Adv. Hort. Sci., 2019 33(3): 403-408 DOI: 10.13128/ahs-25648

AHS
Advances in Horticultural Science

Barcoding assessment of the *Citrus* species cultivated in eastern Afghanistan

M. Gori [1], S. Pecchioli [1], E. Giordani [1], M.A. Saeedi [2], F.H. Wafa [2], S. Biricolti [1 (*)]

[1] *Dipartimento di Scienze e Tecnologie Agrarie, Alimentari, Ambientali e Forestali, Sezione Colture Arboree, Università degli Studi di Firenze, Viale delle Idee, 30, 50019 Sesto Fiorentino, Italy.*

[2] *Afghanistan National Horticulture Development Organization (ANHDO), Taimani Street, 9, Kabul, Afghanistan.*

Keywords: Afghanistan, citrus, *psb*A-*trnH* spacer.

Abstract: The establishment of a modern fruit culture in developing countries requests an accurate evaluation of the preexisting germplasm and its health status. This to prevent the possibility to introduce new germplasm which can be easily prey of the endemic diseases carried by asymptomatic host plants. Therefore, after the identification of cases of citrus plants affected by Tristeza virus, a survey of the germplasm cultivated in the Nangarhar valley and some nearby regions was run. The survey was focused on the identification of the main *Citrus* species widely cultivated using barcoding analysis of conserved sequences located in the plastid DNA. The sequences of *matK* and *rbcl* genes did not show any discriminatory ability while the analysis of the the non-coding *psb*A-*trnH* intergenic spacer (*psbA-trnH)* showed a robust single nucleotide polymorphism (sNP) discriminating *C. aurantium* from *C. sinensis* in all analysed samples. These non-coding regions have no known function; thus, much of the variation may result from the spread of mutations unconstrained by selection. Because nucleotide variation in the *psbA-trnH* spacer region is high, mutational hot spots may be useful to detect species-level variations. According to our study, the afghan citrus germplasm belong to the *C. aurantium* species which is commonly used in citrus culture as a rootstock. In Afghanistan it is widely cultivated for fresh consumption, without topworking with selected varieties and this may be the reason why symptoms are often mild and cultivation can be anyhow carried on.

(*) Corresponding author:
stefano.biricolti@unifi.it

Citation:
GORI M., PECCHIOLI S., GIORDANI E., SAEEDI M.A., WAFA F.H., BIRICOLTI S., 2019 - *Barcoding assessment of the Citrus species cultivated in eastern Afghanistan*. - Adv. Hort. Sci., 33(3): 403-408

Data Availability Statement:
All relevant data are within the paper and its Supporting Information files.

Competing Interests:
The authors declare no competing interests.

Received for publication 28 June 2019
Accepted for publication 1 August 2019

1. Introduction

In developing countries, farmers are growing landraces or older improved varieties that are not optimized for today's climate or production systems. Therefore, the replacement of the local germplasm traditionally cultivated for long should be carried on by introducing elite varieties in order to complying with the rules of a modern cropping system. However, this change rises concerns about the adaptation of the elite

germplasm to the new conditions which may poses the farmers revenue at risk. The choices to be undertaken are very delicate and consequences may be devastating if not based on preliminary scientific investigations.

Nangarhār (Pashto: رامرګنن ; Persian: رامرګنن), Laghman and Kunar are three of the 34 provinces of Afghanistan, located in the eastern part of the country. The lowlands in those regions benefit from a semi-tropical climate and have the highest proportion of high cropping intensity irrigated land in the country. The riverine farms, situated along valley bottoms of varying widths, produce a range of crops throughout the year. Semitropical crops such as citrus, sugar canes and henna are produced around Jalalabad. High potential for fruits production occurs in those regions despite the still ongoing conflict. Promoting commercial orchards, establishment through professional nurseries, farmers investment, extension service with technical input from neighboring countries and research are therefore welcome to develop the potential for fruit production and marketing in the Afghan eastern region (Giordani *et al*., 2014). A screening of the health status of the Afghan germplasm of *Citrus* was undertaken to ensure multiplication of not only the best-selected varieties or ecotypes but also to avoid reproduction and distribution of virus-infected fruit tree. In 2012 a first report showed the occurrence of Citrus Tristeza Virus (CTV) in plants sampled in the National Collection Experimental Farm in Jalalabad (Nangarhar Province) (Rehman *et al*., 2012). A successive investigation showed that CTV was detected in several samples collected in some farms located in the Nangarhar valley, a warning that CTV could spread rapidly if sensitive citrus cultivars would be introduced in the area. However, citrus cultivation is still thriving in Nangarhar valley despite the widespread presence of CTV. Different possibilities are therefore to be considered: 1) the local germplasm is made up of low reactive host species which do not heavily show the symptoms of the disease; 2) mild CTV isolates established in the area; 3) an unknown source of resistance to CTV is occurring in the area. Before introducing new rootstocks resistant to such disease, another survey was conducted in order to identify the main species of *Citrus*. This is to address the choices to be undertaken for the establishment of a modern cropping system according to the market demand and to prevent the diffusion of quarantine diseases. Usually, identification of the cultivated species relies on the use of phenotypic descriptors (UPOV). Phenotypic descriptors are often subjected to the observer's opinion and the environment can deeply affect the behavior of the plant, determining errors in the species attribution.

