

DOI: 10.1177/0003702819859940

Raman Spectroscopic and Chemometric Investigation of Lipid–Protein Ratio Contents of Soybean Mutants

Gulce Ogruc Ildiz,^{a,d,*} Ozge Celik,^b Cimen Atak,^b Ayberk Yilmaz,^c Hayrunnisa Nur Kabuk,^a Ersin Kaygisiz,^c Alp Ayan,^b Sinan Meric^b and Rui Fausto^d

^a Department of Physics, Science and Letters Faculty, Istanbul Kultur University, Atakoy Campus, Bakirkoy 34156, Istanbul, Turkey

^b Department of Molecular Biology and Genetics, Science and Letters Faculty, Istanbul Kultur University, Atakoy Campus, Bakirkoy 34156, Istanbul, Turkey

^c Department of Physics, Science Faculty, Istanbul University, 34134, Vezneciler, Istanbul, Turkey

^d CQC, Department of Chemistry, University of Coimbra, P-3004-535 Coimbra, Portugal

***Corresponding author email:** g.ogruc@iku.edu.tr

ABSTRACT

Seeds belonging to fourth generation mutants (M₄) of Ataem-7 cultivar (A7) variety and S04-05 (S) breeding line salt-tolerant soybeans were investigated by Raman spectroscopy, complemented by chemometrics methods, in order to evaluate changes induced by mutations in the relative lipid–protein contents, and to find fast, efficient strategies for discrimination of the mutants and the control groups based on their Raman spectra. It was concluded that gamma irradiation caused an increase in the lipid to protein ratio of the studied Ataem-7 variety mutants, while it led to a decrease of this ratio in the investigated S04-05 breeding line mutants. These results were found to be in agreement with data obtained by reflectance spectrum analysis of the seeds in the full ultraviolet to near-infrared spectral region, and suggest the possibility to develop strategies where gamma irradiation can be used as a tool to improve mutant soybean plants targeted to different applications, either enriched in proteins or in lipids. Ward's clustering and principal component analysis showed a clear discrimination between mutants and controls and, in the case of the studied S-type species, discrimination between the different mutants. The grouping scheme is also found to be in agreement with the compositional information extracted from the analysis of the lipid–protein contents of the different samples.

Keywords: Food analysis, food composition, soybean, soybean mutant, Raman spectroscopy, principal component analysis, PCA, cluster analysis, chemometrics

Introduction

Soybean (*Glycine max* (L.) Merr.) is an important oilseed crop that is economically the most significant bean in the world.¹ It has been assuming a prominent role in human and animal diets, due to its high nutritional values, as high protein and oil content, and versatility for targeting different uses. Around 400 byproducts of soybean are valuable for various industries, where it is used as ingredient for production of adhesives, biofuel, soaps, doughs, alcohol, paints, fertilizers, among others.^{2–10} Because soybean contains no starch, it is also a good source of protein for diabetics.

Genetic transformation applications have been widely used as attractive advancement for soybean breeding programs, allowing to improve agronomical qualities of soybean.^{11–13} Induced mutation techniques are useful to increase the frequency of the mutations to generate genetic variations. Developing new variations to solve different agricultural crop problems are usually successfully achieved by mutation breeding in a shorter period, comparing to classical breeding processes. Gamma radiation is well known to cause rapid and reliable physicochemical and biological alterations in organisms.^{14–16} In mutation breeding studies, the next crucial step is to select the desired genotypes for new cultivars. Due to the fact that the selected mutants from M₂ and M₃ generations differ for several traits, it is important to discriminate the desired genotypes to improve new cultivars. Therefore, the molecular structural evaluation is advantageous to determine the traits of the mutants.

Numerous methods have been developed for the last couple of decades to measure the protein and oil content of crop seeds. Conventional analysis techniques are mainly destructive, expensive and labor intensive,^{9,17,18} and the use of alternative, cheaper techniques of essentially nondestructive nature, have been acquiring increasing importance for this purpose. Raman spectroscopy has been used along with infrared (IR) spectroscopy, in both mid-IR (MIR) and near-IR (NIR) regions, as reliable analysis tools to probe and characterize biomolecules in complex systems, since these techniques are particularly well suited for identification of functional groups belonging to proteins, lipids and other biochemically relevant molecules, through characteristic spectral features they originate.^{10,19–21} Raman spectroscopy has several advantages over infrared that makes it a more effective technique for compositional analysis of biochemically relevant systems. First, Raman spectra are

considerably less perturbed by the presence of water (a ubiquitous constituent of biological materials) in the sample, compared to infrared spectra (including the NIR region). In fact, water gives rise to bands of very weak Raman intensity, while in infrared water bands are very intense and overshadow wide spectral regions which are relevant for investigation of other species.²² Second, Raman spectroscopy requires only very minor preprocessing of the samples, hence being a very practical, easy to use method, and an essentially nondestructive technique. Finally, Raman spectra of biomolecules usually show bands with narrower profiles compared to infrared spectra, thus allowing a more detailed discrimination of different spectral features and facilitating identification of the characteristic bands ascribable to functional groups belonging to different constituents present in the sample.¹⁰

Both Raman and infrared spectroscopies have been used in the last decade for composition analysis of soybean seeds, in particular protein, oil, carbohydrates and water contents.^{9,10,18} Recently, compositional alterations in second generation soybean mutants obtained by harvesting gamma irradiated soybeans, as well as conformational alterations of their protein isolates, were investigated by means of infrared spectroscopy.^{23,24} In the present study, salt tolerant mutant soybean seeds were investigated. The original seeds (obtained from Black Sea Agricultural Research Institute, Samsun, Turkey) were subjected to gamma irradiation (4.29 Gy/min dose rate) to induce random mutations, allowing subsequent selection of the salt tolerant mutants. The seeds were generated and harvested for several generations to test stability of salinity tolerance of the selected mutants. A detailed description of the procedures followed to generate and select the mutants and evaluate their salt tolerance ability has been provided elsewhere.^{25,26} The present article focuses on the determination of the changes in the lipid–protein ratio contents of Ataem-7 soybean line (A7) and S04-05 soybean breeding line (S) M₄ generation mutants by Raman spectroscopy, together with multivariate data analyses of the spectral data. This study aims to demonstrate that the Raman analysis allows a reliable, fast determination of this fundamental compositional feature, while it may also be used successfully to perform discrimination between the mutants and control groups (non-irradiated soybean seeds of the same variety).

Materials and methods

In this study, seeds of Ataem-7 soybean variety (A7) and S04-05 breeding line (S) provided by the Black Sea Agricultural Research Institute, Samsun, Turkey, were irradiated with a Cs-137 gamma source by 150, 200, and 250 Gy (4.29 Gy/ min dose rate). Six mutants (one

mutant derived from the Ataem-7 variety, five mutants derived from S04-05 soybean breeding line) induced by this irradiation were selected among M₄ generation seeds that were tested under both in vitro and in vivo conditions and proven to be 90 mM NaCl tolerant. Samples were named as A7-C (control) and A7-M (mutant) in the case of Ataem-7 variety, and S-C (control) and S1-M, S2-M, and S3-M. S4-M and S5-M (mutants) for the S04-05 breeding line. The lipid and protein contents of the soybean seeds were determined by UV-Vis-NIR reflectance spectrum data analysis between 800–2500 nm wavelengths, with 1 nm resolution, by using Foss 6500 NIR system, in order to obtain reference values for testing the results obtained by Raman spectroscopy.²⁷ The total lipid and protein contents were calculated in accordance with calibration programme based on dry weight with respect to control samples and represented as percentage values.

