**Orobanche** species identification through DNA barcoding in tomato crop in uplands of Balochistan, Pakistan

Nazir Ahmed¹, Muhammad Ali¹, Hidayatullah ²,  
Zahid Mustafa¹, Tabeel Tariq Bashir¹, Farooq Shehzad¹

¹ Center for Advanced Studies in Vaccinology & Biotechnology (CASVAB), University of Balochistan, Pakistan  
² Agriculture Research Institute (ARI), Quetta, Pakistan

**Abstract**

**Background/Objectives:** *Orobanche* is one of the important parasites of solanaceae crops and causes heavy yield losses. This study aimed to use DNA barcoding technique for identification of *Orobanche* species. **Methods:** Surveillance study was conducted in the year 2018 on tomato crops grown in four different districts in Balochistan. From these tomato fields, 15 *Orobanche* plant samples were collected and put in polyethylene bags containing silica gel for drying. The dried samples were processed for DNA extraction using CTAB method. The extracted DNA was confirmed through gel electrophoresis which were further amplified on PCR using rbcL, matK primers. Then, the PCR products were processed for DNA sequencing using Sanger method. Further, the DNA sequences were edited and aligned with BioEdit and MegaX software and the aligned sequences were matched on NCBI for BLAST. The FASTA results of both partial and/or complete genome were run on MegaX for phylogenetic analysis. **Findings:** The success rates for PCR amplification was 100% for rbcL primer, while it was 68% for matK primer. However, the resolution power of matK was higher than that obtained from rbcL as manifested in the rate of DNA acquisition sequence. The DNA barcode results of 15 samples revealed 99.16% rate of DNA acquisition sequence for identification of *P. ramosa* and 96.41 for *P. purpurea*. While, Rbcl rate of DNA sequencing show no results. **Application/Improvement:** from this study it was inferred that matK gene performed better than rbcL in *Orobanche* species identification and proved as a potential gene for other pseudogene plants species identification. **Keywords:** Orobanche; species; DNA barcode; rbcL; matK; tomato

1 Introduction

The family Orobanchaceae contains 15 obligate parasitic genera with 250 species and is considered closely related to the Scrophulariaceae(1,2). The family has a worldwide distribution with major appearance in the Mediterranean, Western and Central...
Asia, North Africa and North America. In the past, the hemi-parasitic members of this family were classified as a part of Scrophulariaceae (the figwort family), while the holoparasitic members were included in Orobanchaceae (the broomrape family). The genus Orobanche comprises approximately 170 species that inhabit mainly the northern hemisphere. It is restricted to Europe, Middle East and North Africa. Orobanche plants lack chlorophyll and depend entirely on their hosts for their supply of carbon, nitrogen and inorganic solutes. Orobanche species parasitize and cause invasive damage to a wide range of economically important plants, e.g. legumes, sunflower, tomato and tobacco, enhancing production of some very useful constituents like phenylpropanoids and derivatives; such as coumarins, lignin, suberin, cutin, and tannin. These constituents are involved in the resistance of the host, as a defense mechanism against broomrape. These compounds are formed along the connection zones of the haustoria within the host.

The plant parasite Orobanche was considered as medicinal plant in Europe and Russia during 16th century which is used for the treatment of diseases like kidney and bladder’s stone and further this plant was also utilized for detoxification of wine. Apparently, the taste of Orobanche is bitter and very sharp. Plants of Orobanchaceae are famous for the accumulation of large amounts of flavonoids, iridoids, phenyl-propanoid and their glycosides that could explain the various reported biological effects such as; antibacterial, hypoglycemic, hypotensive, anti-inflammatory, anti-obesity, hepatoprotective, cytotoxic, sedative and neuroleptic in addition to the stimulation of immune response.

