*ISSN 1392 – 1231. CHEMINĖ TECHNOLOGIJA. 2012. Nr. 4 (62)*

# The application of hierarchical cluster analysis for clasifying horseradish genotypes (*Armoracia rusticana* L.) roots

**L. Tomsone, Z. Kruma, I. Alsina**
*Latvia University of Agriculture,*
*Liela Str. 2, LV-3001 Jelgava, Latvia;*
*E-mail: lolita.tomsone@llu.lv*

**L. Lepse**
*Pure Horticultural Research Centre,*
*Abavas Iela 2, LV-3124 Pure, Latvia*

crossref http://dx.doi.org/10.5755/j01.ct.62.4.3410

*Received 3 December 2012; Accepted 31 December 2012*

Horseradish (*Armoracia rusticana* L.) is a perennial herb belonging to the *Brassicaceae* family; it contains biologically active substances such as phenolic compounds.

The aim of the present research was to clasify horseradish root genotypes, based on the total phenol content and antioxidant properties, using the hierarchical cluster analysis (HCA), and to compare them with clusters obtained from data of the molecular random amplified polymorphic DNA (RAPD) analysis.

Plant phenolic compounds are among the most important primary antioxidants. The phenolic composition of plants is affected by different factors such as variety, genotype, climate, harvest time, storage, processing. Nine genotypes of horseradish roots harvested at three different times in the period from August to November 2011 were used. Several statistical methods can be used to assess differences in the horseradish genotypes. Using a univariate statistical analysis and standard deviations for each analyzed variable does not help to get a complete insight into the complex analysis. Multivariate statistical methods are appropriate tools for the analysis of a complex data matrix. The hierarchical cluster analysis (HCA) used in the current research is a simple way of grouping the set of available data by their similarities according to a set of selected variables. No similarities were found by clustering the genotypes according to the content of biologically active compounds and molecular analyses.

**Keywords:** horseradish, genotype, total phenols, DPPH˙, hierarchical cluster analysis

## Introduction

Many spices and vegetables possess antioxidant properties, so they can be used in food to help prevent oxidation processes. Free radicals in the human body can be formed by heat, radiation, ultraviolet radiation, tobacco smoke, and the influence of alcohol [1]. Some scientists believe that the destruction of free radicals may contribute to the fight with cancer, heart disease, and stroke [2]. Studies show a different antioxidant activity for each plant type, stimulated by the antioxidant components such as α-tocopherol, β-carotene, vitamin C, selenium, and phenolic compounds [3]. Polyphenols are a large, important and diverse class of antioxidants beneficial to both plants and humans. Extensive studies on the functions and role of polyphenols in humans began in the last century and are continued today [4]. Phenolic compounds are known to be very effective antioxidants [5–7]. Plant phenolic compounds are among the most important primary antioxidants, so it is essential to investigate the quantities of these compounds in plants. Phenolic compounds commonly found in spices are biologically active substances having antiseptic, vitamin activity, etc. properties [4, 8].

The phenolic composition of plants is affected by different factors such as variety, genotype, climate, harvest time, storage, processing [9, 10].

Horseradish (*Armoracia rusticana* L.) is a perennial herb belonging to the *Brassicaceae* family and cultivated in temperate regions of the world mainly for the culinary value of its roots. Since horseradish has long been used as a spice for meat and fish products, the Food and Drug Administration (FDA) approved it as a seasoning, spice, and flavouring and affirmed it as generally recognized as safe (GRAS) [11]. Scientists are interested in horseradish because it is a rich source of peroxidase, a heme-containing enzyme that utilizes hydrogen peroxide to oxidize a wide variety of organic and inorganic compounds [12]. Several authors have reported that also the chemical composition of *Brassicaceae* plants varies depending on the stage of development [13], growing conditions [14, 15], and harvest time [16]. This plant is indigenous to eastern and northern Europe and the Mediterranean, with a particularly pungent flavour, rich in glucosinolates and usually consumed as a pickled vegetable. It is also cultivated in central Europe, but not very broadly. Horseradish contains about 0.2 to 1.0 g 100 $g^{-1}$ of essential oil, mainly sinigrin, sinigrin-derived allylisothiocyanate, diallylsulfide, phenylpropyl, and phenethylthiocyanate. The myrosinase enzyme acts on sinigrin to give allylisothiocyanate, which gives horseradish its burning taste. Horseradish has a high vitamin C content (302 mg 100 $g^{-1}$) [1]. Its leaves are considered to prevent food-spoiling processes. Although glucosinolates, with their antioxidant properties, play an

