Bangladesh J. Bot. **48**(1): 17-24, 2019 (March)

IN SILICO ANALYSIS OF PUTATIVE POLYPHENOL OXIDASES IN OLIVE USING BIOINFORMATICS TOOLS

EMRE SEVINDIK*

Faculty of Agriculture, Department of Agricultural Biotechnology, Adnan Menderes University, South Campus, Cakmar, Aydin, Turkey

Keywords: In silico, Putative polyphenol oxidase, *Olea europaea*, Bioinformatics tools

Abstract

Sequence, physicochemical, and three-dimensional structure properties of putative polyphenol oxidase proteins in olive (*Olea europaea*) using various bioinformatics tools were analyzed. The amino acid length varied from 469 to 582 amino acids. The molecular weights of the proteins (ALG62778.1, AFS28698.1 and AFS28697.1) were 65294.67 Da, 53324.79 Da and 53349.48 Da, and isoelectric points (pI) were 7.58, 7.24 and 6.86, respectively. Instability index values were 40.68, 37.52 and 36.89 while aliphatic index values were 72.08, 73.39 and 73.39 respectively. The GRAVY values were –0.540, –0.580 and –0.578, respectively. The most abundant amino acid was Asp (8.5%) while the least abundant one was Cys (1.4%). The putative phosphorylation sites of the polyphenol oxidase proteins were determined by NetPhos 2.0 and NetPhos 3.1. Based on the phylogenetic analysis, the tree constructed using polyphenol oxidase proteins is composed of two main clades. To predict the three dimensional (3D) structures of these proteins, Py MOL was used. The results of the present study provide insight into fundamental characteristics of putative polyphenol oxidase proteins in *Olea europaea*.

Introduction

Bioinformatic analysis might offer great help in designing better alternatives of enzymes in silico (Sheth and Thaker 2014). Browning and spoiling may occur either as a result of mechanical damage on fruits and vegetables during harvest or transport; or during various activities such as chopping, crushing and patting applied for processing of a product. The reason of this change is the activity of polyphenol oxidase (PPO) enzyme. Polyphenol oxidase belongs to the group of oxidoreductases with 2 copper ions located in the active center (Yemenicioğlu and Cemeroğlu 1998, Tran *et al.* 2012, Demir 2013). These enzymes oxidize ortho-diphenols to ortho-diquinones using molecular oxygen. Some polyphenol oxidases also convert monophenols to ortho-diphenols (Constabel and Barbehenn 2008). Polyphenol oxidase is an important enzyme because it has a positive impact on plant resistance against different biotic and abiotic stresses (Mahmood *et al.* 2015). This enzyme is localized in plants' chloroplast thylakoids membranes. Polyphenol oxidase is expressed more than required for the purpose of defense in case of an infection led by pathogen bacteria. In addition, a similar defense activity can be observed during insect invasion. This enzyme creates a defense system against insects. At the same time, it is asserted that polyphenol oxidase is being connected with auron biosynthesis and phenylpropanoid. Furthermore, polyphenol oxidase is reported to be functional in the Mehler reaction at chloroplast mesophiles (Thipyapong *et al.* 2004). It is known that quinones, which emerge as a result of polyphenol oxidase activity, form dark-colored water-insoluble polymers through a polymerization reaction. Injured tissues filled with polymers set an obstacle against the dissemination of the infection (Turan 2005).

Olea europaea (olive) is an important fruit tree in the world and it is one of the most frequently used food source for its valuable oil, fruits, flowers, leaves and wood (Coşkun and

*Author for correspondence: <ph.d-emre@hotmail.com>.

Parlak 2013). *Olea europaea*, with more than 2,600 cultivars, is one of the oldest plants cultivated in the Mediterranean area where it is the most important oil-producing crop (Kaya 2015). Olive oil, an important commodity, was restricted to the Mediterranean basin in the past, but nowadays it has become a widespread product in the world due to its economical importance as a food source, fuel for lighting and also as an ointment (Sanchez and Spener 2002). The basis of Mediterranean diet is olive oil. It has positive effects on human health through healthy nutrients such as monounsaturated fatty acids, high oleates and significant linoleate contents, vitamins, minerals, dietary fibers, carbohydrates and other minor components (Harwood and Yaqoob 2002, Wahrburg *et al.* 2002). The aim of this study was to generate predicted 3D structures of putative polyphenol oxidases by using comparative homology modeling. Also, primary and secondary structure analyses were performed with various bioinformatics tools.

