Academic Journal of Nawroz University (AJNU), Vol.11, No.4, 2022

https://doi.org/10.25007/ajnu.v14n11a1194

Molecular Identification and Phylogenetic Relationship of Some Plant Species In Ashdagh Mountain, Sangaw, Kurdistan Of Iraq

Mohammed Ibrahim Khalil

Department of animal science, college of agricultural engineering sciences, University of Garmian

ABSTRACT

Phylogenetic and molecular perspective of plant identification can reveal the plant composition of areas. Many new species have been identified based on morphological characters in different area of Kurdistan, but few have been identified through DNA barcoding method. In this study, Internal transcribed spacers 1 and 2 (ITS1 and ITS2) were used to identify some species in Ashdagh Mountain. The identified taxa were belonged to different families and clades and not recorded in any studies. The nucleotide sequences of identified species were 98-100% identical with species in the genebank database. The phylogenetic analysis showed diversity in phylogenetic relationships among coexisting species. In addition, the existing of Chara sp. in the acidic stream of Awa spi is a new record and not observed before. The finding of this study show the importance of using molecular and phylogenetic technique to identify and understand phylogenetic relationship among coexisting species because Kurdistan is a hot spot and need more emphasis. In addition, new species occurrence data has been added to flora of Kurdistan.

KEY WORDS: Phylogenetic, DNA barcoding, ITS, Chara, Kurdistan.

1. Introduction

Molecular based species classification is considered an accurate method to identify species (Almerekova *et al.*, 2020; Li *et al.*, 2015). It has ability to tease out the intraspecific variation among more morphologically resemble species (Álvarez & Wendel, 2003). In addition, looking at the genetic sequences will help to understand how species evolve (i.e. convergent or divergent evolutions) (Ray, 2014). Recently, the taxonomic placement of many species have been changed after doing molecular analyses for them. The use of molecular data enables taxonomists to classify species at species level using DNA barcoding (short DNA sequence)(Hebert, Cywinska, Ball, & deWaard, 2003). Internal transcribed spacers 1 and 2 (ITS1 and ITS2) are part of ribosomal DNA. They are consider important DNA markers in molecular plant identification (Cheng *et al.*, 2015; Li *et al.*, 2015; Mishra *et al.*, 2016; Qin *et al.*, 2017). ITS regions are part of rDNA (ribosomal DNA) and existing between 18S and 28S genes. They have high evolution rates and easy to amplify in the lab (Coleman, 2007). These characteristics enable the taxonomists to use them to identify the species within and among taxa and endangered species (Hebert *et al.*, 2003; Hollingsworth, Graham, & Little, 2011). In the phylogenetic analysis, the use of ITS sequences will tell us the accurate placement of species along phylogenetic tree of coexisting species in certain area (Zhu *et al.*, 2018).

In Kurdistan, most of the classification studies were done based on morphological characteristics. So, incorporating of molecular phylogeny using DNA barcoding into species identification and classification processes will help identify and classify species more accurately than relying only on morphological identification (Nilsson *et al.*, 2008; Yao *et al.*, 2010).

The aim of this study is to do molecular identification for some species in the Sangaw area, Kurdistan. Results of this study will serve as a field guide for researchers to do more taxonomical, ecological and

physiological studies on plant composition structure since there are limited or no floral classification studies done in this area. Especially, the area characterized by distinct geological and geographical characteristics which make it more suitable for high floral diversity (Saman *et al.*, 2013; Sami *et al.*, 2019).

2. MATERIAL AND METHODS

2.1 Study area and sample collection

Sangaw is a subdistrict of Chamchamal district and it located in the northeastern part of Iraq (35.2851° N, 45.1799° E), about 35 km from southwest of Sulaimani city and 6 km west of Qara Dagh. More specifically, west of Sagerma Mountain. The region characterized by distinct geological structures (Kharajiany, 2013). A total of 18 taxa, representing 17 plant species and one algal species (Table 2), were collected from different locations on Ashdagh mountain, while the algal sample was collected from downstream of Awaspi river. The stream is an acidic stream, and it running through Darzilla village, Sangaw. The samples were collected on February 2-5, 2021. The leaves samples (n=10 leaves/plant) were stored in a zipper bag and transferred to the lab for molecular analysis (Polymerase chain reaction).