The mechanism of damage to host plants by Orobanche is caused by special Pseudo roots called Haustoria that directly penetrate into host plant and make contact with food reserves in phloem and xylem. In this way the host plant is exploited for nutrients, moisture and other food resources leading to stunted growth of host plant. When the infection is initiated in roots of host plant, then crops succumbed to damage before diagnosis and emergence of the parasite. The silent features of this parasitic weed is that none of the management practices are fruitful due to its close binding or association with roots of the host plant and deprived off host plant from the available growth promoting resources. It is further narrated that approximately five hundred thousand seeds are produced by a single Orobanche plant and its vigor remain for a long time in the soil. Due to this genetic makeup, the broomrape has developed resistance against harsh environmental conditions such as herbicides and other cultural management practices and has now adopted ecological fluctuation that occur in various regions where plants of Solanaceae are grown. The presence of the parasite strongly reduced the aerial biomass by acting as a competing sink for assimilate, but more importantly, by compromising the efficiency of carbon assimilation via a reduction in leaf chlorophyll content and photosynthetic rate.

The classification of plants during the 250 years has been carried out through the differentiation in their physical structure and phenotypic traits which may also be called as morphological features which finds its roots from comparative anatomy. Generally, it is common practice in plant taxonomy to identify plants of various species through their morphological distinctions. It is difficult for most of taxonomist to decide unanimously one trait of a species which differ from others and due to this difficulty latest information regarding species identification are changeable for defining and redefining a species. According to plant taxonomy that classification of plant species based on their morphological distinguishes possess some drawback that reflects gap of conventional agreement on such trait which liable to change with the passage to time leading to the unsustainability of taxonomic knowledge. One reason of the failure of conventional method of plant species identification is that with the passage to time it becomes multi opinion subject. Another reason is that most plant species change their morphological appearance according to the specified environment and with help of some defense mechanism they camouflage themselves from other species and it became difficult to make difference between them.

Presently, the subject of genetics have revolutionized the way of understanding the knowledge of taxonomy and anatomy and now scientists are using tool of biotechnology and molecular biology to identify species of animal and/or plant origin. The use of DNA tool almost relinquish the multi-opinion controversies over species identification among taxonomist to make boundaries between species on the basis of certain features and further, the method of DNA extraction for species identification is quick and easy as compared to conventional method. The past, there is also indication of the application of DNA barcoding for plant species identification at phylogenetic level but to a very limited level. Basically, the tool of DNA barcoding is used for animal species identification globally in the recent era. However, the use of this molecular tool for plant species identification was not recognized and accepted by plant taxonomists for many years. The detail study of gene region found in mitochondria, plastid and nucleus reflected four basic gene area including rbcL, matK, trnH-psbA and ITS and suggested as a base of DNA barcode for plant species identification. Keeping in view the economic importance of Orobanche. The proposed study will be conducted to identify its species that occurring in upland of Balochistan.

2 Materials and Methods

This study was conducted to identify the Orobanche species occurring in tomato crop in four district of Balochistan such as Mastung, Pishin Killa Saifulla and Quetta. Orobanche plants samples were collected from tomato crop during their vegetative
and reproductive growth stage during 2017 in the month of August. Sampling was carried out randomly from 18 locations across four districts and GPS coordinates each sample was recorded (Table 1). The flowers of *Orobanche* plants were plucked and made 3-4 g by cutting pieces as composite sample and put in plastic bags already containing granules of silica gel that help in drying (24). The sample bags were labeled properly and delivered to biotechnology laboratory at BUITEMS Quetta for DNA barcoding and species identification.

For morphological identification, the picture of sampled *Orobanche* plant was taken and then the same plant was uprooted intact with flower, pseudo stem and roots for voucher preparation. Plant flower were cleaned, and dried using spacers between layers of blotting papers and finally pasted on the Herbarium sheet with proper labeling. The preserved plants in the herbarium were used for taxonomic study keeping in view the prominent traits such as plant habitat, nodal characteristics of stem, leaf shape, leaf type, leaf arrangement, floral type, floral appearance, flower shape (sepal and petal), fruits characteristics etc.

