

## 근적외선분광광도계(NIRS)를 이용한 국내산 콩과 수입콩의 판별분석

안형균\* · 김용호\*<sup>†</sup>

\*충남 아산시 신창면 읍내리 순천향대학교 의료생명공학과

### Discrimination of Korean Domestic and Foreign Soybeans using Near Infrared Reflectance Spectroscopy

Hyung-Gyun Ahn\* and Yong-Ho Kim\*

\*Dept. of Medical Biotechnology, Soonchunhyang Univ., Asan-si, 336-745, Korea

**ABSTRACT** Discrimination of geographic origin of agricultural products is a important issue in Korea because the price difference between Korean domestic and imported cereals is a key among some reasons. NIRS (Near Infrared Reflectance Spectroscopy) has been applied to classify the geographical origin of soybeans. Total 135 samples (Korean domestic 92 and foreign 43) were used to obtain calibration equation through 400~2,500 nm wavelength. The math treatment with 1st derivative and 4 nm gap and the modified partial least squares(MPLS)regression was outstanding for calibration equation. The standard error of calibration and determination coefficient in calibration set(n=115) was 6.65 and 0.98, respectively. And it showed that the extra 20 samples for validation equation were identified their authentication correctly. This study describes that the application of NIRS would be possible for discrimination of geographical origin between Korean domestic and imported soybeans.

**Keywords** : soybean, NIRS, geographical origin, math treatment, MPLS regression

#### NIRS(Near Infrared Reflectance Spectroscopy)

는 근적외선분광분석법을 응용하는 분석기기로 1960년대 Karl Norris에 의해 곡류의 수분, 지방, 단백질의 분석에 이용된 이후 현재는 농산물뿐만 아니라 여러 분야에서 이용되고 있다. NIRS는 시료의 추출이나 화학반응 등의 조작이 필요하지 않으므로 비파괴분석이 가능하며, 동시에 여러 분석항목을 신속하게 측정할 수 있을 뿐만 아니라 측정된 시료는 원래의 상태로 회수가 가능하므로 반복적인 사용으로

재현성 있는 분석이 가능하다는 장점이 있다. 최근에는 근적외선분광분석법에 chemometrics가 접목되어(Blanco & Villarroya, 2002) 정량분석 외에도 다방면에서 응용되고 있는데 NIRS는 식품, 농산물, 의약품의 성분 분석(Kim *et al.*, 2008; Son *et al.*, 2008; Kim *et al.*, 2004; Cha *et al.*, 2002) 뿐만 아니라 원산지 판별에도 이용되고 있다.

원산지 판별 관련 연구로는 Downey(1996)이 8개의 주성분 분석으로 커피 품종 및 혼합물을 구별할 수 있다고 하였으며, Bertan *et al.*(2000)은 올리브유의 원산지 판별분석을 하였고 Li *et al.*(2007)은 chineeses bayberry의 원산지 분석을 하였다. 국내에서는 Kim *et al.*(2003)은 쌀의 원산지 판별, Woo *et al.*(1998)은 생약의 원산지 판별을 시도한 바 있으며, Chun and Cho(2007), Kim *et al.*(1997)은 NIRS를 이용하여 국내산과 일본산 녹차제품을 구분하였고, Lee *et al.*(2003)은 송이버섯의 원산지 판별을 보고한 바 있다. 이 밖에도 아스파라거스(Kim *et al.*, 2003), 고추(Kwon *et al.*, 1999), 참기름(Kwon and Cho, 1998) 등의 원산지 판별이 보고된 바 있다.

본 논문에서는 NIRS를 이용하여 국내산 콩과 수입콩의 판별분석이 가능하였기에 그 결과를 보고하고자 한다.