Material and Methods

The protein sequences of putative polyphenol oxidase (PPO) (Accession no: ALG62778.1, AFS28698.1 and AFS28697.1) for *Olea europaea* were retrieved from NCBI (https://www.ncbi.nlm.nih.gov/protein). The physicochemical analysis to determine its isoelectric point (pI), molecular weight (Mw), instability index (II), aliphatic index (Al), and GRAVY values of the proteins were done using ExPASy's ProtParam (http://web.expasy.org/protparam/) (Gasteiger *et al.* 2005). The average amino acid rates were determined by MEGA 6.0 (Tamura *et al.* 2013). The putative phosphorylation sites of the putative polyphenol oxidase proteins were detected by NetPhos 2.0 (http://www.cbs.dtu.dk/services/NetPhos/) (Blom *et al.* 1999). Amino acid sequences of putative polyphenol oxidase proteins from 11 plant species were aligned using MEGA 6.0 software and a phylogenetic tree was successfully constructed by neighbor-Joining (NJ) method while reliability of each node was determined by bootstrap calculation (1000 replicates) using MEGA 6.0 (Saitou and Nei 1987, Tamura *et al.* 2013). Secondary structure predictions were performed by using SOPMA server (http://npsa-pbil.ibcp.fr/). Subcellular localizations were predicted using CELLO v.2.5, a multi-class SVM (support vector machine) classification system (http://cello.life.nctu.edu.tw/) (Yu *et al.* 2006). To predict the 3D structure of the putative polyphenol oxidases, homology models were used following PSIPRED v.3.3 method options (http://bioinf.cs.ucl.ac.uk/psipred/) (Buchan *et al.* 2013). The results were checked and verified by a Ramachandran plot analysis in RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) (Lovell *et al.* 2003), which determined the best predicted models. Finally, 3D comparative analyzes were performed using PyMOL (https://www.pymol.org/).

Results and Discussion

The physicochemical analysis of the putative polyphenol oxidase proteins was performed using ExPASy-ProtParam and results are shown in Table 1. The amino acid length varied from 469 to 582 amino acids. The molecular weights of the proteins (ALG62778.1, AFS28698.1 and AFS28697.1) were 65294.67 Da, 53324.79 Da and 53349.48 Da, and isoelectric points (pI) were 7.58, 7.24 and 6.86, respectively. Instability index values were 40.68, 37.52 and 36.89 while aliphatic index values were 72.08, 73.39 and 73.39. The GRAVY values were –0.540, –0.580 and –0.578, respectively (Table 1). The prediction of protein subcellular localization (PSL) focuses on determining localization sites of unknown proteins in a cell (Su *et al.* 2007). It was determined (using CELLO v. 2.5) that polyphenol oxidase localizes in chloroplast. The average amino acid rates were determined by MEGA 6.0. Average amino acid composition of putative polyphenol oxidases revealed the highest ratio (8.5%) for Asp, and the lowest (1.4%) for Cys (Fig. 1). The putative phosphorylation sites were determined using the NetPhos 2.0 and NetPhos 3.1 server

based on a score above 0.8. As a result, it was determined that the putative polyphenol oxidase proteins are the serine residue of the most frequent phosphorylation site. The highest serine residue was identified as *Olea europaea* ALG62778.1 (Fig. 2). For phylogenetic analysis, MEGA

Table 1. The physiochemical properties of the putative polyphenol oxidase proteins from *Olea europaea* ssp.

Index	*O. europaea* (ALG62778.1)	*O. europaea* (AFS28698.1)	*O. europaea* (AFS28697.1)
Amino acids	582	469	469
Molecular weight	65294.67	53324.79	53349.48
Theoretical pI	7.58	7.24	6.86
Instability index	40.68	37.52	36.89
Aliphatic index	72.08	73.39	73.39
GRAVY	-0.540	-0.580	-0.578

Fig. 1. Theaveragae amino acidratios of theputativepolyphenoloxidasesfrom*Oleaeuropaea*spp. numbers on Y axisrepresentthepercentages.

