Indian J. Genet., (2021); 81(4): 566-574
doi: 10.31742/ISGPB.81.4.9

ISSN: 0975-6906

RESEARCH ARTICLE

www.isgpg.org

Inheritance of oleic acid trait and high throughput non-destructive phenotyping for nutritional traits in groundnut kernels

Dnyaneshwar B. Deshmukh[1,3,§], Murali T. Variath[3,*,§], Nagesh Patne[3], Praveen Kona[2], Balram Marathi[1] and Janila Pasupuleti[3]

Abstract

The nutritional quality and food use of groundnut (*Arachis hypogaea* L.) is mainly governed by oil, fatty acids, protein, and moisture content of kernels. The breeding for higher proportion of oil, protein, and oleic acid in the kernels is an important objective, which needs a non-destructive, rapid, and reliable method for routine estimation in relatively large breeding populations. The present study reports the development of calibration equations in near-infrared reflectance spectroscopy (NIRS) for rapid and non-destructive estimation of kernel quality. Mode of inheritance pattern of a high oleic trait in groundnut was also studied. The best equation for each trait was selected based on the coefficient of determination in calibration and for cross-validation. The current equation gave high fidelity with the reference to biochemical value as indicated by high values of coefficient of determination in external validation (r^2) for oleic acid (r^2= 0.96), linoleic acid (r^2= 0.96), moisture (r^2= 0.96) and moderate for oil (r^2= 0.89), protein (r^2= 0.83) and palmitic acid (r^2 = 0.80). The study further developed an efficient NIRS equation to deploy in groundnut breeding. The high oleic trait inheritance pattern was studied in $F_{2:3}$ population derived from a cross between Spanish bunch normal oleic ICGV 06420 and high oleic SunOleic 95R parents. The results showed duplicate recessive inheritance pattern with a segregation ratio of 15: 1 (normal oleic: high oleic). The outcomes from the inheritance study helps to breed groundnut cultivars for high oleic trait.

Keywords: Fatty acids, groundnut, NIRS calibrations, oil, protein

Introduction

Groundnut or peanut (*Arachis hypogaea* L.) is the world's sixth-largest oilseed legume crop in production after soybean, oilpalm, rapeseed, cottonseed, sunflower, and second important source of vegetable protein after soybean (FAO 2018). On a global basis, 49% of the groundnut produce is used to extract edible oil, 41% for food use, 10% is used as feed and seed (Singh and Nigam 2016). Groundnut seed contains high-quality edible oil (~50%), easily digestible protein (~25%) and 13 fatty acids, of which 90% comprises of oleic acid (36–81%), linoleic acid (4–40%) and palmitic acid (8–10%) (Janila et al. 2016a).

Nutritional and storage quality of groundnut is mainly governed by the relative proportion of palmitic, oleic, and linoleic acids. Groundnut with a higher ratio of oleic to linoleic (O/L) acid is the preferred choice as it provides health benefits to the consumers (Janila 1 et al. 2016a). High O/L groundnut kernels are preferred by the food processing industry as it retains shelf life longer by 8-10 times than normal oleic groundnut (Wang, 2018). Palmitic acid is known to cause an adverse effect on human health and increase the risk of developing cardiovascular diseases (Hu et al. 2001). Polyunsaturated fatty acids, especially higher linoleic acid, are vulnerable to oxidation, leading to off-flavours

International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Hyderabad, India;

[1]Professor Jayashankar Telangana State Agricultural University, Rajendranagar, Hyderabad, 500 030, Andhra Pradesh, India;

[2]ICAR-Directorate of Groundnut Research. Junagadh, 362 001, Gujrat, India

Corresponding Author: Murali T. Variath, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502 324, Hyderabad, India, E-Mail: murali.tv2006@gmail.com

How to cite this article: Deshmukh D. B., Variath M. T., Patne N., Kona P., Marathi B. and Pasupuleti J. 2021. Inheritance of high oleic acid content and high throughput non-destructive phenotyping for nutritional traits in groundnut kernels. Indian J. Genet., **81**(4): 566-574.

Source of support: Nil

Conflict of interest: None.

Received: July 2021 **Revised:** Oct. 2021 **Accepted:** Nov. 2021

(Braddock et al. 1995). For edible oil extraction, the oil millers prefer groundnuts with high oil content (Janila 2 and Nigam 2013). For the confectionery industry, it is desirable to have low oil content, high oleic to linoleic ratio and high protein content. To meet these diverse market demands, the breeder needs to develop and screen many breeding populations and select the genotypes that adhere to market criteria for quality parameters and higher yield.

The high oleic groundnut mutants (F435-1 and F435-2 with 80.0% oleic acid, and 2.0% linoleic acid) were identified in the field screening studies at the University of Florida (Norden et al. 1987). This discovery led to the availability of high oleic sources for studying genetics and breeding new cultivars. The inheritance studies on high oleic traits showed that duplicate recessive genes are responsible for accumulating high oleic content in the kernels (Moore and Knauft 1989). The molecular studies revealed two homoeologous mutant alleles of *ahFAD2A* and *ahFAD2B* on linkage group A09, and B09, respectively. This mutation results in the substitution of G:C→ A:T in the open reading frame (ORF) at 448 base pair (bp) positions on 9th chromosome of A genome and insertion of adenine at 442 bp in the ORF on 9th chromosome of B genome (Moore and Knauft 1989; Chu et al. 2009). These mutations result in the elevated accumulation of high oleic acid as the conversion of oleic to linoleic is stalled and thus decreases linoleic acid. Although SNP markers are available for confirmation of high oleic trait in early generation progenies, the knowledge of genetics and inheritance pattern of high oleic trait gives the precision to estimate the probability of getting high oleic trait and combining it with other economic traits like short duration, drought tolerance, foliar fungal disease resistance, kernel quality traits etc. (Deshmukh et al. 2020).