#### 재료 및 방법

##### 공시 재료

실험에 사용된 콩의 총 시료 수는 135개이며, 이중 국내산이 92개, 수입콩이 43개(중국 22, 미국 6, 대만 6, 호주 3, 러시아 2, 캐나다 2, 브라질 1, 우크라이나 1)이다. 국내산은

<sup>†</sup>Corresponding author: (Phone) +82-41-530-1281 (E-mail) yohokim@sch.ac.kr

<Received 30 July , 2012; Revised 3 September, 2012; Accepted 6 September, 2012>

작물과학원과 경기도농업기술원에서 2009년에 재배된 콩을 분양받았으며, 수입콩은 2009년에 수입된 콩을 국립식물검역소와 (주)풀무원 식문화연구원에서 공급받아 사용하였다. 모든 공시재료는 황색종피를 가진 일반콩들이었다.

### NIRS 분석

콩 20 g을 30~40°C에서 24시간 건조 후 분쇄기(Ika, A11-basic)로 40 mesh 크기로 조절하고 분말 10 g을 근적외선분광광도계(NIR spectrum Model 6500, Foss)를 사용하여 400~2,500 nm 범위에서 2 nm 간격으로 spectrum을 얻었다. Spectrum은 중첩으로 인해 생기는 noise와 bias를 줄이기 위해 수처리를 하게 된다. 수처리는 우선 spectrum을 대상으로 1차 및 2차 미분을 수행한 후 미분된 각 spectrum을 다시 1차 스무딩과 spectrum gap을 몇 가지 처리로 조합하여 수처리 하였으며 2차 스무딩은 사용하지 않았다. 국내산 콩과 수입콩의 판별 검량식 작성은 Intrasoftware International Software WINISI II 프로그램을 사용하였으며, NIR loading value를 국내산 콩과 수입콩을 달리 지정한 후 변형부분최소자승법(MPLS: modified partial least square)를 사용한 회귀분석법에 의하여 작성되었다. 검량식의 정확도는 calibration equation과 validation equation을 통해 확인을 하였다. Calibration에 사용된 시료 수는 115점으로 이중 국내산이 80점, 수입콩이 35점이었고, validation에 사용된 시료는 총 20점으로 국내산 12점, 외국산 8점이 포함되었다.

## 결과 및 고찰

### 국내산 콩과 수입콩의 NIRS 스펙트럼

국내산 콩과 수입콩의 NIRS 흡수 스펙트럼의 평균적인 모습은 Fig. 1과 같았다. 바탕선의 차이 등에 따른 스펙트럼의 변이는 있었으나 전체적인 스펙트럼의 양상은 국내산과 콩과 수입콩이 유사함을 알 수 있었다.

일반적으로 NIRS 스펙트럼은 시료의 형태에 관계없이 측정할 수 있으나 그 흡수대가 겹치거나 측정하는 물질의 화학적 성분, 입자의 크기 및 밀도 같은 물리적 영향에 의해 바탕선의 변화가 일어나는데 이러한 오차를 줄이고 겹쳐 있는 파장을 분리하기 위해 미분법을 통한 수처리를 한다. Fig. 2는 Fig. 1의 스펙트럼을 1차 미분하여 수처리한 스펙트럼이다. 미분된 스펙트럼에서는 수처리 전의 스펙트럼과는 다르게 파장별로 어느 정도 흡광도의 차이를 볼 수 있는데, 1,162, 1,418, 1,700, 1,906, 2,049 nm 부근에서 흡광도의 차이가 크게 나타났다. 한편, 가시광선영역에서의 변이

가 근적외선 영역의 변이보다 크게 나타났는데 이는 콩 종피의 색도 차이에 기인하는 것으로 판단된다. 이와 같이 원래의 스펙트럼을 수처리를 통해 1차 미분함으로써 바탕선의 변화 제거 및 겹친 흡수대의 분리 등으로 스펙트럼 간의 차이를 확인할 수 있었다.