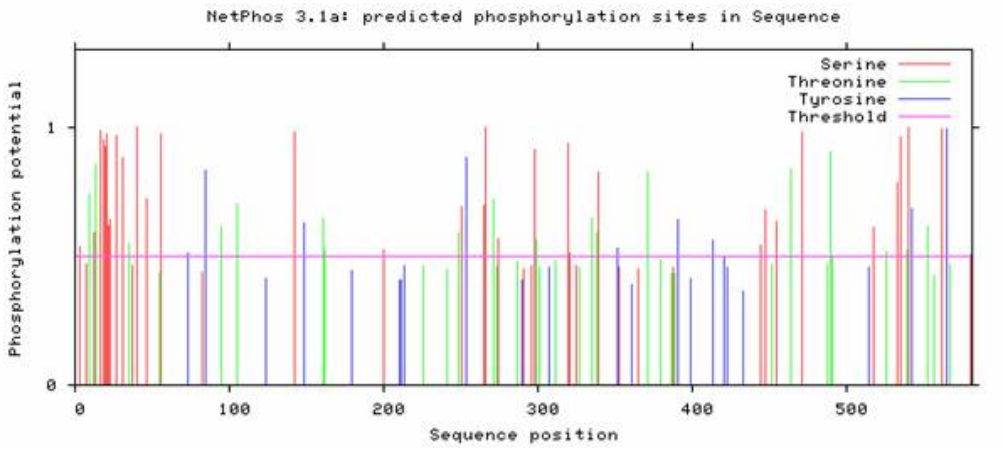

Fig. 2. Putative phosphorylation sites in *Olea europaea* ALG62778.1 predicted by NetPhos 3.1 Server.

Table 2. Pairwise distances data of the putative polyphenol oxidase proteins obtained using MEGA 6.0.

Taxa	1	2	3	4	5	6	7	8	9	10	11
Olea europaea ALG62778.1	-										
Olea europaea AFS28698.1	0.015	-									
Olea europaea AFS28697.1	0.015	0.000	-								
Prunus mexicana	0.696	0.716	0.716	-							
Prunus dulcis	0.706	0.721	0.721	0.023	-						
Pyrus communis subsp. *caucasica*	0.696	0.711	0.711	0.143	0.149	-					
Vauquelinia californica	0.716	0.732	0.732	0.146	0.152	0.046	-				
Sorbus californica	0.691	0.706	0.706	0.140	0.146	0.023	0.041	-			
Fallugia paradoxa	0.661	0.676	0.676	0.135	0.137	0.188	0.197	0.185	-		
Spiraea densiflora	0.721	0.737	0.737	0.244	0.244	0.290	0.280	0.277	0.270	-	
Spiraea cantoniensis	0.721	0.737	0.737	0.238	0.238	0.277	0.267	0.264	0.277	0.041	-

Table 3. Secondary structures of the putative polyphenol oxidase proteins (obtained using SOPMA)..

	Alpha helix (Hh) %	310 helix (Gg)	Pi helix (Ii)	Beta bridge (Bb)	Extended strand (Ee) %	Beta turn (Tt) %	Bend region (Ss)	Random coil (Cc) %	Ambiguous states
O. europaea (ALG62778.1)	20.79	0.00	0.00	0.00	18.04	2.92	0.00	58.25	0.00
O. europaea (AFS28698.1)	24.31	0.00	0.00	0.00	20.68	3.84	0.00	51.17	0.00
O. europaea (AFS28697.1)	24.73	0.00	0.00	0.00	19.62	4.05	0.00	51.60	0.00

6.0 program was used. Together with *Olea europaea* polyphenol oxidase protein sequences, those of *Prunus mexicana, Prunus dulcis, Pyrus communis* subsp. *caucasica, Vauquelinia californica, Sorbus californica, Fallugia paradoxa, Spiraea densiflora* and *Spiraea cantoniensis* were also retrieved from NCBI and used to construct a phylogenetic tree. The phylogenetic tree based on Neighbor-Joining method displays two large clades. *Olea europaea* species are placed in one clade while the other species are in the second one (Fig. 3). Additionally, pairwise distances analysis involving PPO sequences of *Olea europaea* and other species was performed using MEGA 6.0.