important role in the human diet, they have not been systematically investigated [17]. Several horseradish genotypes are included in the collection of vegetable genetic resources of Latvian origin in the Pure Horticultural Research Centre. Until now, biologically active horseradish substances have not been studied in the Latvian collection. Polish researchers investigated the antioxidant properties of leaf and root extracts originated from four different types of horseradish [17]. The tested types were cultivated in two different regions of Poland. A. Majewska et al. [17] reported that leaf and root extracts derived from four Polish horseradish types did not exhibit strong antioxidant properties, but the different environmental conditions of plant growth affected these properties significantly.

Several statistical methods can be used to evaluate the differences of horseradish genotypes. Using a univariate statistical analysis and standard deviations for each analyzed variable does not help to get a complete insight into the complex analysis. Multivariate statistical methods are appropriate tools for the analysis of the complex data matrix. The hierarchical cluster analysis (HCA) used in the current research is a simple way of grouping the set of available data by their similarities according to a set of selected variables. In other words, one can have a cluster of samples or of variables, depending on what one is looking for in each situation.

A succesful use of HCA for the analysis of foodstuff, raw materials and ingredients [18–20] was reported in several publications. SPSS offers all the tools for obtaining similarity dendrograms, including several distance options, cluster methods and the means of transforming the original data.

The aim of the current research was to clasify horseradish root genotypes, based on the total phenol content and antioxidant properties, using the hierarchical cluster analysis (HCA), and to compare them with the clusters obtained from data of the molecular random amplified polymorphic DNA (RAPD) analysis.

## Materials and methods

*MATERIALS.* Nine horseradish (*Armoracia rusticana* L.) genotypes (Table 1) were used for this study. Horseradish roots were collected at the Pure Horticultural Research Centre (latitude 57° 03′ N, longitude 22° 91′ E) in the period from August to November 2011: 29 August (I); 29 September (II), and 14 November (III). Fresh roots were washed, peeled, and homogenized (for 5 min). All samples of one horseradish type (300 g) were homogenized together in order to obtain a representative sample for the analysis of biologically active substances. For the RAPD analysis, young leaves were collected in June for DNA extraction.

**Table 1.** Characterization of horseradish genotypes

| Collection No | Place of origin | Abbreviations | Total phenol content, mg GAE 100 $g^{-1}$ DW [21] | $DPPH^{\cdot}$, % [21] |
|---|---|---|---|---|
| 1 | Valmiera region, Latvia | G1 | 217.80 ± 5.81 | 13.91 |
| 2 | Belarus | G2 | 237.83 ± 3.74 | 2.19 |
| 3 | Jelgava region, Latvia | G3 | 194.94 ± 4.68 | 12.30 |
| 12B | Preili region, Latvia | G12B | 224.52 ± 2.01 | 5.60 |
| 26B | Malnava region, Latvia | G26B | 160.14 ± 1.41 | 11.27 |
| 105 | Kuldiga region, Latvia | G105 | 273.90 ± 1.04 | 12.64 |
| 106 | Koknese region, Latvia | G106 | 307.52 ± 5.35 | 14.72 |
| 280 | Malnava region, Latvia | G280 | 503.54 ± 5.86 | 29.68 |
| 281 | Malnava region, Latvia | G281 | 360.64 ± 3.08 | 20.70 |

*EXTRACTION PROCEDURE.* For the extraction of phenol compounds, the conventional extraction was used. Five grams of a homogenized sample were extracted with 50 ml of an ethanol/water (80/20 v/v) solution in a conical flask with a magnetic stirrer for 1 h at room temperature. The extraction process was done in triplicate.

