

## Research Article

# Exploring Parthenium weed biotypes by chloroplast DNA barcode analysis

Tehmina Anjum<sup>a\*</sup>, Rasheda Jabeen<sup>b</sup>, Steve Adkins<sup>c</sup>, Waheed Akram<sup>d</sup>

<sup>a</sup> University of the Punjab, Lahore, Punjab, Pakistan; <sup>b</sup> Lahore Garrison University, Lahore, Punjab, Pakistan; <sup>c</sup> University of Queensland, Brisbane, Australia; <sup>d</sup> Guangdong Academy of Agricultural Sciences, Guangzhou, China.

## INFORMATION ARTICLE

Received: February 16, 2018

Accepted: May 28, 2018

### Keywords:

*Parthenium hysterophorus*  
variations  
genetic diversity  
MEGA  
haplotypes

### \*Corresponding author:

<[tehminaanjum@yahoo.com](mailto:tehminaanjum@yahoo.com)>

### Cite this article:

Anjum T, Jabeen R, Adkins S, Akram W.  
Exploring Parthenium weed biotypes by  
chloroplast DNA barcode analysis. *Planta  
Daninha*. 2020;38:e020191623.  
<https://doi.org/10.1590/S0100-83582020380100052>

### Conflict of Interest:

The authors declare that there is no  
conflict of interest regarding the  
publication of this manuscript.

**Copyright:** This is an open-access article  
distributed under the terms of the Creative  
Commons Attribution License, which  
permits unrestricted use, distribution, and  
reproduction in any medium, provided that  
the original author and source are credited.



## HIGHLIGHTS

- *psbA-trnH* intergenic region were used as DNA barcode to assess biotypes of Parthenium weed samples.
- During sampling the parthenium samples showed morphological differences between the samples.
- Two clades of genetically diverse populations with 12 haplotypes were found in the sequence alignments.

## ABSTRACT

**Background:** Parthenium weed (*Parthenium hysterophorus* L.) is an invasive weed that has invaded vast regions of Pakistan in a relatively very short period of a decade or two, threatening the crop fields of the agrarian fed country. *Parthenium hysterophorus* L. is native of central South America and Gulf of Mexico, has now turned out to be a weed of global significance due to its alarming invasions and profuse spread in approximately all parts of the world. Its invasion is probably due to the contamination of its seeds in the imported grains from other countries of the world.

**Objective:** During comprehensive sampling from Pakistan and Australia, it was observed that parthenium weed accessions exhibited several distinct morphological features present at different geographical regions. Therefore this study focuses on the use of plastid DNA barcodes (*psbA-trnH*) to evaluate the extent of variations in nucleotide sequences between the parthenium weed sampled accessions.

**Methods:** The variability or genetic diversity was evaluated through sequencing of the amplified products and data was subjected to phylogenetic analysis in Molecular Evolutionary Genetic Analysis (MEGA; version 6.06) software.

**Results:** In Maximum Likelihood tree, mainly two clades with three subdivisions are evident which showed increased heterogeneity. The results of sequence based markers showed 12 haplotypes of *P. hysterophorus* populations (having two parsimony informative sites) with 10 indels and a few SNPs (single nucleotide polymorphisms).

**Conclusions:** The results advocate that there have been multiple introductions of parthenium weed into Pakistan.

## 1 INTRODUCTION

*Parthenium hysterophorus* L. commonly named parthenium weed is described as the most serious

weed of 21<sup>st</sup> century that has attained a major weed status in a relatively shorter time in most of the tropical and subtropical parts of the world due to its invasive nature. *Parthenium hysterophorus* is

notorious due to its adverse and detrimental effects on human beings causing allergic rhinitis, hay fever, severe asthma and contact dermatitis. Native to Central America, it has now achieved substantial invasion and has extended to eastern and Southern Africa, Southern USA, South Asia, the South China Region, Vietnam, Australia and the South Pacific (Navie et al., 1996a; Tamado and Milberg, 2000). In Pakistan, the extent of its spread was demonstrated by an eye opening survey done on the local flora of Islamabad, federal capital of Pakistan and its neighboring city Rawalpindi, by Shabbir and Bajwa (2007) who reported *Cannabis sativa*, another exotic invasive species, as the main dominant weed genus with *Parthenium hysterophorus* L., *Senna occidentalis* L., *Malvastrum coromandelianum* L. Garcke and *Lantana camara* L. Whereas, just after the two years *Parthenium hysterophorus* has achieved the major weed status and recent situation is that no location in Islamabad is without the menace of parthenium invasion (Hassan and Amin, 2009). Although various control measures have been in practice globally for the last two decades, success in management of this weed has not yet been achieved. In order to find out an eco-friendly, comprehensive and effective management option in Pakistan, genetic analysis of the weed using molecular markers can provide extremely useful information. The evolution in genetic diversities of invasive weeds is a major contributing factor in successful invasion of the invasive species (Prentis et al., 2008). Furthermore, this knowledge of invasion histories is extremely important for selecting improved and better management options of the exotic weeds in the invaded areas (Prentis et al., 2010). Keeping in view the importance of genetic analysis a preceding study by Jabeen et al. (2015) elucidated high genetic diversity and increased invasive potential of parthenium weed within and among populations based on ISSR fingerprinting markers which showed that high genetic diversity is related to parthenium weed intraspecific hybridizations.

