Original Article

DISCRIMINANT ANALYSIS OF *SHODHANA* (PROCESSING) ON *BALIOSPERMUM MONTANUM* MUELL (*DANTI*) ROOT SAMPLES BASED ON NEAR INFRARED SPECTROSCOPY AND MULTIVARIATE CHEMOMETRIC TECHNIQUE

SIBA PRASAD ROUT[1*], RABINARAYAN ACHARYA[2], JAYANTA KUMAR MAJI[3]

[1,2]Department of Dravyaguna, IPGT and RA, Gujarat Ayurved University, Jamnagar, Gujarat, India 361008, [3]Department of Pharmaceutical Laboratory, IPGT and RA, Gujarat Ayurved University, Jamnagar, Gujarat, India 361008
Email: drrout.sibaprasad@gmail.com

ABSTRACT

Objective: To establish a noticeable and a justifiable identification system to assess the impact of *shodhana* (processing) on various levels of *Baliospermum montanum (Danti)* root samples obtained through *shodhana* (processing technique) in quality agreement based on near-infrared-spectroscopy.

Methods: Authenticated raw *Danti* (R. D) root and various *Danti* root samples obtained after *shodhana* (processing) such as water processed *Danti* root (WPDR), *Kusha* processed *Danti* root (KPDR) and classical processed *Danti* root (CPDR), were dried, pulverized and shifted through eighty meshes. The samples were subjected to NIR spectral detection from 750 to 2500 nm at the interval of 1 nm. The multivariate analysis, principal component analysis (PCA) and hierarchical cluster analysis (HCA) analyzed with the help of Unscrambler and Matlab software.

Results: Direct spectral analysis indicated the existence of significant numerical and graphical differences between *Danti* root samples containing different treatments during processing in respect to CH, OH and NH functional groups. The multivariate PCA algorithom plot allowed a clear segregation of the *Danti* root samples after various data preprocessing technique onto the hotelling T2 95% confidence limit for principal component 1 and 2. The cluster analysis had shown the extra information on the metabolite profiling of the complex purificatory environment.

Conclusion: The present study demonstrates a generic, non-destructive solution to discriminate qualitatively in the sample matrix all the differently pretreated samples in favor of the NIR-sensitive functional group.

Keywords: *Danti*, Multivariate, PCA, Discrimination. Functional groups

INTRODUCTION

Baliospermum montanum Muell (*Danti*) root, which is sold in the name of "*Danti* mula" by drug dealers in the market is one of the most recommended drugs reported to be used as a purgative due to its resin and starch contain [1]. Medicinal plants like *Danti* (*Baliospermum montanum* Willd); root, *Vacha* (*Acorus calamus* Linn.); rhizome, *Vriddhadaru* (*Argyreia speciosa* Sweet.); seed, *Hingu* (*Ferula foetida* Bioss.); exudate,, *Kampillaka* (*Mallotus philippensis* (Lamk.) Muell-Arg.);fruit and *Guggulu* (*Commiphora mukul* (Hook. ex. Stocks) Engl.);gum resin etc. have not been categorized under poisonous drugs in API, still they have been recommended to pass through the specific *shodhana (*processing*)* process before internal administration[2]. According to ayurvedic principle, drug act as a whole (hostile environment), they are more attached with synergistic approach due to the presence of various primary and secondary plant metabolites other than modern science reductionist approach. On the other hand, near-infrared spectroscopy (NIR) encompasses acquisition of a large number of absorbencies in a spectral region typically ranging from 1000 to 2500 nm, where organic matter has distinct spectral fingerprints related to functional groups [3]. The forthcoming trend, NIR spectroscopy has taken wide acceptance within the pharmaceutical industry for raw material testing, intermediate product, finished product quality control and process monitoring [4, 5]. Along with the availability of chemometric evaluation procedures for qualitative discrimination and quantitative determination is new opportunity to overcome data evaluation. It fulfills all the regulatory demands as any other "chemically "raw materials and finished product materials 21 CFR 11 guidelines, food and pharma GMP guidelines, USP chapter 1119 [6]. In these circumstances, the NIR spectra taken from a sample are representative of the sample as a whole. If drug are used as a whole that there is no need for extensive sample preparation like as HPLC-PDA, LC-MS etc. NIR analysis is useful because a sample may be

rapidly tested without destroying its integrity which retains the ayurvedic philosophy drug act as a whole. Applying multivariate chemometric cluster analysis easily separated materials of raw burdock root powder and concentrated burdock root extract [7].