DNA was extracted in a routine procedure by using a “Fermentas” DNA purification kit. The PCR reactions contained 20–50 ng DNA, 10 x Taq buffer with KCl–$MgCl_2$, 2.5 mM $MgCl_2$, 0.2 mM dNTP Mix, 0.5 u Taq DNA polymerase (recombinant), 0.2 µM primer in a final volume of 25 µl. The PCR was conducted in the following conditions: denaturation at 94 ºC for 2 min, 35 cycles of denaturation at 94 ºC for 45 sec, annealing at 51 ºC for 45 sec, elongation at 72 ºC for 45 sec, final elongation at 72 ºC for 5 min. PCR products were separated by using electroforesis and visualized with ethidium bromide in UV light.

*DETERMINATION OF TOTAL PHENOL CONTENT (TPC).* The TPC of the root extract was determined according to the Folin–Ciocalteu spectrophotometric method [22] with some modifications. Total phenols were expressed as gallic acid equivalents (GAE) per 100 $g^{-1}$ dry weight (DW) of a sample.

*DETERMINATION OF $DPPH^{\cdot}$ RADICAL SCAVENING ACTIVITY.* Antioxidant activity in the plant extracts was measured on the basis of the scavenging activity of the stable 2,2-diphenyl-1-picrylhydraziyl ($DPPH^{\cdot}$) radical as outlined by Yu [23]. The inhibition of $DPPH^{\cdot}$ in percent (1 %) of each extract sample was calculated. A lower absorbance of the reaction mixture indicates a higher free radical scavenging activity [24].

For horseradish roots, moisture content was determined according to the ISO 6496:1999 standard, and TPC results are expressed to a dry basis.

*DETERMINATION OF DNA POLYMORPHISM.* DNA polymorphism was assessed by scoring bands on gels as present (1) or absent (0) for all the primers and genotypes studied. DNA fragments of a similar length (the same band) were described as the same loci.

*STATISTIC ANALYSIS.* Data obtained in the analysis of horseradish roots were analyzed by means of multivariate analysis, employing a hierarchical cluster analysis. The method used was between-groups linkage. The distances between samples were calculated using square Euclidean distances. As the pre-treatment of data, transform values of variables (average zero and standard deviation 1) called Z scores was carried out. The dendrogram similarity scales generated by the SPSS program ranged from zero (greater similarity) to 25 (lower similarity). The similarities between the analyzed samples were presented in the dendrograms for each harvest period. For DNA fingerprinting analysis, EXCEL software was used to obtain the matrix of similarity coefficients according to Nei and Li [25]. STATISTICA software was used to compile the dendrogram.

## Results and discussion

A hierarchical cluster analysis (HCA) was performed on the TPC, DPPH˙ and genetic markers to distinguish similar or close genotypes.

Hierarchical cluster analysis was applied to a data set of two variables (TPC and DPPH˙) and nine genotypes of horseradish roots. The content of total phenols varied from 160.14 mg GAE 100 $g^{-1}$ DW to 503.54 mg GAE 100 $g^{-1}$ DW, but the DPPH˙ scavenging activity ranged from 2.19 % to 29.68 %. The dendrogram (Fig. 1) shows that the samples of horseradish roots are quite homogeneous and most of samples also tend to be distributed in a homogeneous group, with the exception of a few more scattered, contrary, do not from homogeneous groups based on the variety. According to the hierarchical cluster analysis, at the distance of three, horseradish genotypes can be grouped as follows:

cluster A: genotypes 1, 105, 106 and 2;
cluster B: genotypes 3, 26B and 12B;
cluster C: genotype 281;
cluster D: genotype 280.