The *Parthenium hysterophorus* infestation in Pakistan shows high degree of variation in the morphological features in varying environmental conditions at different geographical regions in Pakistan. The same findings were observed in Australia where two biotypes of parthenium weed are known to be present due to the two separate introductions from USA which are morphologically distinct in a way that the plants of biotype from central Queensland is more aggressive, taller, produce larger diaspores in contrast to the plants of second

biotype which are limited in their spread (Navie et al., 1996b). Morphological variations are due to the changes found at DNA level. To elucidate the variations at molecular level the present study was done which is novel in Pakistan regarding the use of DNA barcodes to identify different biotypes/haplotypes of *Parthenium hysterophorus*. Employing small, standard DNA regions or DNA barcodes to genomes is an authentic tool for plant identification with many advantages and potential over the conventional morphology based taxonomy (Okuyama and Kato, 2009; Costion et al., 2011; Yang et al., 2012; Phukunsi et al., 2016). DNA barcodes are usually slow evolving regions that are used to address basal angiosperm relationships at high resolution and have successful applications in plant invasion biology (Comtet et al., 2015). It gave insight into biotypes/haplotypes that is the number of times of introductions of the weed (multiple introductions) and can be exhibited by the differences in the nucleotide sequences. Using DNA barcodes in recent researches can surmount several limitations in detection of morphologically cryptic plant species that show high degree of phenotypic plasticity in which a genotype may acquire diverse phenotypes in response to the changing environment (Valentini et al., 2008).

Earlier workers has recommended ITS sequence as a useful barcode due to its high discriminatory power than other barcodes but sequencing and amplifying whole ITS sequence in various species is a problem (Lee et al., 2011). However, ITS 2 region gained less attention due to the presence of multiple copies in the genome which usually illustrate high levels of differentiation. Moreover, ITS sequences exhibited more inter and intra specific distances than intra generic which may lead to inaccurate results (Yamaguchi et al., 2006). More recent researches proposed the idea of using whole plastid genome as a barcode in identifying the plants in comparison to single or two locus barcodes (Nock et al., 2011; Yang et al., 2014). However, this initiative is still in the waiting list of approval because of high sequencing costs and documented as restricted in neglecting specie boundaries. Some more recent proposals are being known to overcome the problems associated with DNA barcodes to increase the extent of differentiation in plants such as target enrichment and genome skipping (Hollingsworth et al., 2016). The present study was performed by using *psbA-trnH* intergenic region as DNA barcode to evaluate *Parthenium* weed samples of two pre identified biotypes of Australia and samples collected from Pakistan.

## 2 MATERIALS AND METHODS

### 2.1 Sampling

Ten different samples of two discrete populations of *Parthenium* in Australia (Toogoolawah and Clermont) were collected in June 2011 from the School of Land and Food Sciences University of Queensland, Brisbane Australia. In Pakistan parthenium weed sampling was done from Khyber Pakhtun Khwa (KPK) and Punjab province. In KPK, twenty different *Parthenium* weed samples were taken from Peshawar valley (including Mardan, Charsadda, Swabi, Attock and Peshawar) in Feb-March 2012; the remaining samples were collected in March-May 2012 from wide eco-geographical areas in the Punjab Province of Pakistan. A survey of the southern Punjab region was done in May-June, 2012 but parthenium weed population could not be collected due to sparse and reduced growth of parthenium weed at that time of year. A detailed list

of the sample sizes and sampling sites is given in Table 1. In addition, respective latitudes, longitudes (in degrees), elevations, mean annual temperatures and mean annual rainfall of the sampling sites were also determined.

For each population, either ten dried leaf samples in silica gel were collected or ten fresh leaf samples were collected, stored in zip lock bags, held in ice containing cooler and stored at -80 °C refrigerator until used for DNA extraction. In some situations where the fresh leaves were unavailable, weed seed samples were collected (Australian samples).

### 2.2 Morphological observations of the weed during sampling

During sample collection various features were found. *Parthenium* resembles to the introduced weeds of the same family for example its leaves resemble leaves of *Artemisia* sp. and the whole plant