**Table 1.** GPS Coordinates of the sampled area in four districts of Balochistan

| District   | Location | Longitude | Latitude | Elevation |
|------------|----------|-----------|-----------|-----------|
| Killa Saifullah | KS1 30° 50,071 067° 48,657 | 1743m |
|             | KS2 30° 50,701 067° 56,532 | 1668m |
|             | KS3 30° 51,902 067° 56,348 | 1670m |
|             | KS4 30° 50,547 067° 56,569 | 1667m |
|             | KS5 30° 51,122 068° 01,804 | 1635m |
| Pishin      | PN1 30° 31,677 066° 57,801 | 1519m |
|             | PN2 30° 32,853 066° 50,396 | 1534m |
|             | PN3 30° 33,336 066° 58,651 | 1539m |
|             | PN4 30° 25,507 067° 02,954 | 1589m |
|             | PN5 30° 30,774 067° 12,025 | 1749m |
| Mastung     | MG1 29° 56,861 066° 53,164 | 1718m |
|             | MG2 29° 51,892 066° 50,911 | 1637m |
|             | MG3 29° 47,184 066° 49,619 | 1659m |
|             | MG4 29° 46,111 066° 49,736 | 1674m |
| Quetta      | QTA1 30° 05,445 066° 57,953 | 1710m |
|             | QTA2 30° 04,785 066° 57,740 | 1715m |
|             | QTA3 30° 16,816 066° 55,564 | 1516m |
|             | QTA4 30° 23,942 066° 57,566 | 1527m |

*KS1 (Killi Haji Juma-M. Bagh), KS2 (Killi Mullah Abdullah, Bandat Nasai), KS3 (Killi Noorullah Bandat), KS4 (Killi Mullah Saido Nasai), KS5 (Killi Zafarullah Killa Saifullah); PN1 (Killi Sardar Ahmed Khan), PN2 (Killi Malik Moheeb Khan), PN3 (Killi Mohammad Hassan), PN4 (Killi Haji Gulam Mohammad), PN5 (Killi Malik Naseem Khan baba); MG1 (Killi Gunjdori), MG2 (Killi Paringabad), MG3 (Killi Detho), MG4 (Killi Chotho); QTA1 (Sariab Road Quetta), QTA2 (Sheikh Zahid Hospital Lukpas Quetta), QTA3 (Baleli(Quetta), QTA4 (Jalogir Mor Kuchlak, Quetta)

**DNA Extraction**

The samples were processed for DNA extraction by modified CTAB method as described by Doyle and Doyle (25) and Li et al. (26). In the first step, 0.5 g dried samples were grinded manually in the presence of liquid nitrogen using porcelain made mortar and pestle. The grind samples were then transferred into 2 ml Eppendorf tubes, added 1 ml CTAB (Cetyl Trimethyl Ammonium Bromide) extracting buffer and homogenized with plastic pestle. Further added 1 μl RNase A enzyme to facilitate cell lysis and placed in hot water bath at 37° C for 15 minutes. The samples were taken out from water bath, added proteinase-K and placed again in water bath at 65-70° C for one hour incubation. Later on, the samples were kept at room temperature for few minutes and then processed further as second step. Added 1 ml buffer solution of chloroform and iso-amyl alcohol which were prepared in the ratio of 24:1, centrifuged at 12,000 rpm under 4 °C for 5 minutes. The supernatant of 700 μl was taken in 2ml Eppendorf tubes, added 5M NaCl solution, chilled out on ice cube for 15 minutes. After chilling, added 70 μl sodium acetate, 500 μl chilled isopropyl alcohol and mix well by inverting many times to make DNA precipitation for settling in the bottom of the tube, preserved at -20° C for 40 minutes and further centrifuged at 12000 rpm under 4 °C for 5 minutes. Supernatants were discarded and the pellets were kept for air drying in 10-15 minutes, wash it with 500 μl concentrated ethanol and then centrifuged at 12000 rpm under 4 °C for 5 minutes and get pure DNA thread. Discarded the alcohol and get purified DNA pellet which was further dissolved in 50 μl PCR water and preserved in refrigerator at -20° C.
Gel Electrophoresis