Raman spectra were acquired using a B&W Tek iRaman Plus-785 model micro-Raman system, equipped with a 785 nm excitation laser. The laser power at the sample was set as 280 mW. Before used, all soy seeds were peeled and gently divided into half. No further treatment was applied to the seeds for spectra collection. Each sample was placed on an aluminum covered thin glass slide and Raman spectra recorded from different parts of the seeds, in the 400–3000 cm⁻¹ spectral range. An integration time of 30 s was used, and 32 scans were averaged for subsequent spectral and statistical analysis.

Spectra post-processing, such as baseline correction, normalization and calculation of derivative spectra, as well as other corrections such as multiplicative scatter correction (MSC), standard normal variate (SNV), and the multivariate analyses were performed using OMNIC (Thermo Fisher Scientific) and The Unscrambler (CAMO Software). Band areas were calculated using OMNIC in a supervised way and used to estimate the differences in the lipid–protein ratios in the different soybean mutants relatively to the controls. The estimated uncertainty in the percent changes in lipid–protein ratios is smaller than 2%. Cluster analysis and principal component analysis (PCA) were performed using the registered raw spectra for A7 soybeans, only subjected to baseline correction. For the analysis of the S soybeans, which included five different mutants, the spectra were first vector normalized, SNV and MSC corrections were applied, and then the second derivative spectra were obtained and used to perform the analyses. Ward's method using squared Euclidean distances was used to construct the dendrograms.

Results and discussion

Soybean has a complex composition, which includes complete proteins, carbohydrates and lipids, as well as vitamins and minerals.²⁸ The usual absolute amounts of the three main types of nutrients in soybean, proteins, carbohydrates and lipids are approximately 40%, 30%, and 20% of the total weight, respectively.²⁹ It has been shown that gamma irradiation within the doses used in the present investigation gives rise to mutant soybean species where essentially the relative contents of lipids and proteins is affected, the remaining constituents being essentially unaffected by this procedure, including trace constituents.²³

The vector normalized average Raman spectra of the control and mutant groups of the two investigated soybean lines, A7 and S, are depicted in Figures 1 and 2. The spectra exhibit differences in the relative intensities of the bands, demonstrating that the composition of the nutrients of the various groups is dissimilar. A general assignment of the observed bands is given in Table I, following the data in the literature.^{10,20,30} Above 1200 cm^{-1} , the Raman spectra are ruled by bands having dominant contributions from lipids ($1200\text{--}1500\text{ cm}^{-1}$ range, ascribable mainly to the CH_2 , twisting, wagging and scissoring modes, and CH_3 deformation vibrations; $1720\text{--}1750\text{ cm}^{-1}$, $\text{C}=\text{O}$ stretching; and $2800\text{--}3000\text{ cm}^{-1}$, CH_2/CH_3 stretching modes) or proteins ($1500\text{--}1650\text{ cm}^{-1}$, amides I and II).¹⁰ Contribution of other constituents of the soybeans to the total band intensity in these spectral regions can be expected to be minor. Besides, since the total amount of nutrients others than lipids and proteins can be expected to be nearly constant in all groups studied,²³ the observed intensity differences can be attributed to the relative amount of lipids and proteins in the different groups. A good test for this assumption, in the case of the lipid-related spectral ranges, consists in plotting the ratios of the bands observed in the different spectra in the three lipid-related regions against each other. As expected, straight lines were obtained, confirming that the band intensities in these spectral regions have essentially contributions from the lipids (a few outliers were identified, most probably resulting from artifacts due to the performed baseline correction treatment, which were removed from the dataset). These four spectral regions were then used in the present study to estimate the variations in the lipid–protein ratios of the various groups of mutants compared to the corresponding control groups. The results are summarized in Table II.

The striking result is that the studied salt tolerant soybean mutants developed from A7, S plants were found to have responded distinctively to the performed gamma irradiation, with the lipids contents increasing in A7 mutants and decreasing in S mutants. Indeed, as shown in Table II, the lipid–protein ratio increases by *ca.* 6% in A7 mutants, compared to the

control group, while for S mutants the general trend is the opposite, though the different mutants were found to respond in a somewhat different manner to the gamma irradiation treatment: in S3-M and S1-M the lipid–protein ratio decreases by 9% and 7%, respectively, in relation to the control group, being the mutants where the changes are larger; the decrease in the lipid–protein ratio for both S4-M and S5-M are more modest (2% and 3%, respectively); S2-M is predicted to have a lipid–protein ratio practically unchanged when compared to the control group. The identification of the reasons for the different behavior of the soybean mutants or within the different mutants of the S line is beyond the scope of this investigation, but the present results clearly demonstrate the great potential of gamma irradiation to develop mutant soybean plants targeted to different applications, either enriched in proteins or in lipids.

In Table II, the results obtained from the Raman data are compared with quantitative data determined using ultraviolet–visible–near-infrared (UV–Vis–NIR) analysis, the agreement between the two sets of data being very good. Considering the costs of the two types of analyses, it is noteworthy to see that the cheap and faster Raman analysis can provide such a good estimation of the lipid–protein ratio variation in the studied samples.

A cluster analysis of the different soybean groups, according to the Ward's method using squared Euclidean distance, was performed for both A7 and S mutants. The results are presented in Figures 3 and 4, respectively. For A7 variety, since only one type of mutant was investigated, the raw spectra (only with baseline correction) were used, allowing a good clustering of the samples, as shown in Figure 3. The clustering of the different types of mutants of the S line soybean required a more sophisticated treatment. After baseline correction and vector normalization, the second derivative spectra were calculated and used to perform the clustering. As shown in Figure 4, mutants S1-M, S3-M, S4-M, and S5-M group together, while mutant S2-M groups with to the control group. These results are in full agreement with those described above regarding the change in the relative values of the lipid–protein contents of the different S line mutant species, with the S1-M, S3-M, S4-M and S5-M mutants exhibiting a lipid–protein ratio substantially smaller in relation to the control, and the S2-M mutant showing a lipid–protein ratio practically unchanged when compared to the control group.

