UDC 575
https://doi.org/10.2298/GENSR1802603S
Original scientific paper

GENOME-WIDE IDENTIFICATION AND ANALYSIS OF RUBISCO LARGE SUBUNIT PROTEINS IN *Morus* L. (MORACEAE) SPECIES

Emre SEVİNDİK

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

Sevindik E. (2018): *Genome-wide identification and analysis of rubisco large subunit proteins in Morus L. (Moraceae) species*.- Genetika, Vol 50, No.2, 603-616.

In this study, we performed sequence, physicochemical, phylogenetic, and comparative 3D (three-dimensional) analyses of RuBisCO large subunit proteins of the *Morus* species using various bioinformatics tools. The sequence lengths of the RuBisCO proteins varied between 140 and 475 amino acids with molecular weights (MW) ranging from 15463.40 to 52667.95 Da. The most acidic protein sequences were in *M. australis, M. rotundiloba, M. serrata*, and *M. cathayana* (*pI*=4.77). The extinction coefficients of the proteins at 280 nm ranged from 21430 to 69830 M^{-1} cm^{-1}, the instability index (*II*) values ranged from 26.06 to 42.17, and the GRAVY values ranged from -0.355 to -0.215. The most abundant amino acid of RuBisCO proteins was Gly (9.5%) while the least abundant ones were Cys and Trp (1.5%). Based on the phylogenetic analysis, the tree constructed using RuBisCO proteins is composed of three main clades. A RAMPAGE analysis revealed that 96.1% - 98.5% of residues were located in the favored region in RuBisCO proteins. To predict the three dimensional (3D) structures of the RuBisCO proteins PyMOL was used. The highest number of Gly residues was detected in *M.mongolica, M. indica, M.notabilis, M. alba* var. *atropurpurea* and *M. alba* var. *multicaulis*, while the least Gly residues were detected in *M. australis, M.rotundiloba, M. serrata and M.cathayana*. The results of our study provide insights into fundamental characteristic of RuBisCO proteins in *Morus* species.

Key words: genome-wide analysis, *Morus*, phylogenetics, RuBisCO

Corresponding author: Emre Sevindik, Faculty of Agriculture, Department of Agricultural Biotechnology, Adnan Menderes University, South Campus, Cakmar, Aydin, Turkey, e-mail: ph.d-emre@hotmail.com

INTRODUCTION

The 37 genera of *Moraceae* have a striking array of inflorescence forms, pollination syndromes, and breeding systems (DATWYLER and WEIBLEN, 2004). The mulberry belongs to the *Morus* genus of the *Moraceae* family. There are 24 species of *Morus*, with at least 100 known varieties (ORHAN and ERCISLI, 2010). Mulberry is distributed in a wide area of tropical, subtropical, and temperate zones in Asia, Europe, North America, South America, and Africa. The trees historically have been used for sericulture especially in east, central, and south Asia (KAFKAS *et al.*, 2008). Mulberry grown as a perennial tree or shrub is an economically important plant used for sericulture, as it is the sole food source for the domesticated silkworm, *Bombyx mori* (AWASTHI *et al.*, 2004; WEIGUO *et al.*, 2005). In most European countries, mulberries are grown for fruit production (ARFAN *et al.*, 2012). Recently, red and black mulberry have gained an important place in food industry due to their content of anthocyanin (ÖZRENK *et al.*, 2010). *Morus* species have been found to contain high levels of anthocyanins which are flavonoids and are contained in fruits, leaves and other plant organs. These are anthocyanins and flavonol. Due to their antioxidant ingredients, studies have reported that consumption of fresh and fermented mulberry fruits may be beneficial for human health (ZAFAR *et al.*, 2013; PEREZ-GREGORIO *et al.*, 2011).

Photosynthesis can be defined as the reduction of CO_2 by green plants under the catalysis of special pigment molecules with the help of light energy in the presence of water to form various organic substances. In this system, solar energy is stored as chemical energy in organic matter (TÜRK and ÇELIK, 2006). RuBisCO, the most abundant enzyme on earth, catalyses the atmospheric CO_2 uptake in plants (PORTIS, 1992; SUBRAMANI and HWA, 2010). This enzyme transforms the carbon dioxide and ribulose-1,5-bisphosphate (RuBP) into two molecular 3-phosphoglyceric acid, and catalyzes the first reaction of carbon dioxide fixation in photosynthetic dark reaction. Also, RuBisCO catalyzes the reaction of oxygen and RuBP to phosphoglyceric acid and phosphoglycolic acid, which is the first reaction of photorespiration (ANDERSSON *et al.*, 1989; ZHANG *et al.*, 2011). In higher plants, it is composed of 8 large subunits and 8 small subunits both arranged in 4 dimers. The large subunits of RuBisCO (*rbc*L) are encoded in the chloroplast whereas the small subunits (*rbc*S) are encoded in the nucleus (ELLIS, 1979; MUKHERJEE *et al*., 2015). In addition to attracting interest because of its importance in photosynthesis, studies of RuBisCO continue to make major contributions in our understanding of plant diversity, chloroplast biogenesis, the coordination and regulation of gene expression in the chloroplast and nucleus, and chloroplast import and processing of nuclear-encoded proteins (PORTIS, 2009). In this study, RuBisCO large subunit proteins were analyses in economically valuable *Morus* species with respect to physicochemical, phylogenetic and 3D structure properties utilizing bioinformatics tools.

MATERIAL AND METHODS

The RuBisCO large subunit protein sequences of *Morus* species were retrieved in FASTA format from the NCBI (https://www.ncbi.nlm.nih.gov/protein). The NCBI accession numbers are given in Table 1. The physicochemical analysis were analyzed by ExPASy's ProtParam (http://web.expasy.org/protparam/) to determine isoelectric point (*pI*), molecular weight (MW), total number of positive (+R) and negative (-R) residues, extinction coefficient (EC), instability index (II), aliphatic index (Al), and GRAVY values (GASTEIGER *et al.*, 2005).