A more reliable system to identify species is based on the genetic analysis and particularly on the DNA barcoding procedure which consists in the comparison of highly conserved sequences located in the ribosomal nuclear (Sun *et al*., 2015) or plastid DNA (Taberlet *et al*., 1991; Penjor *et al*., 2010, 2013). Being highly conserved, such fragments accumulate mutations slowly and it is possible to design universal primers which can be used to distinguish plenty of species. Furthermore, many databases (PubMed, GenBank, OMIM) collect and store millions of sequences, which can be compared by means of specific software (BLAST) with unknown sequences (Zhang *et al*., 2000). When properly queried a DNA database provides, along with the alignment, a similarity index which can be useful for identifying species or even sub-species, depending on the proximity of the taxonomic entities. Therefore, due to the restrictions to accessibility to a conflicting area, a barcoding analysis has been carried out in order to identify the species of *Citrus* which are commonly grown in the eastern area of Afghanistan in order to better address the choices for a modern fruit culture.

2. Materials and Methods

Plant material

Afghan citrus seeds have been collected in three areas where citrus fruit trees are intensively cultivated (Laghman, Kunar and Nangarhar) in Eastern Afghanistan (Fig. 1). We used seeds because they can

Fig. 1 - Map of Afghanistan showing the provinces where citrus production is located and where the survey was carried on.

be easily transported. Furthermore, most *Citrus* species are characterized by adventitious embryony and seeds originate individuals which are true clones of the mother plant because rarely zygotic embryo survives. Most sequences for DNA barcoding, being located in the plastid or mitochondrial DNA, are inherited only from the maternal line, therefore germinating seeds produce plants whose organelle genomes are not affected by the pollen donor plant, but are expected to be exactly the same as the seed donor plant. In each area, some plants were selected and batch of seeds have been collected from single fruit. In Table 1 the origin of the seeds is reported. Each batch was labeled and delivered to the Plant Pathology Department of the University of Bologna.

Table 1 - List of the samples coming from Afghanistan

Sample	Province	District	Village	Variety
AF-1 a	Laghman	Mehtarlam	Chardehi	Local
AF-2 a	Laghman	Mehtarlam	Chardehi	Local
AF-5 a	Laghman	Mehtarlam	Chardehi	Local
AF-7 a	Kunar	Asadabad	Landi Tesha	Local
AF-8 a	Kunar	Asadabad	Landi Tesha	Local
AF-9 a	Kunar	Asadabad	Landi Tesha	Local
AF-11 a	Kunar	Asadabad	Landi Tesha	Local
AF-18 a	Nangarhar	Surkhroad	Naghrak	Local
AF-19 a	Nangarhar	Surkhroad	Naghrak	Local
AF-21 a	Nangarhar	Surkhroad	Sabzabad	Local
AF-22 a	Nangarhar	Surkhroad	Sabzabad	Local
AF-limon	Nangarhar		PHDC-Center	*Citrus volkamar*

The seeds were germinated in pots and leaf samples were collected. For each seed batch, a single seedling has been selected for DNA barcoding analysis. Before starting the present work, *Citrus* barcoding sequences (ITS 1 and 2, *matK*, *rbcl*, *psbA-trnH* intergenic spacer) were retrieved from GenBank. The dataset was used to compare with the results of sequencing and to assign the analysed samples to a *Citrus* species. Since the available sequences are few and showed heavy discrepancies, casting doubts about the reliability of the data retrieved online (Bengtsson-Palme *et al*., 2016), we decided to provide robust references by analyzing accessions whose origin was absolutely certain. Therefore we have sampled sour orange along with other *Citrus* species from private and public collections (Vivai Oscar Tintori, Orto Botanico "Giardino dei Semplici" of the University of Florence, Istituto Agronomico per l'Oltremare) in order to be compared with the samples coming from Afghanistan (Table 2).