Improved varieties with desirable oil, protein and fatty acid profiles meet the needs of food and oil processing value chains of groundnut, and thus, breeding varieties to meet this need has been a key breeding objective globally. In breeding for improvement of quality traits, the steps that are of critical importance are: availability of significant genetic variability for the trait of importance enabling the selection of suitable parents for use in crossing program; knowledge of genetics of the trait involved allowing the selection of an appropriate breeding and selection strategy for achieving trait improvement; availability of markers that helps in the selection process in the early generations through identification of progenies with presence or absence of the marker and faster, reliable and robust phenotyping methods which can help in the selection and advancement process for trait improvement.

Sufficient genetic variability exists for quality traits, but the genetic gains for quality parameters are often limited as the breeders have to use laborious and expensive estimation methods. Often this is done later in the breeding cycles when the variability has become narrow. This selection strategy may be good for observable traits. However, for traits such as quality phenotyping needs to be done from the early stages of advancement so that only the superior performing lines are carried forward. Non-destructive and robust methods of estimating the quality parameter facilitates the early generation selection in breeding cycles and thus enhance the genetic gain for these traits.

Near-infrared reflectance spectroscopy (NIRS) offers a simple, non-destructive, economical and fast screening tool for estimation of quality parameters by utilizing the near-infrared rays (750–3000 nm) of the electromagnetic spectrum (Sundaram et al. 2009). The NIRS has proven its applicability in quantitative measurements of protein, carbohydrates, moisture, oil, fatty acids, acidity, and total soluble solids in oilseed crops (Velasco et al. 1999; Sundaram 1 et al. 2010, Vollmann and Jankowicz-Cieslak 2017). Due to its quick non-destructive assay (~80 seconds/sample), favorable economics, absence of sample preparation, and chemicals, the NIRS has been successfully used to estimate biochemical compounds in several oilseeds crops such as soybean (Pazdernik et al. 1997) and sunflower (Fassio et al. 2004). In groundnut, NIRS has been successfully deployed to predict oil content (Fox and Cruickshank 2005), fatty acids (Wang 1 et al. 2014; Lee et al. 2016), protein (Chaudhury et al. 2016; Cheng et al. 2018), and moisture (Govindarajan et al. 2009). The present study was, therefore, attempted to develop individual equations for estimating oil, protein, fatty acids, and moisture content in groundnut kernels and use them to predict the different traits in a single scan.

Materials and methods

Groundnut samples

Groundnut genotypes with variable content of oil, protein, and fatty acids were used for wet lab analysis to develop calibration equations for using NIRS. A set of 264 $F_{2:3}$ population from the cross, ICGV 06420 × SunOleic 95R developed at International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, India was used for developing calibration equations for fatty acids, especially palmitic, oleic, and linoleic acid. The population was grown in two replications during 2013-14 season. Bed to bed distance was 60 cm, the distance between rows of the bed was 30 cm, and plant to plant distance within a row was 10 cm. The parent ICGV 06420 is high oil (>50%) genotype with normal oleic (43.4%) and linoleic acid (36.7%) content. SunOleic 95R is a high oleic acid (80.6%) and low linoleic acid (2.8%) variety released by USA (Gorbet and Knauft 1997). The fatty acid profile data obtained from 161 F_2 -derived F_3 progenies were studied for oleic acid inheritance trait. A set of advanced/elite breeding population and released cultivars comprising 156 lines for oil, 148 lines for protein and 8 genotypes with 80 different data points (10 data

points/genotype) for moisture content were used in the calibration set.

NIR spectrum collection

All the groundnut samples (separate sets for oil, protein and fatty acids) except moisture content, were first scanned in a NIRS monochromator (model XDS, Foss, Denmark) and the scanned samples were sent for wet lab analysis. In the case of moisture content, the scanning was carried out at different time intervals coinciding with the moisture meter recordings.

About 50-60g of sample was loaded into a rectangular cup, and reflectance spectra (log1/R) from 400 nm to 2498 nm were recorded at 2 nm intervals. The cup was filled up sufficiently to allow good absorption of the incident light. Each sample was subsequently scanned 32 times and the average spectrum was collected.

Oil extraction

The oil content of 156 groundnut samples was determined at Charles Renard Analytical Laboratory (CRAL), ICRISAT by Soxhlet method (Warra et al. 2011). About 70 g of powdered seed samples were placed in a porous thimble and added in a Soxhlet extractor (Robert Krups, Germany). Oil extraction was carried out using 150 cm^3 of n-hexane as an extracting solvent for 6 hours. The oil was obtained after removing the solvent under reduced temperature and pressure and reflux at 70°C to remove excess solvent used in the oil. The crude oil of each sample was measured with the formula and expressed as total oil content (%).

$$\text{Total oil content (\%)} = \frac{(W_2 - W_1)}{T} \times 100$$

Where W_1= Weight of empty thimble; W_2= Weight of thimble with the sample after oil extraction; T= Weight of sample after oil extraction

Protein estimation

Protein estimation of 148 groundnut samples was carried out at CRAL, ICRISAT by Skalar colorimetric method (Chaudhary et al. 2016). About 0.3 g of finely ground groundnut samples were digested with 2.5 mL of concentrated sulphuric acid-selenium mixture (0.4% Se, v/w, was heated to dissolve Se), and the total Nitrogen content in the digest was determined using Skalarautoanalyzer. The estimated Nitrogen (N %) was converted into protein by multiplying with a correction factor of 5.46.

$$\text{Protein content (\%)} = \frac{\left(\frac{\text{Sample} \times 250}{1000}\right) \times 5.46 \times 100}{W \times (100 - WL)}$$

Where 1000= conversion factor from ppm to percentage Nitrogen; W= Weight of the sample utilized for estimation; WL= Weight loss on drying.