The putative phosphorylation sites of the RuBisCO proteins were detected by NetPhos 2.0 and NetPhos 3.1 (http://www.cbs.dtu.dk/services/NetPhos/). All of the protein sequences were aligned using MEGA 6.0 (TAMURA *et al.,* 2013). A phylogenetic tree of *Morus* species' RuBisCO proteins were constructed using Maximum likelihood method with MEGA 6.0 and bootstrap values were calculated using 1000 replicates. To predict the 3D structure of the RuBisCO proteins, homology models were performed using PSIPRED server (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 detected the best predicted models. 3D comparative analyses were performed using PyMOL.

RESULTS AND DISCUSSION

The physicochemical analysis of the predicted RuBisCO large subunit proteins were performed using ExPASy Protparam (Table 1). The amino acid sequence length ranged from 140 to 475. The shortest amino acid sequences were in *M. australis*, *M. rotundiloba*, *M. serrata*, *M. cathayana* (140 aa.) while the longest amino acid sequences were in *M. mongolica*, *M. indica*, *M. notabilis*, *M. alba* var. *atropurpurea* and *M. alba* var. *multicaulis* (475 aa.) species. The lowest and highest MW values were determined as 15463.40 Da and 52667.95 Da, respectively.

Table 1. The physiochemical properties of RuBisCO proteins from Morus species

Taxa	NCBI Accession Number	Sequence Lenght (aa)	Mw	*pI*	-R	+R	EC	II	Al	GRAVY
M. australis	ADA70010.1	140	15463.40	4.77	19	13	21430	31.82	79.43	-0.266
M.mesozygia	AJP33337.1	184	20333.20	6.57	21	21	29005	26.06	81.63	-0.215
M.rotundiloba	ADA70015.1	140	15463.40	4.77	19	13	21430	31.82	79.43	-0.266
M. serrata	ADA70014.1	140	15463.40	4.77	19	13	21430	31.82	79.43	-0.266
M.mongolica	AIX11670.1	475	52667.95	6.00	61	53	69830	42.17	80.53	-0.245
M.cathayana	ADA70012.1	140	15463.40	4.77	19	13	21430	31.82	79.43	-0.266
M.nigra	AFU65414.1	184	20332.17	6.57	21	21	29005	28.56	81.63	-0.230
M.bombycis	SHO42454.1	407	44993.26	6.57	48	45	54235	42.03	82.46	-0.217
M. macroura	AKP97540.1	206	23162.23	5.36	27	24	34630	33.47	75.78	-0.355
M.celtidifolia	AFI21303.1	184	20288.08	6.57	21	21	30495	29.81	79.51	-0.274
M. indica	ABB20966.1	475	52667.95	6.00	61	53	69830	42.17	80.53	-0.245
M. rubra	AAA21230.1	465	51606.77	6.00	60	52	69830	40.14	82.04	-0.225
M.notabilis	AKF34051.1	475	52585.80	6.00	61	53	69830	40.67	82.78	-0.231
M. alba var. *atropurpurea*	ANE10809.1	475	52667.95	6.00	61	53	69830	42.17	80.53	-0.245
M. alba var. *multicaulis*	AMY59119.1	475	52667.95	6.00	61	53	69830	42.17	80.53	-0.245

AVCI and TEZCAN (2016) have also conducted a silico analysis of RuBisCO proteins of Asteraceae species, and determined the lowest and the highest MW values as 18341.8 Da and 52895.1 Da, respectively. Similarly, ZHANG *et al*., (2011) performed bioinformatic analysis of

RuBisCO proteins of different plants. In this study, the lowest and the highest MW values were determined as 52.52 Da and 54.04 kDa, respectively. The most acidic and basic protein sequences were detected between *pI*=4.77 and *pI*=6.57 respectively. AVCI and TEZCAN (2016) have reported that the most acidic and basic protein sequences as between 5.33 and 8.65, respectively. The EC of RuBisCO proteins at 280 nm ranged from 21430 to 69830 M^{-1} cm^{-1}. The highest EC values were determined in *M. mongolica*, *M. indica*, *M. rubra*, *M. notabilis*, *M. alba* var. *atropurpurea* and *M. alba* var. *multicaulis* while the lowest ones were found in *M. australis*, *M. rotundiloba*, *M. serrata*, and *M. cathayana*. AVCI and TEZCAN (2016) have reported EC values between 27515 and 69830. The II values for the RuBisCO proteins ranged from 26.06 to 42.17. According to the study conducted by AVCI and TEZCAN (2016), II values were determined between 24.41 and 40.65, respectively. In their study, ZHANG *et al.* (2011) have found II values between 37.47 and 43.50. The AI of proteins of thermophilic bacteria was found to be higher and it was determined that the index could be used as a measure of thermostability of proteins. This index is directly related to the mole fraction of Ala, Ile, Leu and Val in proteins (IDICULA-THOMAS and BALAJI, 2005). The AI values ranged from 75.78 to 82.78. The highest AI value was detected in *M. notabilis* while the lowest AI value was found in *M. macroura* (Table 1). AVCI and TEZCAN (2016) detected AI values between 72.56 and 81.13. When the total ratios of aliphatic amino acid contents were compared with the minimum and maximum AI values in the proteins, it was clearly seen that AI values increased in parallel with the increase in aliphatic amino acid content (Table 2). The GRAVY values of RuBisCO proteins ranged from -0.355 to -0.215 (Table 1). AVCI and TEZCAN (2016) detected GRAVY values between -0.394 and -0.179. The most abundant amino acid in the RuBisCO proteins was Gly (9.5%) while the least amount of amino acids were Cys and Trp (1.5%) (Fig1). The putative phosphorylation sites were determined on a score above 0.8 through NetPhos 2.0 and NetPhos 3.1 applications (Table 3).