PCR amplification and primers

Total DNA was extracted from the selected plants following the guidelines of the DNA Invisorb DNA extraction kit producer (Stratec Italy). DNA has been electrophoresed on an agarose gel to validate quality and quantity. Universal primers for PCR amplification, used in the present study have been retrieved in literature (Chen *et al.,* 2010; Luo *et al* 2010; Penjor *et al*., 2013; Mahadani and Ghosh, 2014; Uchoi *et al*., 2016; Wattoo *et al*., 2016; Bailey *et al*., 2018; Zhao *et*

Table 2 - *Citrus* species used as control, coming from public and private collections in Tuscany

Sample no.	Species or cultivar	Province	Origin of the accession
1	*Citrus sinensis* (Washington Navel)	Pescia (Italy)	Oscar Tintori
2	*Citrus sinensis* (ovale calabrese)	Pescia (Italy)	Oscar Tintori
3	*Citrus sinensis* (Tarocco)	Pescia (Italy)	Oscar Tintori
4	*Citrus aurantium*	Lucca (Italy)	Botanical garden
5	*Citrus aurantium*	Firenze (Italy)	IAO
6	*Citrus aurantium* (Foetifera)	Pescia (Italy)	Oscar Tintori
7	*Citrus aurantium* (tipo)	Pescia (Italy)	Oscar Tintori
8	*Citrus aurantium* (dolce del Gargano)	Pescia (Italy)	Oscar Tintori
9	*Citrus aurantiifolia* (Philippines Red lime)	Pescia (Italy)	Oscar Tintori
10	*Citrus aurantiifolia* (Mexico)	Pescia (Italy)	Oscar Tintori
11	*Citrus aurantium*	Firenze (Italy)	Botanical garden
12	*Citrus limon*	Firenze (Italy)	Botanical garden
13	*Citrus mitis*	Firenze (Italy)	Botanical garden
14	*Citrus histrix*	Firenze (Italy)	Botanical garden
15	*Citrus decumana*	Firenze (Italy)	Botanical garden
16	*Citrus grandis*	Firenze (Italy)	Botanical garden
17	*Citrus reticulata*	Firenze (Italy)	Botanical garden
18	*Citrus lumia*	Firenze (Italy)	Botanical garden
19	*Citrus medica*	Firenze (Italy)	Botanical garden

al., 2018) and are designed on the sequence of the internal transcribed sequence 1 (ITS1) and 2 (ITS2) of the nuclear ribosomal DNA, of the maturase K (*matK*), the RuBisCo large chain unit (*rbcl*) and of the predominantly non-coding *psb*A-*trnH* intergenic spacer (Table 3). PCR analyses were performed with a thermal cycler Primus 96 (PeqLab) in a 25 µl volume containing 25 ng total DNA, 1x Taq buffer, 1,5 mM $MgCl_2$, 1 unit GoTaq (Promega) and 200 nM of each primer. PCR conditions were 95°C for 5 min, then 35 cycles at 95°C for 30 seconds followed by 30 seconds at a temperature ranging from 50 to 60°C depending on the chosen primer pair and an extension cycle at 72°C for 50 sec. A final extension at 72°C for 5 minutes was carried on. The amplification products were purified with the Qiaquick PCR purification tubes (Qiagen) and then sequenced. Even though the primers for barcoding are called "universal", the presence of conserved flanking sites complementary to the primers sequence near the variable part and the locus copy number have been shown to account for much of the variability of amplification success. Therefore, the all primers have been tested by PCR, amplifying and sequencing the amplicons using a set of the samples as templates.

To perform a phylogenetic analysis we have used the Mega7 software package (Tamura *et al*., 2013) which compares the sequence calculating a similarity matrix, transforms similarity coefficients into distances and makes a clustering using the Unweighted Pair Group Method with Arithmetic mean (UPGMA) algorithm. The final output is represented by a dendrogram which clusters the samples.