**Table 1** - Ninety five accessions of *Parthenium hysterophorus* L. populations analyzed

Accession	Site	Latitude (in °C)	Longitude (in °C)	Elevation (ft above sea level)	Mean anual temperature	Mean anual rain fall (mm)
phchar pop1	Charsadda	34.15	71.73	990	22.5 °C	460
Phswa	Swabi	34.11	72.46	1094	22.2° C	637
Phmar	Mardan	34.19	72.06	1002	22.2° C	559
Phpes	Peshawar (GT Road)	34.26	71.54	1158	22.7° C	384
Phpes	Peshawar cantonment	34.00	71.47	1235	as above	as above
Phpes	Kabul River	34.09	71.70	953	as above	as above
Phpes	Attock	33.80	71.72	2003	22.2 °C	537
phisl pop2	Islamabad (I-8)	33.67	73.07	1715	21.3 °C	941
Phisl	Chakshehzad	33.16	73.14	1636	as above	as above
Phisl	KaralChowk	33.60	73.13	1602	as above	as above
Phisl	Daman-e-Koh	33.74	73.05	2391	as above	as above
Phisl	Rawalpindi	33.66	73.16	1658	21.5°C	941
phchk pop4	Chakwal	32.93	72.85	1628	22.3°C	519
Phdin	Dina	33.02	73.58	854	as above	as above
phjhl pop5	Jhelum GTroad	32.93	73.71	762	23.6° C	847
Phjhl	Jhelum Cantt.	32.91	72.10	768	as above	as above
phshk pop6	Shakargarh	32.26	75.15	879	26.2 °C	722
phsia pop7	Sialkot	32.49	74.53	837	32.6°C	957
phlhr pop3	Gujranwala	32.15	74.18	748	23.9 °C	581
Phlhr	Lahore	31.96	74.54	715	24.2 °C	607
Phlhr	Walton	31.48	74.35	707	as above	as above
Phlhr	Changamanga	31.09	73.96	660	as above	as above
Phlhr	Barqi road	31.49	74.48	709	as above	as above
Phlhr	Ittefaq hospital	31.51	74.32	714	as above	as above
Phlhr	Minare Pakistan	31.59	74.32	696	as above	as above
phchi pop8	Chiniot	31.72	72.97	597	24.1 °C	336
Phfsd	UAF	31.44	73.07	606	24.2 °C	346
Phfsd	Ayub agriculture Research Institute	31.41	73.06	618	as above	as above
Phjhg	Jhang sadder	31.27	72.32	514	24.7° C	248
Phjhg	Jhang city	31.30	72.35	507	as above	as above
phokr pop9	Okara	30.80	73.44	604	24.5 °C	296
Phshw	Sahiwal	30.66	73.10	568	24.7° C	279
phtogw pop10	Kilcoy	26.94	152.56	12259	24.5 °C	620
phclm pop11	Emerald	23.50	148.16	589	29.4 °C	644

Collected from Pakistan and Australia, respective latitudes, longitudes, altitudes, mean annual temperatures, mean annual rainfall and their respective sample size (N).

resembles the Chrysanthemum in its earlier/rosette stage. For these reasons the sampling was done with care at different locations identifying the *P. hysterophorus* properly according to the plant characteristics described by Tiwari et al. (2005). Photographs were taken with a digital camera. Two varying growth phases of *P. hysterophorus* were also noted along with the tallness of the weed plant, pinnately or bipinnately dissected leaves, length of peduncles, flowering prospective (number of flowers) and seed production potential.

### 2.3 DNA extraction, PCR and PCR purification

Plant leaf DNA extraction was done using Doyle and Doyle, (1990) protocol, with some modifications. The modifications involved the addition of 0.2 mL of proteinase K in the DNA extraction buffer, two times extractions i- with phenol, chloroform, Isoamylalcohol (25:24:1) ii- with chloroform: Isoamylalcohol (24:1) along the cold isopropanol incubation period of DNA from 30 min 3 days at -20 °C refrigerator. The DNA samples were run on the 1.0% agarose gel for examining the quality and quantity of the extracted DNA. The *psbA-trnH* non coding intergenic chloroplast DNA region was amplified with universal primers *psbA* forward (forward: 5'-GGAAGTTATGCATGAACGTAATGCTC-3') and *trnH* reverse (reverse: 5'-CGCGCATGGTGGATTCACAA-3'). The PCR reaction was undertaken in 25  $\mu$ L containing: 1X Taq Buffer, 3 mM MgCl<sub>2</sub>, 0.2 mM each dNTPs, 0.8  $\mu$ M each forward and reverse primers (1  $\mu$ L of 10 pmol/ $\mu$ L), 1.0  $\mu$ L Taq polymerase and from 10-30 ng genomic DNA. PCR amplification conditions were: 94 °C for 5 minutes, then 30 cycles of 94 °C (30 seconds) denaturation, 58 °C (30 seconds) annealing, 72 °C (2 minutes) extension, and 72 °C (10 minutes) final extension. After 1.0% agarose gel electrophoresis, PCR product (20  $\mu$ L) was purified with the Isolate PCR and Gel kit (DNA purification Kit, Bionline).

Purified and amplified DNA fragments were sent to the Advanced Biosciences International (ABI) sequencing facility, Malaysia for sequencing. The sequences were retrieved from the company through electronic mail.

### 2.4 Sequence assembly, alignments and phylogenetic tree

The obtained sequences were first analyzed with the Chromas LITE (version 2.1) software program. Bio Edit program (Version 7.2.5) (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>) and Contig assembly program were used for detailed sequence analysis

and consensus sequences were obtained. For phylogenetic analysis of *psbA-trnH*, sequences were multiple aligned using CLUSTALW executed in MEGA 6.0.