### 국내산 콩과 수입콩의 판별 검량식 작성

수처리된 스펙트럼은 MPLS 회귀분석법을 사용하여 검량식이 작성되었다. NIRS의 검량식 작성에는 여러 가지 회귀분석법이 가능하나 본 연구에서는 MPLS 회귀분석법이 가장 우수하였다. MPLS 회귀분석법은 cross validation을 이용하여 NIRS의 전체파장(400~2,500 nm)과 lab data와의 상관관계를 유도하는 과정에서 최적의 factor를 선정하여 overfitting을 방지하고 정확도를 높여주기 때문에 스펙트럼에 나타나는 기준선 변화, 산란, 중첩 등의 영향을 최소화하여 비교적  $R^2$  값이 높은 검량식을 유도할 수 있다. Kim *et al.*(2003)은 NIRS를 이용한 쌀의 원산지 판별에서 MPLS 회귀분석법을

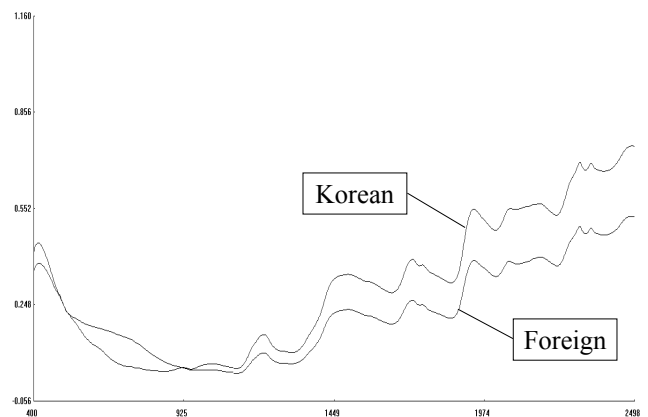


Fig. 1. Mean value of NIRS spectra of Korean domestic and foreign soybeans.

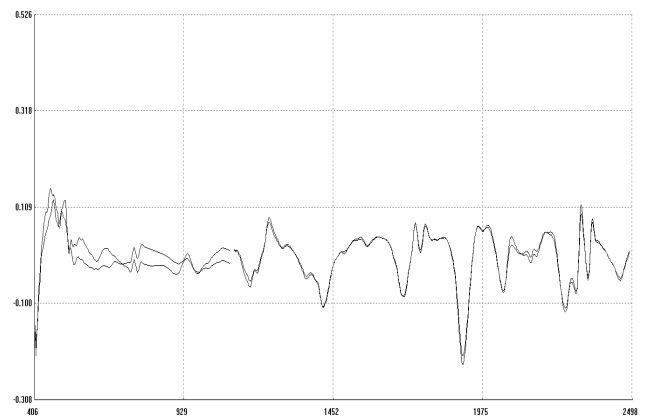


Fig. 2. The 1st derivative NIR spectra of soybeans.

이용하여 예측모델을 개발하였으며, Lee *et al.*(2003)은 송이버섯의 원산지 판별을 MPLS법을 사용하여 수행한 바 있다.

Table 1은 MPLS 회귀분석법으로 도출된 검량식 결과이다.

국내산 콩 loading value를 ‘100’으로 하고, 수입콩 loading value를 ‘1’로 설정한 다음, 1차 미분한 스펙트럼을 MPLS 회귀분석법을 적용하여 국내산과 수입콩을 분석을 하였다. 검량식 작성 후에는 다시 CV(cross validation, 교차검증)을 사용하여 내부검증을 하게 된다. 교차검증은 검량식 작성에 사용된 시료를 제외한 나머지 시료를 이용하여 도출된 검량식을 예측 검증하는 방법으로 1-VR과 SECV 값으로 정확도를 검증한다. 이러한 검량식 작성에서 가장 어려운 결정 중의 하나는 최적화된 factor(파장요인)의 수를 결정하는 일이다. 일반적으로 factor에 따른 R<sup>2</sup>, SEC 및 SECV를 보면 factor의 수가 많아질수록 R<sup>2</sup> 과 SEC 값이 좋아지나 반면에 SECV가 증가하므로 적절한 factor를 선택해야 over fitting

**Table 1.** Statistics of calibration and validation equations for discrimination between Korean domestics and foreign soybeans.