Determination of fatty acid content

The fatty acid composition of 322 groundnut samples was determined at CRAL, ICRISAT. The quality of analysis was assured by regularly monitoring and analyzing the standard samples received from the International Plant and Soil Analytical Exchange Laboratory, Wageningen Evaluating Programmes for Analytical Laboratories, located in the Netherlands. A standard fatty acid methyl ester (FAME) mixture (Nupack) was used to identify fatty acids in the samples (Metcalfe et al. 1996). Approximately 3 g of groundnut was macerated in a Krups grinder (Krups International), and about 300 mg of the ground seed powder sample was suspended in 15 mL of petroleum ether. The contents were centrifuged at 4000 rpm and 5 mL supernatant was transferred into a small culture tube. Residual solvent was evaporated under a stream of nitrogen gas. To this, about 1.3 mL of 0.5 N NaOH in methanol was added and kept in a boiling water bath. The contents were allowed to cool and 2 mL of Boron Trifluoride in Methanol was added and kept for 5 min in a water bath. The solution was suspended with 2 mL saturated NaCl and shaken on a tube rotator for 10 min, followed by 2 mL of petroleum ether. The entire mixture was shaken on a tube rotator for 5 min and centrifuged at 4000 rpm for 5 minutes. The supernatant containing the FAME was transferred into the glass vial and stored in a freezer for gas chromatography.

Fatty acid profiles were determined using a Shimadzu GG-9A GLC unit (Tokyo, Japan) equipped with programmable oven, flame ionization detector and connected to the integrator, and Chromasorb W-AW (Supelco Inc) glass column (6 feet long with 3 mm inner diameter). The Initial column temperature was set at 195°C and held for 5 min followed by a step-up of 10°C/min to reach a final temperature of 250°C and retained for 2 min. The flame ionization detector was maintained at 250°C and the injector at 220°C. Helium was a carrier gas and its flow rate was maintained at 50 mL/min (primary pressure 6kg/cm^2). The hydrogen gas flow and airflow rates were maintained at 0.6kg/cm^2 and 0.5kg/cm^2, respectively. A Shimadzu CR4A chromatopac integrator (Shimadzu Corporation, Kyoto, Japan) was used to record fatty acid retention time and peak area (palmitic, oleic and linoleic acids).

The inheritance pattern of high oleic traits in the population was studied by Chi-square analysis of the observed ratio of normal, and high oleic plants to test the goodness of fit into the assumed phenotypic ratio.

Moisture estimation

Eight groundnut genotypes (ICGVs 00441, 92195, 00440, 00308, 91114, 87846, 03043 and 89280) developed at ICRISAT were used to establish moisture calibration using NIRS. The seeds (ca. 800 g) of eight genotypes were soaked overnight in water and dried in hot air oven at 60° C. Moisture content was estimated from the samples after every two hrs of drying using a moisture meter (National Instruments, Vadodara, India; capacity: 0% to 40% moisture). Simultaneously, the samples were also scanned in NIRS and the moisture value

was incorporated into the calibration file for developing the equation. Drying was continued in all the genotypes for 23 hrs till a stable moisture content range of 3-5% was obtained. The moisture content in the calibration set ranged from 3.71-29.60 (Table 1). Eleven groundnut genotypes (ICGVs 91114, 00350, 03043, 16526, 16540, 16542, 16551, 16553, 170047, 171145, and ICGS 76) were utilized for external validation of moisture equation. The moisture gradient was created as per the method mentioned above (ca. 80 data points) and moisture content was measured using a moisture meter.

Spectral data analysis

NIR spectral data for all target traits were collected, processed and calibrated using WinISI 4 software (version 4.3, Infrasoft International, Port Matilda, Pennsylvania, USA). Before proceeding to develop the calibration equation, the groundnut sample set for the different traits was categorized into calibration and external validation set (Table 1) using the select algorithm available in WinISI 4 software.

To develop the regression models, calibrations were performed between spectral data and wet lab reference values using the modified partial least squares (MPLS) regression method. The fatty acids, oil, protein, and moisture content were dependent variables for the data analysis. To minimize or remove unwanted sources of variability in the sample data sets, data pre-treatment algorithms such as mathematical treatment and scatter correction algorithms were used. The mathematical treatment uses the raw spectra, or first or second derivatives to enhance spectral variability by removing background noise in combination with gap sizes in data points over which the derivative is calculated, including a smoothing algorithm to reduce random noise in the spectral data. For example, in the treatment 1,4,4,1 the first number, i.e., 1 is the first derivative of log 1/R; the second number i.e. 4 is the gap (the length in nm); the third number i.e., 4 represents the number of data points (segment length) used in the first smoothing and the fourth number which is usually set as 1 indicates the number of data points in the second smoothing (Barnes et al. 1989).

Derivation and smoothing were used in combination with scatter correction algorithms. Scatter correction helps to reduce differences in the spectra arising from the sample's physical characteristics, such as particle size and path length of reflectance from the particle surface (Barnes et al. 1989; Shenk and Westerhaus 1993). Calibrations were performed with five different mathematical treatments (1,2,2,1; 1,4,4,1; 1,8,8,1; 2,4,4,1 and 2,8,8,1) using Standard Normal Variate + Detrend (SNV + D) scatter correction factor (Table 2).