The confirmation of DNA and PCR product was carried out using 1 and 2% agarose gel. Both solutions were prepared by adding 0.5 and 1.0 g agarose gel in 50ml chilled T.B.E buffer and then heated for 2 minutes in a micro wave oven. The gel was then allowed to cool down, but not to the extent that it would get solidified at room temperature, added 3μl Ethidium Bromide to both agarose gel products under Ultra Violet light. The gel was then poured into a casting tray and combs were placed to make wells for genomic DNA/ PCR product, it was then allowed to get solidified at room temperature. 7μl of DNA sample and 3μl of Bromophenol blue xylene Cyanol loading dye were mixed and poured into the wells of gel. Now gel was allowed to run for 45 min at 110 volts. The results from the gel were then confirmed under the UV light.

PCR

PCR product of the extracted DNA was prepared by adding 10μl PCR mix, 1.0 μl template DNA, 0.5μl forward primer and 0.5μl reverse primer. In this study two chloroplast (rbcl and mat-k) were evaluated by using the specific primers for the stated regions (Table 2). E.g. 60 slot PCR was used for gene amplification (Bio-RAD T100, Singapore). The PCR products were placed in thermocycler and Hot Start PCR Protocol for rbcl and matK was followed by adjusting temperature for cycle steps such as initial denaturation (94°C), Denaturation (94°C), Annealing (50-60°C for rbcl and 52°C for matK), Extension (72°C) and Final extension (72°C) at 5 minutes, 60 sec., 45 sec., 60 sec. and 5 minutes with total cycle of 35. PCR amplification success rates and sequencing rates were calculated following the method of Kress et al. (27).

Table 2. The primer used to amplify DNA barcodes and the amplification protocol.

| Primer       | Sequence (5’-3’)                          | Reference                   |
|--------------|-------------------------------------------|-----------------------------|
| rbclLa-F     | ATGTCAACACAAACAGACTAAAGC (26 mer)         | Levin, 2003                 |
| rbclLa-R     | GAAAAATCAAAGTCCACCRCG (20 mer)            | Kress & Erickson, 2007      |
| MatK-1R-Kim  | ACCGATCTCATGGAATCCTTGTTTC (27 mer)        | Ki-Joong Kim, pers. comm.   |
| MatK-3F-Kim  | CGTACAGTACTTTGTGTTTACAGAG (25 mer)        | Ki-Joong Kim, pers. comm.   |

DNA Sequencing and phylogenetic analysis

The PCR products of 18 samples were sent to TsingKe Biological Technology, Hong Kong (http://www.tsingke.net) for DNA sequencing through Pakistani company viz. Pakistan Hospital and Industrial Laboratory Co. (P.O. Box 1546.32/Abkari road, New Anarkali, Lahore-54400 Pakistan). Further, the DNA sequences were edited and aligned with BioEdit and MegaX software and the aligned sequences were matched on NCBI for BLAST (http://www.ncbi.nlm.nih.gov/package). The FASTA results of both partial and/or complete genome were run on MegaX for phylogenetic analysis (28,29).

