 두 방법론을 비교하였다. 편차의 제곱합을 계산한 결과 제안된 방법은 0.0210, 기존 방법은 0.0364였다. 기존 방법이 성분 조성비의 유지에 강조를 두는 방법임에도 불구하고 오히려 제안한 방법이 전반적으로 시너지 효과가 유지되는 시료들을 선정하는 것을 확인하였다.

Figure 12C는 어떤 방법이 시료간의 편차(inter-sample deviation)을 줄여주고 12D는 개별 성분 측면에서 어떤 방법을

이용했을 때 시료들의 상대면적비가 더 1에 가까운지를 보여주는 플롯이다. 예를 들면, 성분 3번에서 제안된 방법으로부터 선정된 22개 시료들의 상대면적비의 편차는 15.68%만큼 기존 방법에 의해 선정된 시료들보다 시료 간 편차가 적었다 (Figure 12C). 그리고 그 상대면적비의 평균값은 0.0358만큼 1에 더 가까웠다 (Figure 12D). 이와 마찬가지로 제안된 방법의 긍정적인 효과는 다른 성분들 대다수에서 나타났다. 위 결과를 종합하면 본 연구에서 개발한 품질 관리 체계를 이용했을 때 선정된 시료들의 다성분(multi-components) 성분 조성비율과 개별 성분의 농도가 기존 시료의 것들과 더 동등했으며 시료들 간의 편차도 더 줄어들었음을 확인하였다. 31개 시료 중 단지 9개의 시료만 제거하여 두 방법론을 비교한 것을 감안하면 이는 상당히 유의한 결과임을 알 수 있다.



**Figure 12.** Comparative analysis of the samples qualified by the proposed approach and the conventional approach. (A) Relative peak areas (sample/reference) on log scale. (B) Averages of the relative peak areas of the 22 samples at the individual components. (C) The height of the bars shows how much smaller the inter-sample deviation (RSD) of the relative peak areas is in some approach and so which approach is favorable. (D) The height of the bars shows how much the average of the relative peak areas of the 22 samples is closer to 1 in some approach and so which approach is favorable. Blue indicates the proposed approach, and red indicates the conventional approach.

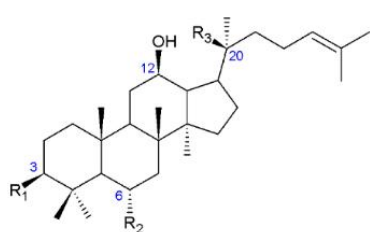
## 2. 상대정량을 이용한 품질 평가법

### 2.1. 이동상 조성이 검출 감도에 미치는 영향

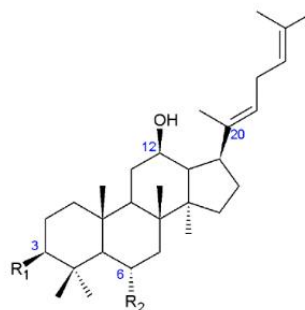
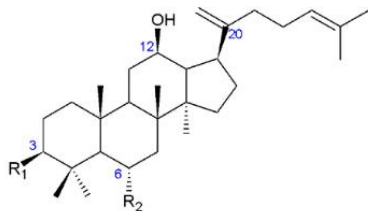
Ginsenoside Rg1, Re, Rb1, Rc, Rb2, Rd는 *P. ginseng*과 *P. quinquefolius*의 주요 Ginsenoside로 국내뿐만 아니라 여러 나라의 공정서에 지표 성분으로 등록되어 있다. Ginsenoside Rg3, Rk1, Rg5는 인삼을 열로 가열했을 때 발생하는 독특한 주요 Ginsenoside로 홍삼, 가공인삼의 주성분이며 활성성분이다. 따라서 본 연구에서는 위 9가지 Ginsenoside를 실험 대상으로 선정하였다 (Figure 13). 먼저 ELSD가 보편적으로 알려진 것처럼 (non-specific detector) 동일한 이동상, flow injection 조건에서 이 물질들에 대해 유사한 검출 감도를 보여주는지 확인하였다. 그리고 다양한 이동상 조건에서 같은 실험을 반복함으로써 이동상 조성이 검출 감도에 어떠한 영향을 미치는지 확인하였다.

Figure 14는 다른 이동상 조성에서의 9개의 Ginsenoside의 반응 곡선(response curve)을 나타낸다. 먼저 Figure 14A는 이동상 조성이 40% 아세트니트릴일 때의 9개 Ginsenoside의 반응 곡선을 보여준다. 예상했던 대로 이 물질들은 구조와 분자량 차이가 크지 않으며 ELSD의 검출 감도는 이런 물질의 화학적 특성에 영향을 거의 받지 않기 때문에 9개의 반응 곡선은 상당히 유사했다. 여기서 오차 막대를 통해 나타난 기기오차는 최대 7.5% (Relative standard deviation, RSD,  $n=3$ )에 불과했다. Figure 14B는 4개의 이동상 조건(20, 40, 60, 80% 아세트니트릴)에서 9개의 Ginsenoside의 평균 검출 감도로 반응 곡선을 나타낸 것이다. 이동상이 유기 용매인 아세트니트릴의 함량이 높아질수록 ELSD의 검출 감도는 향상된 것을 확인할 수 있었다. 80% 아세트니트릴을 용리 했을 때의 검출 감도는 20% 아세트니트릴을 용리 했을

때보다 2배 가량 컸다. 이것은 ELSD의 검출 감도가 이동상 조성에 의존성이 크다는 것을 의미한다. 여기서 오차 막대를 통해 나타난 9개 Ginsenoside 감도 사이의 편차는 최대 9.5%(RSD)에 불과했다. 요약하면 분석 대상의 화학적 특성보다는 이동상 조성이 ELSD의 검출 감도에 큰 영향을 미쳤으며 이동상의 극성이 감소할수록, 휘발성이 증가할수록, 점성이 감소할수록 검출 감도가 증가한다는 것을 알 수 있었다. 이 경향성은 기존에 보고된 결과들과 일치했고 두 가지 해석으로 이해될 수 있다 [33, 46]. 먼저, 이동상의 휘발성 증가는 분사된 액체방울이 drift tube의 벽면에 응결하는 것을 방지하고 분석체(analyte)의 수송 효율을 증가시키기 때문에 검출 감도가 증가하는 것이라고 이해할 수 있다. 두번째로, 이동상의 점성 감소는 분사 가스에 의하여 이동상이 분사될 때 더 작은 에어로졸(aerosol)을 만들게 하며 이것은 역시 분사된 액체방울이 drift tube 벽면에 응결하는 것을 막는다.