Four cycles of outlier elimination were performed to eliminate samples with very high Global H (GH) and Neighborhood H (NH) values. Global H or GH calculates the distance of each sample in the population from centre of the sample set, while the Neighborhood H or NH calculates the distance between a sample and its neighbor. GH defines the population boundaries of a complete set and NH the boundaries of a subset. In the present study, the GH and NH values were set at 3 for all the target traits. This limit resulted in the removal of the maximum of 32 outliers for oil, 18 outliers for protein, 20 outliers for palmitic and linoleic acid, 15 outliers for oleic acid and 5 outliers for moisture content in the calibration set using the different mathematical treatments. The different calibration models were assessed for their predictive efficiency based on the values for standard error of calibration (SEC), the coefficient of determination in calibrations (RSQ), the standard error of cross-validation (SECV), and the coefficient of determination in cross-validation (1-VR) (Shenk and Westerhaus 1993).

For validation studies, calibration equations with the highest RSQ and 1-VR values and the lowest number of outlier eliminations, SEC and SECV values were selected. The relative predictive determinant for calibration (RPD_c), the ratio of the standard deviation to SECV (SD/SECV), was used to evaluate the calibration equations' performance.

External validation

The selected calibration equation was evaluated for its accuracy and precision of prediction for the different quality parameters using the external validation set. The external validation set consisted of 26 samples for oil, 28 for protein, 50 for the three fatty acids and 30 for moisture content (Table 1). The indicators for external validation were: standard error of prediction (SEP), the coefficient of determination in external validation (RSQ) and relative

Table 1. Reference data of oil, protein, fatty acids, and moisture content in groundnut seeds to calibrate equations in NIRS

	Calibration set				External validation set			
Constituent (%)	N	Range (%)	Mean (%)	SD	N	Range (%)	Mean (%)	SD
Oil	129	40.76-56.37	48.57	2.60	27	43.60-54.14	49.37	0.32
Protein	120	19.71-35.59	27.65	2.66	28	20.40-30.81	27.22	0.63
Palmitic acid	208	6.97-15.06	11.47	1.49	50	7.58-13.69	11.39	0.47
Oleic acid	208	25.31-77.65	51.48	8.72	50	37.51-79.32	52.23	0.21
Linoleic acid	208	4.16-50.89	27.53	6.92	50	3.11-41.81	26.85	0.23
Moisture	80	3.71- 29.60	10.38	6.51	43	3.47-29.98	16.5	0.33

N=Number of samples and SD=Standard deviation

Table 2. Statistical parameters of the calibration set and internal validation used to calibrate the NIRS model for oil, protein, fatty acids, and moisture content in groundnut

Constituent (%)	Maths treatment	Calibration						Cross-validation		RPD_c
		N	Range (%)	Mean (%)	SD	RSQ	SEC	SECV	1-VR	
Oil	1,2,2,1	103	41.08-56.22	48.65	2.52	0.81	1.07	1.22	0.76	2.07
Oil	1,4,4,1	100	40.76-56.37	48.57	2.60	0.84	1.02	1.17	0.79	2.22
Oil	1,8,8,1	97	40.69-56.29	48.49	2.60	0.83	1.05	1.17	0.79	2.22
Oil	2,4,4,1	101	40.77-56.28	48.52	2.58	0.82	1.06	1.24	0.76	2.08
Oil	2,8,8,1	104	40.68-56.51	48.59	2.64	0.82	1.09	1.26	0.76	2.10
Protein	1,2,2,1	106	20.54-35.29	27.92	2.46	0.86	0.91	1.41	0.67	1.74
Protein	1,4,4,1	104	20.54-35.23	27.89	2.45	0.82	1.03	1.41	0.67	1.74
Protein	1,8,8,1	102	21.25-34.93	28.09	2.28	0.81	1.00	1.30	0.67	1.75
Protein	2,4,4,1	112	19.71-35.59	27.65	2.65	0.91	0.77	1.41	0.72	1.88
Protein	2,8,8,1	113	19.69-35.68	27.69	2.66	0.86	1.00	1.50	0.68	1.77
Palmitic acid	1,2,2,1	190	7.18-15.83	11.51	1.44	0.88	0.50	0.59	0.83	2.44
Palmitic acid	1,4,4,1	191	7.18-15.84	11.52	1.44	0.87	0.53	0.61	0.82	2.36
Palmitic acid	1,8,8,1	189	7.36-15.71	11.54	1.39	0.86	0.52	0.60	0.81	2.32
Palmitic acid	2,4,4,1	196	6.97-15.97	11.47	1.50	0.89	0.50	0.61	0.83	2.46
Palmitic acid	2,8,8,1	188	7.27-15.97	11.50	1.41	0.87	0.51	0.60	0.82	2.35
Oleic acid	1,2,2,1	200	24.58-78.54	51.56	8.99	0.95	2.04	2.35	0.93	3.83
Oleic acid	1,4,4,1	200	24.04-79.34	51.69	9.22	0.95	1.98	2.21	0.94	4.17
Oleic acid	1,8,8,1	198	25.33-77.28	51.31	8.66	0.94	2.13	2.39	0.92	3.62
Oleic acid	2,4,4,1	195	25.31-77.65	51.48	8.72	0.96	1.66	2.21	0.94	3.95
Oleic acid	2,8,8,1	193	25.83-76.85	51.34	8.50	0.96	1.66	2.13	0.94	3.99
Linoleic acid	1,2,2,1	195	5.19-50.38	27.79	7.53	0.95	1.62	1.93	0.93	3.90
Linoleic acid	1,4,4,1	193	6.64-49.45	28.05	7.14	0.95	1.57	1.83	0.93	3.90
Linoleic acid	1,8,8,1	193	5.86-49.97	27.92	7.35	0.95	1.66	1.89	0.93	3.89
Linoleic acid	2,4,4,1	195	4.16-50.89	27.53	7.79	0.97	1.39	1.73	0.94	4.50
Linoleic acid	2,8,8,1	188	7.38-49.08	28.23	6.95	0.97	1.30	1.80	0.95	3.86
Moisture	1,2,2,1	77	3.74-29.80	10.51	6.43	0.99	0.49	0.61	0.99	10.54
Moisture	1,4,4,1	77	3.87-29.80	10.51	6.43	0.99	0.50	0.61	0.99	10.54
Moisture	1,8,8,1	75	3.70-29.38	10.34	6.34	0.99	0.42	0.56	0.99	11.32
Moisture	2,4,4,1	77	3.71-29.60	10.38	6.41	0.99	0.33	0.55	0.99	11.65
Moisture	2,8,8,1	76	3.90-29.16	10.23	6.31	0.99	0.39	0.58	0.99	10.88