Table 2. The RuBisCO proteins with aliphatic index (AI) values and their corresponding number of aliphatic residues

Taxa	Ala (A)	IIe (I)	Leu (L)	Val (V)
M. australis	9.3	5.0	7.1	7.9
M.mesozygia	8.7	4.3	8.7	7.6
M.rotundiloba	9.3	5.0	7.1	7.9
M. serrata	9.3	5.0	7.1	7.9
M.mongolica	9.5	4.6	8.4	6.9
M.cathayana	9.3	5.0	7.1	7.9
M.nigra	8.7	4.3	8.7	7.6
M.bombycis	9.1	4.7	8.8	7.1
M. macroura	8.3	3.9	8.7	6.3
M.celtidifolia	9.2	3.3	8.7	8.2
M. indica	9.5	4.6	8.4	6.9
M. rubra	9.5	4.7	8.6	7.1
M.notabilis	9.5	4.8	8.6	7.2
M. alba var. *atropurpurea*	9.5	4.6	8.4	6.9
M. alba var. *multicaulis*	9.5	4.6	8.4	6.9

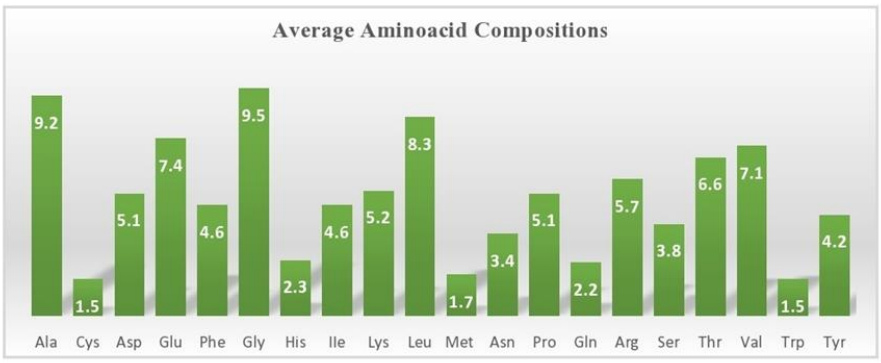

Fig 1: The average amino acid composition of RuBisCo proteins from *Morus ssp*.

Table 3. Putative phosphorylation residues in RuBisCO proteins of Morus species with a score above 0.8

Taxa	Ser	Thr	Tyr	Total
M. australis	2	1	1	4
M.mesozygia	3	1	1	5
M.rotundiloba	2	1	1	4
M. serrata	2	1	1	4
M.mongolica	9	3	4	16
M.cathayana	2	1	1	4
M.nigra	3	1	1	5
M.bombycis	6	2	4	12
M. macroura	3	1	2	6
M.celtidifolia	3	1	2	6
M. indica	9	3	4	16
M. rubra	8	2	4	14
M.notabilis	8	3	4	15
M. alba var. *atropurpurea*	9	3	4	16
M. alba var. *multicaulis*	9	3	4	16

The most phosphorylated sites were found in *M. mongolica*, *M. indica*, *M. alba* var. *atropurpurea* and *M. alba* var. *multicaulis* (Fig 2, Table 3). The confidence rates of these phosphorylation sites were above the threshold (0.5) and output score was given in a range between 0.0 - 0.1. In previous studies such as the one by AVCI and TEZCAN (2016), the most putative phosphorylation was detected in *Gymnarrhena micrantha*. SUBRAMANI and HWA (2010) have performed in silico studies of RuBisCO proteins of some C3 plants belonging to the

Poaceae family. In their study, the most putative phosphorylation was detected in *Portulaca oleracea*.

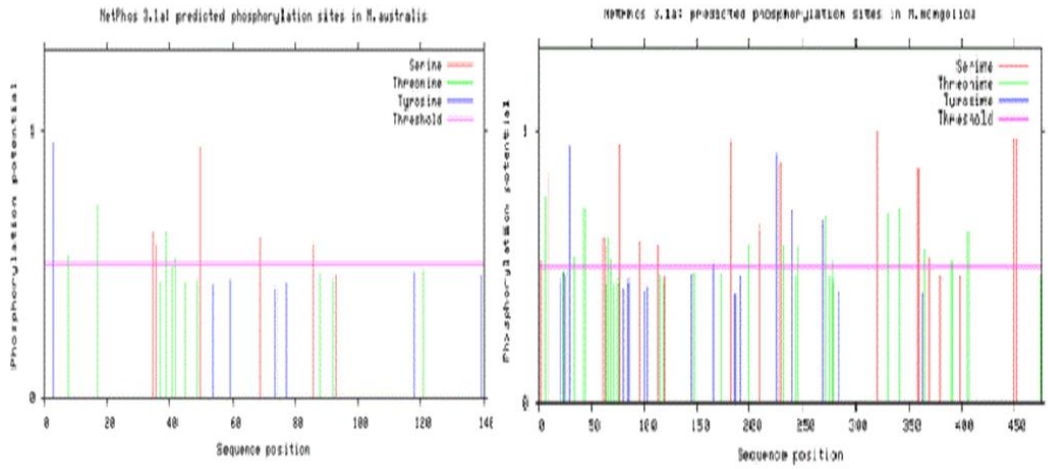

Fig 2: RuBisCO proteins in *Morus* ssp. determined by a score above a threshold of 0.5