Math treatment	Factor	SEC	R <sup>2</sup>	SECV	1-VR
1, 4, 4, 1*	1	21.69	0.75	22.79	0.72
	2	19.93	0.80	21.16	0.77
	3	16.62	0.86	19.13	0.81
	4	11.92	0.93	15.82	0.87
	5	10.14	0.95	13.66	0.91
	6	9.31	0.96	13.00	0.91
	7	7.92	0.97	12.13	0.92
	8	7.26	0.97	11.73	0.93
	9	7.26	0.97	11.73	0.93
	10	6.54	0.98	11.16	0.94

\* Math treatment 1, 4, 4, 1 : 1 = derivative funtion, 4 = gap, 4= 1st smooth, 1 = 2st smooth

SEC = standard error of calibration; R<sup>2</sup> = squared coefficient of multiple determination in calibration; SECV = standard error of validation; 1-VR = squared coefficient of multiple determination in validation

**Table 2.** Validation statistics values for discrimination between Korean domestics and foreign soybeans.

Math treatment	Set No.	SEP	Mean	r <sup>2</sup>
1, 4, 4, 1	20	10.59	52.35	0.94

SEP = standard error of performance; r<sup>2</sup> = squared coefficient of multiple determination in validation

이 일어나지 않고 정확한 검량식을 유도할 수 있다.

Table 1에서 보는 바와 같이 factor가 증가됨에 따라 calibration의 SEC는 감소하고 상관값인 R<sup>2</sup>이 증가하며 validation의 상관값인 1-VR도 증가하였다. Factor 가 10일 때, R<sup>2</sup>는 0.98로서 우수한 검량식을 얻을 수 있었으며, 이 검량식을 cross validation한 결과 SECV는 11.1574, 1-VR는 0.94를 보여 검량식의 정확도가 높음을 알 수 있었다. Table 2는 validation 결과로써 loading value의 평균값이 52.35로 나타났으며 상관값이 0.94를 보여 Table 1 검량식의 정확성을 확인할 수 있었다. 따라서 NIRS에서 본 검량식을 이용하면 국내산 콩과 수입콩의 판별이 용이하게 수행될 수 있을 것으로 판단되었다.

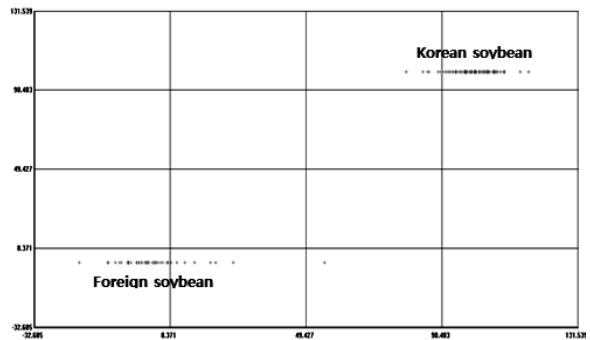
Table 3은 작성된 검량식을 사용하여 국내산 콩과 수입콩을 실제 판별한 결과이다. 국내산 콩 70점과 수입콩 25점을 사용하여 판별한 결과 Table에서 보는 바와 같이 국내산 콩은 국내산 콩으로 수입콩은 모두 수입콩으로 인지하였으며, 잘못되거나 불명확하게 판별되는 것은 없었다.

Fig. 3은 loading value를 국내산 콩은 ‘100’, 수입콩은 ‘1’로 설정한 후 MPLS 회귀분석법으로 도출된 판별식에 의한 상관도를 나타낸 것이다. 국내산 콩과 수입콩은 그룹별로 뚜렷하게 구분됨을 알 수 있었다.

한편, 농산물의 원산지 판별은 Howard(1992)에 의해 정립된 discrimination equation에 의해서도 가능하다. Discrimination equation은 주성분분석(principle component analysis, PCA)

**Table 3.** Results of discriminant analysis between Korean domestics and foreign soybeans by NIRS.

Calibration equation	Kor. cal	Foreign.cal	Total	Misses	Uncertain
Kor.cal	70	0	70	0	0
Foreign.cal	0	25	25	0	0



**Fig. 3.** Relationship between Korean domestic and foreign soybeans by MPLS regression on NIRS.

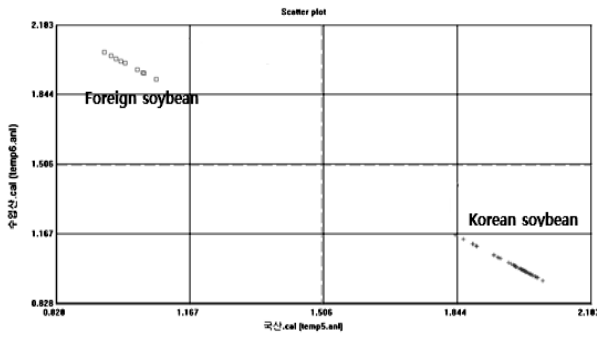


Fig. 4. Relationship between Korean domestic and foreign soybeans by discrimination equation on NIRS.