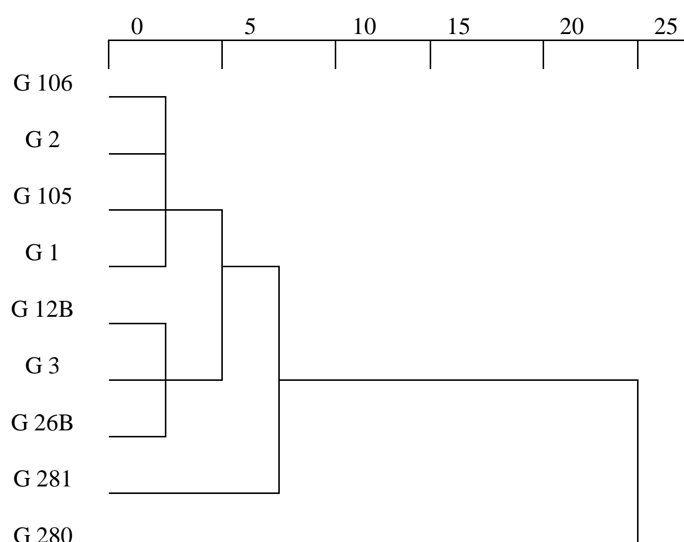


**Fig. 1.** Dendrogram obtained by hierarchical cluster analysis using means of TPC and DPPH ˙
G 106: genotype 106; G 2: genotype 2; G 105: genotype 105; G 1: genotype 1; G 12B: genotype 12B G 3: genotype 3; G 26B: genotype 26B; G 281: genotype 281; G 280: genotype 280

Such a distribution of clusters can be done if we take step 3 as presented in the Chinese scientists' investigation [19, 26]. Results of the analysis showed that TPC and DPPH˙ in horseradish genotypes belonging to clusters A and B were similar, ranging from 192.50 to 307.52 mg GAE 100 $g^{-1}$ DW and 160.14–248.72 mg GAE 100 $g^{-1}$ DW, respectively. A similar tendency for DPPH˙ scavenging activity could be observed. It can be concluded that it is more useful to use distance 6 to distinguish genotypes for the extraction of natural antioxidants. Also, other authors used different steps to interpret the HCA results [18]. Genotypes 280 and 281 showed the highest results of TPC and DPPH˙ in all periods of the analysis. Therefore, they can be segregated as separate groups.

Clustering according to DNA analysis shows a different grouping of genotypes (Fig. 2).

The cluster developed on the basis of the molecular analysis does not correspond to the cluster based on the results of TPC and DPPH. This leads to the assumption that the loci randomly amplified by used primers are not connected with the loci responsible for the TPC and DPPH levels. Some similarities of grouping were found according to the taste properties of the roots and root size (data not shown). The most distinct genotype 12 B was evaluated as the most pungent, and also the root size of this genotype was the biggest.