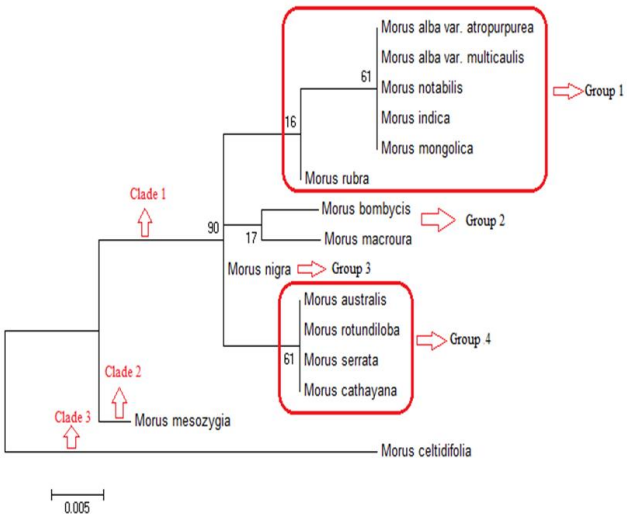

Fig.3: Phylogenetic tree of *Morus* ssp. RuBisCO proteins constructed using Maximum likelood method with MEGA 6.0

For phylogenetic analysis, MEGA 6.0 was used. Maximum likelihood tree with bootstrap analysis was constructed in order to identify the relationships among *Morus* species. Phylogenetics has become a powerful tool and a starting point in many areas of biology, such as developmental genetics, genomics, taxonomy and biogeography (WEI *et al.*, 2014). Recently, many molecular markers have been developed to overcome the limitations of morphological and

biochemical markers in plant phylogenetics (AGARWAL *et al.*, 2008; SUN *et al.*, 2015). With the development of molecular biology; RAPD (VIJAYAN, 2004; ORHAN *et al.*, 2007), AFLP (KAFKAS *et al.*, 2008; SHARMA *et al.*, 2000), PCR-RFLP (HU *et al.*, 2014), ISSR repeat (AWASTHI *et al.*, 2004; VIJAYAN *et al.*, 2004; ZHAO *et al.*, 2006), SSR (AGGARWAL *et al.*, 2004; ZHAO *et al.*, 2005), nrDNA (ITS) and cpDNA (*trn*L-F, *trn*D-*trn*T and *ndh*F sequences) (NEPAL and FERGUSON, 2012; WEIGUO *et al.*, 2005; HU *et al.*, 2014; DATWYLER and WEIBLEN, 2004), molecular markers have been used in studies on the genetic diversity and phylogenetic analyses of mulberry plants.

A phylogenetic analysis was performed with the RuBisCO protein sequences which were found to comprise three main clades. Clade 1 received a 90% bootstrap support and divided into 4 groups. Group 1 consists of *M. alba* var. *atropurpurea*, *M. alba* var. *multicaulis*, *M. notabilis*, *M. indica*, *M. mongolica* and *M. rubra*. Group 2 consists of *M. bombycis* and *M. macroura*. Group 3 consists of only *M. nigra* while group 4 contains *M. australis*, *M. rotundiloba*, *M. serrata* and *M. cathayana* with a 61% bootstrap support. Clade 2 consists of only *M. mesozygia* and clade 3 consist of only *M. celtidifolia* (Fig 3). Pairwise distance method based on *Morus* RuBisCO large subunit protein sequences set was performed with MEGA 6.0. The lowest pairwise distance was 0.000 while the highest was 0.052 (Table 4).

Table 4. Pairwise distance among Morus species for RuBisCO proteins. Data were generated using MEGA 6.0

Taxa	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
M. australis	-														
M.mesozygia	0,022	-													
M.rotundiloba	0,000	0,022	-												
M. serrata	0,000	0,022	0,000	-											
M.mongolica	0,007	0,022	0,007	0,007	-										
M.cathayana	0,000	0,022	0,000	0,000	0,007	-									
M.nigra	0,007	0,015	0,007	0,007	0,007	0,007	-								
M.bombycis	0,007	0,022	0,007	0,007	0,007	0,007	0,007	-							
M. macroura	0,007	0,022	0,007	0,007	0,007	0,007	0,007	0,007	-						
M.celtidifolia	0,052	0,044	0,052	0,052	0,052	0,052	0,044	0,052	0,052	-					
M. indica	0,007	0,022	0,007	0,007	0,000	0,007	0,007	0,007	0,007	0,052	-				
M. rubra	0,007	0,022	0,007	0,007	0,007	0,007	0,007	0,007	0,007	0,052	0,007	-			
M.notabilis	0,007	0,022	0,007	0,007	0,000	0,007	0,007	0,007	0,007	0,052	0,000	0,007	-		
M. alba var. *atropurpurea*	0,007	0,022	0,007	0,007	0,000	0,007	0,007	0,007	0,007	0,052	0,000	0,007	0,000	-	
M. alba var. *multicaulis*	0,007	0,022	0,007	0,007	0,000	0,007	0,007	0,007	0,007	0,052	0,000	0,007	0,000	0,000	-

In the model validation, the Ramachandran plot analysis conducted through the RAMPAGE application revealed that the 96.1% - 98.5% of the residues were found in the favored regions, 1.5% - 3.9% of the residues were located in the allowed regions, and 0% - 1.1% of the residues were found in the outlier regions. Two *Morus* RuBisCO large subunit proteins revealed the results of RAMPAGE (Fig 4) indicating that the 3D models were fairly good in quality. In the research of AVCI and TEZCAN (2016) on homology models, between 95.0% and

98.4% of the residues were found in the favored regions, between 1.0% and 3.6% of the residues were located in the allowed regions, and between 0% and 1.6% of the residues were found in the outlier regions. NAEEM *et al*. (2012) performed a similar study through an in silico study of the RuBisCO protein of *Triticum aestivum*. In their study, 96.8% of the residues were found in the favored region, 2.7% of the residues were located in the allowed regions, and 0.4% of the residues were found in the outlier region.