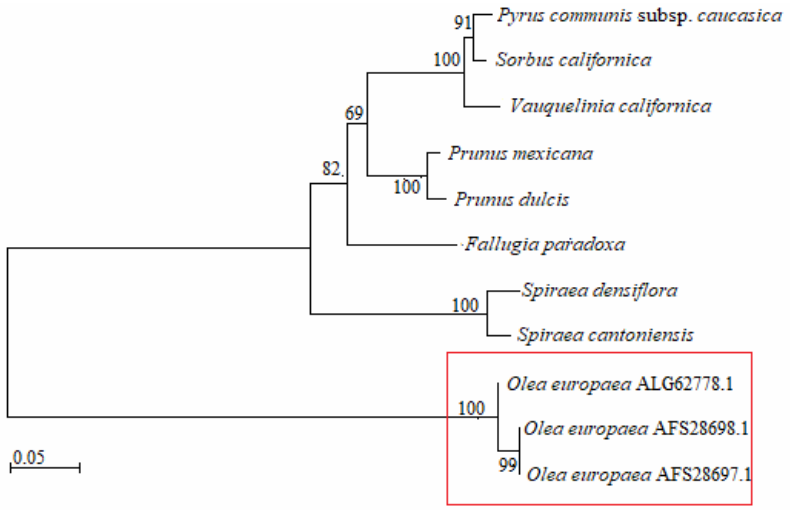

Fig. 3. Molecular phylogram analysis of putative polyphenol oxidase protein sequences from other plants. Phylogenetic trees were constructed by Neighbor-Joining (NJ) as well as the bootstrap values were showed on branch using MEGA 6.0 software. The GenBank accession numbers of the protein sequences used for the phylogenic analysis; *Prunus mexicana* (ABI96945.1), *Prunus dulcis* (ABI96944.1), *Pyrus communis* subsp. *caucasica* (ABI96946.1), *Vauquelinia californica* (ABI96951.1), *Sorbus californica* (ABI96948.1), *Fallugia paradoxa* (ABI96932.1), *Spiraea densiflora* (ABI96950.1) and *Spiraea cantoniensis* (ABI96949.1).

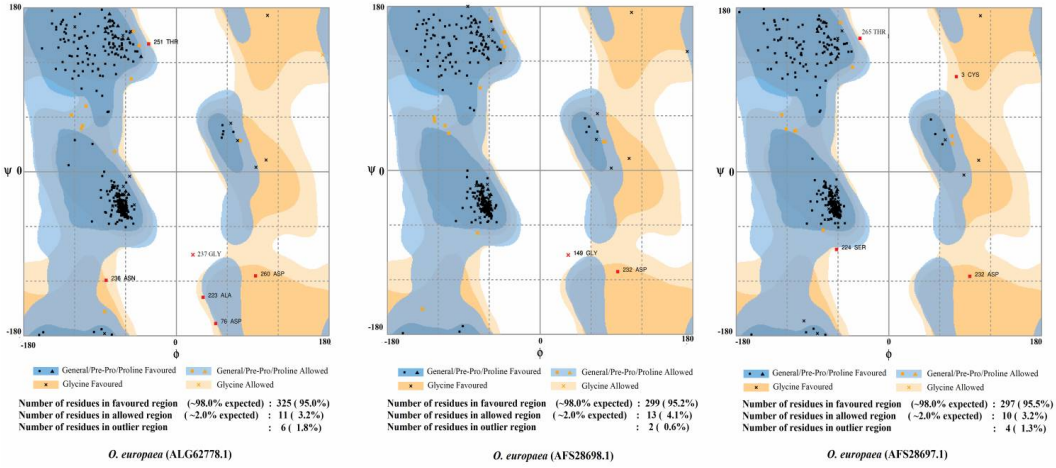

Fig. 4. RAMPAGE values indicating number of residues in favored, allowed and outer regions.