This paper presents a detailed study on the multivariate fingerprinting comparison of various processed *Danti* samples in favor of nondestructive method i.e. near infrared spectroscopy. The objective of the study was to determine diffused reflectance on the electromagnetic spectrum ranges from 700-2500 nm (12821-4000 cm-1) of authentify *Shodhana Danti* medicament; analyze highly absorbing functional group (alkyl, phenyls, amines, thiols, hydroxyls, acids, esters etc.) on favor of dried pulverized fine powder of respective contents, which are classical processed *Danti* root (C. P. D. R), *Kusha* processed *Danti* root (K. P. D. R) and water processed *Danti* root (W. P. D. R), raw *Danti* root (R. D); determine distinguishing different material to compute principal component analysis after preprocessing of the raw diffuse reflection data using mathematical transformation; establish in library which significantly influence real-time quality monitoring and discuss easy clear-cut identifying tool to qualify respective *Danti* samples. NIR spectroscopy is not only a fingerprint of chemical composition but also a signature of the physical state with the order of material under investigation.

MATERIALS AND METHODS

Plant materials and chemicals

Roots of *Baliospermum montanum,* Muell were collected from Bolangir forest area of Odisha in the month of February 2016. Then the plant herbariums were authenticated by taxonomist of pharmacognosy laboratory of I. P. G. T and R. A, Gujarat Ayurved University, Jamnagar (voucher specimens No. PHM/6208/2015-16) for further references. Other accessory raw drugs used for processing of *Danti* root were

collected, then dried properly under sun drying and stored in an air-tight glass container for further processing.

Standard operating procedure for *Shodhana* of *Baliospermum montanum* Muell (*Danti*) in different environment

Collected raw *Danti* (R. D) root being washed properly was smeared with a thin layer of paste prepared from powder of *Pippali* (*Piper longum* Linn.) and honey then wrapped with leaves of *Kusha* (*Desmostachya bipinnata* Stapf). The resultant was coated with mud and fomented with steam at temperature 125 °C for 3 h [8].

This process was adopted for three times obtaining 3 batches of classical processed *Danti* root (CPDR). Then the obtained three batches of roots were dried in sun rays, then grinded with a mechanical grinder, assembled, finally, the coarse powders were separated by sieving using 80 meshes and stored in an airtight glass container for further use. Individual powders at different levels of *Danti Shodhana* like *Kusha* processed *Danti* root (KPDR) group and Water processed *Danti* root (WPDR) group was also obtained in the same way.

Near infrared spectroscopy

Samples were analyzed using a bench top Perkin Elmer Lambda 19 UV-VIS-NIR spectrophotometer system in the range between 1000-2500 nm in diffused reflectance mode. The various *Danti* powders were placed in a closed rotating sample lead sulfide cell cup at a resolution of 2 nm^{-1} wavelength accuracy and acquisition rate were specified as<0.05 nm and 2 scans per second, respectively. The contact probe was placed against a block surface, and spectral data were collected. The spread between the spectra of each material is characteristic of reflectance spectra of *Danti* powders. The particle size of these spectra is represented by the numerical values corresponding to the reflectance of the material at each wavelength (one-nanometer interval). Instrumentation selected standard for the wavelength and absorption calibration, validation traceable to NIST [9]. The NIR spectroscopy study was done NABL accredited laboratory Sciart at Anand, Gujarat.