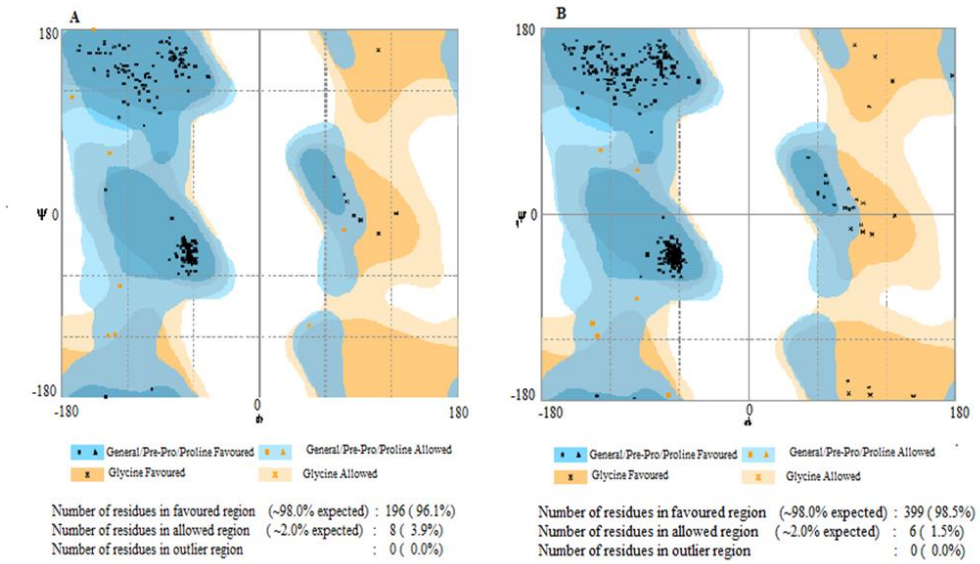

Fig. 4: The Ramachandram plots of two *Morus* RuBisCO proteins produced using RAMPAGE according to the results of PSIPRED. A-) *Morus macroura*, B-) *Morus.bombycis* These RuBisCO proteins have the least (A) and most (B) number of residues located in the favored regions, respectively.

The three-dimensional structures of the RuBisCO large subunit proteins were constructed using the PyMOL program, and the secondary structures were demonstrated (Fig 5). The three-dimensional structure of the proteins contributes to the understanding of protein function and active regions facilitating drug design (FILIZ and KOÇ, 2014).

The Gly residue is unique among the amino acids in that its variable group is a hydrogen atom. Its conformation has greater freedom so that it can provide flexibility for adjacent residues. Therefore, it is not surprising that Gly plays a special role in enzyme structure and function (YAN and SUN, 1997). In this study, the highest amount Gly residues was detected in *M.mongolica, M. indica, M.notabilis, M. alba* var. *atropurpurea* and *M. alba* var. *multicaulis*, while the least amount was detected in *M. australis, M.rotundiloba, M. serrata and M.cathayana* (Fig. 6).

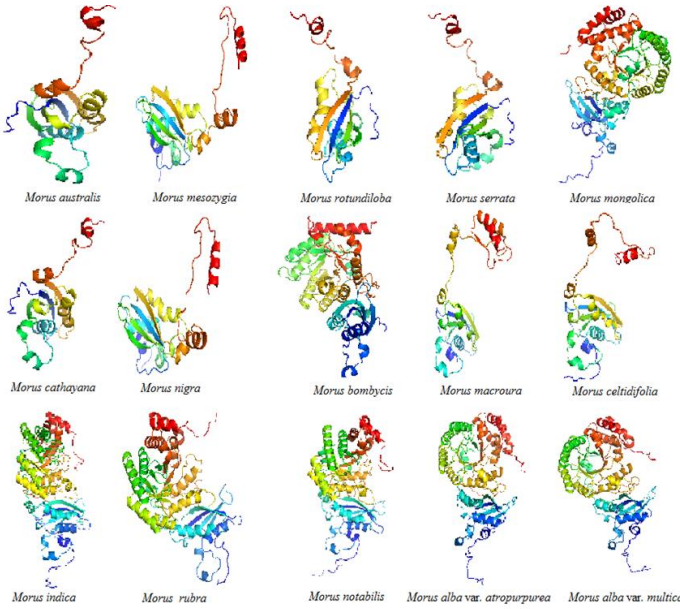

Fig.5: The compherensive 3D structures of *Morus* ssp. RuBisCO proteins. The visual data were obtained from PSIPRED software and analyzed using PyMOL.

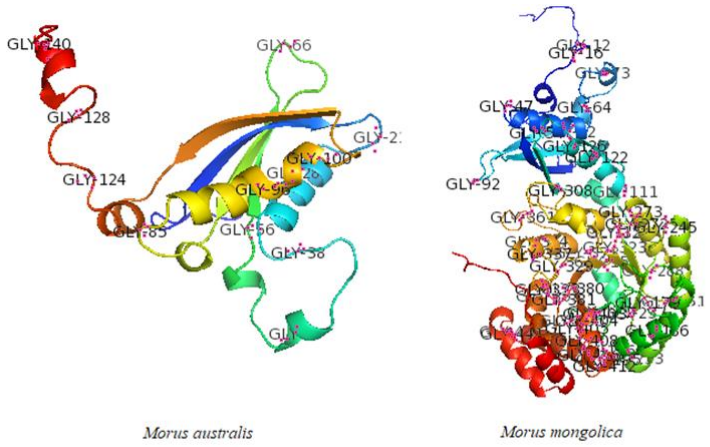

Fig. 6: Conserved Gly residues in the *Morus australis* and *Morus mongolica* RuBisCO protein sequences