Phylogenetic tree was constructed based on the maximum likelihood by Bootstrap re sampling (100 replications) method which was used to measure the reliability of individual nodes in each phylogenetic tree. Phylogenetic inferences were done and conclusion was drawn.

## 3 RESULTS AND DISCUSSION

During surveys the Parthenium plant was generally found with strong taproot (Figure 1F), erect branched stem and alternately arranged pinnately and bipinnately dissected leaves. Parthenium weed showed two growth phases during its life cycle; the vegetative/juvenile stage in which large, green, pinnate leaves form rosette like arrangement on the top of ground which was devoid of flowering (Figure 1A). The larger lower leaves did not allow any



**Figure 1** - Various morphological features of Parthenium weed. **A:** rosette stage or juvenile stage of Parthenium weed. **B:** the hairy, octangular stem with trichomes. **C:** the inflorescence of parthenium weed, a capitulum measured with Vernier caliper. **D:** fork like peduncles possessing capitula on the top. **E:** Microscopic view of the Parthenium seed. **F:** the strong tap root of the weed.

other vegetation to grow beneath their carpet forming pattern and ultimately forms area of its own. The other stage was the adult stage which originated from the juvenile stage as an erect, profusely branched, octangular and hairy stem (Figure 1B). The stem was woody in the mature plants. It was grooved along its length reaching to the height of 2 meters or above in the soils where water was in adequate amounts usually during the monsoon weather when the Khyber Pakhtun Khwa (KPK) and northern Punjab get maximum rainfall during the year. Each capitulum was whitish to creamish in color, pentangular, about 4-5 mm in diameter produced on the top of the plant on fork shaped peduncles which arised from the leaves (Figure 1C and D). The heterogamous capitula consisted of 4-5 pistillate ray florets and staminate central disc florets which were more than 20 in number. The ray florets produced single basal seed. The weed seeds were minute, light in weight, wedge shaped and each flower usually produced 4-5 seeds about 1.5-2 mm in length with white scales. The microscopic view of the structure of the parthenium seed is presented in Figure 1E.

During sampling the parthenium samples showed morphological differences between the samples. The samples collected from KPK and northern Punjab showed longer stems with larger number of leaves,

branches and flowers. These flowers were found to be produced on elongated peduncles with bracts (Figure 2B) protruding out of the plant (Figure 2A) to get more advantage of the amphiphilous pollination. As the number of flowers was more therefore these plants were hypothesized to produce larger number of seeds which also contributed to its high invasion potential and larger infestation (Figure 2C) all along the Peshawar valley and northern and central Punjab Province.

In contrary, at Sahiwal and Okara, the regions in the south of Lahore division (towards the southern Punjab region) another biotype of *Parthenium* weed was found that possessed distinct morphology. In these areas, the *Parthenium hysterophorus* infestation was found to be less invasive as compared to the aforementioned biotype. These plants were not so much long, have shorter stems and therefore less number of fleshy green alternately arranged broad leaves with inflorescence producing on the short peduncles with reduced or no bracts (Figure 3A and B) possibly capable of only self-pollination. Only a few capitula/flowers were noticed to be found in the central top portion of the plant and hence, these weed plants produced lesser number of seeds and have longer generation time. This biotype was confined only at the specified area and thus exhibited weaker



**Figure 2** - The morphological features of taller biotype of *Parthenium* weed. Whereas, A: A mature plant of *Parthenium* weed. B: large number of flowers with longer peduncles and bracts. C: The violent infestation of the taller biotype.



**Figure 3** - Morphological features of shorter biotype of Parthenium weed. Whereas, A: The second shorter biotype with a few capitula on the top. B: the shorter

potential to spread. The farmers of the area were investigated about the cultivation of crops there which was found to be imported Australian wheat that was being cultivated in that region.

The former biotype possess increased invasive potential which was found to occur mostly in all the regions of north eastern Pakistan and exhibited distinct morphological features which not only enhance its maximum growth but are major contributors for its increased competitive ability in contrast to the other shorter biotype. Taller stems, more number of leaves and branches and particularly augment production of capitula on long peduncles and provide bright chances for efficient amphiphilous mode of pollination and faster growth rates. Faster growth rates are correspondingly high due to increased capitula production and consequently generation of large number of viable, vigorous seeds due to increased intra specific hybridizations and cross pollinations. This more aggressive biotype evidenced in our study can be related to the Clermont biotype of Australian parthenium population described by Navie et al. (1996b) and Hanif et al. (2012) in their respective studies.

The second biotype which although is invasive but not to the extent to the former one can be linked to

the Toogoolawah Australian parthenium populations Navie et al. (1996b) and Hanif et al. (2012), exhibited peculiar morphological features which render this biotype less competitive and confined only to specified regions of Okara and Sahiwal. These weed plants have shorter stems, less number of branches and leaves with a few capitula produced on short peduncles only on the top portion of the plants. These biological structures provide more chances of self-pollination in contrast to the cross-pollination and high production of viable seeds in the aforementioned biotype.