Data analysis

Data were manipulated into two forms: 1) Data were exported as 2D ASCII files at all NIR wavelength region at 1 nm interval to a corresponding reflectance value to build a matrix respective four samples i.e., wavelengths (λ) are one direction and samples in other direction; 2) Data were exported as ASCII files to reduced reflectance value of whole wavelength (a single profile by four coded samples giving a matrix). All data operations (preprocessing, PCA, HCA) were performed using Unscrambler and MATLAB R2008 (Math works) on a computer Intel Pentium 4 processor containing 500 MB RAM and running Microsoft Windows seven.

Data preprocessing

Raw reflectance is converted to absorbance using the function, Absorbance =-log (10)* Reflectance, commonly referred to as log (1/R). Raw data were submitted to chemometric analysis with first order derivative preprocessing for multivariate distribution. Several preprocessing methods, including standard normal variates (SNV) [10], 2nd derivative [11], smoothing (Savitzy goloy derivative) [12], baseline correction (BC) [13] were applied to the dataset before principal component analysis.

Chemometrics technique

Principal component analysis (PCA)

PCA is a mathematical method which decomposes a set of observations (set of spectra) into a set of principal components (PC$_S$). Concerning spectroscopy, a set of spectra can, therefore, the expressed by several orthogonal 'factor spectra' called loadings. Furthermore, each spectrum can be recombined by a linear combination of those loadings using individual scores. Only few PC$_S$ are necessary to characterize a spectral system it is, therefore, possible to visualize differences between spectra by plotting the scores of those PC$_S$ against each other. Thus the position of a sample in a score plot describes the characteristics of the measured sample. PCA is a method for transforming data based on eigenanalysis. PCA algorithms were used in this study for reducing the high-dimensional data by constructing a linear combination of the original variable into a few orthogonal principle components which contain most of the variability of the data set. [14] The mathematical formula-

$A = T_K V_K^T + \in$ [T_K is the n × k matrix of the Principal component score, V_K is the m× k matrix of eigenvectors, loading, $\in$ = residual error].

HCA

Hierarchical clustering is the visualization tool as a supervised approach (classification approach). It is used usually based on object wise similarities or distances. One of the most valid arguments is that most clustering methods are heuristic in nature with solid statistical foundation. Dendrogram plotting can handle visualize more than three independent variables using a hierarchical family tree like structure. The algorithm performs n-1 steps where n is the sample size. In all the steps it calculates the Euclidean distance of the two closest samples to a new one with, for example, single coordinate, complete coordinate (single linkage, complete linkage algorithom). After n-1 steps, a dendrogram can be drawn to visualize hierarchically in which samples are most probably comparable even considering many independent variables. This tool is very useful if classification using PCA along a sample set needs to consider more than three factors [15]. Each distance is generally converted value s_{ik} = 1-d_{ik}/d_{max} [s_{ik} is the measure of similarity between samples i and k, d_{ik} is the Euclidean distance between samples i and k, d_{max} is the distance between the two most dissimilar distance].

RESULTS AND DISCUSSION

Multivariate analysis

NIR fingerprints of *Danti* powders

The NIR spectrum displayed characteristic "fingerprints'' for various treated *Danti* samples (fig. 1: A) along with directed functional groups of organic compound (table 4)[16]. The multivariate distribution used as a secondary measure of plant chemistry NIR sensitive mainly (C=H, O=H, N=H, S=H) due to the interaction of the electromagnetic radiation of the samples. Solid forms of powders are measured by reflectance-mode to generate the discriminating criteria of the different treatment in the face of the extraneous source of variation (various treatment of *Danti* sample during processing) in the measured spectra.

Particularly for NIR Spectra of the resultant materials usually show noise and baseline fluctuation but data processing such as after derivatization also removes the significant amount of signal, resulting in a lower signal to noise ratio in the derivative curve. The different baseline offset among the spectra, which are typically due to differences in the particle size along with purification effect of the measured spectra. The first derivative of the spectra (fig. 1: C) highlights the key differences between the respected samples.