However, Pro residues were detected neighboring to Gly. Gly appears to be restricted in its conformational freedom when followed by a Pro. Thus the Gly-Pro unit nearly always adopts the extended conformation (YAN and SUN, 1997). In this study, a triplet of Gly-Pro-Pro (For example, *M.indica* Gly 150-Pro 151-Pro 152) was found in all RuBisCO large subunit proteins (Fig 7). Additionally, Phe and Cys dipeptide residues were detected. This peptide has been reported in RuBisCO LSU derived from various photosynthesising plant-based foods and edible microalgae, and has been predicted to be bioactive; the two amino acid residues, Phe220 and Cys221, can make the peptide an antioxidant (JE *et al.*, 2015). This dipeptide was found to be located within the *Morus* RuBisCO large subunit proteins (Fig 8).

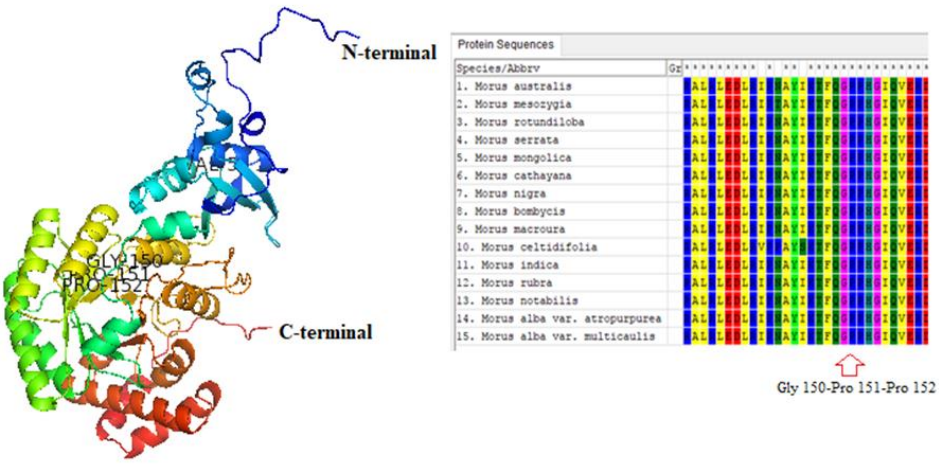

Fig. 7: The location of the *Morus indica* Gly 150-Pro 151-Pro 152 RuBisCO residue

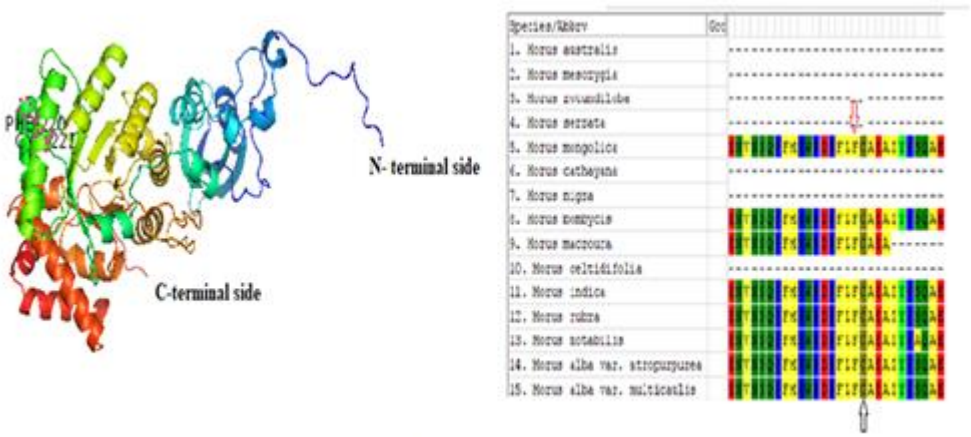

Fig. 8: The location of the *Morus mongolica* Phe220-C221 RuBisCO residues

CONCLUSIONS

In this study, genome-wide analysis was carried out using bioinformatics tools such as ExPASy's ProtParam, MEGA 6.0, NetPhos 2.0, NetPhos 3.1, PSIPRED v3.3, RAMPAGE, and PyMOL for the analysis of RuBisCO large subunit protein in *Morus* species. In this study, the physicochemical properties of the RuBisCO proteins of *Morus* species were also determined. Comparison of the phylogenetic analyses to the previous studies will guide the future phylogenetic studies on RuBisCO as well as on similar proteins. With three-dimensional structures of proteins, the presence and importance of the key amino acid residues have been emphasized. The results of this study will be useful for further research on RuBisCO of different plant species.

Received, December 31st, 2017
Accepted, April 18th, 2018

REFERENCES

AGARWAL, M., S., NEETA, P., HARISH (2008): Advances in molecular marker techniques and their application in plant sciences. Plant Cell Rep., *27*: 617-631.

AGGARWAL, R.K., D., UDAYKUMAR, P.S., HENDER, A., SARKAR, L.A., SINGH (2004): Isolation and characterization of six novel microsatellite markers for mulberry (*Morus indica*). Mol. Ecol. Resour., *4*: 477-479.

ANDERSSON, I., S., KNIGHT., G., SCHNEIDER., G., SCHNEIDER., Y., LINDQVIST., T., LUNDQVIST., C.I., BRÄNDÉN., G.H., LORIMER (1989): Crystal structure of the active site of ribulose-1,5- bisphosphate carboxylase, Nature, *337*(19): 229-234

ARFAN, M., R., KHAN., A., RYBARCZYK., R., AMAROWICZ (2012): Antioxidant activity of mulberry fruit extracts. Int. J. Mol. Sci., *13*(2): 2472-2480.