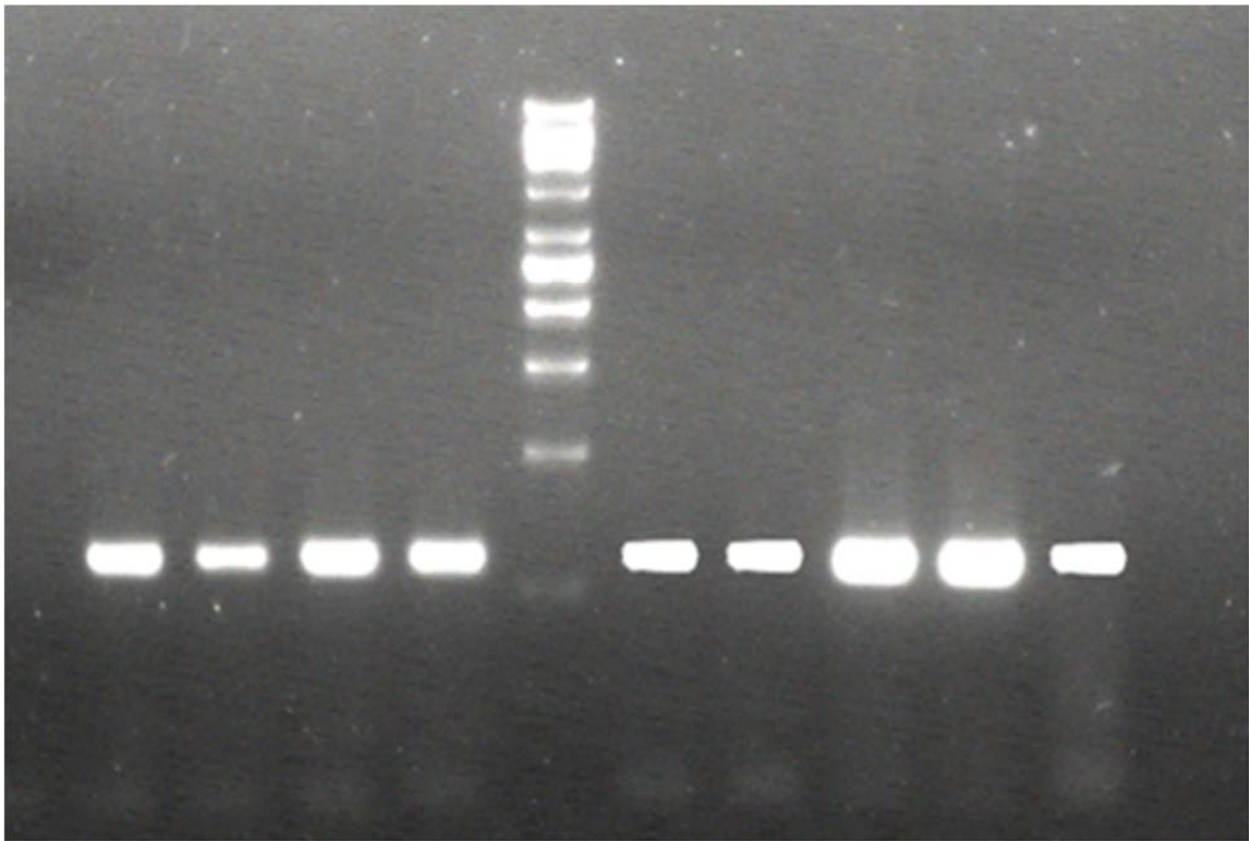
These varying morphological features of the two different biotypes provided the urge and strong recommendation for the molecular genetic analysis of these putative two biotypes in Pakistan. This notion of performing molecular characterization is also of great importance because molecular data can be employed in the form of DNA sequences and thus easily used in comparing the living beings, demonstrating their process of evolution. In the present study, the collected *Parthenium hysterophorus* samples were investigated through the molecular characterization among Australian parthenium samples.