A key feature corresponding relative intensity difference of the respective *Danti* samples powder on the various generic functional group at approximately (1410, 1415, 1417, 1420, 1440, 1446, 1450,1471,1483, 1490,1500,1510, 1620, 1685, 1695, 1705, 1860, 1900, 1908, 1920, 2030 nm) which are depicted on [table 2].[17]. On the other hand, the remaining frequencies give more or less give same response pattern in the context of a functional group (alkyl, phenyl, amines, thiols, hydroxyl, Acids) (table 1).

Actually, various band described to free, and terminal OH and NH groups of the polymeric species may be differentiated in the NIR region. Larger anharmonicity, band ascribed to the first overtone of OH and NH stretching modes of monomeric species are enhanced compared with the corresponding bands arising from poly component species. Since NIR spectra are typically composed of broad overlapping ill-defined absorption bands containing chemical and physical information of all sample components.
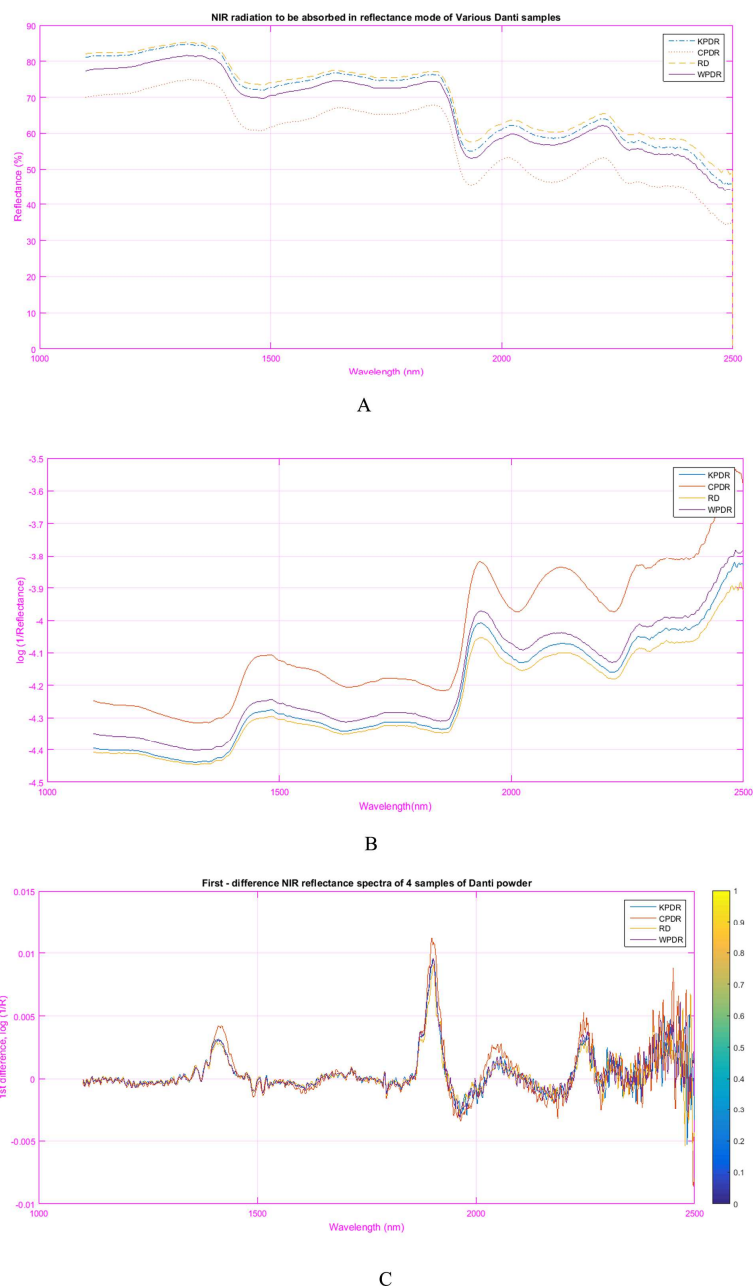
A