## 적 요

국내산 콩과 수입콩의 판별에 NIRS를 도입함으로써 보다 빠르고 정확한 식별분석을 하고자 실험을 수행하였다. NIRS를 사용하여 400~2,500 nm 범위에서 콩 분말의 파장을 측정하였으며, 측정된 spectrum은 WINISI II program을 이용하여 수처리와 회귀분석을 하였다. 검량식 작성을 위한 수처리는 spectrum을 1차미분 및 4 nm gap으로 조정하는 것이 가장 적합하였으며, 회귀식은 변형부분최소자승회귀법(Modified partial least squares regression)이 우수하였다. MPLS 회귀분석시 원산지 판별을 위해 loading value를 국내산 콩은 '100', 수입콩은 '1'로 처리하여 검량식을 작성하고 그 적합성을 검증한 결과 factor가 10일 때 도출된 calibration equation의 상관값이 0.98, 교차검증의 상관값이 0.94를 나타내어 상관도가 높음을 알 수 있었다. 따라서 NIRS를 이용한 국내산 콩과 수입콩의 판별분석이 가능할 것으로 판단되었다.

## 인용문헌

의 중앙 평균에서의 절대거리를 측정함으로써 시료의 동질성 여부를 판단하는 것으로 유도된 PCA 값으로 two-block PLS(partial least square regression)방법에 의해 집단을 2개로 구분한 후 집단간의 PLS score에 의하여 집단의 차이를 판별한다. Fig. 4는 국내산을 '2', 수입콩을 '1'로 loading value를 설정한 다음 discrimination equation로 분석된 상관도를 나타낸 것이다. 그림에서 보는 바와 같이 기준치 1.506으로 국내산 콩과 수입콩의 판별이 가능함을 알 수 있었는데 본 연구에 사용된 국내산 시료는 1.84 이상을, 수입콩은 1.16 이하의 값을 나타내었다.

국내에서 NIRS를 이용한 discrimination equation로 농산물의 원산지 판별을 수행한 결과를 보면 Chun & Cho(2007)는 한국산과 일본산 녹차 판별에서 79.5%의 정확도를 보고한 반면 Kim *et al.*(1997)은 97.9%의 정확도를 보고한 바 있다. 이러한 결과의 차이는 주성분분석에 사용된 PC의 수에 영향을 받은 것으로 사료된다. 따라서 면밀한 주성분분석이 뒷받침 된다면 NIRS를 이용한 discrimination equation 방법도 농산물의 원산지 판별에 이용될 수 있을 것으로 판단된다.

현재 농산물의 원산지 판별은 미량원소 분석법, 동위원소 이용, DNA 분석 등 다양한 방법이 시도되고 있는데 이들 방법들이 비교적 긴 시간과 실험방법이 복잡한 반면, 본 연구의 결과 등에서 보는 바와 같이 NIRS를 사용하면 손쉽게 원산지 판별이 가능한 것으로 판단된다. 그러나 원산지 판별에서는 정확성이 무엇보다도 중요하지만 NIRS 분석법이 아직까지는 다른 방법들에 비해 정확도가 떨어지는 것은 사실인 것 같다. 그러나 NIRS를 이용한 원산지 판별에 다양한 품종 및 산지가 다른 시료 등 충분한 유전자원이 확보되어 검량식이 유도된다면 정확하고 신속한 원산지 판별이 가능할 것으로 판단된다.