### 3.1 Molecular analysis: DNA barcoding

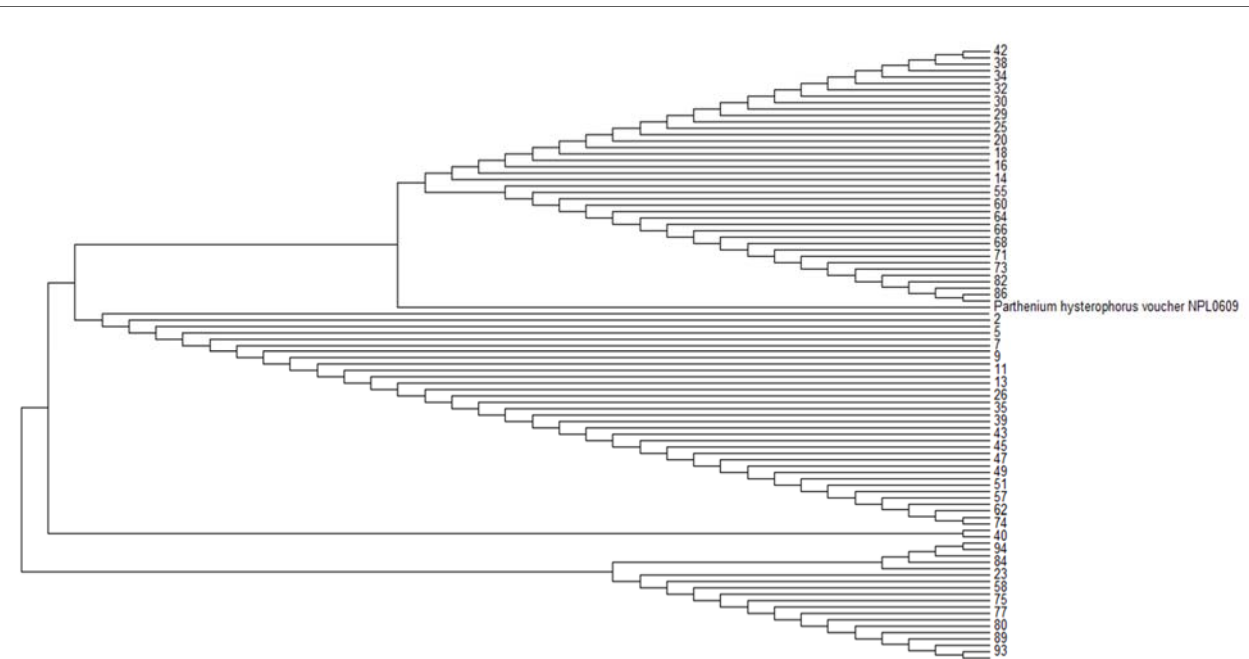
**DNA extraction and PCR amplification of *psbA-trnH* chloroplast DNA** - The extracted DNA of the parthenium leaf samples is shown in Figure 4. In PCR, amplification band of about 500 base pairs was obtained using specific *psbA-trnH* barcode primers on the extracted DNA of the 95 samples (Figure 4). The purified PCR product after running on the agarose gel was sent to sequencing facility in Malaysia.

### 3.2 Sequencing and sequence phylogenetic analysis of the amplified fragments

After three weeks the nucleotide sequences were retrieved through electronic mail which were subsequently viewed and analyzed in Chromas LITE software. Consensus sequences were produced by CAP 3 (Contig Assembly Program) free online software program. The consensus sequences were transformed into FASTA format and saved in a Word notepad file. The subsequent multiple alignments of all the contigs were carried out by CLUSTALW in Molecular MEGA (version 6.06) sequence analysis program. Barcode differentiations were visualized using character based Maximum Likelihood (ML) tree method and then bootstrapped with 100 replicates. The ML tree can be visualized in the Figure 5. In ML



**Figure 4** - PCR amplification of psbA-trnH intergenic barcode with VC 1Kb DNA ladder in middle.



**Figure 5** - Maximum likelihood phylogenetic tree of *Parthenium hysterophorus* L. accessions obtained after the alignments.

tree phylogenetic analysis, mainly two clades with three subdivisions was obtained by the division of first clad into two. There was a high level of genetic heterogeneity in the first clad. In total, there were 12 haplotypes of *P. hysterophorus* samples have been obtained by editing the sequence data. The

*Parthenium hysterophorus* voucher NL0609 was also included in the alignments. Two parsimony informative sites were found with 14 indel polymorphisms. In our study increased heterogeneity was found in the collected *P. hysterophorus* samples. The sequences can be provided on demand.

The collection of geographically distinct *Parthenium* populations took about more than a year and the respective sampling sites were documented using GPS along respective latitudes, longitudes and altitudes.

Non coding sequences of the chloroplast genome are a principal source of data for population genetics and molecular systematic studies of plants however comparatively little is known about points of variation amongst diverse non coding areas of the chloroplast genome. Many researchers have focused on the use of combination of nuclear ITS sequences and different regions of the plastid genome protein coding regions as the potential plant barcodes (Wang et al., 2010; Srirama et al., 2010).

The utility of *psbA-trnH* intergenic spacer non coding region at lower taxonomic levels i.e. intra specific differentiation is used in the present study in some preliminary experiments with the weed which explored the variations efficiently after sequencing the region in 95 accessions of *Parthenium* weed. In our study we have not used the universal barcodes *rbcL+matK* as recommended by CBOL plant working group in (Peter et al., 2009) because in some studies it is shown that this barcode in combination or each one marker have less discriminatory power (less than 80%) in taxonomically complex groups such as Angiosperms (Zhu et al., 2014). Hao et al. (2012) ranked *psbA-trnH* chloroplast non coding spacer third most efficient region after *psbD-trnT* and *psbJ-petA* plastid spacers in discriminating plant taxa due to the high degree of divergence. The *psbA-trnH* region has proved to be efficient in determining heterogeneity at the nucleotide level and appears to have greater intra specific differences making it popular candidate for population studies. The *psbA-trnH* also proved to be superior in inferring the mean inter specific distance in plants and in examining the barcoding gap. Kress and Erikson (2007) in another similar study, regarded *psbA-trnH* spacer region as the most efficient single locus plant barcode due to its high level of genetic variability. Tang et al. (2016) documented the use of *psbA-trnH* intergenic region with four other barcodes (*matK*, *rbcL*, ITS and ITS2) to identify a Chinese medicinal herb genus *Uncaria* (Gouteng). Their study declared *psbA-trnH* as the best applicable barcode region with ITS2.