B

C

Fig. 1: NIR reflectance spectra of various *Danti* samples, Raw (A), Pre-processed (B) and (C) the first derivative

Table 1: Major analytical bands and relative peak positions for prominent real-infrared absorption region respective functional group on combination and overtone mode

Highly absorbing functional group	Overtone mode (NIR region) nm	Combination mode(NIR region)nm
Alkyl	C-H 1st overtone-(1650-1750) C-H 2nd overtone(1100-1150) C-H 4th overtone (700-750)	C-H (1350-1450) C-H+C. H (2200-2300) C-H+C. C (2300-2400).
Phenyl (Aromatic)	1143 2nd overtone, CH stretch 1st overtone	1417,1446–CH, 1492 Aromatic NH combination
Amines	N-H 1st overtone(1450-1500), N-H 2nd overtone(1000-1050) N-H 3rd overtone(800-850)	N-H combination (2050-2600)
Thiols	S-H 1st overtone (1740)	
Hydroxyl	O-H 1st overtone-(1400-1450) O-H 2nd overtone(900-950) O-H third overtone (700-750)	O-H combination(1900-2000)
Acids (Acetamide)	C==O (1920) stretch second overtone	1980,Asym N−−H stretch/amide IIb combination, N−−H/Amide IIb or 2050,CONH N−−H/Amide IIIb or combination

Table 2: Selected wavelength region where after 1st order difference of samples is discriminated

| Wavelength(nm)	Bond vibration	Structural assignment
1410	O-H first overtone	ROH
1415	C-H combination	. CH_2
1417	C-H combination	Aromatic
1420	O-H first overtone	ArOH
1440	C-H combination	. CH_2
1446	C-H combination	Aromatic
1450	O-H strech first overtone	Starch
1471	N-H stretch first overtone	CONHR
1483	N-H stretch first overtone	. $CONH_2$
1490	O-H,N-H stretch first overtone	Cellulose, Urea
1500	N-H stretch first overtone	. NH
1510	N-H stretch first overtone	Protein
1620	C-H strech first overtone	Aromatic
1685	C-H strech first overtone	Aromatic
1695	C-H strech first overtone	. CH_3
1705	C-H strech first overtone	. CH_3
1860	C-Cl strech sixth overtone	C-Cl
1900	C=O stretch fourth overtone	-COOH
1908	O-H strech first overtone	P-OH
1920	C=O stretch second overtone	-CONH
2030	C=O stretch second overtone	Urea
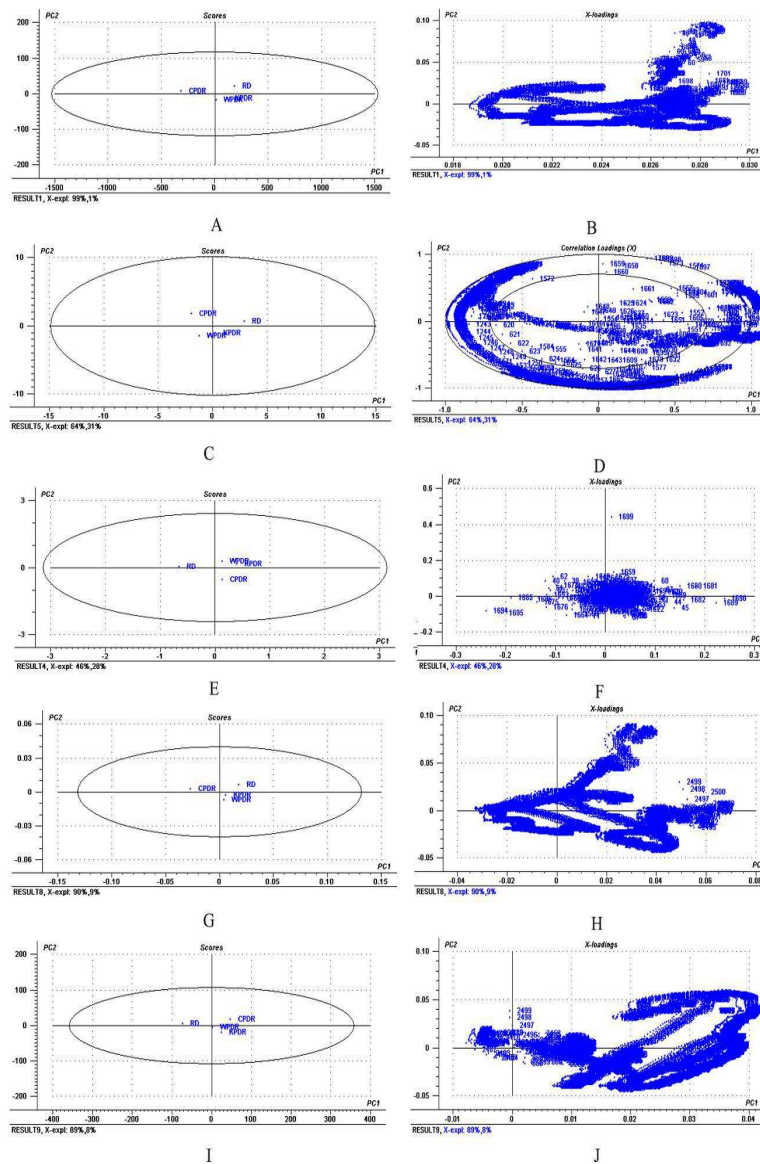
Fig. 2: PCA plot of the four type of *Danti* samples after various preprocessing technique, A and B (raw data), C and D (standard normal variates), E and F (2nd Derivative), G and H (Savitzy goloy derivative), I and J (Baseline correction)