3. Results

The primers constructed on the sequence of the internal transcribed sequence 1 (ITS1) and 2 (ITS2) resulted in a faint amplification and were therefore discarded, while primers designed on the sequence of the maturase K (*matK*) and RuBisCo large chain unit (*rbcl*) genes were correctly amplified and sequenced. Unfortunately, such sequences did not show any discriminatory ability being completely overlapping for most analysed samples (Mahadani and Ghosh, 2014). On the contrary, the primers of the non-coding *psb*A-*trnH* intergenic spacer showed either good amplification (a fragment of about 525 bp) and bidirectional sequencing output either a good discriminatory capacity. We found a robust single nucleotide polymorphism (sNP) in the *psbA-trnH* spacer (Fig. 2) which is capable to discriminating *C. aurantium* (sour orange) from *C. sinensis* showing this mutation in all analysed samples independently from the origin area.

Despite the limited region of the genome analysed containing, as expected, a low number of mutations, all the samples of *C. aurantium*, independently from their origin, clustered together with a reasonable certainty (Fig. 3). The same occurs for the three *C. sinensis* samples which resulted in a well-

Table 3 - Universal primers sequence for barcoding used in this study

Name	5' → 3' primer sequence	References
matK f	CGTACAGTACTTTTGTGTTTACGAG	Jeanson *et al*., 2011
matK r	ACCCAGTCCATCTGGAAATCTTGGTTC	Jeanson *et al.*, 2011
Rbcl_F1	ATGTCACCACAAACAGAGACTAAAGC	Uchoi *et al.*, 2016
Rbcl_R634	GAAACGGTCCCTCCAACGCAT	Jeanson *et al*., 2011
Rbcl_R724	TCGCATGTCCCTGCAGTAGC	Kress *et al*., 2005
ITS1_5F_746	GGAAGTAAAAGTCGTAACAAGG	Cheng *et al*., 2016
ITS1_4R_746	TCCTCCGCTTATTGATATGC	Cheng *et al.*, 2016
ITS2_S2_F497	ATGCGATACTTGGTGTGAAT	Chen *et al*., 2010
ITS2_S3R_497	GACGCTTCTCCAGACTACAAT	Chen *et al.*, 2010
psbA-trnH_Fw	CGCGCATGGTGGATTCACAATCC	Zhao *et al*., 2018
psbA-trnH_Rev	GTTATGCATGAACGTAATGCTC	Zhao *et al*., 2018
matK1F	ACCGTATCGCACTATGTATC	Penjor *et al*., 2013
matK1R	GAACTAGTCGGATGGAGTAG	Penjor *et al*., 2013
matK2F	ACGGTTCTTTCTCCACGAGT	Penjor *et al*., 2013
matK3F	GGTCCGATTTCTCTGATTCT	Penjor *et al*., 2013
matK2R	AGAATCAGAGAAATCGGACC	Penjor *et al*., 2013
matK3R	ACTCGTGGAGAAAGAACCGT	Penjor *et al*., 2013

Species/Abbrv	Group Name	********** **********
1. Citrus sinensis (Washington Navel)	sinensis	GCGCTAATACTACTAATAAATTACTAAATTTCTAATTTTATTATTAGATTATTAA
2. Citrus sinensis (Ovale Calabrese)	sinensis	GCGCTAATACTACTAATAAATTACTAAATTTCTAATTTTATTATTAGATTATTAA
3. Citrus sinensis (Tarocco)	sinensis	GCGCTAATACTACTAATAAATTACTAAATTTCTAATTTTATTATTAGATTATTAA
4. Citrus limon	limon	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
5. Citrus aurantium (Lucca botanical garden)	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
6. Citrus aurantium (Foetifera)	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
7. Citrus aurantium (dolce del Gargano)	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
8. Citrus aurantium (Tipo)	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
9. Citrus aurantium (Firenze botanical garden	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
10. 9a afghan sample	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
11. 5a afghan sample	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
12. 2a afghan sample	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
13. 21a afghan sample	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
14. 1a afghan sample	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA
15. 11a afghan sample	aurantium	GCGCTAATACTACTAATAAATTACTAAATTTATAATTTTATTATTAGATTATTAA

Fig. 2 - Two haplotypes of *C. aurantium* and *C. sinensis* originated by a single nucleotide polymorphism getting from the sequencing of *psbA-trnH* intergenic spacer (*psbA-trnH*).