Pang et al. (2012) has evaluated the use of *psbA-trnH* intergenic spacer in the meta analysis in which it is documented that frequent intra-population inversions are present in the spacer which could mislead to the increased measure of diversities

among species. This region is although short i.e. 500 bp in length ranges from 119-1,000 base pairs in angiosperms (Cowan et al., 2006). Our results showed 12 haplotypes with two parsimony informative sites and 14 indel polymorphisms. Although *psbA-trnH* is more polymorphic than *matK* or *rbcL*, its use as core barcode is limited due to several reasons. Relatively short length of the spacer with extensive prevalence of inversions and insertions within species and sequencing errors due to long polystructures make it problematic. In addition, due to frequent indels it can also explore excessive polymorphisms (Dong et al., 2015). For such population genetic structure analysis the use of further gene regions as barcodes would prove to be more advantageous. However, in the present study more regions were not practiced due to the shortage of time and funding.

In present study a phylogenetic tree was constructed for the final consensus sequence data by the Maximum Likelihood method. This was constructed by using MEGA 6.0 software (Tamura et al., 2011). In figure 5, two clades of genetically diverse populations with 12 haplotypes were found along 2 parsimony informative sites in the sequence alignments in MEGA 6.06 software. Ten indels and a few SNPs (single nucleotide polymorphisms) were observed. The analysis in the present study showed that the weed has experienced multiple chances of introductions in the country. In the ML tree, all the sampled accessions were included with Australian and Pakistan *Parthenium hysterophorus* samples. There are two main divisions where the first division is further divided into two, ultimately with total three subdivisions appearing in the phylogenetic tree. The two main divisions might have showed the two varying biotypes of parthenium weed as evidenced in the Australia, while the first division is further grouped into two showing the high level of heterogeneity in this group which included most samples of the first biotype. The Australian samples were assembled in second group along the weed samples collected from okara and other southern Punjab samples are grouped. There are many reasons for high genetic heterogeneity in the parthenium weed sampled accessions. The multiple introductions of the weed in Pakistan plausibly have formulated the basis for genetically diverse populations which got the chance to prevail and spread in the country at alarming rate. Conversely, single introduction of genetically diverse founder populations is another possibility which may perhaps has caused increased genetic diversity in the *P. hysterophorus* populations.



## 4 CONCLUSIONS

The results of our analytical study showed that *Parthenium hysterophorus* got entry in the country through multiple introductions and has now established in vast agricultural areas which is an alarming situation.

## 5 CONTRIBUTIONS

TA: designed the project and provided financial assistance. Also provided scientific assistance in performing trials and write-up. RJ: PhD scholar; performed all necessary experiments. SA: co-supervisor; provided research facilities and guidance. WA: helped in write up.

## 6 ACKNOWLEDGEMENT

We are highly obliged to Dr. Wajiha Irum for improving the manuscript and Dr. Peter Prentis for providing his useful suggestions during the work. This work was supported by Higher Education Commission Pakistan as a split PhD program in Australia and Pakistan.