2.2 DNA extraction, amplification, and sequencing

AddPrep Genomic DNA Extraction Kit was used to extract the total genomic DNA. Two universal nuclear regions named internal transcribed spacers 1 and 2 (ITS1 and ITS2) primers were used to identify the unknown species and do phylogenetic analysis. These two primers are widely used in plant phylogenetic studies due to their ability to teasing out the intra- and interspecific variation among species (Almerekova *et al.*, 2020; Li *et al.*, 2015). Primers used in this study are listed in Table 1.

In the standard PCR, each reaction contained (MyTaqTM HS Mix-Bioline, USA) master mix (10 µL) 10 pmol of each primer and 20ng template DNA in a final volume of 20µL. PCR was performed using a three-step cycling protocol: initial denaturation (95°C/5 min) 1 cycle; [(denaturation (95°C/ 30sec), annealing (57°C/ 30 sec) (extension 72°C/ 30 sec) 40 cycle] and final extension (72°C/5) 1 cycle (Bio-Rad C1000 Thermal Cycler, USA). The PCR products were analyzed in 1% agarose gel (Only 1g of agarose was dissolved in 100 mL (1X TAE buffer) TAE: Tris-acetate EDTA) stained with Ethidium bromide (Figure 2).

Table 1. Universal plant ITS primers (Cheng *et al.*, 2015) used in this study.

DNA markers	Sequences (5′–3′)
	Forward primers
ITS1	GGAAGKARAAGTCGTAACAAGG
ITS2	CAWCGATGAAGAACGYAGC
	Reverse primers
ITS1	GCGTTCAAAGAYTCGATGRTTC
ITS2	RGTTTCTTTTCCTCCGCTTA

2.3 Blast, sequence alignment, and phylogenetic tree

Forward and reverse sequences of internal transcribed spacers regions per species have been edited and combined using Bioedit software (Hall, 1999). Nucleotide blast (Blastn) function in NCBI (National Center for Biotechnology Information) database was used to identify each species. Identification of species was based upon blast (% identity) and distance tree (fast minimum evolution and neighbor-joining) in NCBI online database.

Sequences were aligned using the ClustalW function in Bioedit with some manual adjustments (Figure 2). The length of the combined regions (ITS1 and ITS2) included 1430 base pairs. Mega X (Kumar *et al.*, 2018) was used to build Neighbor-Joining (NJ) tree. The use of both ITS1 and ITS2 in the phylogenetic analysis will provide better understanding of phylogenetic relationship among coexisting species (Qin *et al.*, 2017). The algal species (*Chara sp.*) was selected as an outgroup.

3. RESULTS

The PCR analysis showed clear band of 18 samples

(Table 2), with a band of 300-500bp corresponding to the amplification of the forward and revers ITS1 and ITS2 sequences. The nucleotide sequences of identified species were 98-100% identical with species in the genebank database. While, other taxa showed low identical percentage (<90%) with other species of same taxa in the database and marked as "sp.". The identified taxon have belonged to 12 families, and most of them are members of the rosids clade (Table 2). The sequence alignments showed significant variation in the base pairs among species (Figure 2). The existing of such variation is due to that ITS regions are not conserved and its subjected serial of addition and deletion of base pairs.

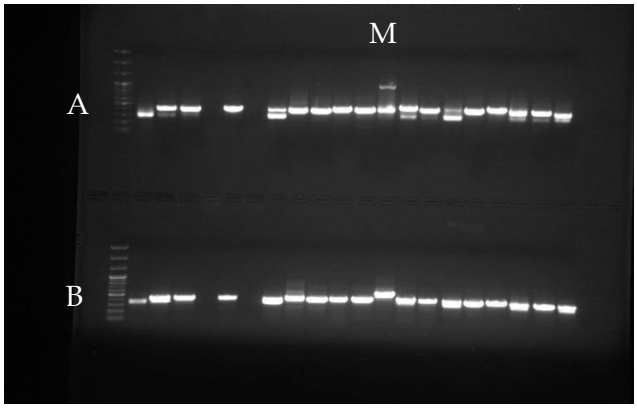

Figure 1: Agarose electrophoresis image, showing the PCR results for the collected species of ITS1 and ITS2 regions at 300-500 bp PCR product size. (A) Forward and reverse sequence of ITS1. (B) Forward and reverse sequence of ITS2. Lane M: Marker 100bp.

Table 2. List of the identified species in this study with their family names and clades.