N = Number of samples; SD = Standard deviation; RSQ = Coefficient of determination of calibration; SEC = Standard error of calibration; SECV = Standard error of cross-validation;
1-VR = Coefficient of determination of cross-validation and RPDc = Relative predictive determinant for calibration

predictive determinant of external validation (RPD_v) which is the ratio of the standard deviation to SEP (SD/SEP).

Results and discussion

Variability for oil, protein, fatty acids and moisture content

The relative values for the mean, standard deviation and the range of oil, protein, individual fatty acids (palmitic, oleic and linoleic acid), and moisture content in the calibration and validation sets are shown in Table 1. The wide range of groundnut quality attributes (oil, protein, fatty acids and moisture) and relatively even values distribution favored the development of robust and reliable NIRS calibration models. Wide range of variability was observed in both calibration and validation sets for oil content (40.76-56.37% in calibration, 43.60-54.14% in validation), protein content (19.71-35.59% in calibration, 20.40-30.81% in validation), oleic acid (25.31-77.65% in calibration, 37.51-79.32% in validation), linoleic acid (4.16-50.89% in calibration, 3.11-41.81% in validation), moisture content (3.71-29.60% in calibration, 3.47-29.98% in validation); whereas palmitic acid (6.97-15.06% in calibration, 7.58-13.69% in validation), had the narrow range.

Breeding for quality in groundnut has mainly focused on improving the oil and fatty acid profile and several studies have indicated the existence of variability for these traits (Janila et al. 2016a). The genetics involved in the inheritance of these traits has also been discussed (Moore and Knauft, 1989) and QTL mapping studies were done using the population indicated in the present study for high oleic acid (Shasidhar et al., 2017). Recently two high oil-containing lines GJG 32 and GJG 33, with >50% oil content was released for cultivation in India (Janila et al., 2016b). For oil quality two high oleic varieties, Girnar 4 and Girnar 5 with more than 78% oleic content, were released for cultivation in India for the first time (Bera, 2020).

NIRS spectral analysis

The D^2 log (1/R) spectra and mean standard deviation spectrum of calibration samples were obtained by using the entire wavelength range of 400-2498 nm (Fig. 1(a)). The high scatter coefficient of NIRS allows for excellent diffuse reflectance spectra of solids, thereby improving its prediction efficiency. The peaks and valleys in the NIR region represent the absorption of electromagnetic energy (Fig.1(b) & 1(c)). The untreated log (1/R) spectra of the groundnut samples with high and low oleic acid are shown in Fig. 1(d). NIR spectroscopy directs electromagnetic energy on the sample, and detects the transmittance and/or reflectance energy from the vibrational response of C–H, O–H, and N–H bonds present in biochemical samples(Sundaram et al. 2010). Three chemical bonds: C-H, corresponding to fats and oils; O-H bond present in water; and N-H bonds corresponding to protein are responsible for the absorbance bands detected in NIR spectra (Pereira et al. 2008). Other types of chemical bonds may appear in weak overtone bands in the NIR region and are ignored in the analysis of complex food crops such as groundnut.

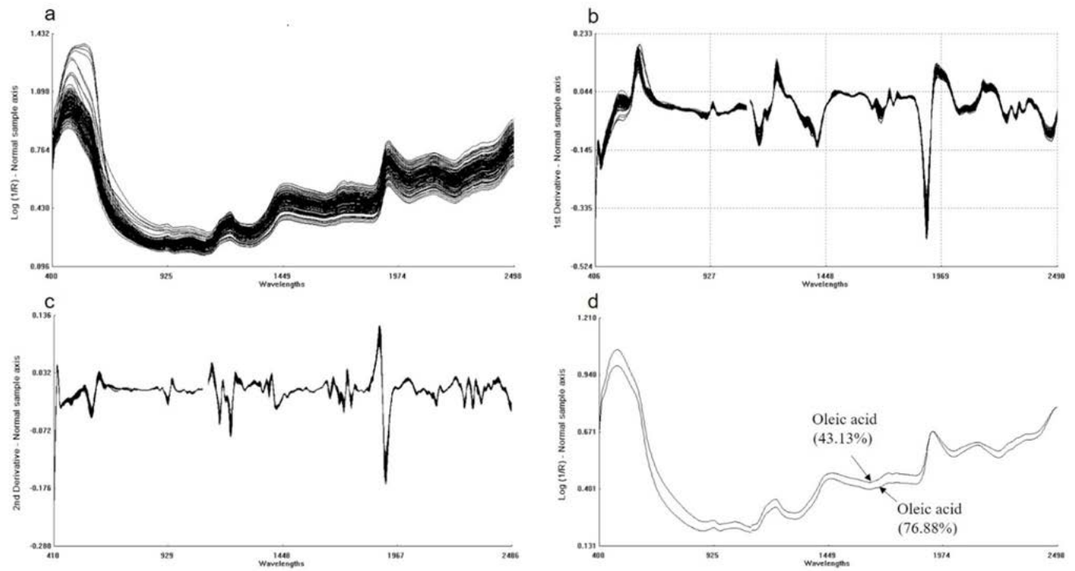

Fig.1. Pre - and post processing spectra of groundnut seeds for oil and oleic acid contents: a. Wavelength spectra of groundnut seeds for oil content in the range of 400 to 2498 nm; b. 1,4,4,1 treated spectra for oil content; c. 2,4,4,1 treated spectra for oleic acid content and d. Spectra showing variation for oleic acid content with high and low oleic acid samples.