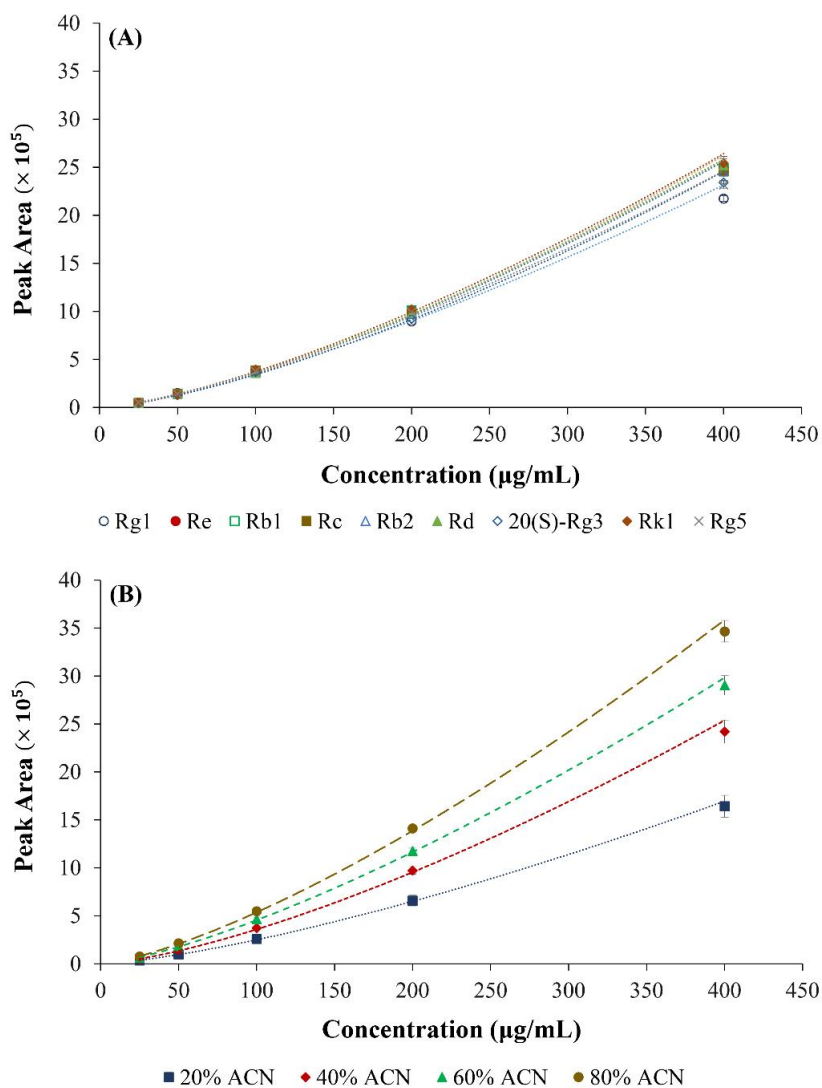
Ginsenoside	Formula	MW	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>
Ginsenoside Rg1	C <sub>42</sub> H <sub>72</sub> O <sub>14</sub>	801.01	OH	Oglc	Oglc
Ginsenoside Re	C <sub>48</sub> H <sub>82</sub> O <sub>18</sub>	947.15	OH	Oglc(2-1)rha	Oglc
Ginsenoside Rb1	C <sub>54</sub> H <sub>92</sub> O <sub>23</sub>	1109.29	Oglc(2-1)glc	H	Oglc(6-1)glc
Ginsenoside Rc	C <sub>53</sub> H <sub>90</sub> O <sub>22</sub>	1079.27	Oglc(2-1)glc	H	Oglc(6-1)ara(f)
Ginsenoside Rb2	C <sub>53</sub> H <sub>90</sub> O <sub>22</sub>	1079.27	Oglc(2-1)glc	H	Oglc(6-1)ara(p)
Ginsenoside Rd	C <sub>48</sub> H <sub>82</sub> O <sub>18</sub>	947.15	Oglc(2-1)glc	H	Oglc
Ginsenoside Rg3	C <sub>42</sub> H <sub>72</sub> O <sub>13</sub>	785.01	Oglc(2-1)glc	H	OH



Ginsenoside	Formula	MW	R <sub>1</sub>	R <sub>2</sub>
Ginsenoside Rk1	C <sub>42</sub> H <sub>70</sub> O <sub>12</sub>	766.99	Oglc(2-1)glc	H

Ginsenoside	Formula	MW	R <sub>1</sub>	R <sub>2</sub>
Ginsenoside Rg5	C <sub>42</sub> H <sub>70</sub> O <sub>12</sub>	766.99	Oglc(2-1)glc	H

**Figure 13.** Structures and molecular weights of the selected ginsenosides. ara(f), arabinose (furanose); ara(p), arabinose (pyranose); glc, glucose; rha, rhamnose.