No.	Species	Family	Clade
1.	*Linum mucronatum*	Linaceae	Rosids
2.	*Onobrychis ptolemaica*	Fabaceae	Rosids
3.	*Xanthium strumarium*	Asteraceae	Asterids
4.	*Teucrium oliverianum*	Lamiaceae	Asterids
5.	*Euphorbia Craspedia*	Euphorbiaceae	Rosids
6.	*Glaucium flavum*	Papaveraceae	Eudicots
7.	*Alcea sp.*	Malvaceae	Rosids
8.	*Onopordum tauricum*	Asteraceae	Asterids
9.	*Reseda aucheri subsp. aucheri*	Resedaceae	Rosids
10.	*Alcea pallida*	Malvaceae	Rosids
11.	*Umbilicus sp.*	Crassulaceae	Core eudicots
12.	*Haplophyllum sp.*	Rutaceae	Rosids
13.	*Verbascum sp.*	Scrophulariaceae	Asterids
14.	*Thymbra sp.*	Lamiaceae	Asterids
15.	*Chara sp.*	Characeae	Charophyta

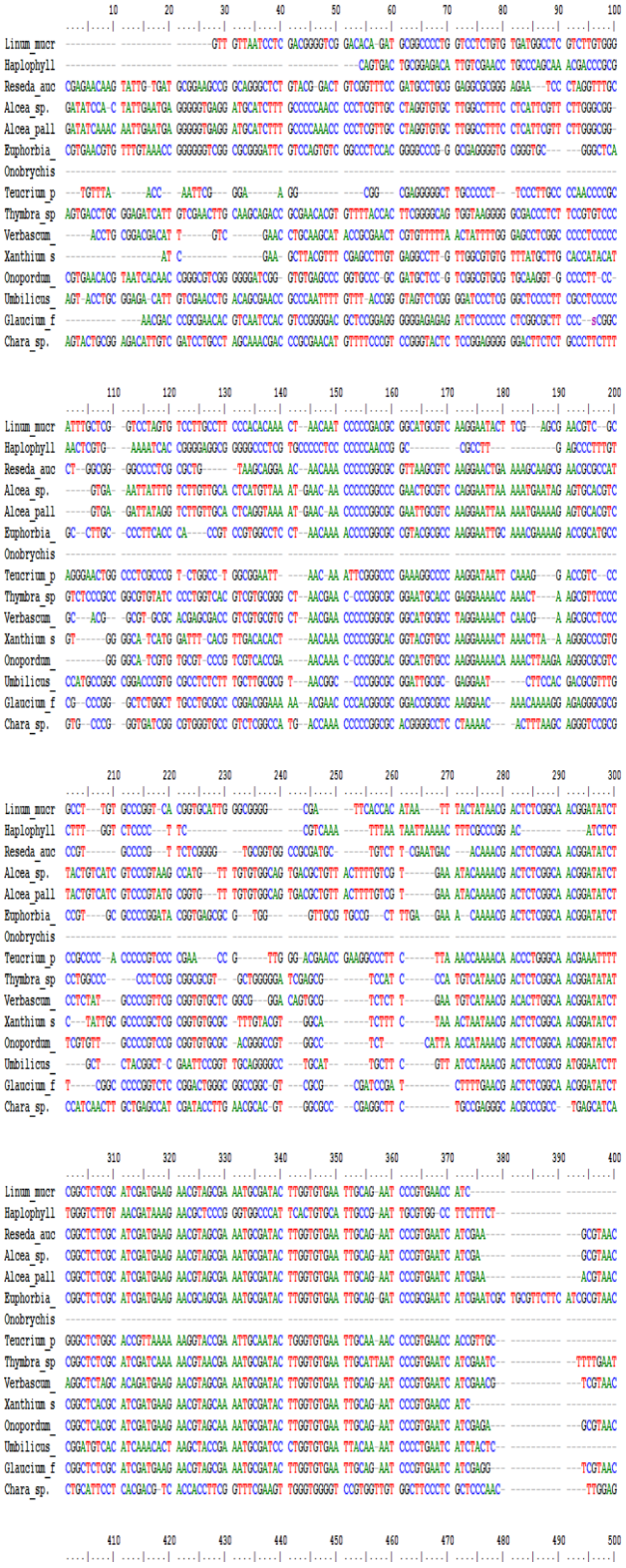

Figure 2: Clustal W multiple sequence alignment of ITS1 and ITS2 regions of the collected species.

The NJ tree is well resolved (bootstrap value >50%)

except for the *Euphorbia Craspedia,* in which the bootstrap value was below 50% (i.e., 48%) (Figure. 3).