Development of multivariate calibration models and internal cross-validation

In order to select the best calibration model, calibration and internal cross-validation was performed using MPLS regression method with five different mathematical treatments (1,2,2,1; 1,4,4,1; 1,8,8,1; 2,4,4,1; 2,8,8,1) and Standard Normal Variate + Detrend (SNV + D) scatter correction factor (Table 2).

For protein, fatty acids (palmitic, oleic and linoleic acid) and moisture content the treatment 2,4,4,1 gave optimum RSQ and 1-VR values, while for oil content the treatment 1,4,4,1 gave the best results (Table 2). The spectra with 1,4,4,1 and 2,4,4,1 mathematical treatments are depicted in Fig. 1(b)&1(c), respectively. The RPD_c values were used to assess the robustness of each equation. RPD_c values >3 imply equation with excellent calibration; between 2 and 3 indicates equations with very reliable predictions; between 2 and 1.5 imply limited prediction effects, and <1.5 indicates

equations with unreliable correlations. In the present study, the equation for moisture content showed high RSQ (0.99) and 1-VR (0.99) with low SEC (0.33) and SECV (0.55) and exhibited a high RPD_c value (11.65) and was the best equation. Oleic acid and linoleic acid also exhibited high RSQ (0.96 and 0.97, respectively), 1-VR (0.94) and RPD_c (3.95 and 4.50, respectively), along with low SEC (1.66 and 1.39, respectively)and SECV (2.21 and 1.73, respectively). Palmitic acid and oil content had RSQ of 0.84 to 0.89; 1-VR of 0.79 to 0.83 and RPD_c of 2.22 to 2.46, indicating an equation with reliable prediction efficiency. Protein content had an RSQ of 0.91 and 1-VR of 0.72, an SEC of 0.77 and SECV 1.41 and an acceptable RPD_c of 1.88.

External validation of the calibrated NIRS equation

The performance of the calibrated NIRS equation was tested by external validation which involves the test of prediction for target traits. The predictive ability of the selected calibration equations with the best RSQ,1-VR and RPD_c values was further assessed to predict the oil, protein, fatty acids and moisture content using an external validation set (Table 3). The best equation was identified based on low standard errors of prediction (SEP) and high coefficients of determination for external validation (r^2). High values were observed for the equations of oil ($r^2 = 0.89$), protein ($r^2 = 0.83$), oleic acid ($r^2 = 0.96$), linoleic acid ($r^2 = 0.96$), palmitic acid content ($r^2 = 0.80$) and moisture content ($r^2 = 0.96$). The biases and slopes of the selected equations recorded values close to '0' and '1', respectively. The RPD_v values for the external validation set ranged from 2.0 to 5.6, indicating a good fit.

The regression plots of reference versus NIRS predicted values for the oil, protein, palmitic acid, oleic acid, linoleic acid and moisture content are depicted in Fig. 2. The regression plots for oleic, linoleic moisture and oil content had a higher degree of prediction accuracy as indicated by their high r^2 and SEP values. For palmitic acid and protein content, the prediction was less accurate. In the present study, based on interpretation of RPD_v values, the equations for oleic, linoleic, moisture and oil content can be applied reliably for quality assurance and research applications as well as sample screening for breeding programs; while those for protein and palmitic acid content can be used only in screening. Also, in the present study, the high variability for oleic, linoleic, moisture and oil content was attributed to achieve a good calibration equation (Table 1). Calibration equations were also developed for predicting main fatty

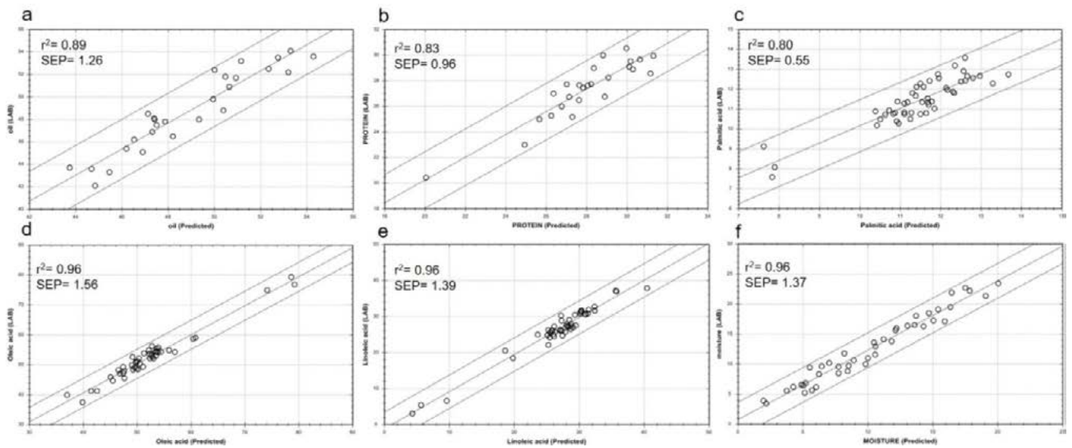

Fig. 2. Regression plots for oil, protein, palmitic acid, oleic acid, linoleic acid and moisture contents as determined by the reference method and NIRS prediction for the external validation set: a. Oil, b. Protein, c. Palmitic acid, d. Oleic acid, e. Linoleic acid, f. Moisture. The r^2 is the coefficient of determination and SEP is the standard error of prediction between the reference method and the NIRS method.