Principal component analysis

PCA technique was executed to remove the background physical effect and normalize the data at various spectral pre-processing (i. e standard normal variates, 2nd derivative, savitzy goloy derivative, baseline correction) phases. PCA was performed to each data matrix (after various preprocessing) such as it is illustrated in fig. 2, where it is compared the score and loading plot. The spectra from various *Danti* samples were projected onto the Hotelling T2 95% confidence limits for principal components 1 and 2. PCA successfully showed possible trends in their values and emphasize the similarities and differences between samples on a score plot. The score plot in fig. 2 (A) showed that the CPDR was grouped in the upper left quadrant of the score plot, though KPDR, WPDR scattered in the lower right quadrant, except for the sample RD which appeared in the upper right quadrant. However, the sample WPDR and KPDR may create cluster due to the similarity in NIR spectral response. In similar pattern was seen after savitzy goloy derivative in fig. 2 (G). On the other hand baseline correction and standard normal variates preprocessing, PCA output was reflected that the RD was grouped in the upper left quadrant of the score plot, though CPDR scattered in the lower right quadrant, except for the sample WPDR, KPDR which appeared in the upper right quadrant in fig. 2 (E, I). From the loading plot in fig. 2 (B, D, F, H and J), it appeared that the NIR sensitive functional group contributing to the grouping of *Danti* samples. These wavelength variables corresponded to the PC1 and PC2 which explaining about 100,95,74,99,97% total variance simultaneously without preprocessing, with preprocessing standard normal variates, 2nd derivative, savitzy goloy derivative, baseline correction. It is possible to determine what preprocessing data is able to provide the better grouping of *Danti* samples. Generally, diffuse reflectance is the light that is reflected by the chemical components present in the particles, where some of the light is absorbed by the chemical components present in the particles. It should be noted that CPDR samples are differentiated from other samples by their higher spectral relative intensity in favor of NIR-sensitive generic organic functional group may be addition organic load from long papper and honey during *Shodhana* process. In fig. 2 (D) from the loading point, it appeared the correlation for the variables to the principal components. Center position variables are less informative due to consider of the average data set. Variables (1100-1243 nm) closely correlated and shown similarities among samples. Anticorrelated with variables (1600-1550) shown dissimilarity. Above explaining variable mainly indicated aromatic, C=O, starch, an aromatic NH_2 structural assignment usually exemplified in the presence of resin and starch component [17].