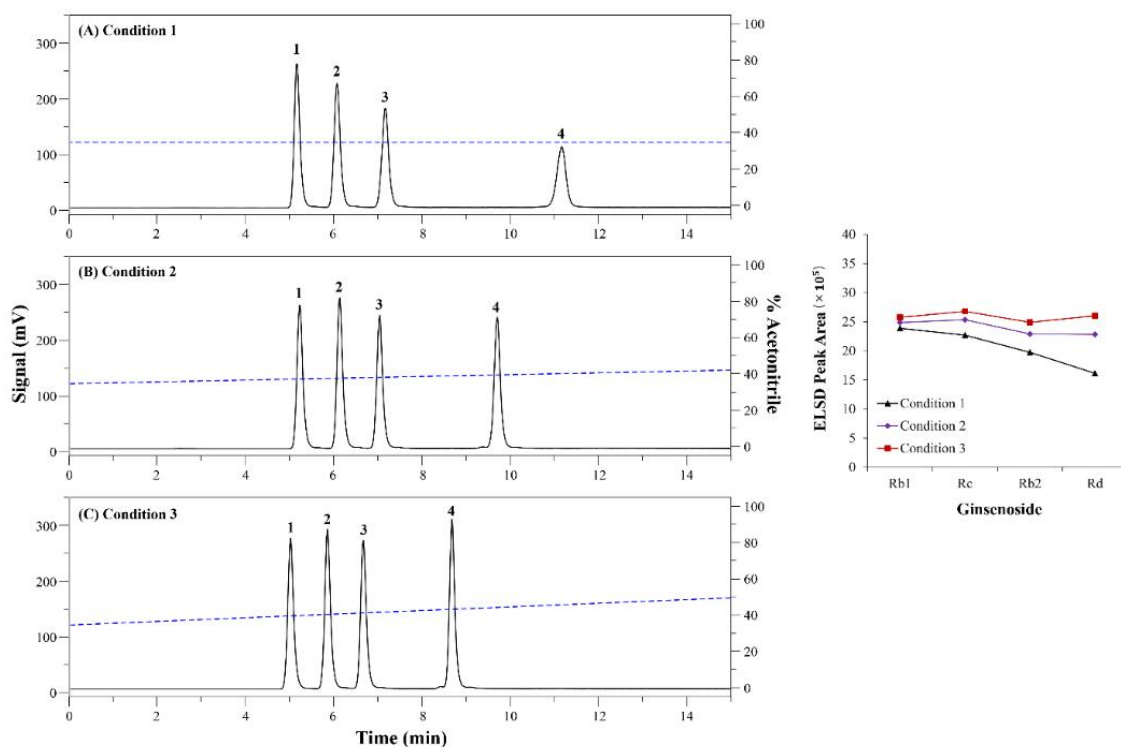


**Figure 14.** ELSD response curves for the nine ginsenosides in flow injection analysis. Flow injection analysis was performed in triplicate. (A) The response curves for the individual ginsenosides using 40% acetonitrile as a mobile phase composition (The error bars indicate the standard deviation of repeated analyses) and (B) The response curves constructed from average responses of the nine ginsenosides using four different mobile phase compositions: 20, 40, 60, and 80% acetonitrile (The error bars indicate the standard deviation of responses of the nine ginsenosides).

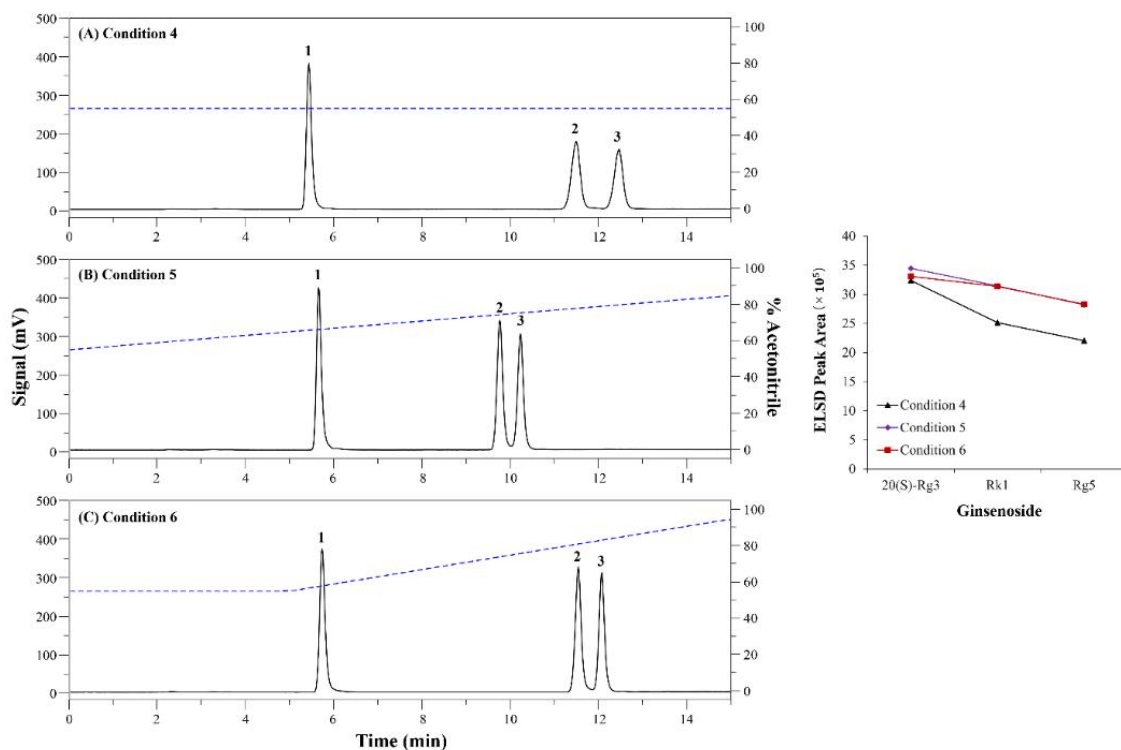
## 2.2. Peak bandwidth가 검출 감도에 미치는 영향

Peak bandwidth가 ELSD의 검출 감도에 미치는 영향을 확인하기 위하여 PPT, PPD, LP 그룹으로 9개의 Ginsenoside를 구분하였으며 각 그룹 Ginsenoside의 화학적 특성에 맞게 22, 35, 55% 아세토니트릴의 isocratic조건을 적용하여 실험을 진행하였다. PPD (Figure 15A), LP (Figure 16A) 그룹에서 고정상에 의해 강하게 유지된 분석체의 경우 isocratic 조건에서 넓은 피크의 형태로 용리 되었다. 9개의 Ginsenoside가 flow injection 실험에서 비슷한 검출 감도를 보여줬음에도 불구하고 두 실험에서 모두 Peak bandwidth가 넓어짐에 따라 검출 감도가 감소하는 것을 확인할 수 있었다. PPD 그룹 실험에서 Ginsenoside Rd의 검출 감도는 Ginsenoside Rb1의 검출 감도의 67%였다. 마찬가지로 LP 그룹 실험에서 Ginsenoside Rg5의 검출 감도는 Ginsenoside Rg3의 검출 감도의 68%만큼 줄어들었다. 이 결과는 아마도 분석체가 검출되는 단계에서 광산란 메커니즘들의 경쟁을 통해 발생된 결과로 보여진다. 몇몇 메커니즘들 중, 레일리 산란(Rayleigh scattering)은 입자의 직경이 파장의 0.1배 이하일 때 광산란에 크게 영향을 미치는 것으로 알려져 있으며 특히 이 때의 광산란 정도는 입자 사이즈의 6제곱에 비례하는 것으로 알려져 있다 [47]. 또한 분사 단계에서의 분사체의 농도는 정비례하게 검출 단계의 입자 사이즈를 결정하는 것으로 알려져 있다 [48]. 따라서 넓은 피크에서는 단위 시간당 낮은 분사체의 농도가 형성되고 최종적으로 검출 단계에서 더 작은 입자 사이즈를 만들며 결국 이 때의 광산란 정도는 레일리 산란의 영향에 의하여 크게 감소한다. PPT 그룹에 속하는 Ginsenoside Rg1과 Re는 해당 조건에서 거의 비슷한 Peak bandwidth로 용리 되기 때문에 그 결과를 생략하였다.