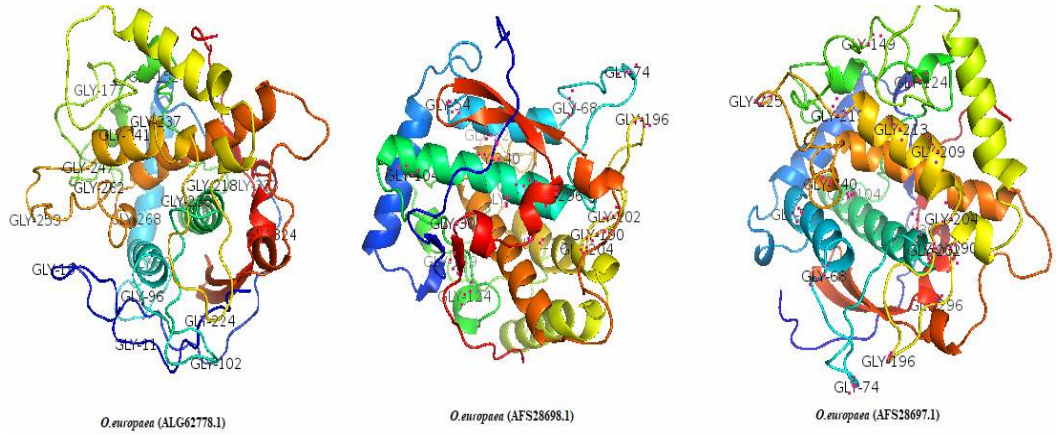

Fig. 5. Conserved Gly residues of *Olea europaea* ssp. polyphenol oxidases depicted on three dimensional models.

The lowest pairwise distance was 0.000 while the highest pairwise distance was determined as 0.737 (Table 2). The secondary structures of the proteins were predicted using SOPMA server (Geourjon and Deléage 1995). It was observed that the alpha helix rates of the proteins (ALG62778.1, AFS28698.1 and AFS28697.1) were 20.79, 24.31 and 24.73%, extended strand rates were 18.04, 20.68 and 19.62%, beta turns were 2.92, 3.84 and 4.05%, and the predominant random coils were 58.25, 51.17 and 51.60% (Table 3). Random coils have important functions in proteins for flexibility and conformational changes such as enzymatic turnover (Filiz and Koç 2014). In the model validation, the Ramachandran plot analysis using the RAMPAGE server showed that 95.0, 95.2, and 95.5% were in the favoured region; 3.2, 4.1, and 3.2% in the allowed region; and 1.8, 0.6 and 1.3 % were in the outlier region in *Olea europaea* ssp. (Fig. 4). The three-dimensional structures of the putative polyphenol oxidases were constructed using the PyMOL program, and the alpha helix and beta sheet structures were demonstrated. Gly residue is unique among amino acids since all side chains of it are hydrogen atoms. Its conformation has more freedom, so it can provide flexibility for adjacent residues. Therefore, it is not surprising that Gly plays a special role in the structure and function of the enzyme (Yan and Sun 1997). In this study, Gly residues are presented in Fig. 5. In addition, the three-dimensional structures of the proteins contribute to the understanding of protein function and active regions thereby facilitating drug design (Filiz and Koç 2014). Bioinformatics is a field based on mathematics and computer science to understand biology. In the post genome era, research on protein structures and functions is the focus of molecular biology. Today, a number of computing software and online servers are rapidly being developed for the identification and characterization of proteins and coding nucleotide sequences. Physiochemical properties and biological functions of proteins can be studied better by bioinformatics methods (Li *et al.* 2017). As a results of this study, in silico analysis was carried out using bioinformatics tools such as ExPASy's ProtParam, NetPhos 2.0, NetPhos 3.1 MEGA 6.0, CELLO v.2.5., PSIPRED v3.3, RAMPAGE and PyMOL for the putative polyphenol oxidases in *Olea europaea*. The results of this study has the potential to pave the way for further research on the putative polyphenol oxidases in different plant species, and to shed light on future in silico studies on this enzyme

References

Blom N, Gammeltoft S and Brunak S 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites1. J. Mol. Biol. **294**: 1351-1362.

Buchan DW, Minneci F, Nugent T C, Bryson K and Jones D T 2013. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res. **41**: 349-357.

Constabel CP and Barbehenn R 2008. Defensive roles of polyphenol oxidase in plants. In Induced Plant Resistance to Herbivory. Dordrecht: Springer. 253-270.