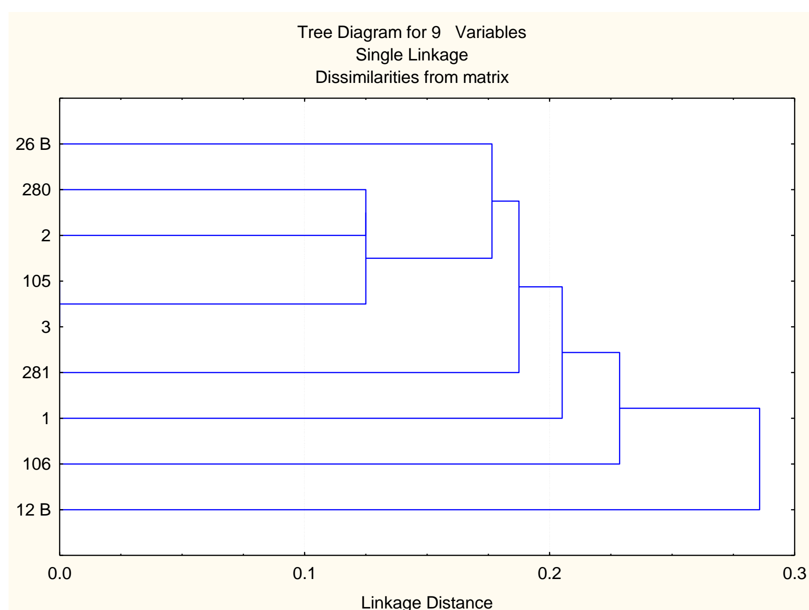


**Fig. 2.** Genotypes grouped according to molecular (RAPD) analysis

## Conclusion

The results obtained in the present study have revealed that the hierarchical cluster analysis can be used to differentiate the genotypes according to certain properties, but they not always correspond to the grouping according to other data. In our case, grouping according to the results of molecular analysis did not show any congruence with the grouping according to biochemical analysis. Nevertheless, both dendrograms segregate most distinct genotypes according to data used in clustering. According to the content of phenolic compounds, genotypes 280 and 281 are separated as the most valuable, but according to differences in genome segments covered by the used primers, genotype 12 B is segregated as most distinct.

## Acknowledgement

Authors acknowledge the Pure Horticultural Research Centre for supplying horseradish roots.

## References

1. **Raghavan Uhl S.** Handbook of Spices, Seasonings, and Flavorings. London, New York, Washington, 2000.
2. **Forristal A., Wang L. J., Shiow Y.** // World & I. 2002. Vol. 17. P. 122.
3. **Ismail A., Marjan Z. M., Foong C. W**. // Food Chemistry. 2004. Vol. 87. P. 581–586. http://dx.doi.org/10.1016/j.foodchem.2004.01.010
4. **Rappoport Z.** The Chemistry of Phenols. 2003.
5. **Shahidi F., Wanasundara P. K. J. P. D.** // Food Science and Nutrition. 1992. Vol. 32. P. 67–103.
6. **Tapeiro H., Tew K. D., Nguyenba G., Mathe G.** // Biomedicine and Pharmacotherapy. 2002. Vol. 56. P. 200–207. http://dx.doi.org/10.1016/S0753-3322(02)00178-6
7. **Shahidi F., Naczk M.** // Journal of Chromatography A. 2004. Vol. 1054. P. 95–111.
8. **Daayf F., Lattanzio V.** Recent Advances in Polyohenol Research. Wiley-Blackwell, 2008.
9. **Kreutzmann S., Christensen L.P., Edelenbos M**. // LWT - Food Science and Technology. 2008. Vol. 41. P. 193–205. http://dx.doi.org/10.1016/j.lwt.2007.02.024
10. **Marrelli M., Menichini F., Statti G.A., Bonesi M., Duez P., Menichini F., Conforti F.** // Food and Chemical Toxicology. 2012. Vol. 50. P. 726–733. http://dx.doi.org/10.1016/j.fct.2011.12.025
11. Food and Drug Administration, // Food and Drug Administration. 2008. Everything added to food in the United States (EAFUS). <http://www.fda.gov/Food/FoodIngredientsPackaging/ucm115326.htm>
12. **Veitch N. C.** // Phytochemistry. 2004. Vol. 65. P. 249–259. http://dx.doi.org/10.1016/j.phytochem.2003.10.022
13. **Björkman M., Klingen I., Birch A. N. E., Bones A. M., Bruce T. J. A., Johansen T. J., Meadow R., Mølmann J., Seljåsen R., Smart L. E., Stewart D.** // Phytochemistry. 2011 Vol. 72. P. 538–556. http://dx.doi.org/10.1016/j.phytochem.2011.01.014
14. **Podsędek A.** // LWT - Food Science and Technology. 2007. Vol. 40. P. 1–11. http://dx.doi.org/10.1016/j.lwt.2005.07.023
15. **Kusznierewicz B., Bartoszek A., Wolska L., Drzewiecki J., Gorinstein S., Namieśnik J**. // LWT - Food Science and Technology. 2008. Vol. 41. P. 1–9. http://dx.doi.org/10.1016/j.lwt.2007.02.007
16. **Koh E., Wimalasiri K. M. S., Chassy A. W., Mitchell A. E.** // Journal of Food Composition and Analysis. 2009. Vol. 22. P. 637–643. http://dx.doi.org/10.1016/j.jfca.2009.01.019
17. **Majewska A., Bałasińska B., Dąbrowska B.** // Folia Horticulturae. 2004. Vol. 16/1. P. 15–22.
18. **Ferraz da Silva Torres E. A., Garbelotti M. L., Moita Neto J. M.** // Food Chemistry. 2006. Vol. 99. P. 622–629. http://dx.doi.org/10.1016/j.foodchem.2005.08.032