## 7 REFERENCES

- Comtet T, Sandionigi A, Viard F, Casiraghi M. DNA metabarcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. *Biol Invas*. 2015;17:905-22.
- Costion C, Ford A, Cross H, Crayn D, Harrington M, Lowe A. Plant DNA barcodes can accurately estimate species richness in poorly known floras. *PLoS One*. 2011;6(11):e26841.
- Cowan RS, Chase MW, Kress WJ, Savolainen V. 300,000 species to identify: problems, progress, and prospects in DNA barcoding of land plants. *Taxon*. 2006;55:611-6.
- Dong W, Xu C, Li C, Sun J, Zuo Y, Shi S, et al. *ycf1*, the most promising plastid DNA barcode of land plants. *Sci Reports*. 2015;5(8348)
- Doyle JJ, Doyle JL. Isolation of plant DNA from fresh tissue. *Focus*. 1990;12:13-5.
- Hanif Z, Adkins SW, Prentis PJ, Navie SC, O'Donnell C. Characterization of the reproductive behavior and invasive potential of *Parthenium* weed in Australia. *Pak J Weed Sci Res*. 2012;18:767-74.
- Hao DC, Xiao PG, Peng Y, Dong J, Liu W. Evaluation of the chloroplast barcoding markers by mean and smallest inter specific distances. *Pak J Bot*. 2012;44:1271-4.
- Hassan G, Amin A. First annual Technical Report. HEC International Linkage Project with University of Queensland, Australia, on *Parthenium* Weed; 2009.
- Hollingsworth PM, Li DZ, van der Bank M, Twyford AD. Telling plant species apart with DNA: from barcodest to genomes. *Philos Trans Royal Soc B*. 2016;371:1-9.
- Jabeen R, Prentis P, Anjum T, Adkins SW. Genetic Structure of Invasive Weed *Parthenium hysterophorus* in Australia and Pakistan. *Int J Agric Biol*. 2015;17:327-33.
- Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One*. 2007;2(6):e508.
- Navie SC, McFadyen RE, Panetta F, Adkins SW. A comparison of the growth and phenology of two introduced biotypes of *Parthenium hysterophorus*. In: Proceedings of the Eleventh Australian Weeds Conference. Melbourne: Weed Science Society; 1996b. p.313-6.
- Navie SC, McFadyen RE, Panetta FD, Adkins SW. The biology of Australian weeds, *Parthenium hysterophorus* L. *Plant Prot Quart*. 1996a;11:76-88.
- Nock CJ, Waters DL, Edwards MA. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol J*. 2011;9:328-33.
- Okuyama Y, Kato M. Unveiling cryptic species diversity of flowering plants: successful biological species identification of Asian *Mitella* using nuclear ribosomal DNA sequences. *BMC Evol Biol*. 2009;9:105.
- Pang X, Liu C, Shi L, Liu R, Liang D, Li H. Utility of the *trnH-psbA* intergenic spacer region and its combinations as plant DNA barcodes: A Meta-Analysis. *PLoS One*. 2012;7(11):e48833.
- Peter M, Hollingsworth LL, Forrest JL, Spouge MH. CBOL Plant Working Group. A DNA barcode for land plants. *Proc Natl Acad Sci USA*. 2009;106(31):12794-7.
- Phukuntsi MA, Brettschneider H, Dalton LD, Kearney T, Badenhorst J, Kotze A. DNA barcoding for identification of cryptic species in the field and existing museum collections: a case study of *Aethomys* and *Micaelamys* (Rodentia: Muridae). *Afr Zool*. 2016;51(1):69-76.
- Prentis PJ, Wilson JR, Dormontt EE, Richardson DM, Lowe AJ. Adaptive evolution in invasive species. *Trends Plant Sci*. 2008;13:288-94.
- Prentis PJ, Woolfit M, Thomas-Hall SR, Ortiz-Barrientos D, Pavasovic A, Lowe AJ, et al. Massively parallel sequencing and analysis of expressed sequence tags in a successful invasive plant. *Ann Bot*. 2010;106:1009-17.
- Shabbir A, Bajwa R. *Parthenium* invasion in Pakistan- a threat still unrecognized. *Pak J Bot*. 2007;39:2519-26.
- Srirama R, Senthilkumar U, Sreejayan N, Ravikanth G, Gurusurthy BR, Shivanna MB, et al. Assessing species admixtures in raw drug trade of *Phyllanthus*, a hepatoprotective plant using molecular tools. *J Ethnopharma*. 2010;130(2):208-15.
- Tamado T, Milberg P. Weed flora in arable fields of eastern Ethiopia with emphasis on the occurrence of *Parthenium hysterophorus*. *Weed Res*. 2000;40:507-21.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011;28:2731-9.
- Tang Y, Wu Y, Huang R, Chao N, Liu Y, Xu P. Molecular identification of *Uncaria* (*Gouteng*) through DNA barcoding. *Chinese Med*. 2016;11:1-14.

Tiwari S, Adhikari B, Siwakoti M, Subedi K. An inventory and assessment of invasive alien plant species of Nepal. Nepal: IUCN –The World Conservation Union; 2005. 115p.

Valentini A, Pompanon F, Taberlet P. DNA barcoding for ecologists. *Trends Ecol Evol.* 2008;24:110-7.

Wang W, Wu Y, Yan Y, Ermakova M, Kerstetter R, Messing J. DNA barcoding of the Lemnaceae, a family of aquatic monocots. *BMC Plant Biol.* 2010;10:205.

Yamaguchi A, Kawamura H, Horiguchi T. A further phylogenetic study of the heterotrophic dinoflagellate genus, *Protoperidinium* (Dinophyceae) based on small and large subunit ribosomal RNA gene sequences. *Phycol Res.* 2006;54:317-29.

Yang JB, Li DZ, Li HT. Highly effective sequencing whole chloroplast genomes of angiosperms by nine novel universal primer pairs. *Mol Ecol Res.* 2014;14:1024-31.

Yang JB, Wang YP, Moller M, Gao LM, Wu D. Applying plant DNA barcodes to identify species of *Parnassia* (Parnassiaceae). *Mol Ecol Res.* 2012;12(2):267-75.

Zhu X, Meyer L, Gopurenko D, Weston PA, Gurr GM, Callaway RM, et al. Selection of DNA barcoding regions for identification and genetic analysis of two *Echium* invaders in Australia: *E. plantagineum* and *E. vulgare*. In: Conference: Nineteenth Australasian Weeds Conference, At Hobart, Tasmania: 2014.