Coşkun F and Parlak S 2013. Molecular phylogenetic analysis of *Olea europaea* L. subsp. *europaea* cultivars grown in the Marmara Region, Turkey. Sains Malaysiana. **42**: 1357-1364.

Demir D 2013. Zeytinde Polifenol Oksidaz Enzimi (Derleme). ZZ Dergisi. **29**: 68-73.

Filiz E and Koç İ 2014. In silico sequence analysis and homology modeling of predicted beta-amylase 7-like protein in *Brachypodium distachyon* L. J. BioSci. Biotech. 3: 61-67.

Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel R D and Bairoch A 2005. Protein identification and analysis tools on the ExPASy server. In The proteomics protocols handbook, Humana Press. 571-607.

Geourjon C and Deléage G 1995. SOPMA: Significant improvement in protein secondary structure prediction by consensus prediction from multiple alignments, Computer Application in the Bioscience. CABIOS. **11**: 681-684

Harwood JL and Yaqoob P 2002. Nutritional and health aspects of olive oil. Eur. J. Lipid Sci. Technol. **104**: 685-697.

Kaya E 2015. ISSR Analysis for Determination of Genetic Diversity and Relationship in Eight Turkish Olive (*Olea europaea* L.) Cultivars. Not. Bot. Hortı. Agrobo. **43**: 96-99.

Li G, Lan X, Shui X, Huang S, Chen C and Lei W 2017. In silico Analysis of 4CL Family in Scutellaria baicalensis through Biocomputational Tools and Servers. Am. J. Biochem. Biotechnol. **13**: 27.33

Lovell SC, Davis IW, Arendall III WB, De Bakker PI, Word JM, Prisant MG and Richardson DC 2003. Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. Proteins: Protein Struct Funct Genet. **50**: 437-450.

Mahmood T, Rehman MU, Aziz E, Ali I and Shinwari ZK 2015. In silico analysis, mapping of regulatory elements and corresponding dna-protein interaction in polyphenol oxidase gene promoter from different rice varieties. Pak. J. Bot. **47**: 2321-2327.

Saitou N and Nei M 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol. Evol. **4**: 406-425.

Sanchez J and Spener F 2002. Editorial: Olive oil - a special issue. Eur. J. Lipid Sci. Technol. **104**: 557-558.

Sheth BP and Thaker VS 2014. In silico analyses of RuBisCO Enzymes from different classes of Algae. Int. Res. J. Biological Sci. **3**: 11-17.

Su ECY, Chiu HS, Lo A, Hwang JK, Sung TY and Hsu WL 2007. Protein subcellular localization prediction based on compartment-specific features and structure conservation. BMC Bioinformatics. **8**: 330.

Tamura K, Stecher G, Peterson D, Filipski A and Kumar S 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. **30**: 2725-2729.

Thipyapong P, Melkonian J, Wolfe DW and Steffens JC 2004. Supression of polyphenol oxidase increases stress tolarence in tomato. Plant Sci. **167**: 693-703.

Tran LT, Taylor JS and Constabel CP 2012. The polyphenol oxidase gene family in land plants: Lineage-specific duplication and expansion. BMC Genomics **13**: 395.

Turan P 2005. Ocimum basilicum L. Polifenol Oksidazının Saflastırılması, Karakterizasyonu ve Kinetik Özelliklerinin Incelenmesi, Yüksek Lisans Tezi, Balıkesir Üniversitesi Fen Bilimleri Enstitüsü, Kimya Anabilim Dalı, Balıkesir.

Wahrburg U, Kratz M and Cullen P 2002. Mediterranean diet, olive oil and health. Eur. J. Lipid Sci. Technol. **104**: 698-705.

Yan BX and YQ Sun 1997. Glycine Residues Provide Flexibility for Enzyme Active Sites. J. Biol. Chem. **272**: 3190-3194.

Yemenicioğlu A and Cemeroğlu B 1998. Hale Haven şeftalilerinde polifenol oksidaz enzimlerinin bazı nitelikleri. Tr. J. Agric. For. **22**: 261-265.

Yu CS, Chen YC, Lu CH and Hwang JK 2006. Prediction of protein subcellular localization. Proteins: Protein Struct Funct Genet. **64**: 643-651.

(Manuscript received on 1February, 2018; revised on 9 September, 2018)