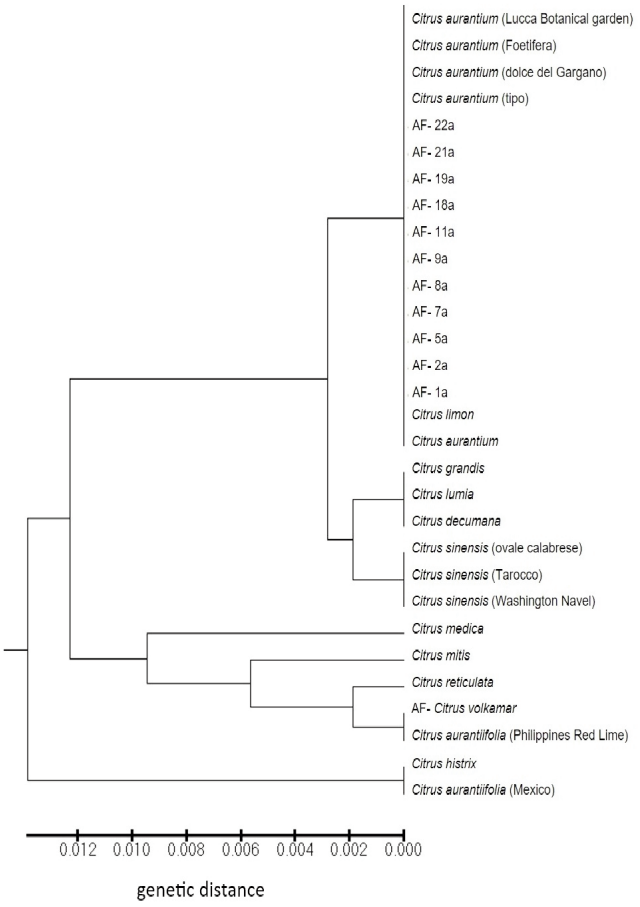

Fig. 3 - UPGMA dendrogram representing the genetic distances among the analysed samples.

separated group while the sample *C. limon* need to be better characterized with additional sequence analysis (Penjor *et al*., 2010, 2013).

4. Discussion and Conclusions

The barcoding analysis of the afghan germplasm and *Citrus* species has been carried out with a set of primers targeting nuclear ribosomal DNA or chloroplast genes (ITS1 and 2, *mat*K, *rbcl* and *psb*A-*trnH*). Apart from ITS 1 and 2, all the primers enabled amplification and sequencing of all the samples. Most of the *Citrus* species were correctly identified even if the comparison of the analysed DNA fragments with the data available in the databases showed several discrepancies. This observation suggested to introduce samples of certain origin in barcoding analysis in order to avoid biases due to errors occurring in the sequences uploaded on the database. According to the results of our analysis, we can state with reasonable certainty that all the afghan *Citrus* samples are sour orange, which is commonly used as a rootstock.

The wide diffusion of sour orange in those regions of Afghanistan is due to the fact that its fruits are consumed fresh (http://anhdo.org.af/wp-content/uploads/2017/06/Citrus-Market-Trend.pdf). The plants are not topworked, as usual, with selected *C. sinensis* varieties. Grafting would have shown heavily the symptoms of Tristeza disease, while this does not occur in ungrafted *C. aurantium* (Gómez-Muñoz *et al*., 2017). In conclusion, stepwise replacement of the orange germplasm in Afghanistan with CTV resistant rootstocks is advisable before grafting with selected orange varieties to prevent CTV spreading. However, the replacement of sour orange, a rootstock characterized by highly desirable agronomic features, with CTV resistant rootstocks should be carried out after devising the whole production chain, starting from the adoption of proper cropping

system (irrigation, fertilization, soil management, etc.) in order to prevent that other endemic diseases can affect the newly introduced germplasm.

References

BAILEY J.B., LAMB M., WALKER M., WEED C., CRAVEN K.S., 2018 - *Detection of potential fungal pathogens* Fusarium falciforme *and* F. keratoplasticum *in unhatched loggerhead turtle eggs using a molecular approach.* - Endang. Species Res., 36: 111-119.

BENGTSSON-PALME J., BOULUND F., EDSTRÖM R., FEIZI A., JOHNNING A., JONSSON V.A., KARLSSON F.H., PAL C., BUONGERMINO PEREIRA M., REHAMMAR A., SANCHEZ J., SANLI K., THORELL K., 2016 - *Strategies to improve usability and preserve accuracy in biological sequence databases*. - Proteomics, 16(18): 2454-2460.

CHEN S., YAO H., HAN J., LIU C., SONG J., SHI L., ZHU Y., 2010 - *Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species*. - PLoS One, 5 (1) e8613: 1-8.