Figures 5 and 6 show the PCA score plots (PC1 versus PC2) grouping the different investigated soybean species. The same data used to construct the cluster dendrograms shown in Figures 3 and 4 were used in the PCA study. As seen in the figures, control and mutant

groups for both A7 and S soybean seeds were successfully discriminated through PCA. In the case of A7 variety, PC1 expresses 78% of the variance in the dataset and clearly separates the control samples (which have positive PC1 scores) from the mutants (negative scores). The variance expressed in PC2 is small (8%) and the control and mutant samples distribute identically along this axis. For S line soybeans, the total variance accounted by PC1 and PC2 is 73%, with 45% variance in PC1 and 28% in PC2. Both PC1 and PC2 principal component axes show some discriminative power, with PC1 separating the control and S2-M mutant groups from the remaining mutant species (S1-M, S3-M, S4-M, and S5-M), the first having positive scores and the later negative. Along PC1 there is also a trend to group the two mutant lines exhibiting a larger decrease in the lipid–protein ratio relative to the control group (S1-M and S3-M) in the more negative scores range, while S4-M and S5-M, which were shown to undergo more modest percent changes in the lipid–protein ratio relative to the control group, show fewer negative scores. The PC1 axis is then clearly in S soybean lines a measure of the relative amount of lipids versus proteins. Since in the PCA scores plot the control appears with positive scores, larger negative scores correlate with a larger lipid contents relative to protein contents. It shall be stressed that for A7 soybean line, PC1 expresses essentially the same as for S line (i.e., it is a measure of the relative amount of lipids versus proteins), but in this case larger negative scores correlate with a smaller lipid–protein ratio (see Figure 5), since the control group appears in the plot with positive scores. Note that the principal component axes absolute directions are in fact arbitrary as they are eigenvectors, so that the relevant information about the lipids versus proteins contents expressed in PC1 in both Figures 4 and 5 is given by the relative control-mutant positions along the axes. In the plots shown in Figure 4 and Figure 5, both control groups show positive PC1 scores. With the exception of S2-M, the mutants show negative PC1 scores; however, A7 mutants have larger lipid–protein ratios than their control samples, whereas the S mutants' seeds have smaller ratios than their control samples.

On the other hand, PC2 separates the S2-M mutant group from the control in a clear way, as well as S1-M, S3-M, and S5-M from S4-M mutant groups. Since no correlation can be established between these trends and the relative contents of lipids and proteins in the different groups, PC2 must express other characteristic features of the samples. Comparison of the Raman spectra of the samples with more negative and more positive scores along PC2 (Figure 7) shows that, in consonance with this conclusion, the main differences in the two spectra occur below 1200 cm^{-1} , i.e., in the fingerprint spectral region where, with a few

exceptions, the bands cannot be assigned to a single class of compounds but receive significant contributions from several species present in the sample (see Table I). On the other hand, the lipid and protein dominant spectral regions are rather similar in these two spectra, as expected, since they correspond to points in the scores plot with relatively similar PC1 scores (see Figure 6). An interesting observation resulting from the comparison of the Raman spectra shown in Figure 7 is the fact that in the spectrum of the mutant the characteristic band of phenylalanine at about 1003 cm^{-1} is considerable more intense than in the spectrum of the control, indicating that the amino acids' contents of the two samples is different, at least in what concerns the relative amount of phenylalanine, which is well known to exist in considerable amounts in soybean (up to ~2% of the total weight).¹⁰

On the whole, the principal component analyses performed on both A7 and S salt tolerant soybean mutants are in very good agreement, in one hand with the trends extracted from the cluster analysis regarding the similarities/dissimilarities of the various S line species studied, and on the other hand with the results obtained from the Raman intensities on the effects of the gamma irradiation on the relative contents of lipids and proteins of the soybeans. The conjugation of the Raman with the chemometrics analyses is then a convenient strategy for discrimination between the mutants and control groups, besides being also a fast, reliable method for evaluation of compositional parameters, specifically the lipid–protein ratio.

Conclusion

In this study, the relative lipid–protein contents of Ataem-7 variety and S04-05 breeding line salt-tolerant soybean seeds belonging to fourth generation mutants (M_4), together with the corresponding controls, were investigated by means of Raman spectroscopy and compared with the data obtained using the UV–Vis–NIR reflectance spectroscopy detection method. Conspicuously, both techniques showed that gamma irradiation caused a noticeable increase in the lipid to protein ratio of the studied Ataem-7 variety mutant, while it caused a decrease of this ratio in the investigated S04-05 breeding line mutants. These results suggest the possibility to develop strategies where gamma irradiation can be used as a tool to improve mutant soybean plants targeted to different applications, either enriched in proteins or in lipids. The Raman data also indicated an increase in the characteristic bands of phenylalanine of S-mutants, compared to the control, which indicates an increase in the

relative amount of this nutritionally relevant amino acid contents in the genetically modified plants' seeds.

Dendrograms for both studied soybean species, obtained by Ward's cluster analysis, showed a clear discrimination between mutants and controls. The grouping scheme is also in good general agreement with the compositional information extracted from the analysis of the lipid–protein contents of the different samples. Similar results were obtained through principal component analysis, with the obtained PCA scores associated with the major principal component axes (PC1 and PC2, expressing more than 83% of total variance for both types of soybean samples) allowing a good discrimination between mutants and controls, and grouping the different mutants in a way that is in consonance with the results obtained from the Ward's cluster analysis and also with the compositional information extracted from the Raman bands' intensity analysis and reflectance spectroscopy measurements.

Acknowledgments

Authors acknowledge Dr. Mehmet Erdoğan (Black Sea Agricultural Research Institute, Samsun, Turkey) for the analysis of oil and protein contents. This project is supported by General Directorate of Agricultural Research and Policies by the project number TAGEM/16/AR-GE/56. The Coimbra Chemistry Centre (CQC) is supported by FCT, through the project UI0313/QUI/2013, also co-funded by FEDER/COMPETE 2020-EU. This work is also supported by the Scientific and Technological Research Council of Turkey-TUBITAK 1001 Project (Project No.: 116Z294) (also MATIS).

References

1. P.D. Goldsmith. "Economics of Soybean Production, Marketing, and Utilization. In, Soybeans: Chemistry, Production, Processing, and Utilization". In: L.A. Johnson, P.J. White, R. Galloway editors. Urbana, Illinois: AOCS Press, 2008. Chap. 5.
2. D.K. Agarwal, S.D. Billore, A.N. Sharma, B.U. Dubare, et al." Soybean: Introduction, Improvement, and Utilization in India Problems and Prospects". Agric. Res. 2013. 2(4): 293–300.
3. R.J. Singh, R.L. Nelson, G.H. Chung. "Soybean (*Glycine max*(L.) Merr.), Genetic Resources, Chromosome Engineering, and Crop Improvement". In: R.J. Singh, editor. Genetic Resources, Chromosome Engineering, and Crop Improvement: Oilseed Crops, Volume 4. Boca Raton, Florida: CRC Press, 2007. Chap. 2, Pp. 13–50.