Table 3. Statistical parameters of external validation set for predicting oil, protein, fatty acids, and moisture content in groundnut seed

Constituent (%)	Maths treatment	Number	SD	Bias	r^2	SEP	Slope	RPD_v
Oil	1,4,4,1	27	3.35	-0.17	0.89	1.26	1.12	2.9
Protein	2,4,4,1	28	2.42	-0.62	0.83	0.96	0.84	2.0
Palmitic acid	2,4,4,1	50	1.17	-0.02	0.80	0.55	0.84	2.0
Oleic acid	2,4,4,1	50	7.83	0.21	0.96	1.56	0.95	4.4
Linoleic acid	2,4,4,1	50	7.00	-0.32	0.96	1.39	1.01	4.5
Moisture	2,4,4,1	43	5.01	-0.02	0.96	1.19	1.03	5.6

SD = Standard deviation; r^2 = Coefficient of determination of calibration; SEP = Standard error of prediction and RPDv = Relative predictive determinant of external validation

Table 4. Segregation of oleic acid trait in $F_{2:3}$ population

Cross	Category	Frequency		O-E	$(O-E)^2$	$(O-E)^2/E$	χ^2 value	P value
		O	E					
ICGV 06420 x SunOleic 95R	High oleic	12	11	1	1.60	0.15	0.16	0.75-0.5
	Normal oleic	149	150	-1	1.60	0.01		

O = Observed and E = Expected

acids, oil and protein content in bulk seeds/ single seeds of groundnut using NIRS (Sundaram et al. 2009, Wang et al. 2014). However, none of the studies attempted to compare the effect of different mathematical treatments and scatter correction methods on the calibration model and prediction accuracy in groundnut. In one of the earlier reports, involving seven different math treatments with and without spectral corrections and using different regression models, the treatment 2,6,4,1 was most suitable for measuring fatty acid content in single seeds of groundnut (Nadaf and Hanchinal 2014). In the present study, comparison of different mathematical treatments showed the 2,4,4,1 model along with SNV + D scatter correction to be the most suited for predicting protein, palmitic, oleic, linoleic and moisture content; while 1,4,4,1 was most appropriate for the prediction of oil content in whole groundnut seeds.

Inheritance of high oleic trait

The oleic acid content ranged from 35.79 to 81.77% and linoleic acid was ranged from 2.52 to 41.81% in the studied $F_{2:3}$ populations. Based on O/L ratio the population was classified as normal oleic (<10 O/L ratio) and high oleic (>10 O/L ratio). Out of 161 F_2:3, 12 F_2:3 were found to be high oleic and the rest 149 fell in the normal oleic category (Table 4). Based on observed and expected values, the segregation pattern of oleic trait showed 15:1 (normal oleic:high oleic). The P-value for the Chi-square test confirmed a good fit for the expected inheritance ratio of duplicate recessive inheritance, which is in agreement with earlier reports (Moore and Knauft 1989).

Authors' cobntribution

Conceptualization of research (JP, MTV, BM); Designing of the experiments (JP, MTV); Contribution of experimental materials (JP); Execution of field/lab experiments and data collection (DBD, MTV, NP, PK); Analysis of data and interpretation (DBD, MTV); Preparation of the manuscript (DBD, MTV)

References

Barnes R. J., Dhanoa M. S. and Lister S. J. 1989. Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. Appl. Spectrosc., 43(5): 772-777.

Bera S. K. 2020. Girnar 4 and Girnar 5: the first ever indigenous high-oleic groundnut varieties, Technical Report. December 2020 DOI - 10.13140/RG.2.2.27314.43200 ER

Braddock J. C., Sims C. A. and O'keefe S. F. 1995. Flavor and oxidative stability of roasted high oleic acid peanuts. J. Food Sci., 60 (3): 489-493.

Chaudhury S., Sahrawat K. L., Srinivasu K., Wani S. P. and Puppala N. 2016. Comparative evaluation of protein content in groundnut samples by near infrared reflectance spectroscopy and Skalar colorimetric methods. Curr. Sci., 110 (12): 2219-2220.

Cheng J. H., Jin H. and Liu Z. 2018. Developing a NIR multispectral imaging for prediction and visualization of peanut protein content using variable selection algorithms. Infrared Phys. Technol., 88: 92-96.

Chu Y., Holbrook C. C. and Ozias-Akins P. 2009. Two alleles of ahFAD2B control the high oleic acid trait in cultivated peanut. Crop Sci., 49: 2029-2036.

Deshmukh D. B., Balram M., Sudini H. K., Murali T. V., Chaudhari S., Manohar S. M., Durgarani Ch. V., Pandey M. K. and Janila P. 2021. Combining high oleic acid trait and resistance to late leaf spot and rust diseases in groundnut (*Arachis hypogaea* L.). Front. Genet.,514 (11)**:** 1-15.

Food and Agriculture Organization (FAO). Statistical Database FAOSTAT. 2018, URL (http://www.fao.org/faostat/en/#data/QC) (Accessed date: 25th February 2020).

Fox G. and Cruickshank A. 2005. Near infrared reflectance as a rapid and inexpensive surrogate measure for fatty acid composition and oil content of peanuts (*Arachis hypogaea* L.). J. Near Infrared Spectrosc., 13 (5): 287-291.