HCA

Cluster analysis was performed to classify samples on the basis of the similarities of the NIR diffuse spectral response. A hierarchical agglomerative procedure was employed to establish clusters. The results obtained are shown as a dendrogram, at a similarity level of 0.5. The dendrogram in fig. 3 shows the clustering pattern of the data set contains various purificatory samples of *Danti*. Additionally, it can be said that group 1,3 was the most outlying group because its averaged euclidean distance to the rest of the samples was very high. Finally, it was noted that hierarchical clustering empowered strong classification power along the different processed *Danti* data set. Especially by interpreting the dendrogram in fig. 3, (A, B, C) It was observed that respective four samples create two prominent clusters. Samples KPDR, RD groups have similarity NIR spectral response. But CPDR has a strong influence on the NIR spectra and therefore on the functional group addition during purification process as previously discussed in [17]. The finding shows that PCA in combination with HCA is perfectly suitable for discriminating between a different kind of treatment schedule during *Shodhana* (processing).
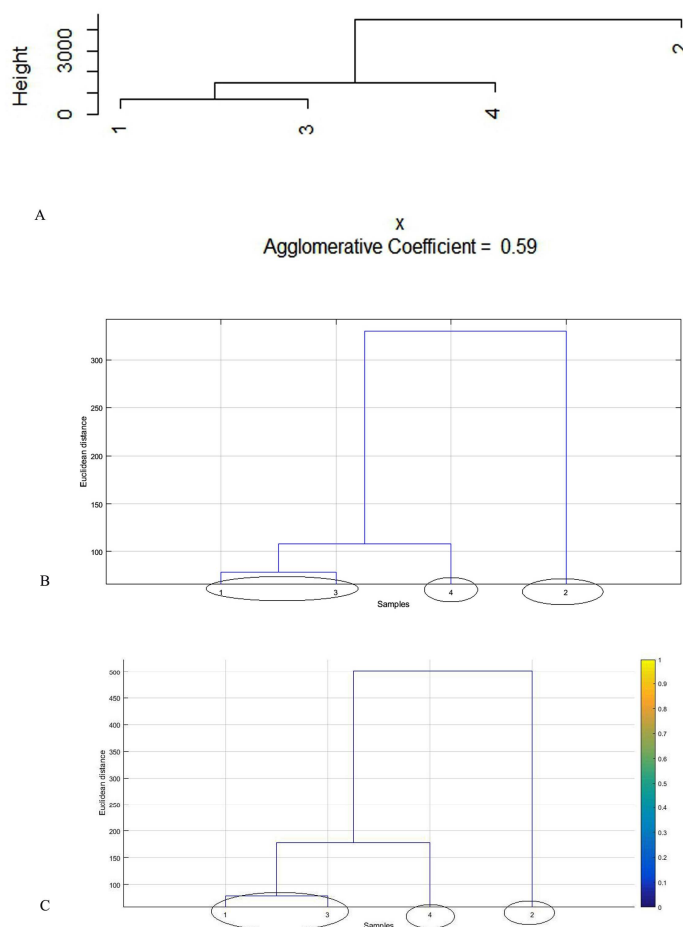
Fig. 3: Cluster analysis for the discrimination of four *Danti* samples, A-Manhattan distance with single linkage, a B-Euclidean distance of single linkage, C-Euclidean distance with complete linkage. 1-KPDR, 2-CPDR, 3-RD, 4-WPDR