4. K.I. Chen, M.H. Erh, N.W. Su, W.H. Liu, et al. "Soyfoods and Soybean Products: From Traditional Use to Modern Applications". *Appl. Microbiol. Biotechnol.* 2012. 96(1): 9–22.
5. G.S. Anastácio, K.O. Santos, P.A.Z. Suarez, F.A.G. Torres, et al. "Utilization of Glycerin Byproduct Derived From Soybean Oil Biodiesel as a Carbon Source for Heterologous Protein Production in *Pichia Pastoris*". *Biores. Technol.* 2014. 152: 505–510.
6. L. Borsatti, S. Vieira, C. Stefanello, L. Kindlein, et al. "Apparent Metabolizable Energy of By-products from the Soybean Oil Industry for Broilers: Acidulated Soapstock, Glycerin, Lecithin, and Their Mixture". *Poultry Science.* 2018. 97(1): 124–130.
7. S. Li, D. Zhu, K. Li, Y. Yang, et al. "Soybean Curd Residue: Composition, Utilization, and Related Limiting Factors". *ISRN Indust. Eng.* 2013: 1–8.
8. J. Peña, S. Vieira, L. Borsatti, C. Pontin, et al. "Energy Utilization of By-products from the Soybean Oil Industry by Broiler Chickens: Acidulated Soapstock, Lecithin, Glycerol, and Their Mixture". *Ver. Bras. de Cienc. Avic.* 2014. 16(4): 437–442.
9. H. Lee, B.K. Cho, M.S. Kim, W.H. Lee, J. Tewari, H. Bae, S.I. Sohn, H.Y. Chi H. "Prediction of Crude Protein and Oil Content of Soybeans Using Raman Spectroscopy". *Sensors and Actuators B.* 2013. 185, 694–700.
10. M.V. Schulmerich, M.J. Walsh, M.K. Gelber, R. Kong, et al. "Protein and Oil Composition Predictions of Single Soybeans by Transmission Raman Spectroscopy". *J. Agric. Food Chem.* 2012. 60(33): 8097–8102.
11. Y. Chen, Y. Ci, Q. Meng, X. Wang, et al. "GmSK1, an SKP1 Homologue in Soybean, is Involved in the Tolerance to Salt and Drought". *Plant Physiol. Biochem.* 2018. 127: 25–31.
12. M.S. Homrich, B. Wiebke-Strohm, R.L.M. Weber, M.H. Bodanese-Zanettin. "Soybean Genetic Transformation: A Valuable Tool or the Functional Study of Genes and the Production of Agronomically Improved Plants". *Genetics and Molecular Biology.* 2012. 35(4): 998–1010.
12. B. Zhang, T. Zhang, X. Wu. "Agronomic Traits of Transgenic Soybean Overexpressing TaDREB". *J. Northeast Agric. Univ. (Engl. Ed.)*. 2011. 18(2): 1–6.
14. O. Celik, C. Atak. "Random Amplified Polymorphic DNA Analysis of Salt-tolerant Tobacco Mutants Generated by Gamma Radiation". *GMR, Genet. Mol. Res.* 2015. 14(1): 1324–1337.

15. R.S. Hanafy, S.A. Akladios. "Physiological and Molecular Studies on The Effect of Gamma Radiation in Fenugreek (*Trigonella foenum-graecum* L.) Plants". *J. Genet. Eng. Biotechnol.* 2018. 16(2): 683–692.
16. K. Ulukapi, S.F. Ozmen. "Study of the Effect of Irradiation (^{60}Co) on M1 Plants of Common Bean (*Phaseolus vulgaris* L.) Cultivars and Determined of Proper Doses for Mutation Breeding". *J. Radiat. Res. Appl. Sci.* 2018. 11(2): 157–161.
17. M.M. Bradford. "A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding". *Anal. Biochem.* 1976. 72: 248–254.
18. J. Guo, T. You, V. Prisecaru, D. Costescu, et al. "NIR Calibrations for Soybean Seeds Soy Food Composition Analysis: Total Carbohydrates, Oil, Proteins, and Water Contents". *Nature Precedings.* 2011. 6611: 2.
<http://precedings.nature.com/documents/6611/version/1> [accessed 6 Jun 2019].
19. G. Niaura. "Raman Spectroscopy in Analysis of Biomolecules". In: R.A. Meyers, editor. *Encyclopedia of Analytical Chemistry.* Chichester, UK: John Wiley and Sons, Ltd., 2014. Pp. 1–34.
20. E.C.Y. Li-Chan. "The Applications of Raman Spectroscopy in Food Science". *Trends Food Sci. Technol.* 1996. 7(11): 361–370.
21. J. Qin, M.S. Kim, K. Chao, B.K. Cho. "Raman Chemical Imaging Technology for Food and Agricultural Applications". *J. Biosyst. Eng.* 2017. 42(3): 170–189.
22. D. Bougeard, M. Buback, A. Cao, K. Gerwert, et al. "Infrared and Raman Spectroscopy: Methods and Applications". In: B. Schrader, editor. *Weinheim; New York: VCH Verlagsgesellschaft mbH,* 2008.
23. S. Akyuz, T. Akyuz, O. Celik, C. Atak. "FTIR And EDXRF Investigations of Salt Tolerant Soybean Mutants". *J. Mol. Struct.* 2013. 1044(1): 67–71.
24. S. Akyuz, T. Akyuz, O. Celik, C. Atak. "FTIR Spectroscopy of Protein Isolates of Salt-Tolerant Soybean Mutants". *J. Appl. Spectrosc.* 2017. 84(6): 1019–1023.
25. O. Celik, S.G. Unsal. "Expression Analysis of Proline Metabolism-Related Genes in Salt Tolerant Soybean Mutant Plants". *Plant Omics J.* 2013. 6(5): 364–370.
26. O. Celik, C. Atak, Z. Suludere. "Response of Soybean Plants to Gamma Radiation: Biochemical Analyses and Expression Patterns of Trichome Development". *Plant Omics J.* 2014. 7(5): 382–391.

27. S. Kumar, J.S. Chauhan, A. Kumar. “Screening for Erucic Acid and Glucosinate Content in Rapeseed-Mustard Seeds Using Near Infrared Reflectance Spectroscopy”. *J. Food Sci. Technol.* 2010. 47(6): 690–692.
28. M.A. El-Metwally, A.E.A. El-Saidy, K.M.A. El-Hai. “Role of Some Vitamins in Decreasing of Soybean Seed Deterioration During Storage”. *Res. J. Seed Sci.* 2014. 7(4): 97–115.
29. S. Ciabotti, A.C.B.B. Silva, A.C.P. Juhasz, C.D. Mendonca, et al. “Chemical Composition, Protein Profile and Isoflavones Content in Soybean Genotypes with Different Seed Coat Colors”. *Inte. Food Res. J.* 2016. 23(2): 621–629.
30. M. Baranska, H. Schulz, M. Strehle, J. Popp. “Applications of Vibrational Spectroscopy to Oilseeds Analysis”. In: J.M. Chalmers, P.R. Griffiths, editors. *Handbook of Vibrational Spectroscopy*. Chichester, UK: Wiley, 2010. Chap. ??, Pp. 397–420.

Captions

Figure 1. Raman spectra of control (A7-C) and salt tolerant mutant soybeans (A7-M). The spectra were vector normalized and subjected to baseline correction, as described in Section 2.