CHENG T., XU C., LEI L., LI C., ZHANG Y., ZHOU S., 2016 - *Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity*. - Mol. Ecol. Resour, 16: 138-149.

GIORDANI E., CULLEN G., DEGL'INNOCENTI P., MASINI G., 2014 - *Local fruits and nuts as a tool for the development of Afghanistan*. - JUNCO, 1: 627-635.

GÓMEZ-MUÑOZ N., VELÁZQUEZ K.V., MARÍA C., RUIZ-RUIZ S., PINA J.A., FLORES R., MORENO P., GUERRI J., 2017 - *The resistance of sour orange to Citrus tristeza virus is mediated by both the salicylic acid and RNA silencing defence pathways.* - Mol. Plant Pathol., 18(9): 1253-1266.

JEANSON M.L., LABAT J.N., LITTLE D.P., 2011 - *DNA barcoding: a new tool for palm taxonomists?* - Ann. Bot., 108: 1445-1451.

KRESS W.J., WURDACK K.J., ZIMMER E.A., WEIGT L.A., JANZEN D.H., 2005 - *Use of DNA barcodes to identify flowering plants.* - Proceedings of the National Academy of Sciences (PNAS), 102 (23): 8369-8374.

LUO K., CHEN S.L., CHEN K.L., SONG J.Y., YAO H., MA X.Y., ZHU Y.J., PANG X.H., YU H., LI X.W., LIU Z., 2010 - *Assessment of candidate plant DNA barcodes using the Rutaceae family*. - Sci. China Life Sci., 53(6): 701-708.

MAHADANI P., GOSH S.H., 2014 - *Utility of indels for species-level identification of a biologically complex plant group: a study with intergeneric spacer in* Citrus. - Mol. Biol. Rep., 41: 7217-7222.

PENJOR T., ANAI T., NAGANO Y., MATSUMOTO R., YAMAMOTO M., 2010 - *Phylogenetic relationships of* Citrus *and its relatives based on rbcL gene sequences*. - Tree Genet. Genomes, 6: 931-939.

PENJOR T., YAMAMOTO M., UEHARA M., IDE M., MATSUMOTO N., MATSUMOTO R., NAGANO Y., 2013 - *Phylogenetic relationships of* Citrus *and its relatives based on matK gene sequences*. - PloS One, 8(4) e62574: 1-13.

REHMAN S., AHMAD J., LANZONI C., RUBIES AUTONELL C., RATTI C., 2012 - *First report of Citrus tristeza virus in national germplasm of citrus in Afghanistan.* - Plant Disease, 96(2): 296.

SUN Y.L., KANG H.M., HAN S.H., PARK Y.C., HONG S.K., 2015 - *Taxonomy and phylogeny of the genus* Citrus *based on the nuclear ribosomal DNA ITS region sequence.* - Pak. J. Bot., 47(1): 95-101.

TABERLET P., GIELLY L., PAUTOU G., BOUVET G., 1991 - *Universal primers for amplification of 3 noncoding regions of chloroplast DNA*. - Plant Mol. Biol., 17(5): 1105-1109.

TAMURA K., STECHER G., PETERSON D., FILIPSKI A., KUMAR S., 2013 - *MEGA6: Molecular evolutionary genetics analysis version 6.0.* - Mol Biol Evol., 30(12): 2725-2729.

UCHOI A., MALIK S.K., CHOUDHARY R., KUMAR S., ROHINI M.R., PAL D., ERCISLI S., CHAUDHURY R., 2016 - *Inferring phylogenetic relationships of Indian Citron (*Citrus medica *L.) based on rbcL and matK sequences of chloroplast DNA.* - Biochem. Genet., 54: 249-269.

WATTOO J.I., SALEEM M.Z., SHAHZAD M.S., ARIF A., HAMEED A., SALEEM M.A., 2016 - *DNA Barcoding: Amplification and sequence analysis of rbcl and matK genome regions in three divergent plant species.* - Adv. Life Sci., 4(1): 3-7.

ZHANG Z., SCHWARTZ S., WAGNER L., MILLER W., 2000 - *A greedy algorithm for aligning DNA sequences*. - J. Comput. Biol., 7(1-2): 203-214.

ZHAO L.L., FENG S.J., TIAN J.Y., WEI A.Z., YANG T.X., 2018 - *Internal transcribed spacer 2 (ITS2) barcodes: A useful tool for identifying* Chinese Zanthoxylum. - Appl. Plant Sci., 6(6) e1157: 1-8.