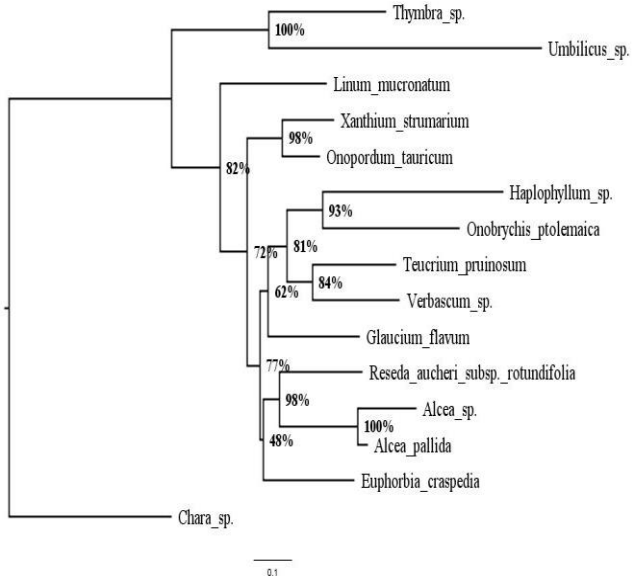

Figure 3. Neighbor-Joining tree based on ITS1 and ITS2 regions. The optimal tree with the sum of branch length = 5.67718577. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (100 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 15 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All ambiguous positions were removed for each sequence pair.

4. DISCUSSION

The distinct geological structure of the regions (Kharajiany, 2013) made it suitable habitat for plants diversity. Most of the identified species were belong to two clades; Rosids and Atersids. To our knowledge, no occurrence data of; *Onobrychis Ptolemaic, Teucrium oliverianum, Euphorbia Craspedia, Onopordum tauricum, Reseda aucheri subsp. aucheri* were recorded in the Sangaw. On the other hand, the possible reason that the species of some taxa was not able to identify is due to that the species of these taxa are new and not exist in the NCBI database (because the study area is neglected before and no molecular studies on floral composition have been done before).

Furthermore, the phylogenetic tree showed the distance among most species is not equal, and this due that the species belong to different families and have different DNA sequences. The existence of phylogenetic distance between closely related taxa such as *Alcea_sp* and *Alcea_pallida* is due to the change in some base pairs (addition or deletion). Such distance is might due to the long-term impacts of abiotic stress in the area such as low precipitation rates and geological structure of the area that cause adaptation and speciation of some species. The phylogenetic distance among coexisting species in the area reveals the existence of phylogenetic diversity among taxa of same family.

Most of the identified species were belong to the rosids clade, and they have different morphology, life span, and reproductive patterns (Folk *et al.*, 2018). Such diversity among members of this clade is considered a critical factor for understanding the evolutionary pattern and ecological dynamics of the angiosperms (Folk, Siniscalchi, & Soltis, 2020). Green algae *(Chara sp.)* has been found in the downstream of the acid stream. This study showed the importance of using DNA barcoding to identify and classify the taxa and species. Relying only on the classical classification methods (i.e., Morphological characteristics) will not be enough to identify and classify the species. In-plant molecular phylogeny, internal transcript spacer 1 and 2 (ITS1 and ITS2) can identify inter and intraspecific variation (Nilsson *et al.*, 2008; Yao *et al.*, 2010). The newly recorded species in different area of Kurdistan and the finding of this study could support that many new other species could evolve (due to species diversification) in the study area and in Kurdistan. This means that more molecular and morphological taxonomic studies are needed because Kurdistan is a hot spot and need more emphasis.

5. CONCLUSION

The difference in the nucleotide sequence and phylogenetic distance among coexisting species in the area reveals that the identified species are not in the

same clade and families and the taxonomic diversity of the area. This study added some new floral taxa found in the area to the flora of Kurdistan. The identified species are new and not recorded before in Sangaw. Therefore, it is important to include molecular methods in species identification process.

6. ACKNOWLEGEMENTS

Many thanks to Dr. Sirwan M. Amin for his assistance with collecting the samples.