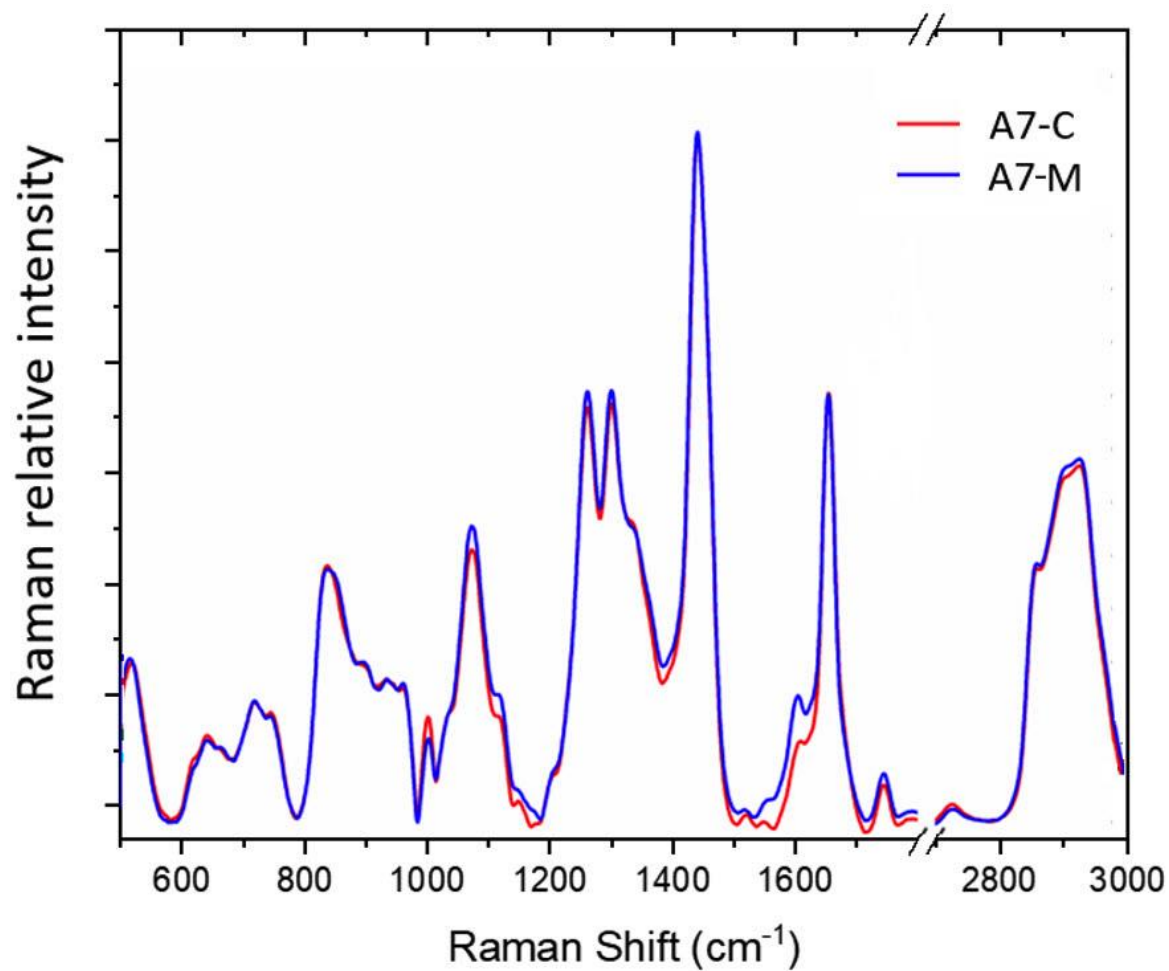


Figure 2. Raman spectra of control (S-C) and different species of salt tolerant mutant soybeans (S1-M, S2-M, S3-M, S4-M, S5-M). The spectra were vector normalized and subjected to baseline correction, as described in Section 2

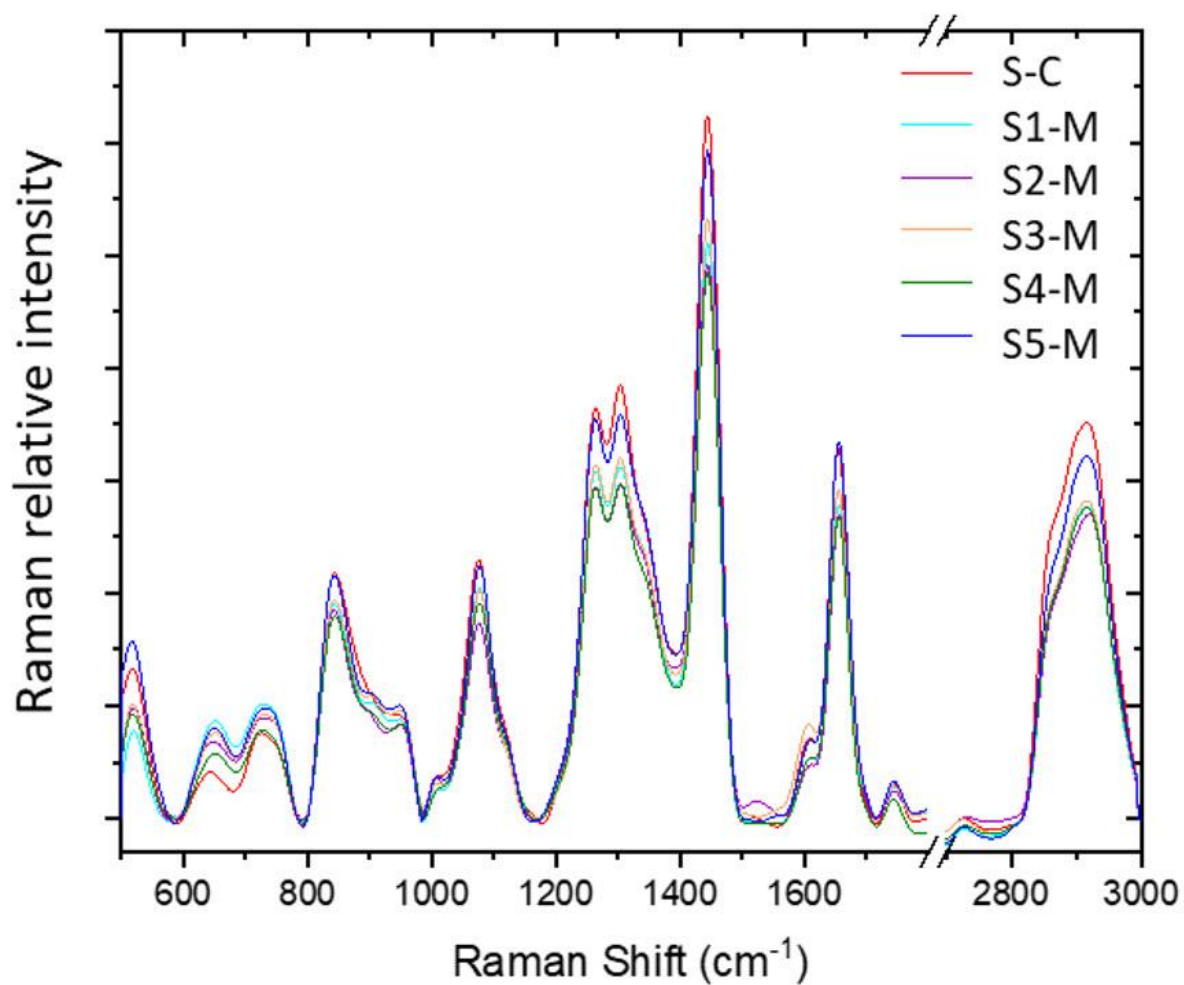


Figure 3. Cluster analysis for A7 soybean variety according to the Ward's method using squared Euclidean distance.