AVCI, M.K., E., TEZCAN (2016): Erratum to: Genome-wide identification and comparative structural analysis of RuBisCo proteins in the asteraceae. Hortic. Environ. Biotech., *5*(57): 529-529.

AWASTHI A.K, G.M., NAGARAJA, G.V., NAIK, S., K., KANGINAKUDRU, THANGAVELU, J., NAGARAJU (2004): Genetic diversity and relationships in mulberry (genus *Morus*) as revealed by RAPD and ISSR marker assays. BMC Genetics, *5*: 1-8.

BUCHAN, D.W., F., MINNECI, T.C., NUGENT, K., BRYSON, D.T., JONES (2013): Scalable web services for the PSIPRED Protein Analysis Workbench. Nuc. Acids Res., *41*(1): 349-357.

DATWYLER, S.L., G.D., WEIBLEN (2004): On the origin of the fig: phylogenetic relationships of Moraceae from *ndh*F sequences. Am. J. Bot., *91*(5): 767-777.

ELLIS, R.J. (1979): The most abundant protein in the world. Trends Bioch. Sci., *4*(11): 241-244.

FILIZ, E., I., KOÇ (2014): In silico sequence analysis and homology modeling of predicted beta-amylase 7-like protein in *Brachypodium distachyon* L. J. BioSci. Biotech., *3*(1): 61-67.

GASTEIGER, E., C., HOOGLAND, A., GATTIKER, S., DUVAUD, M.R., WILKINS, R.D., APPEL, A., BAIROCH (2005): Protein identification and analysis tools on the ExPASy server, In: John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press: 571–607.

HU, D., P., ZHANG, Y.L., SUN, S., ZHANG, Z., WANG, C., CHEN (2014): Genetic relationship in mulberry (*Morus* L.) inferred through PCR–RFLP and *trn*D-*trn*T sequence data of chloroplast DNA. Biotech. Biotech. Eq., *28*(3): 425-430.

IDICULA-THOMAS, S., P.V., BALAJI (2005): Understanding the relationship between the primary structure of proteins and their amyloidogenic propensity: clues from inclusion body formation. Prot. Eng. Des. Sel., *18*: 175-180.

JE, J.Y., Y.S., CHO, M., GONG, C.C., UDENIGWE (2015): Dipeptide Phe-Cys derived from in silico thermolysin-hydrolysed RuBisCO large subunit suppresses oxidative stress in cultured human hepatocytes. Food Chem., *171*: 287-291.

KAFKAS, S., M., ÖZGEN, Y., DOĞAN, B., ÖZCAN, S., ERCIŞLI, S., SERÇE (2008): Molecular characterization of mulberry accessions in Turkey by AFLP markers. J. Am. Soc. Hortic. Sci., *133*(4): 593-597.

LOVELL, S.C., I.W., DAVIS, W.B., ARENDALL, P.I.W., BAKKER, J.M., WORD, M.G., PRISANT, J.S., RICHARDSON, D.C., RICHARDSON (2003): Structure validation by Cα geometry: ϕ, ψ and Cβ deviation. Proteins: Structure, Function, and Bioinformatics, *50*(3): 437-450.

MUKHERJEE, S., C., STASOLLA, A., BRÛLÉ-BABEL, B.T., AYELE (2015): Isolation and characterization of rubisco small subunit gene promoter from common wheat (*Triticum aestivum* L.). Plant Signal. Behav., *10*(2): e989033.

NAEEM, M.K., S., RAUF, H., IQBAL, M.K., NAWAZ SHAH, A., MIR (2012): In silico studies of C3 metabolic pathway proteins of wheat (*Triticum aestivum*). BioMed Res. Int., *2013*: 1-7

NEPAL, M.P., C.J., FERGUSON (2012): Phylogenetics of *Morus* (Moraceae) inferred from ITS and *trn*L-*trn*F sequence data. Syst. Bot., *37*(2): 442-450.

ORHAN, E., S., ERCISLI, N., YILDIRIM, G., AGAR (2007): Genetic variations among mulberry genotypes (*Morus alba*) as revealed by Random Amplified Polymorphic DNA (RAPD) markers. Plant Syst. Evol., *265*: 251-258.

ORHAN, E., S., ERCISLI (2010): Genetic relationships between selected Turkish mulberry genotypes (*Morus* spp) based on RAPD markers. Gen. Mol. Res., *9*: 2176–2183.

ÖZRENK, K., R.I., GAZIOGLU SENSOY, C., ERDINC, M., GULERYUZ, A., AYKANAT (2010): Molecular characterization of mulberry germplasm from Eastern Anatolia. Afr. J. Biotech., *9* (1): 001-006.

PEREZ-GREGORIO, M.R., J., REGUEIRO, E., ALONSO-GONZALEZ, L.M., PASTRANA-CASTRO, J., SIMAL-GANDARA (2011): Influence of alcoholic fermentation process on antioxidant activity and phenolic levels from mulberries (*Morus nigra* L.). LWT - Food Sci. and Technol., *44*(8): 1793-1801.

PORTIS, J.R.A.R. (1992): Regulation of ribulose 1, 5-bisphosphate carboxylase/oxygenase activity. Ann. Rev. Plant Biol., *43(*1): 415-437.

PORTIS A.R. PARRY. Rubisco. eLS. available on 2009; http://onlinelibrary.wiley.com/doi/10.1002/9780470015902.a0001293.pub2/full.

SHARMA, A., R., SHARMA, H., MACHII (2000): Assessment of genetic diversity in a *Morus* germplasm collection using fluorescence based AFLP markers. TAG, *101*: 1049-1055.