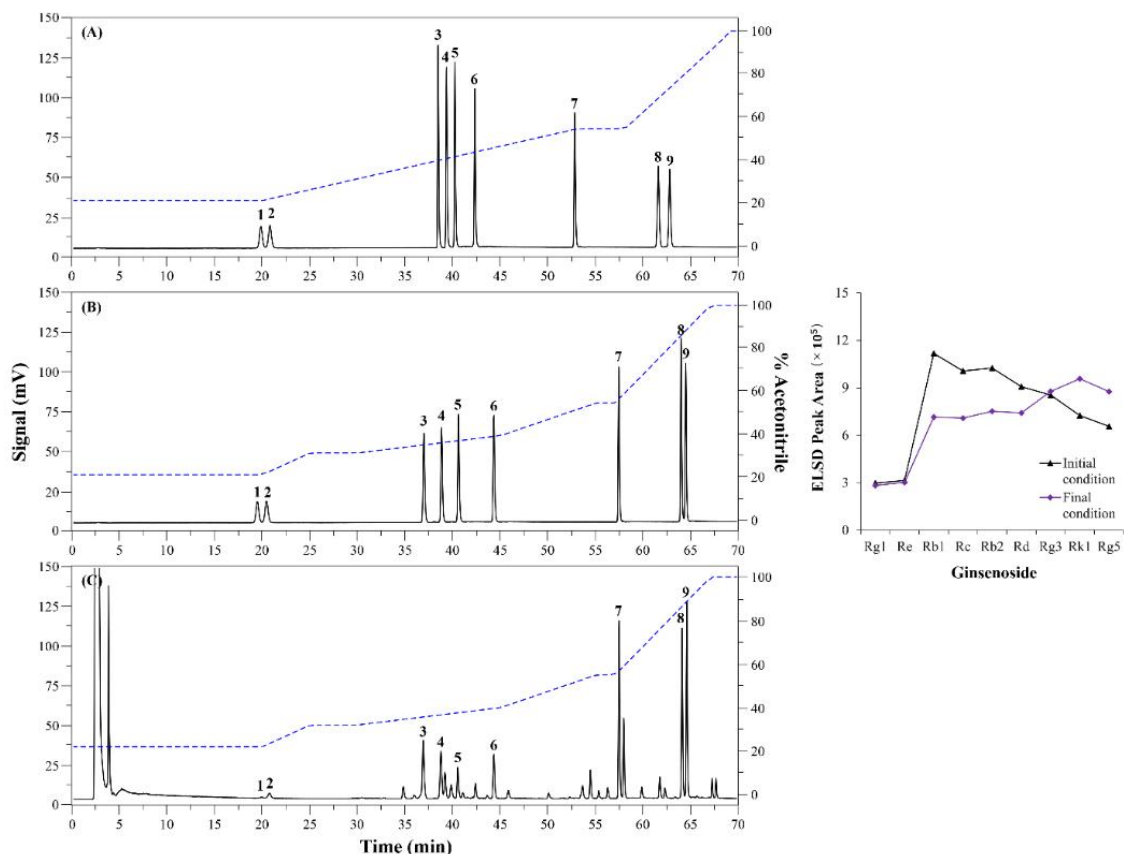
**Figure 15.** Evaluation of elution bandwidth effect on the ELSD response and stepwise changes of the gradient condition for the response normalization in the protopanaxadiol (PPD) group. Chromatographic conditions: (A) 35% acetonitrile for 15 min (Condition 1), (B) increase from 35% acetonitrile to 42.5% acetonitrile for 15 min (Condition 2) and (C) increase from 35% acetonitrile to 50% acetonitrile for 15 min (Condition 3). Elution order: (1) ginsenoside Rb1, (2) ginsenoside Rc, (3) ginsenoside Rb2 and (4) ginsenoside Rd.



**Figure 16.** Evaluation of elution bandwidth effect on the ELSD response and stepwise changes of the gradient condition for the response normalization in the less-polar (LP) group. Chromatographic conditions: (A) 55% acetonitrile for 15 min (Condition 4), (B) increase from 55% acetonitrile to 85% acetonitrile for 15 min (Condition 5) and (C) 55% acetonitrile for 5 min and increase from 55% acetonitrile to 95% acetonitrile for 10 min (Condition 6). Elution order: (1) ginsenoside 20(S)-Rg3, (2) ginsenoside Rk1 and (3) ginsenoside Rg5.

### 2.3. 순차적 Gradient 변경을 통한 분석 조건 최적화

Figure 15, 16는 두 중요 요소를 고려하여 순차적인 Gradient 변경을 통해 검출 감도를 평준화 하는 과정을 보여준다. PPD 그룹에서 조건 3은 (Figure 15C) 강하게 고정상과 상호 작용하던 Ginsenoside Rd의 검출 감도를 특히 많이 상승시켰고 (이동상 극성 상승에 의한 영향 + Peak bandwidth 감소에 의한 영향), 그 결과 4개의 Ginsenoside의 검출 감도를 비슷한 수준으로 맞출 수 있었다. 이 때 4개의 Ginsenoside의 감도 편차는 3.04%에 불과했지만 조건 1에서의 편차는 16.73% 였다. LP 그룹에서는 조건 5처럼 (Figure 16B) 급격한 상승의 Gradient만으로는 검출 감도를 평준화 할 수 없었다. 따라서 Ginsenoside Rg3의 Peak bandwidth는 유지한 채 다른 성분의 검출 감도를 증가시키기 위해 조건 6과 같은 (Figure 16C) 부분적인 Gradient를 사용하였다. 그 결과 검출 감도를 평준화 할 수 있었다. 이 때 3개의 Ginsenoside 감도 편차는 7.86%였으며 반면에 조건 4에서의 편차는 20.02%였다. 각 그룹 마다 최적화된 조건들은 전체 Ginsenoside 및 원료 농축물을 분석하기 위해 연결되었다 (Figure 17A). 하지만, 먼저 용리 되는 Ginsenoside들을 위한 조건이 다음에 용리 될 Ginsenoside들에 영향을 미쳐 만족스러운 결과(검출 감도 평준화와 성분들의 완전한 분리)를 얻지 못하여 미세한 수정을 추가하였다 (Figure 17B). 수정에 대한 세세한 내용은 분석 시스템 환경에 따라 다르기 때문에 본 연구 결과에서 다루지 않는다. 최종적으로 최적화 된 Gradient 조건에서 PPT, PPD, LP 그룹 내 Ginsenoside들의 감도 편차는 각각 4.49, 2.85, 5.14%에 불과했다. 요약하면, ELSD 검출 감도는 실제 분석에서 분석 조건의 수정에 따라 쉽게 조절될 수 있으며 이것은 분석 시스템에서 큰 변화 없이 상대 정량을 가능하게 할 수 있음을 의미한다.