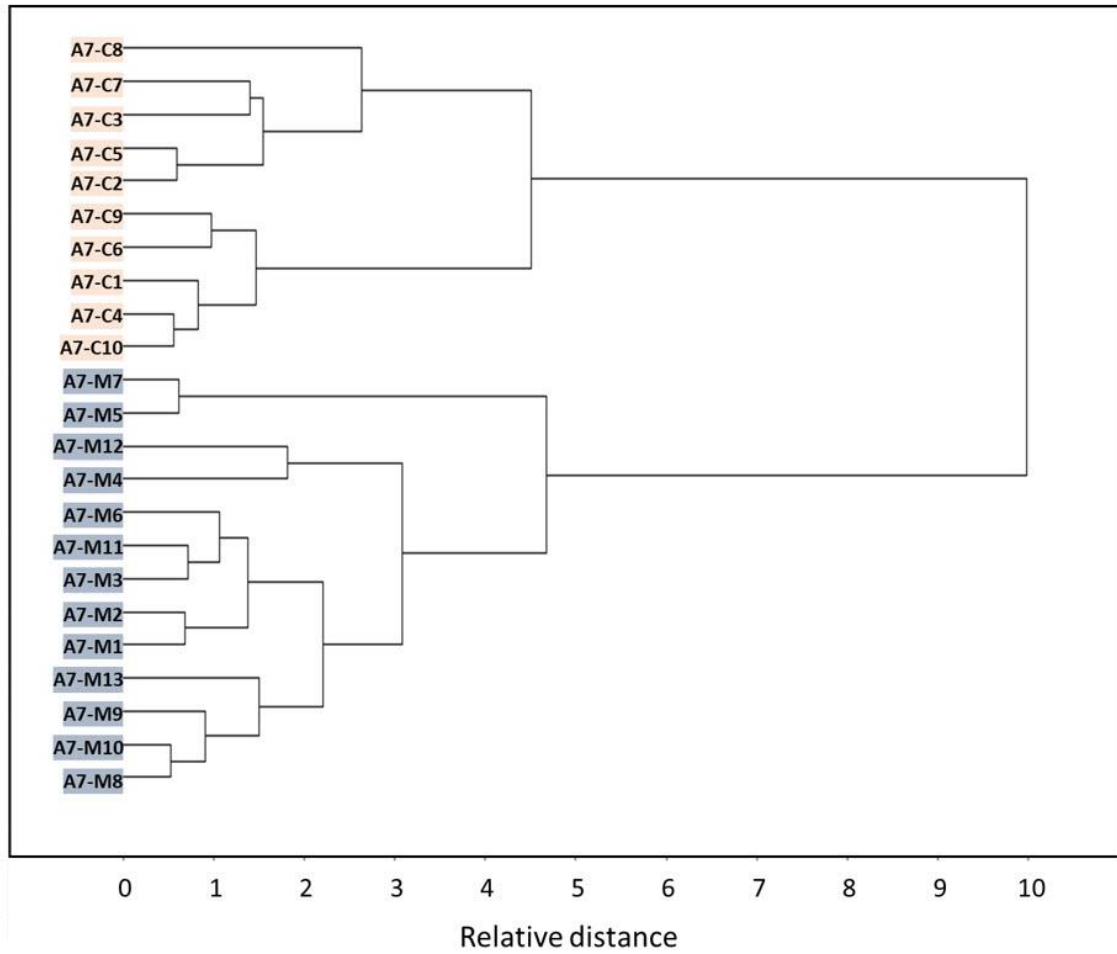


Figure 4. Cluster analysis for S soybean line, according to the Ward's method, using squared Euclidean distance.