7. REFERENCES

1. Almerekova, S., Shchegoleva, N., Abugalieva, S., & Turuspekov, Y. (2020). The molecular taxonomy of three endemic Central Asian species of Ranunculus(Ranunculaceae). *PLOS ONE, 15*(10), e0240121. doi: 10.1371/journal.pone.0240121
2. Álvarez, I., & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution, 29*(3), 417-434. doi: https://doi.org/10.1016/S1055-7903(03)00208-2
3. Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., & Zhou, S. (2015). Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. *Molecular ecology resources*. doi: 10.1111/1755-0998.12438
4. Coleman, A. W. (2007). Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nucleic Acids Research, 35*(10), 3322-3329. doi: 10.1093/nar/gkm233
5. Folk, R. A., Siniscalchi, C. M., & Soltis, D. E. (2020). Angiosperms at the edge: Extremity, diversity, and phylogeny. *Plant, Cell & Environment, 43*(12), 2871-2893. doi: https://doi.org/10.1111/pce.13887
6. Folk, R. A., Sun, M., Soltis, P. S., Smith, S. A., Soltis, D. E., & Guralnick, R. P. (2018). Challenges of comprehensive taxon sampling in comparative biology: Wrestling with rosids. *American Journal of Botany, 105*(3), 433-445. doi: https://doi.org/10.1002/ajb2.1059
7. Hall, T. A. (1999). BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. *Nucleic Acids Symposium Series, 41*, 95-98.
8. Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 270*(1512), 313-321. doi: 10.1098/rspb.2002.2218
9. Hollingsworth, P. M., Graham, S. W., & Little, D. P. (2011). Choosing and Using a Plant DNA Barcode. *PLOS ONE, 6*(5), e19254. doi: 10.1371/journal.pone.0019254
10. Kharajiany, S. (2013). The Middle Oligocene Rock Strata (Tarjil Formation) in Ashdagh Mountain, Sangaw District, Sulaimani Governorate, Kurdistan Region , NE Iraq. *Journal of Zankoy Sulaimani - Part A, 15*, 119-132. doi: 10.17656/jzs.10262
11. Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution, 35*, 1547-1549.
12. Li, X., Yang, Y., Henry, R. J., Rossetto, M., Wang, Y., & Chen, S. (2015). Plant DNA barcoding: from gene to genome. *Biological Reviews, 90*(1), 157-166. doi: https://doi.org/10.1111/brv.12104
13. Mishra, P., Kumar, A., Rodrigues, V., Shukla, A. K., & Sundaresan, V. (2016). Feasibility of nuclear ribosomal region ITS1 over ITS2 in barcoding taxonomically challenging genera of subtribe Cassiinae (Fabaceae). *PeerJ, 4*, e2638-e2638. doi: 10.7717/peerj.2638
14. Nilsson, R. H., Kristiansson, E., Ryberg, M., Hallenberg, N., & Larsson, K.-H. (2008). Intraspecific ITS variability in the kingdom fungi as expressed in the international sequence databases and its implications for molecular species identification. *Evolutionary bioinformatics online, 4*, 193-201. doi: 10.4137/ebo.s653
15. Qin, Y., Li, M., Cao, Y., Gao, Y., & Zhang, W. (2017). Molecular thresholds of ITS2 and their implications for molecular evolution and species identification in seed plants. *Scientific Reports, 7*(1), 17316. doi: 10.1038/s41598-017-17695-2
16. Ray, S. (2014). Molecular Markers in Phylogenetic Studies-A Review. *Journal of Phylogenetics & Evolutionary Biology, 02*. doi: 10.4172/2329-9002.1000131
17. Saman R., A., Mohammad K., M., Hasan H., A., Azhar A., A.-M., & M. S., A.-R. (2013). Fauna and flora of Hawraman mountain (part one) Hawraman lowest zone, Kurdistan province north east of Iraq. *Bulletin of the Iraq Natural Histroy Museum, 12*(4).
18. Sami, Y., Ali, G., Ahmed, M., Honar, M., & Errol, V. (2019). Wild orchids of the Kurdistan Region areas: a scientific window on the unexpected nature of the North-Western Zagros.
19. Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., . . . Chen, S. (2010). Use of ITS2 Region as the Universal DNA Barcode for Plants and Animals. *PLOS ONE, 5*(10), e13102. doi: 10.1371/journal.pone.0013102
20. Zhu, S., Li, Q., Chen, S., Wang, Y., Zhou, L., Zeng, C., & Dong, J. (2018). Phylogenetic analysis of Uncaria species based on internal transcribed spacer (ITS) region and ITS2 secondary structure. *Pharmaceutical Biology, 56*(1), 548-558. doi: 10.1080/13880209.2018.1499780