3 Results and Discussion

The PCR amplified length of rbcl. marker was 600 bp as shown in Figure 1 that revealed 100% resolution power for all 15 samples of broomrape but the resolution power of PCR amplified band of rbcl in six samples (MG1, MG3, PN2, PN3, PN4 and QTA1) were highly visible over others. While, amplified length of matK marker was 800 bp but indicated lower resolution power for all samples except few ones (Figure 2). Both rbcl and matK are universal primers and its PCR based amplification is prerequisite for plant species identification including broomrape through DNA barcoding. The resolution power of matK in 8 samples was good that was clearly visible in 50% samples, faded in 25% samples (Figure 2) but very low and highly faded in the remaining samples (Fig. not shown). The low resolution of matK was might be due to inferior DNA quality, inappropriately adopted protocol and due to temperature fluctuation during amplification. In the PCR amplified bands of matK for 8 samples, four samples included KS1, KS2, MG1 and MG2 manifested clear and visible bands and faded in two samples of PN1 and PN2 but the other two samples (QTA1 and QTA2) were invisible and have no resolution power (Figure 2). These results are in line with findings of Wattoo et al. (30) who used rbcl and matK universal primers for identification of three plant species such as Solanum nigrum, Euphorbia helioscopia and Euphorbia helioscopia. Their PCR amplification results for these plant samples were good. Group et al. (22) used seven primers of plastid regions in species identification of angiosperms, gymnosperms and liverwort which were included rbcl, matK, rpoC1, rpoB, trnH-psbA, atpF- atpH and psbK-psbl respectively. According to Khanuja et al. (31) that quality of DNA plays vital role in PCR based amplification because the existence of some impurities and metabolites affect DNA quality. In such as cases different kind of protocol is need for DNA isolation. Our matK PCR amplification for broomrape samples are not in line with those reported by Awad et al. (32) who conducted DNA barcoding for different 18 Egyptian Triticum accessions and observed 100% PCR amplification for all samples using matK gene. However, Roman et al. (33) used plastid region trnD-trnF for PCR amplification of three Orobanche species collected from south Spain. The respective forward and reverse primers used their study were [tRNA-Asp (GUC)]; 5’-ACC AAT TGA ACT ACA ATC CC-3’ and [tRNA-Thr (GGU)] 5’-CTA

https://www.indjst.org/
CCA CTG AGT TAA AAG GG-3’. Through these plastid region they found a new diagnostic PCR amplified band of 219 bp in Orobanche species (Orobanche crenata) that distinguish from Orobanche minor.

**Fig 1.** PCR amplified bands of rbcl primer from 15 different samples of broomrape

**Fig 2.** PCR amplified bands of matK primer from 15 different samples of broomrape

**DNA Sequencing and phylogenetic tree analysis**

The sequence of plastid region matK gene was decoded to identify the phylogenetic association in 15 broomrape plants. The evolutionary history was inferred using the Neighbor-Joining method\(^{[34]}\). The optimal tree with the sum of branch length = 7.3182 is shown in Figure 3. The evolutionary distances were computed using the Maximum Composite Likelihood method\(^{[35]}\) and are in the units of the number of base substitutions per site. This analysis involved 15 nucleotide sequences. All ambiguous positions were removed for each sequence pair (pairwise deletion option). There were a total of 62303 positions in the final dataset. Evolutionary analyses were conducted in MegaX\(^{[36]}\). The phylogenetic tree shown the 15 samples into two distinct groups indicating two broomrape species i.e. Pramosa and Purpurea. The external nodes of 1, 2, 5, 7, 9, 13 and 11 representing Pramosa while 2, 4, 6, 8, 10, 12 and 14 representing Purpurea. Whereas, external node 15 and 16 showing the two species as out group in this phylogenetic tree. The BLAST analysis manifested accession no. HG803180.1 with 99.16% indent
for $P. ramosa$ and HG515536.1 with 96.41% indent for $P. purpurea$. Both species exhibited 95% Query cover with 0.0 E-value in BLAST results. In addition, the maximum and total score for $P. ramosa$ was 1498 and for $P. purpurea$ was 1371. The same 15 broomrape samples were analyzed for DNA sequencing using rbcl gene which did not perform well and none of the samples was identified through this marker whatever the reasons might be. According to Notredame et al. (37) that matK barcode possesses larger substitution rates within species and has proved a better candidate for plant species evaluation and documentation. Further, Savolainen et al. (38) reported that genetic variability recorded through DNA barcoding using matK gene has shown astonishing results as compared to other barcode candidate like rbcl alone and/or combination with atpB. The scientific studies have revealed that the plastid region matK is substantially important DNA barcode for plant species identification (39–43).