19. **Sun W., Zhao O., Zhao H., Zhao M., Yang B.** // Food Chemistry. 2010. Vol. 121. P. 319–325. http://dx.doi.org/10.1016/j.foodchem.2009.12.031
20. **Chaieb N., González J.L., López-Mesas M., Bouslama M., Valiente M.** // Food Research International. 2011. Vol. 44. P. 970–977. http://dx.doi.org/10.1016/j.foodres.2011.02.026
21. **Tomsone L., Kruma Z., Lepse L.** // Proceeding of Conference Research for Rural Development 2012. Jelgava, 2012. P. 124–130.
22. **Singleton V. L., Orthofer R., Lamuela-Raventos R. M.** // Methods in Enzymology. 1999. Vol. 29. P. 152–178. http://dx.doi.org/10.1016/S0076-6879(99)99017-1
23. **Yu L., Haley S., Perret J., Harris M., Wilson J., Haley S.** // Journal of Agriculture and Food Chemistry. 2003. Vol. 51. P. 1566–1570. http://dx.doi.org/10.1021/jf020950z
24. **Zhao H., Fan W., Dong J., Lu J., Chen J., Shan L., Lin Y., Kong W.** // Food Chemistry. 2008. Vol. 107. P. 296–304. http://dx.doi.org/10.1016/j.foodchem.2007.08.018
25. **Nei M., Li W. H.** // Proc. Natl. Acad. Sci. 1979. Vol. 76. P. 5269–5273. http://dx.doi.org/10.1073/pnas.76.10.5269
26. **Zhao D., Tang J., Ding X.** // LWT - Food Science and Technology. 2007. Vol. 40. P. 439–447. http://dx.doi.org/10.1016/j.lwt.2005.12.002

L. Tomsone, Z. Kruma, I. Alsina, L. Lepse

## HIERARCHINĖS KLASTERINĖS ANALIZĖS TAIKYMAS KRIENŲ (*ARMORACIA RUSTICANA* L.) ŠAKNŲ GENOTIPAMS KLASIFIKUOTI

### **S a n t r a u k a**

Valgomasis krienas (*Armoracia rusticana* L.) yra daugiametis augalas priklausantis bastutinių (*Brassicaceae*) šeimai. Auginamas kaip prieskoninis augalas, kuriame yra biologiškai aktyvių medžiagų, fenolinių junginių.

Šio darbo tikslas suklasifikuoti krienų genotipus, remiantis nusatytu fenolinių junginių kiekiu bei antioksidacinėmis savybėmis taikant hierarchinę klasterinę analizę (HCA), ir palyginti su molekulinės analizės duomenimis (RAPD).

Augalų fenoliniai junginiai yra vienas iš svarbiausių pirminių antioksidantų. Augalų fenolinių junginių sudėtis priklauso nuo daugelio veiksnių – veislės, genotipo, klimato, derliaus nuėmimo, perdirbimo sąlygų. Tyrimuose naudotos devynių krienų genotipų šaknys, kurios rinktos tris kartus skrtingu 2011 m. rugpjūčio–lapkričio mėn. laikotarpiu. Yra keletas statistikos metodų, taikomų augalams klasifikuoti pagal genotipus. Pasirinkus vienmatės statistikos analizės metodus su standartiniais nuokrypiais, sunku įvertinti gautus rezultatus. Analizuojant sudėtingas duomenų matricas, naudojama hierarchinė klasterinė analizės sistema (HCA), kuri, mokslinių tyrimų duomenimis, yra paprastas duomenų grupavimo būdas, pasirenkant vieną kintamąjį. Atlikus krienų genotipų tyrimus, naudojant hierarchinę klasterinę analizės sistemą, tarp biologiškai aktyvių junginių ir molekulinės analizės panašumų nebuvo rasta.