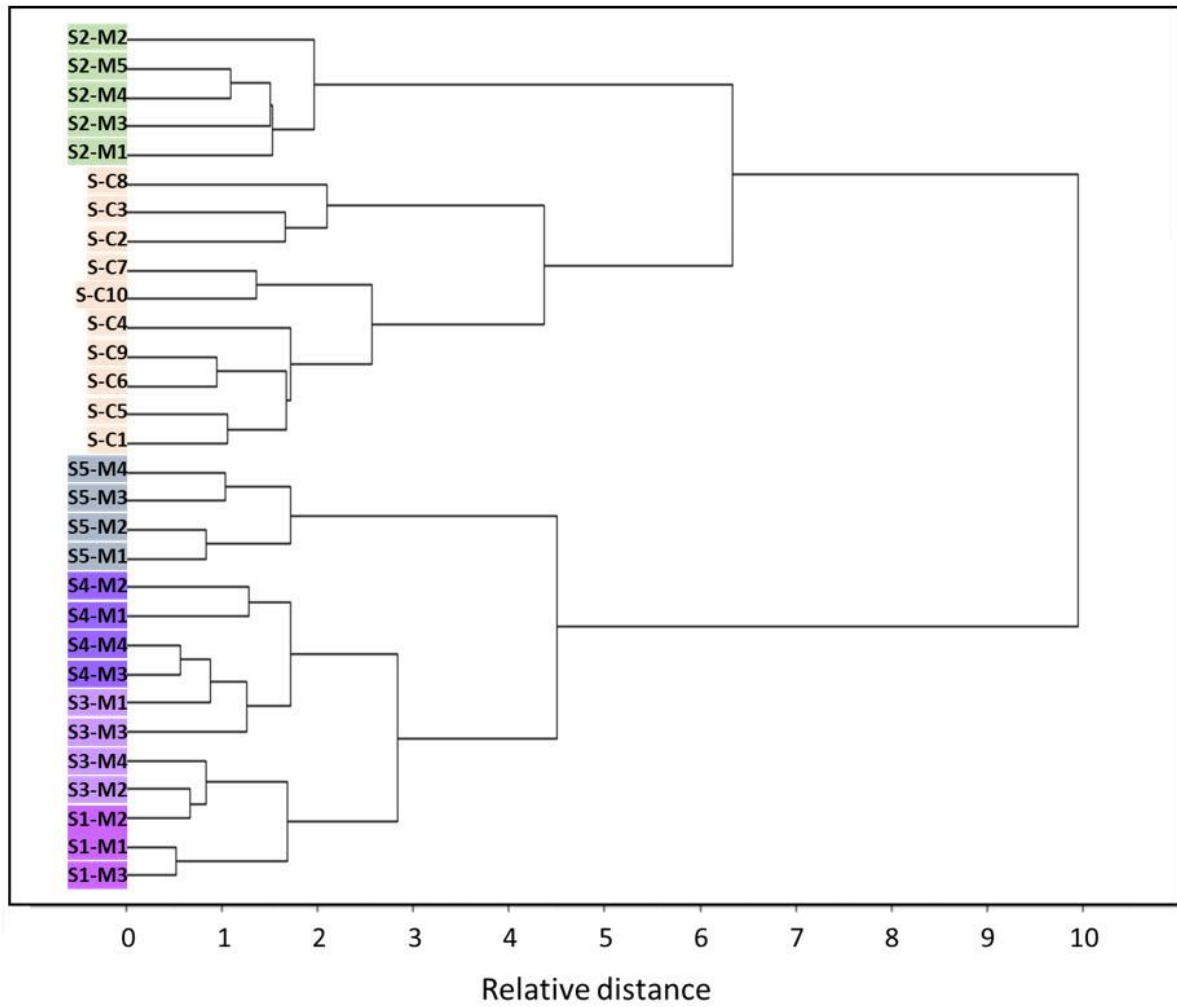


Figure 5. PCA scores plot of the first two principal components (PC1 versus PC2) obtained from PCA of Raman spectra of control (A7-C) and mutant (A7-M) *Ataem-7* soybean seeds. Axes were normalized to unity.

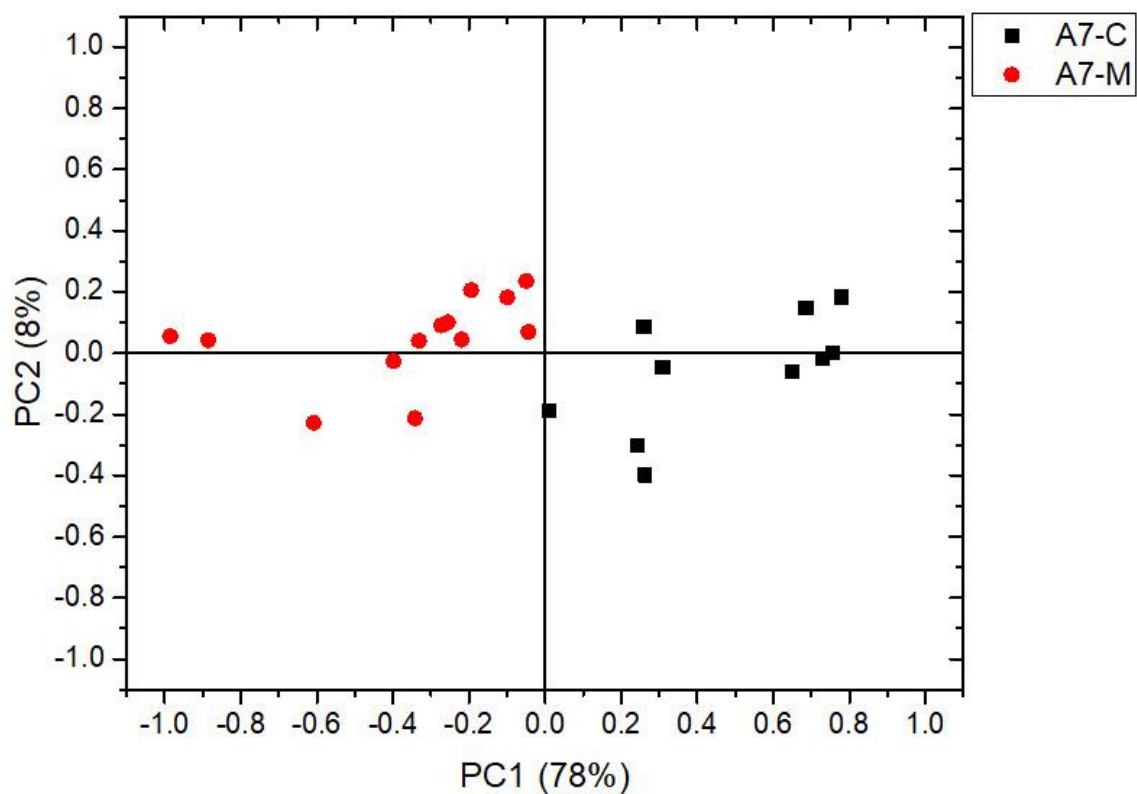


Figure 6. PCA scores plot of the first two principal components (PC1 versus PC2) obtained from PCA of Raman spectra of control (S04-05 soybean breeding line seeds) and mutant (S1-M, S2-M, S3-M, S4-M, S5-M) S soybean seeds. Axes were normalized to unity.