**Figure 17.** HPLC-ELSD chromatograms of (A) mixed standards by the initial condition, simple connection of the three optimized conditions: 0-20 min (22% acetonitrile, PPT group), 20-53 min (22-55% acetonitrile, Condition 3 in Figure 15, PPD group), 53-58 min (55% acetonitrile, Condition 6 in Figure 16, LP group), 58-68 min (55-95% acetonitrile, Condition 6 in Figure 16, LP group), (B) mixed standards by the final condition and (C) processed ginseng concentrate by the final condition. Elution order: (1) ginsenoside Rg1, (2) ginsenoside Re, (3) ginsenoside Rb1, (4) ginsenoside Rc, (5) ginsenoside Rb2, (6) ginsenoside Rd, (7) ginsenoside 20(*S*)-Rg3, (8) ginsenoside Rk1 and (9) ginsenoside Rg5.

## 2.4. 상대 정량법의 유효성 검증

분석법의 Validation 결과는 Table 5에 요약하였다. 직선성( $R^2$ )의 경우 모든 성분이 0.995 이상으로 1에 가까웠다. 실제로 넓은 농도 범위에서 측정하였을 때 ELSD 검량선은 sigmoidal 형태에 가깝다. 이는 레일리 산란, 미 산란, 반사&굴절 이론과 같은 광산란 메커니즘들의 상호작용에 의한 결과이다. 하지만 대부분의 분석에서 높은 농도는 활용하지 않기 때문에 이 부분을 배제하면 검량선의 형태는 지수 함수(power function)의 형태와 가깝다. 따라서 ELSD 검량선의 직선성을 측정 할 때에는 x, y 변수를 로그화 하여 측정하는 것이 일반적이다. LOD와 LOQ의 범주는 각각 0.070 - 0.125  $\mu\text{g}$ , 0.120 - 0.250  $\mu\text{g}$ 였다. 기존 ELSD 연구 결과와 달리 본 실험에서는 용리가 진행 될수록 검출 한계와 정량 한계가 낮아지는 결과를 확인할 수 있었다. 이것은 PPT 그룹의 Ginsenoside를 분석 할 때 낮은 %의 아세토니트릴과 넓은 peak bandwidth를 형성하는 isocratic 조건을 이용해서 감도가 낮았던 반면 LP 그룹의 Ginsenoside를 분석 할 때는 높은 %의 아세토니트릴과 좁은 peak bandwidth를 형성하는 급격한 gradient 조건을 사용해서 감도가 높았기 때문이다. 이것으로부터 ELSD의 검출, 정량 한계 역시 분석 조건의 수정을 통해 개선될 수 있음을 확인하였다. 9개 Ginsenoside를 분석했을 때 분석법의 정확성은 최소 93.77%에서 107.61%였으며 그 편차는 최대 4.47%였다. Table 6은 정밀도의 측정 결과를 요약한 것이며 해당 조건에서 상대 정량을 진행하였을 때 직접 정량을 했을 때와 정밀도 차이가 거의 없음을 보여준다. 분석법의 불확정도는 시료 준비 과정, 표준품 희석 과정, 시스템 편차, 실험자 차이로부터 오는 편차 등을 포함하는 모든 실험변인들을 독립적인 것으로 가정하고 Validation 결과를 바탕으로 측정하였다. Table 7은 확립한 분석 조건에서 상대

정량과 직접 정량을 진행하였을 때 발생하는 개별 불확정도와 전체적 불확정도를 요약한 표다. 전체적 불확정도에 보상계수(coverage factor) 2를 곱함으로써 신뢰도 95% 수준에서 두 가지 방법의 최종 불확정도를 확인할 수 있었다. 그 결과 두 방법 모두 10% 이하의 작은 불확정도를 나타냈다.

**Table 5**

Regression data, LOD, LOQ and recovery for the nine ginsenosides

Ginsenoside	Calibration curve <sup>a</sup>	$R^2$	LOD ( $\mu\text{g}$ )	LOQ ( $\mu\text{g}$ )	Recovery <sup>b</sup> (n=3)	
					Mean (%)	RSD <sup>c</sup> (%)
Rg1	$y=1.6908x+1.6250$	0.9957	0.125	0.250	103.36	1.90
Re	$y=1.7522x+1.4948$	0.9977	0.125	0.250	93.77	3.03
Rb1	$y=1.6652x+1.9421$	0.9965	0.090	0.190	93.83	1.36
Rc	$y=1.6903x+1.9413$	0.9967	0.090	0.190	106.47	3.35
Rb2	$y=1.6621x+2.0059$	0.9970	0.085	0.170	105.38	3.92
Rd	$y=1.6583x+2.0346$	0.9978	0.085	0.170	107.61	4.47
Rg3	$y=1.6099x+2.2404$	0.9970	0.070	0.120	99.13	0.98
Rk1	$y=1.6268x+2.2295$	0.9962	0.070	0.120	102.15	0.51
Rg5	$y=1.6132x+2.2088$	0.9951	0.070	0.120	104.80	2.58

<sup>a</sup> y, log-transformed value of peak area; x, log-transformed value of concentration ( $\mu\text{g}/\text{mL}$ )

<sup>b</sup> Recovery (%) = (observed amount – original amount)/spiked amount  $\times 100\%$

<sup>c</sup> RSD (%) = standard deviation/mean  $\times 100\%$

**Table 6**

Intra- and inter-day precision of the direct and group quantification method for the nine ginsenosides (n=3)