- Blanco, M. and I. Villarroya. 2002. NIR spectroscopy : a rapid -response analytical tool. *Trends in analytical Chem.* 21(4) : 240~250.
- Bertan, E., M. Blanco, J. Coello, H. Iturriaga, S. Maspocho, and I. Montoliu. 2000. Near infrared spectrometry and pattern recognition as screening methods for the authentication of virgin olive oils of very close geographical origins. *J. Near infrared spectrosc.* 8 : 45~52.
- Cha, K. W., K. R. Ze, M. O. Youn, S. J. Lee, H. C. Choi, and H. J. Kim. 2002. Determination of phenobarbital in intact phenobarbital tablets using NIRS. *Analytical Sci. & Technol.* 15(2) : 102~107.
- Chun, J. U. and K. C. Cho. 2007. Discrimination of Korean and Japanese green tea products using NIR Spectroscopy. *J. Kor. Tea Soc.* 13(1) : 93~104.
- Downey, G. 1996. Authentication of coffee bean variety by near infrared reflectance spectroscopy of dried extract. *J. Sci. Food Agric.* 71 : 41~49.
- Howard, M. 1992. Qualitative discriminant analysis. In *Handbook of near infrared analysis*. p. 329~363. Marcel Dekker, Inc. NY.
- Kim, E. Y., J. H. Kim, N. Y. Lee, S. J. Kim, and M. R. Rhyu. 2003. Discrimination of geographical origin for astragalus root (*Astragalus membranaceus*) by capillary electrophoresis and near infrared spectroscopy. *Korean J. Food Sci. Technol.* 35(5) : 818~824.
- Kim, J. S., Y. H. Cho, J. G. Gwag, K. H. Ma, Y. M. Choi, J. B. Kim, J. H. Lee, T. S. Kim, J. K. Cho, and S. Y. Lee. 2008.

- Quantitative analysis of amylopectin and protein content of rice germplasm in RDA-Genbank by near infrared reflectance spectroscopy. *J. Korean. Crop. Sci.* 53(2) : 217~223.
- Kim, S. S., M. R. Rhyu, J. M. Kim, and S. H. Lee. 2003. Authentication of rice using Near-Infrared Reflectance Spectroscopy. *Cereal Chem.* 80(3) : 346~349.
- Kim, Y. S., C. Scotter, M. Voyiagis, and M. Hall. 1997. Potential of NIR spectroscopy for discrimination the geographical origin of green tea from Korean and Japan. *Foods and Biotechnol.* 6(2) : 74~78.
- Kim, Y. H., C. S. Kang, and Y. S. Lee. 2004. Quantification of tocopherol and tocotrienol content in rice bran by near infrared reflectance spectroscopy. *Korean J. Crop Sci.* 49 : 211~215.
- Kwon, H. S., S. J. Lee, S. J. Kim, S. S. Chung, and J. H. Kim. 1999. Discrimination of geographical origin and seed content in red pepper powder by near infrared reflectance spectroscopic analysis. *J. Korean Oil Chemists. Soc.* 16(2) : 155~161.
- Kwon, Y. K. and R. K. Cho. 1998. Identification of geographical origin of sesame seeds by near infrared spectroscopy. *Agric. Chem. Biotechnol.* 41 : 240~246.
- Lee, N. Y., H. R. Bea, and B. S. Noh. 2003. Discrimination of Geographical Origin of Mushroom (*Tricholoma matsutake*) using near infrared spectroscopy. *Korean J. Food Sci. Technol.* 35(2) : 835~837.
- Li X, Y. He, and H. Fang. 2007. Nondestructive discrimination of Chinese bayberry varieties using Vis/NIR spectroscopy. *J. Food Engineering.* 81 : 357~363.
- Son, J. Y., K. J. Lee, S. W. Kang, and W. K. Choi. 2008. Quality evaluation of sugar content for grapes using NIR Spectroscopy. *Food Engineering Progress.* 12(4) : 263~268.
- Woo, Y. A., H. J. Kim, J. H. Cho, H. Chung. 1998. Discrimination of herbal medicines according to geographical original with near infrared reflectance spectroscopy and pattern recognition techniques. *J. Pharm. Biomed. Anal.* 21 : 407~413.