Fassio A. and Cozzolino D. 2004. Non-destructive prediction of chemical composition in sunflower seeds by near infrared spectroscopy. Ind. Crops Prod., 20(3): 321-329.

Gorbet D. W. and Knauft D. A. 1997. Registration of 'SunOleic 95R'peanut. Crop Sci., 37: 1392.

Govindarajan K. N., Kandala C. V. and Subbiah J. 2009. NIR reflectance spectroscopy for non-destructive moisture content determination in peanut kernels. Trans. ASABE, 52(5):1661-1665.

Hu F. B., Manson J. E. and Willett W. C. 2001. Types of dietary fat and risk of coronary heart disease: a critical review. J. Am. Coll. Nutr., 20 (1): 5-19.

Janila P. and Nigam S. N. 2013. Phenotyping for groundnut (*Arachis hypogaea* L.) improvement. In: Phenotyping for Plant Breeding. Springer, New York, USA: 129-167.

Janila P., Pandey M. K., Shasidhar Y., Variath M. T., Sriswathi M., Khera P., Manohar S. S., Nagesh P., Vishwakarma M. K., Mishra G. P. and Radhakrishnan T. 2016a. Molecular breeding for introgression of fatty acid desaturase mutant alleles (ahFAD2A and ahFAD2B) enhances oil quality in high and low oil containing peanut genotypes. Plant Sci., 242: 203-213.

Janila P., Manohar S. S., Patne N., Variath M. T., Nigam S. N. 2016b. Genotype× Environment interactions for oil content in peanut and stable high-oil-yielding sources. Crop Sci., 56: 1-10.

Lee J. M., Pae S. B., Choung M. G., Lee M. H., Kim S. U., Oh E. Y., Oh K. W., Jung C. S. and Oh I. S. 2016. Determination of fatty acid composition in peanut seed by near infrared reflectance spectroscopy. Korean J. Crop Sci.,61(1):64-69.

Metcalfe L. D., Schmitz A. A. and Pelka J. R. 1996. Rapid preparation of fatty acid esters from lipids for gas chromatographic analysis. Anal. Chem., 38(3): 514-515.

Moore K. M. and Knauft D. A. 1989. The inheritance of high oleic acid in peanut. J. Hered., 80: 252-253.

Nadaf H. L. and Hanchinal R. R. 2014. Near infrared reflectance spectroscopy (NIRS) for large scale screening of fatty acid profile in peanut (*Arachis hypogaea* L.). Legume Res., 37(3): 272-280.

Norden A. J., Gorbet D. W., Knauft D. A. and Young C. T. 1987. Variability in oil quality among peanut genotypes in the Florida breeding program. Peanut Sci., 14: 7-11.

Pereira A. F., Pontes M. J., Neto F. F., Santos S. R., Galvao R. K. and Araujo M. C. 2008. NIR spectrometric determination of quality parameters in vegetable oils using iPLS and variable selection. Food Res. Int., 41(4):341-348.

Pazdernik D. L., Killam A. S. and Orf J. H. 1997. Analysis of amino and fatty acid composition in soybean seed, using near infrared reflectance spectroscopy. Agron. J., 89(4):679-685.

Shasidhar Y., Vishwakarma M. K., Pandey M. K., Janila P., Variath M. T., Manohar S. S., Nigam S. N., Guo B. and Varshney R. K. 2017. Molecular mapping of oil content and fatty acids using dense genetic maps in groundnut (*Arachis hypogaea* L.). Front.Plant Sci., 8: 794.

Shenk J. S. and Westerhaus M. O. 1993. Near infrared reflectance analysis with single and multiproduct calibrations. Crop Sci. 33(3): 582-584.

Singh A. K. and Nigam S. N. 2016. Arachis gene pools and genetic improvement in groundnut. In: Gene Pool Diversity and Crop Improvement, vol. I (Eds. V.R. Rajpal, S. R. Rao and S.N. Raina). Springer, Cham.: 17-75.

Sundaram J., Kandala C. V. and Butts C. L. 2009. Application of near infrared spectroscopy to peanut grading and quality analysis: overview. Sens. Instrum. Food Qual. Saf., 3(3):156-164.

Sundaram J., Kandala C. V., Holser R. A., Butts C. L. and Windham W. R. 2010. Determination of in-shell peanut oil and fatty acid composition using near-infrared reflectance spectroscopy. J. Am. Oil Chem. Soc., 87(10):1103-1114.

Velasco L., Möllers C. and Becker H. C. 1999. Estimation of seed weight, oil content and fatty acid composition in intact single seeds of rapeseed (Brassica napus} L.) by near-infrared reflectance spectroscopy. Euphytica, 106(1):79-85.

Vollmann J. andJankowicz-Cieslak J. 2017. Utilising NIRS for Qualitative and Non-destructive Identification of Seed Mutants in Large Populations. In: Biotechnologies for Plant Mutation Breeding (Eds. J. Jankowicz-Cieslak , T. H. Tai, J. Kumlehn and B.J. Till). Springer, Cham.:193-202.

Wang C. T., Wang X. Z., Tang Y. Y., Wu Q., Xu J. Z., Hu D. Q. and Qu B. 2014. Predicting main fatty acids, oil and protein content in intact single seeds of groundnut by near infrared reflectance spectroscopy. In: Advanced Materials Research**.** Trans Tech Publications Ltd. 860: 490-496.

Wang Q. 2018. Relationship between raw material quality and product quality of peanut. In: Peanut Processing Characteristics and Quality Evaluation. Springer, Singapore: 151-209.

Warra A. A., Wawata I. G., Gunu S. Y. andAujara K. M. 2011. Extraction and physicochemical analysis of some selected Northern Nigerian industrial oils. Arch. Appl. Sci. Res. 3(4): 536-541.