Ginsenoside (group <sup>a</sup> )	Direct quantification				Group quantification			
	Intra-day		Inter-day		Intra-day		Inter-day	
	precision		precision		precision		precision	
	Content (mg/g)	RSD <sup>b</sup> (%)	Content (mg/g)	RSD (%)	Content (mg/g)	RSD (%)	Content (mg/g)	RSD (%)
Rg1 <sup>c</sup> (PPT)	1.16±0.05	4.57	1.16±0.04	3.22	1.16±0.05	4.57	1.16±0.04	3.22
Re (PPT)	3.25±0.09	2.76	3.06±0.09	2.85	3.33±0.09	2.64	3.10±0.12	3.78
Rb1 (PPD)	7.02±0.15	2.14	6.83±0.24	3.53	6.47±0.14	2.11	6.25±0.23	3.67
Rc <sup>d</sup> (PPD)	5.61±0.08	1.44	5.49±0.25	4.62	5.61±0.08	1.44	5.49±0.25	4.62
Rb2 (PPD)	3.88±0.10	2.55	3.84±0.10	2.61	3.97±0.10	2.49	4.01±0.12	2.95
Rd (PPD)	4.68±0.08	1.63	4.58±0.08	1.75	4.80±0.08	1.57	4.70±0.09	1.83
Rg3 <sup>e</sup> (LP)	8.13±0.07	0.81	7.92±0.30	3.76	8.13±0.07	0.81	7.92±0.30	3.76
Rk1 (LP)	7.58±0.11	1.50	7.40±0.19	2.63	7.80±0.12	1.51	7.65±0.21	2.78
Rg5 (LP)	8.85±0.15	1.68	8.67±0.18	2.08	8.43±0.14	1.71	8.33±0.15	1.84

<sup>a</sup> Ginsenoside group designations: PPT, protopanaxatriol group; PPD, protopanaxadiol group; LP, less-polar group

<sup>b</sup> RSD (%) = standard deviation/mean × 100%

<sup>c</sup> Representative standard of PPT group in group quantification method

<sup>d</sup> Representative standard of PPD group in group quantification method

<sup>e</sup> Representative standard of LP group in group quantification method



**Table 7**

Individual and global uncertainty of the direct and group quantification method for the nine ginsenosides (calculated at quantification result)

Ginsenoside (group)	Calibration curve		Precision				Accuracy				Global uncertainty		Expanded uncertainty	
	$U_1$		$U_2$		$U_3$		$U_4$		$U_5$		$U$		$2U$	
	DQ <sup>a</sup>	GQ <sup>b</sup>	DQ	GQ	DQ	GQ	DQ	GQ	DQ	GQ	DQ	GQ	DQ	GQ
Rg1 <sup>c</sup> (PPT)	0.017	0.017	0.019	0.019	0.026	0.026	0.011	0.011	0.013	0.013	0.041	0.041	0.081	0.081
Re (PPT)	0.015	0.016	0.016	0.022	0.016	0.015	0.018	0.018	0.021	0.021	0.039	0.042	0.078	0.083
Rb1 (PPD)	0.016	0.015	0.020	0.021	0.012	0.012	0.008	0.008	0.010	0.010	0.031	0.031	0.062	0.063
Re <sup>d</sup> (PPD)	0.015	0.015	0.027	0.027	0.008	0.008	0.019	0.019	0.024	0.024	0.044	0.044	0.088	0.088
Rb2 (PPD)	0.015	0.015	0.015	0.017	0.015	0.014	0.023	0.023	0.028	0.028	0.044	0.045	0.089	0.089
Rd (PPD)	0.015	0.015	0.010	0.011	0.009	0.009	0.026	0.026	0.032	0.032	0.045	0.046	0.091	0.091
Rg3 <sup>e</sup> (LP)	0.012	0.012	0.022	0.022	0.005	0.005	0.006	0.006	0.007	0.007	0.027	0.027	0.054	0.054
Rk1 (LP)	0.016	0.012	0.015	0.016	0.009	0.009	0.003	0.003	0.004	0.004	0.024	0.022	0.049	0.045
Rg5 (LP)	0.018	0.012	0.012	0.011	0.010	0.010	0.015	0.015	0.018	0.018	0.034	0.030	0.067	0.060

<sup>a</sup> Direct quantification

<sup>b</sup> Group quantification

<sup>c</sup> Representative standard of PPT group in group quantification method

<sup>d</sup> Representative standard of PPD group in group quantification method

<sup>e</sup> Representative standard of LP group in group quantification method

## 2.5. 직접 정량법과의 비교

Table 8은 인삼 제품 원료 농축물을 상대 정량과 직접 정량을 진행했을 때 Ginsenoside 함량의 비교 결과를 보여준다. Ginsenoside Rg1과 Rg3는 성분 자체의 연구 빈도(공정서 등), 확보의 용이성 측면을 고려하여 대표 표준물질로 선정하였고 Ginsenoside Rc는 검출 감도 측면에서 PPD 그룹 내 중간에 위치하여 상대 정량 시 오차를 줄일 수 있어 선정하였다. 상대 정량에서는 위 세 가지 표준물질의 검량선들이 각 그룹 내 Ginsenoside들을 정량 하기 위한 검량선으로 사용된다. 정량 결과, 상대 정량을 진행했을 때의 결과가 직접 정량을 진행했을 때의 결과와 큰 차이가 없음을 확인하였다. 가장 큰 차이는 Ginsenoside Rb1에서 나타난 8.26%였으며, 전체 함량의 차이는 0.91%에 불과했다. 각 그룹끼리 비교(Subtotal)했을 때의 차이도 최대 1.67%에 불과했다.

**Table 8**

Comparison of contents for the nine ginsenosides between the direct and group quantification method (n=5)

Ginsenoside (group)	Direct quantification		Group quantification		$\Delta$ Content <sup>b</sup> (%)
	Content A (mg/g)	RSD <sup>a</sup> (%)	Content B (mg/g)	RSD (%)	
Rg1 <sup>c</sup> (PPT)	1.15±0.04	3.35	1.15±0.04	3.35	0.00
Re (PPT)	3.16±0.15	4.64	3.21±0.17	5.42	1.58
Subtotal	4.31±0.19	4.41	4.36±0.21	4.82	1.16
Rb1 (PPD)	6.90±0.22	3.21	6.33±0.22	3.45	8.26
Rc <sup>d</sup> (PPD)	5.52±0.19	3.48	5.52±0.19	3.48	0.00
Rb2 (PPD)	3.87±0.14	3.60	3.98±0.16	4.00	2.84
Rd (PPD)	4.63±0.10	2.10	4.74±0.10	2.10	2.38
Subtotal	20.92±0.65	3.11	20.57±0.67	3.26	1.67
Rg3 <sup>e</sup> (LP)	7.99±0.23	2.88	7.99±0.23	2.88	0.00
Rk1 (LP)	7.47±0.19	2.50	7.70±0.19	2.46	3.08
Rg5 (LP)	8.75±0.19	2.21	8.37±0.16	1.90	4.34
Subtotal	24.21±0.61	2.52	24.06±0.58	2.41	0.62
Total	49.44±1.45	2.93	48.99±1.46	2.98	0.91