Morphologically both species were resembled to $O. aegyptiaca$ but the results of DNA barcoding validated as $P. ramosa$ and $P. purpurea$. Further, the special variability of broomrape species across the sampled area indicated no clear distribution but most of the samples were associated with two species as described in DNA sequencing.

Fig 3. Phylogenetic tree analysis of 15 broomrape samples depending on nucleotide sequence of matK gene using MegaX computer software program

### 4 Conclusion

The identification of broomrape species through DNA barcoding based on rbcl and matK gene was carried out and only matK gene sequence successfully detected two species of broomrape as Phelipanche ramosa and Phelipanche purpurea. But the rbcl gene sequence did not detected any species and its detection power was limited only to PCR bands. From this study, it was inferred that matK gene is a potential candidate for plant species discrimination through DNA barcoding.

### Acknowledgment

Authors are thankful to Department of Biotechnology, Balochistan University of Information Technology, Engineering and Management Sciences (BUTTEMS) Quetta, Balochistan, Pakistan for facilitating this research work.
32) Awad M, Fahmy RM, Mosa KA, Helmy M, El-Feky FA. Identification of effective DNA barcodes for Triticum plants through chloroplast genome-wide analysis. *Computational Biology and Chemistry*. 2017;71:20–31. Available from: https://dx.doi.org/10.1016/j.compbiolchem.2017.09.003.

33) Román B, Rubiales D, Torres AM, Cubero JI, Satovic Z. Genetic diversity in Orobanche crenata populations from southern Spain. *Theoretical and Applied Genetics*. 2001;103:1108–1114. Available from: https://dx.doi.org/10.1007/s001220100644.

34) Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. 1987;4:406–425.

35) Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences*. 2004;101(30):11030–11035. Available from: https://dx.doi.org/10.1073/pnas.0404206101.

36) Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution*. 2018;35(6):1547–1549. Available from: https://dx.doi.org/10.1093/molbev/msy096.

37) Notredame C, Higgins DG, Heringa J. T-coffee: a novel method for fast and accurate multiple sequence alignment 1 Edited by J. Thornton. *Journal of Molecular Biology*. 2000;302(1):205–217. Available from: https://dx.doi.org/10.1006/jmbi.2000.4042.

38) Savolainen V, Chase MW, Hoot SB, Morton CM, Soltis DE, Bayer C, et al. Phylogenetics of Flowering Plants Based on Combined Analysis of Plastid atpB and rbcL Gene Sequences. *Systematic Biology*. 2000;49(2):306–362. Available from: https://dx.doi.org/10.1093/sysbio/49.2.306.

39) Newmaster SG. Ethnobotany genomics – use of DNA barcoding to explore cryptic diversity in economically important plants. *Indian Journal of Science and Technology*. 2009;2(5):1–8. Available from: https://dx.doi.org/10.17485/ijst/2009/v2i5.1.

40) Mattia FD, Bruni I, Galimberti A, Cattaneo F, Casiraghi M, Labra M, et al. A comparative study of different DNA barcoding markers for the identification of some members of Lamiaceae. *Food Research International*. 2011;44(3):693–702. Available from: https://dx.doi.org/10.1016/j.foodres.2010.12.032.

41) Rolland M, Dupuy A, Pelleray A, Delavault P. Molecular identification of Broomrape species from a single seed by high resolution melting analysis. *Frontiers in Plant Science*. 2016. Available from: https://doi.org/10.3389/fpls.2016.01838.

42) Ghorbani A, Saeedi Y, Boer HJD. Unidentifiable by morphology: DNA barcoding of plant material in local markets in Iran. *PLOS ONE*. 2012;12(4).

43) Nishimura A, Fuse S, Tamura MN, Kato H, Takayama K. DNA barcoding reveals evolutionary changes in host specificity of a parasitic plants, Orobanche boninsimae (Orobanchaceae), endemic to the Bonin (Ogasawara) islands. *Pacific Science*. 2020;p. 87–97. Available from: https://www.muse.jhu.edu/article/752377.