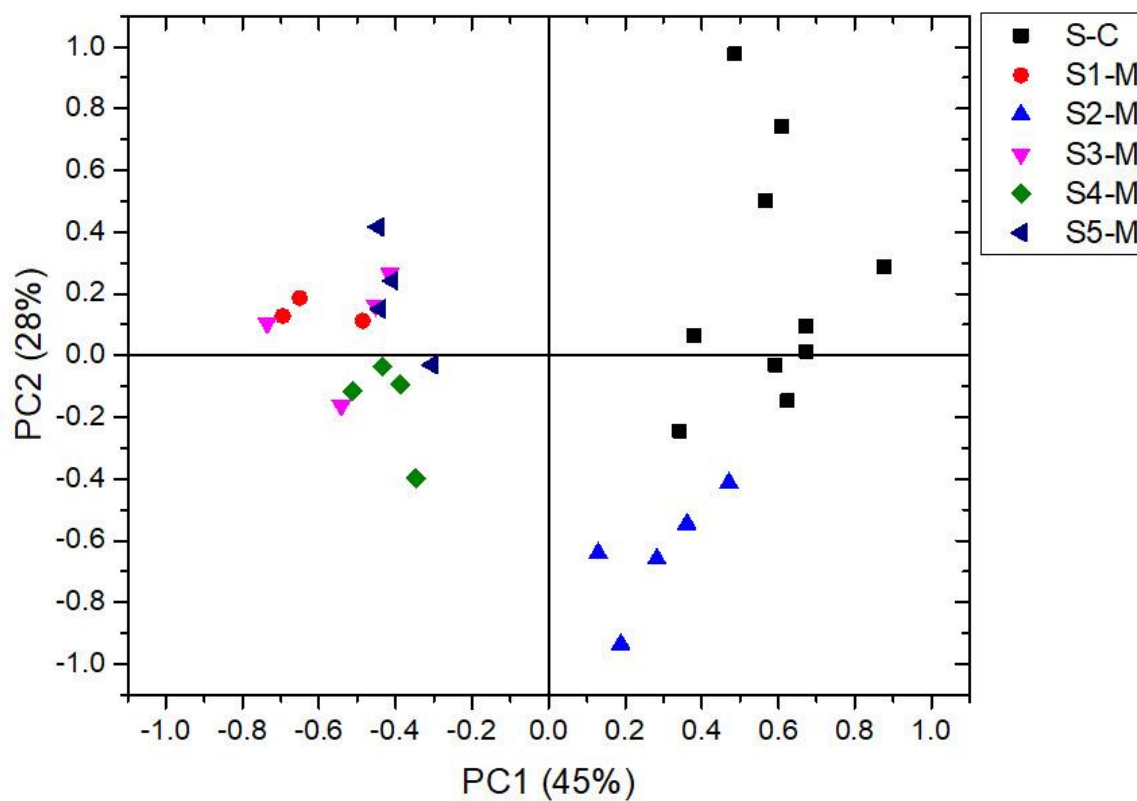
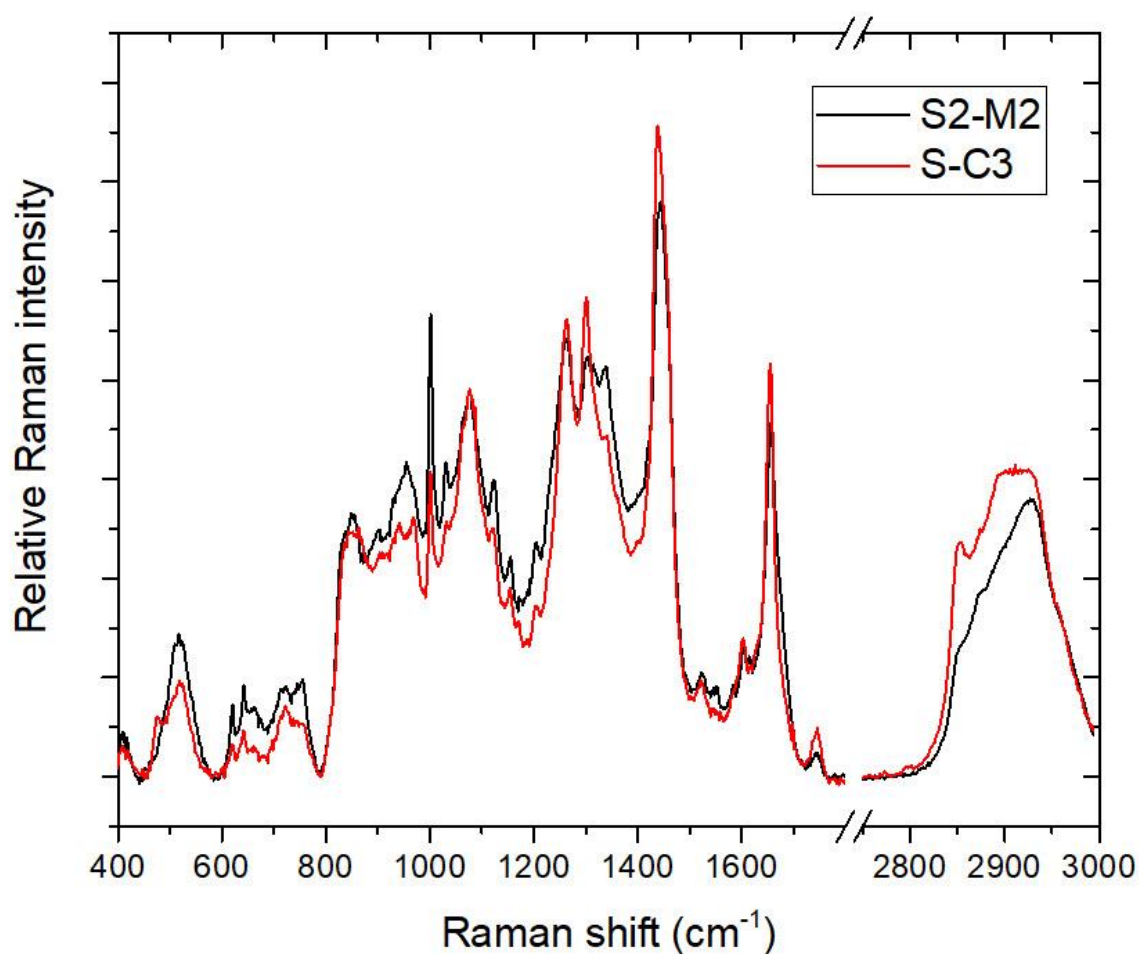


Figure 7. Raman spectra of the studied S04-05 soybean samples exhibiting the most negative and most positive scores along the PC2 axis in the scores' plot shown in Figure 6.



Tables

Table I. Band assignments of Raman spectra of soybean seeds, according to the literature.^{10,20,30}

Raman shift (cm ⁻¹)	Assignment ^a
2924	CH ₂ /CH ₃ asymmetric stretchings of lipids
2855	CH ₂ /CH ₃ symmetric stretchings of lipids
1745	C=O of phospholipids and cholesterol esters
1655	Amide I: Peptide C=O stretching of proteins
1604	C=C stretching
1540	Amide II: Peptide NH bending and C-N stretching of proteins
1441	CH ₂ scissoring/CH ₃ asymmetric bending of lipids
1336	CH ₂ wagging/CH ₃ symmetric bending of lipids
1301	C-N stretching (amide III)/P=O stretchings of phospholipids
1263	CH ₂ twisting of lipids/C-O stretchings phospholipids
1205	CC Aromatic (amino acids, e.g., tryptophan and phenylalanine)
1175	Tyrosine
1151	CH ₃ rocking of lipids

1121	Skeletal C–C, C–N stretching modes
1075	C–O stretching of esters
1003	Phenylalanine (breathing mode)
967	Glycogen/=CH ₂ alkenes
944	CH ₃ rocking of lipids/valine
905	Proline
833	NH rocking (amide IV)/tyrosine
749	=CH ₂ aromatic/tryptophan, tyrosine, phenylalanine
723	CH ₂ rocking
641	OCO bending of esters/phenylalanine

^aAssignments refer to the expected major contributors. In the case of the methylenic vibrations, for example, contributions to polysaccharides and proteins can also be expected

Table II. Lipid–protein ratios in the studied soybean control groups and salt tolerant mutants, and percent variation relative to the controls determined by Raman spectroscopy and quantitative analysis using UV–Vis–NIR reflectance spectroscopy.

Spectral ranges used in the analysis and main contributor	Mutant					
	S1-M	S2-M	S3-M	S4-M	S5-M	A7-M
<i>Raman spectroscopy measurements</i>						
2990–2800 cm ⁻¹ (CH ₂ CH ₃ stretchings) / 1720–1575 cm ⁻¹ (Amide I)	0.912	0.984	0.894	0.974	0.953	1.020
1770–1720 cm ⁻¹ (C=O stretching – phospholipids)/1720–1575 cm ⁻¹ (Amide I)	0.935	0.997	0.938	0.998	0.988	1.114
1500–1200 cm ⁻¹ (CH ₂ deformation modes)/1720–1575 cm ⁻¹ (Amide I)	0.957	0.994	0.910	0.976	0.944	1.042
Average lipid/protein ratio:	0.935	0.992	0.914	0.982	0.969	1.058
Percent variation relative to the controls	-6%	-1%	-9%	-2%	-3%	6%
Average lipid/protein ratio:	-6.2%	1.3%	-9.4%	-5.6%	-0.3%	4.2%
Percent variation relative to the controls						