<sup>a</sup> RSD (%) = standard deviation/mean × 100%

<sup>b</sup>  $\Delta$ Content = |Content A – Content B|/Content A × 100%

<sup>c</sup> Representative standard of PPT group in group quantification method

<sup>d</sup> Representative standard of PPD group in group quantification method

<sup>e</sup> Representative standard of LP group in group quantification method

## IV. 결론

현재 천연물 원료 품질평가를 위해 사용되는 일반적인 지문 비교 방법을 보완하기 위해 복합 유사성 지표 체계인 CAM을 개발하였다. 이 체계를 이용해서 기존 방법과 달리 다각도로 두 지문의 화학적 유사성을 더 정확하게 평가할 수 있었다. CAM값을 바탕으로 커널 밀도를 얻고 역치를 정함으로써 구체적인 시료의 품질을 결정하였다. 이 추정치는 연구나 생산 초기와 같이 품질 평가에 대한 사전 정보나 지식이 없는 상황에서도 객관적인 근거 하에 평가를 가능하게 하는 것이 장점이었다. 전체적인 체계는 *Atractylodes*속 시료에 적용하여 그 유효성을 검증하였다. 기존에 사용되던 체계와 비교하였을 때 본 개발 체계를 통해 품질이 입증된 시료들은 전반적으로 기존 시료와 더 화학적으로 동등함을 확인하였다. 연구 초기부터 엄격하게 원료를 통제하는 방법이 우수한 제품 개발과 일관적인 생산에 핵심 기술임을 감안하면 개발 체계는 의약품, 기능성 식품, 식품 등 관련 산업에 널리 활용될 수 있을 것이다.

또한 지표 성분의 함량을 경제적으로 정량 분석할 수 있도록 ELSD를 이용한 상대 정량법을 개발하였다. 소비자들은 제품에 대한 다양한 성분 정보를 요구하지만 많은 성분들을 정량 분석하려면 고비용의 표준물질 확보가 지속적으로 수반되어야 한다. 본 개발 방법은 추가적인 기기 설치, 첨가물, 통계 모델 등의 필요가 없이 ELSD 검출 감도에 영향을 주는 인자들을 고려하여 간단하게 분석 조건을 수정함으로써 상대 정량을 가능하게 했다는 것에 큰 의미를 갖는다. 이 방법은 인삼 원료 분석에 적용되었고 그 결과 직접 정량과 거의 차이가 없음을 확인하였다. 이 방법론은 다른 원료에도 경제적인 품질 평가를 위해 적용될 수 있을 것이다.

## V. 참고문헌

1. Liang, Y.-Z.; Xie, P.; Chan, K. *J. Chromatogr. B* 2004, 812, 53–70.
2. Li, S.; Zhao, J.; Yang, B. *J. Pharm. Biomed. Anal.* 2011, 55, 802–809.
3. European Medicines Agency, Guideline on quality of herbal medicinal products/traditional herbal medicinal products, 2011, [http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general\\_content\\_000724.jsp&mid=WC0b01ac0580033a9b](http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000724.jsp&mid=WC0b01ac0580033a9b) (accessed 22.03.17).
4. Food and Drug Administration, Botanical Drug Development: Guidance for Industry, 2016, <https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm064979.htm> (accessed 22.03.17).
5. World Health Organization, General Guidelines for Methodologies on Research and Evaluation of Traditional Medicine, 2000, <http://apps.who.int/medicinedocs/en/d/Jwhozip42e> (accessed 22.03.17).
6. Xu, C.; Wang, Y.; Chen, J.; Zhou, Q.; Wang, P.; Yang, Y.; Sun, S. *J. Pharm. Biomed. Anal.* 2013, 74, 298–307.
7. Li, C.-Y.; Lin, C.-H.; Wu, C.-C.; Lee, K.-H.; Wu, T.-S. *J. Agric. Food Chem.* 2004, 52, 3721–3725.
8. Lesiak, A. D.; Cody, R. B.; Dane, A. J.; Musah, R. A. *Anal. Chem.* 2015, 87, 8748–8757.
9. Chen, S.-B.; Liu, H.-P.; Tian, R.-T.; Yang, D.-J.; Chen, S.-L.; Xu, H.-X.; Chan, A. S.; Xie, P.-S. *J. Chromatogr. A* 2006, 1121, 114–119.
10. Tian, R.-T.; Xie, P.-S.; Liu, H.-P. *J. Chromatogr. A* 2009, 1216, 2150–2155.
11. Wei, H.; Sun, L.; Tai, Z.; Gao, S.; Xu, W.; Chen, W. *Anal. Chim. Acta* 2010, 662, 97–104.
12. Sun, X.; Chen, P.; Cook, S. L.; Jackson, G. P.; Harnly, J. M.;