SUBRAMANI, B., K.Y., HWA (2010): In silico Analysis for Enhancing the Rubisco Activity among the C3 Plants of Poaceae Family. Information Technology Convergence and Services (ITCS), 2nd International Conference on. 1-6. IEEE.

SUN, Y.L., H.M., KANG, S.H., HAN, Y.C., PARK, S.K., HONG (2015): Taxonomy and Phylogeny of the Genus *Citrus* Based on the Nuclear Ribosomal DNA ITS Region Sequences. Pak. J. Bot., *47*(1): 95-101.

TAMURA, K., G., STECHER, D., PETERSON, A., FILIPSKI, S., KUMAR (2013): MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol., *30*: 2725-2729.

TÜRK, M., N., ÇELIK (2006): Determination of Photosynthesis-Respiration Compensation Points In C-3 And C-4 Photosynthetic Pathway Plants. SDU. J. Nat. Appl. Sci., *10*(1): 48-51.

VIJAYAN, K., P.P., SRIVASTAVA, A.K., AWASTHI (2004): Analysis of phylogenetic relationship among five mulberry (*Morus*) species using molecular markers. Genome, *47*: 439-448.

VIJAYAN, K. (2004): Genetic relationships of Japanese and Indian mulberry (*Morus* spp.) genotypes revealed by DNA fingerprinting. Plant Syst. Evol., *243*: 221-232.

WEI, W., L., HONGLEI, C., ZHDUAN (2014): Analysis of plastid and nuclear DNA data in plant phylogenetics-evaluation and improvement. Life Sci., *57*(3): 280-286.

WEIGUO, Z., P., YILE, S., ZHANG, X., MIAO, Y., HUANG (2005): Phylogeny of the genus *Morus* (Urticales: Moraceae) inferred from ITS and *trn*L-F sequences. Afr. J. Biotech., *4*(6): 563-569.

YAN, B.X., Y.Q., SUN (1997): Glycine Residues Provide Flexibility for Enzyme Active Sites. J. Biol. Chem., *272*: 3190-3194.

ZAFAR, M.S., F., MUHAMMAD, I., JAVED, M., AKHTAR, T., KHALIQ, B., ASLAM, A., WAHEED, R., YASMIN, H., ZAFAR (2013): White mulberry (*Morus alba*): A brief phytochemical and pharmacological evaluations account. Int. J. Agric. Biol., *15*: 612–620.

ZHANG, B., L., LUO, X., ZHANG, R., LI, Y., SONG, D., ZHANG, ...Y., WEI (2011): Bioinformatics analysis on ribulose-1, 5-bisphosphate carboxylase/oxygenase large subunits in different plants. Mol. Plant Breed., *2*:101-108.

ZHAO, W.G., X.X., MIAO, S.H., JI, Y.L., PAN, Y.P., HUANG (2005): Isolation and characterization of microsatellite loci from the mulberry, *Morus* L. Plant Sci., *168*: 519-525.

ZHAO, W.G., Z.H., ZHOU, X.X., MIAO, Y., ZHANG, S.B., WANG, J.H., HUANG, H., XIANG., Y.L., PAN (2006): Comparison of genetic variation among wild and cultivated *Morus* Species (Moraceae: *Morus*) as revealed by ISSR and SSR markers. Biodivers. Conserv., *16:* 275-290.

GENOME-WIDE IDENTIFIKACIJA I ANALIZA RUBISCO VELIKE PROTEINSKE SUBJEDINICE KOD *MORUS* L. (MORACEAE) VRSTE

Emre SEVİNDİK

Poljoprivredni fakultet, Departman za poljoprivrednu biotehnologiju, Adnan Menderes Univerzitet, Južni kampus, Cakmar, Aydin, Turska

Izvod

U ovom radu izvršili smo sekvencionu, fizičkohemijsku, filogenetsku i komparativnu 3D (trodimenzionalnu) analizu RuBisCO velikih proteinskih subjedinica Morus vrste koristeći različite bioinformatičke metode. Dužina sekvence RuBisCO proteina varirala je od 140 do 475 amino kiselina sa molekulskom težinom od 15463.40 do 52667.95 Da. Najkiselije proteinske sekvence bile su u rodovima *M. australis*, *M. rotundiloba*, *M. serrata* i *M. cathayana* (*pI*=4.77). Koeficijenti ekstinkcije proteina na 280 nm kreću se od 21430 do 69830 M^{-1} cm^{-1}, vrednosti indeksa nestabilnosti (II) kreću se od 26.06 do 42.17, a vrednosti GRAVY-a se kreću od -0.355 do -0.215. Najčešća aminokiselina RuBisCO proteina bila je Gli (9,5%), dok su najmanj zastupljeni bili Cis i Trp (1,5%). Na osnovu filogenetske analize, stablo izgrađeno korišćenjem RuBisCO proteina sastoji se od tri glavne klase. Analiza RAMPAGE pokazala je da je 96,1% - 98,5% ostataka bilo locirano u najboljoj regiji RuBisCO proteina. Za predviđanje trodimenzionalnih (3D) struktura RuBisCO proteina korišćen je PYMOL. Najveći broj Gly ostataka je otkriven u *M.mongolica*, *M. indica*, *M.notabilis*, *M. alba* var. *atropurpurea* i *M. alba* var. *multicaulis*, dok je najmanje Gly ostataka otkriveno u *M. australis*, *M.rotundiloba*, *M. serrata* i *M.cathaiana*. Rezultati našeg istraživanja pružaju uvid u osnovne karakteristike RuBisCO proteina u *Morus* vrsti.

Primljeno 31.XII.2017.
Odobreno 18. IV. 2018.