CONCLUSION

The discriminant analysis (classification) technique was used to determine whether unknown *Danti* spectra matched the reference library or not. It is observed that relative intensity difference of the various *Danti* samples powder at different generic functional group was at approximately (1410 nm, 1415 nm, 1417 nm, 1700 nm wavelength). Sample raw *Danti* root and sample classical processed *Danti* root was the most outlying group to rest of the samples i. e water processed *Danti* root and *Kusha* processed *Danti* root due to their higher averaged euclidean distance. By generating Dendrogram, samples, *Kusha* processed *Danti* root and raw *Danti* root group shows similar NIR spectral response. ROH(alcohol), CH_2(Methylene methyle group), aromatic, ArOH (Aryl)–structural assignment (functional group) present in classical processed *Danti* root discriminant it to raw *Danti* root against 1410 nm onward wavelength on NIR spectral region suggesting the addition of these above-mentioned functional groups. The present observation may be used as a standard referencing library for analyzing the sample in future.

CONFLICT OF INTERESTS

Declared none

REFERENCES

1. KM Nadkarni. Indian Materia Medica. Vol. 1. 3rd ed. Bombay Popular Publication; 1982. p. 708.
2. Santosh Kumar Maurya, Ankit Seth, Damiki Laloo, Narendra Kumar Singh. Śodhana: an ayurvedic process for detoxification and modification of therapeutic activities of poisonous medicinal plants Anc. Sci Life 2015;34:188–97.
3. Zou XB. Variable selection methods in near-infrared spectroscopy. Anal Chim Acta 2010;667:14-32.
4. G Reich. Near–infrared spectroscopy and imaging: basic principles and pharmaceutical applications. Adv Drug Delivery Rev 2005;57:1109-43.
5. NK Shah, PJ Gemperline. The combination of the mahalanobis distance and residual variance pattern recognition technique for classification of near–infrared spectra. Anal Chem 1990;62:465-70.
6. HW Siesler, Y Ozaki, S Kawata, HM Heise. Near-infrared spectroscopy, principles, instruments, applications: Wiley-VCH Weinheim Germany; 2002.
7. Cynthia Kradjel. Practical use of FT-NIR for identification and qualification of botanicals: a fit–for-purpose approach. In: Kurt Reynertson, Khalid Mahmood. Botanicals: methods and techniques for quality and authenticity. 1st ed. Boca Raton, London, New York; 2015. p. 237.
8. Charaka Samhita, Ayurveda-Deepika. Commentary of chakrapanidatta, chowkhamba sanskrit sansthana; Varanasi. Seventh edition. Ed. Reprint, Chaukhamba Surbharati Prakashan, Varanasi; 2007. p. 936.
9. Newly A Burns, Emil W Ciurczak. Handbook of near infrared analysis spectroscopy. 3rd edition. CRC Press; Boca Raton; 2008. p. 67-78.
10. Geladi P, MacDougall D, Martens H. Linearization and scatter correction for near-infrared reflectance spectra of meat. Appl Spectrosc 1985;39:491-500.
11. Brown CD, Vega-Montoto L, Wentzell PD. Derivative preprocessing and optimal corrections for baseline drift in multivariate calibration. Appl Spectrosc 2000;54:1055-68.
12. Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least square procedures. Anal Chem 1964; 36:1627-39.
13. Vogt F, Rebstock K, Tacke M. Correction of background drifts in optical spectra by means of "pseudo principal components. Chemom Intell Lab Syst 2000;50:175-8.
14. Thielemans A, Mass Art DL. The use of principal component analysis as a display method in the interpretation of analytical chemical, biochemical, environmental, and epidemiological data. Chimia 1985;39:236–42.
15. Massart DL, Kaufman L. The interpretation of Analytical Chemical Data by the Use of Cluster Analysis, John Wiley and Sons, New York; 1983.
16. BH Stuart. Infrared spectroscopy: Fundamental and Applications, John Wiley and Sons, New York, USA; 2004.
17. John S Shenk. Application of NIR spectroscopy to agricultural products, handbook of near-infrared analysis. Third Edition. Editor. CRC Press; Taylor and Francis Group, Boca raton, London, New York; 2008. p. 356.

How to cite this article