- Harrington, P. B. *Anal. Chem.* 2012, 84, 3628–3634.
13. Chen, Y.; Zhu, S.-B.; Xie, M.-Y.; Nie, S.-P.; Liu, W.; Li, C.; Gong, X.-F.; Wang, Y.-X. *Anal. Chim. Acta* 2008, 623, 146–156.
  14. Song, W.; Qiao, X.; Chen, K.; Wang, Y.; Ji, S.; Feng, J.; Li, K.; Lin, Y.; Ye, M. *Anal. Chem.* 2017, 89, 3146–3153.
  15. Huang, L.-F.; Wu, M.-J.; Zhong, K.-J.; Sun, X.-J.; Liang, Y.-Z.; Dai, Y.-H.; Huang, K.-L.; Guo, F.-Q. *Anal. Chim. Acta* 2007, 588, 216–223.
  16. Zhai, H. L.; Li, B. Q.; Tian, Y. L.; Li, P. Z.; Zhang, X. Y. *Food Chem.* 2014, 145, 625–631.
  17. Kong, W.-J.; Zhao, Y.-L.; Xiao, X.-H.; Wang, J.-B.; Li, H.-B.; Li, Z.-L.; Jin, C.; Liu, Y. *Anal. Chim. Acta* 2009, 634, 279–285.
  18. Tistaert, C.; Dejaegher, B.; Vander Heyden, Y. *Anal. Chim. Acta* 2011, 690, 148–161.
  19. Goodarzi, M.; Russell, P. J.; Vander Heyden, Y. *Anal. Chim. Acta* 2013, 804, 16–28.
  20. Alaerts, G.; Van Erps, J.; Pieters, S.; Dumarey, M.; Van Nederkassel, A.; Goodarzi, M.; Smeyers-Verbeke, J.; Vander Heyden, Y. *J. Chromatogr. B* 2012, 910, 61–70.
  21. Yang, L.-W.; Wu, D.-H.; Tang, X.; Peng, W.; Wang, X.-R.; Ma, Y.; Su, W.-W. *J. Chromatogr. A* 2005, 1070, 35–42.
  22. Jin, Y.; Liang, T.; Fu, Q.; Xiao, Y.-S.; Feng, J.-T.; Ke, Y.-X.; Liang, X.-M. *J. Chromatogr. A* 2009, 1216, 2136–2141.
  23. Zhu, H.; Wang, C.; Qi, Y.; Song, F.; Liu, Z.; Liu, S. *Talanta* 2013, 103, 56–65.
  24. Xu, C.-J.; Liang, Y.-Z.; Chau, F.-T.; Vander Heyden, Y. *J. Chromatogr. A* 2006, 1134, 253–259.
  25. Gong, F.; Wang, B. T.; Chau, F. T.; Liang, Y. Z. *Anal. Lett.* 2005, 38, 2475–2492.
  26. Xu, X.; Jiang, J.; Liang, Y.; Yi, L.; Cheng, J. *Anal. Methods* 2010, 2, 2002–2010.

27. Choi, K. H.; Jeong, S. I.; Lee, J. H.; Hwang, B. S.; Kim, S. J.; Lee, S.; Choi, B. K.; Jung, K. Y. *Phytomedicine* 2011, 18, 408–413.
28. Song, H.-P.; Li, R.-L.; Chen, X.; Wang, Y.-Y.; Cai, J.-Z.; Liu, J.; Chen, W.-W. *J. Ethnopharmacol.* 2014, 152, 163–172.
29. G. Guiochon; A. Moysan; C. Holley. *J. Liq. Chromatogr.* 1988, 11, 2547–2570.
30. Y. Mengerink; R. Peters; C. Dekoster; S. Van der Wal; H. Claessens; C. Cramers. *J. Chromatogr. A* 2001, 914, 131–145.
31. N. Adnani; C.R. Michel; T.S. Bugni. *J. Nat. Prod.* 2012, 75, 802–806.
32. N.C. Megoulas; M.A. Koupparis. *Crit. Rev. Anal. Chem.* 2005, 35, 301–316.
33. N.C. Megoulas; M.A. Koupparis. *J. Chromatogr. A* 2004, 1057, 125–131.
34. B. Mathews; P. Higginson; R. Lyons; J. Mitchell; N. Sach; M. Snowden; M. Taylor; A. Wright. *Chromatographia* 2004, 60, 625–633.
35. A. De Villiers; T. Gorecki; F. Lynen; R. Szucs; P. Sandra. *J. Chromatogr. A* 2007, 1161, 183–191.
36. M.M. Khandagale; E.F. Hilder; R.A. Shellie; P.R. Haddad. *J. Chromatogr. A* 2014, 1356, 180–187.
37. C.H. Lee; J.-H. Kim. *J Ginseng Res.* 2014, 38, 161–166.
38. S.-K. Kim; J.H. Park. *J Ginseng Res.* 2011, 35, 389–398.
39. L.P. Christensen. *Adv. Food Nutr. Res.* 2008, 55, 1–99.
40. S.W. Kwon; S.B. Han; I.H. Park; J.M. Kim; M.K. Park; J.H. Park. *J. Chromatogr. A* 2001, 921, 335–339.
41. Zhang, S.; You, X.; Jin, Z.; Wu, X. *Expert Syst. Appl.* 2009, 36, 10863–10869.
42. Uncertainty, EURACHEM/CITAC Guide CG 4, EURACHEM/CITAC Guide, Quantifying Uncertainty in Analytical Measurement, third ed., 2012, <http://eurachem.org/>.

43. Tarassenko, L.; Hayton, P.; Cerneaz, N.; Brady, M. In *4th International Conference on Artificial Neural Networks*: Cambridge, 1995, pp 442–447.
44. Tarassenko, L.; Nairac, A.; Townsend, N.; Buxton, I.; Cowley, P. *Int. J. Syst. Sci.* 2000, 31, 1427–1439.
45. Botev, Z. I.; Grotowski, J. F.; Kroese, D. P. *Ann. Stat.* 2010, 38, 2916–2957.
46. F.S. Deschamps; A. Baillet; P. Chaminade. *Analyst* 2002, 127, 35–41.
47. J.H. Seinfeld, S.N. Pandis, *Atmospheric chemistry and physics: from air pollution to climate change*, John Wiley & Sons, 2012.
48. M. Righezza; G. Guiochon. *J. Liq. Chromatogr.* 1988, 11, 1967–2004.



## Abstract

The inconsistent quality of herbal materials caused by biodiversity highlights the importance of a quality evaluation system which can rigorously filter out herbal materials not equivalent to the qualified ones. For this purpose, both the fingerprinting method and the marker determination method are currently used. In the present study, we propose better ways to improve the methods commonly used in each approach. Firstly, we developed a mixture of three similarity indices for the first time, which is termed as CAM, to make up for weak points of some widespread similarity indices in the similarity analysis of chromatographic herbal fingerprints. It was suitable for the fingerprints with large deviations in the data and offered in-depth information on the similarity of the samples to the reference. Forty *Atractylodes* samples were the case, and their three similarity values (CAM values) were fed into a three-dimensional kernel density estimation for determining their qualities. As a result of comparing to the common approach, the samples qualified by the proposed approach were more chemically analogous to the reference. Secondly, we developed a relative quantification method using an evaporative light scattering detector (ELSD) for determining multi-components in herbal materials economically. The potential variables, mobile phase composition and elution bandwidth, affecting the ELSD response in the gradient analysis were investigated. With the consideration of the two important effects, stepwise changes of the gradient condition were carried out, and the inconsistent responses from different analytes could be normalized. The availability was confirmed by comparing the ginsenoside contents in a material of ginseng products determined by the direct and relative quantification method. The largest difference

in the determination results from the two methods was 8.26%, and the difference of total contents was only 0.91%.

**Keywords:** Herbal materials, Chromatographic fingerprints, Relative quantification, Chemical equivalence, Quality assurance

**Student Number:** 2012-21606