# Single Nucleotide Polymorphisms in the *AB096061* Gene of *Phragmites australis*

Submitted by
Brianne A. Gryspeerd

Biology

To
The Honors College
Oakland University

In partial fulfillment of the
Requirement to graduate from
The Honors College

Mentor: Dr. Douglas Wendell, Professor of Biology
Department of Biological Sciences
Oakland University

April 2, 2021

Abstract

*Phragmites australis* exists in two primary strains in North America: native and introduced. Within the *AB096061* gene of both strains, nine single-nucleotide polymorphisms (SNPs) of interest have been noted in Midwestern and Canadian leaf samples, but previous findings might have been indicative of local SNPs. In the following research project, the geographic sampling range was increased to test native, invasive, and hybrid *Phragmites* samples from the Western United States. DNA from Utah, Nevada, and California samples were isolated, prepared, and sequenced according to previous protocols. The sequencing data and results from the project continued to prove the indicative SNPs exist uniformly in native and invasive strains from Western samples. The project also found one new SNP within the *AB09061* gene and showed the higher prevalence of an additional polymorphism in the Western samples. By continuing to distinguish and confirm the differences between strains, the introduced strain can be properly managed by applying information and results to expand and refine control methods.

## Introduction

*Phragmites australis,* hereafter referred to as *Phragmites,* is a wetland grass species which can be found in all U.S. states. According to the U.S. Fish and Wildlife Service (2007), *Phragmites* is a tall, perennial plant with fluffy seedheads. The water reed grows and spreads easily by above-ground seeds or by rhizomes (an underground network of projections, which spread to create new above-ground stems). *Phragmites* is a grass species found on multiple continents, but when a strain from one area is introduced to another, the introduced strain has the potential to become an invasive species.

An invasive species is an introduced organism that may cause harm to native ecosystems. Native *Phragmites* does not grow as dense as introduced *Phragmites* and it grows and allows for species richness in an ecosystem (Bourke & Bickford, 2015). However, introduced *Phragmites* chokes out native *Phragmites* as well as other native plants, leading to decreased species richness. Introduced *Phragmites* is also easily dispersed and thrives in disturbance-rich areas, such as ditches along roads (U.S. Fish and Wildlife Service, 2007). Due to these adaptations, introduced *Phragmites* can easily invade and harm native wetlands. This, paired with the ability to spread by underground rhizomes, creates many issues for controlling the spread of the invading plant.

In the United States, two main strains of *Phragmites* exist in the same habitats: the native, North American strain and the introduced, European strain. These two strains of *Phragmites* are distinct physically and genetically. With respect to physical appearance, the native plant has yellower leaves and red-purple stems (Bourke & Bickford, 2015). The introduced *Phragmites* have bluish leaves and dull green stems (Bourke & Bickford, 2015). Despite these variances, the plants are still difficult to distinguish, so genetic differences allow definite identification by

identifying DNA polymorphisms at the *trnL* and *rbcL* gene loci (Saltonstall *et al.,* 2004). Both strains have these genetic loci, but the DNA sequences are unique to each strain. Therefore, these instances in the genomic sequence are ideal locations for identification. In addition to physical and genetic differences, the two *Phragmites* strains grow differently- the introduced *Phragmites* grow in large, tall, and dense patches with low species diversity (Bourke & Bickford, 2015). This feature lends to the introduced strains' competitive nature, which allows the plant to outcompete its native counterpart, leading to the introduced strains' status as an invasive species.

Common methods of eliminating introduced *Phragmites* include applying herbicides, burning, and cutting patches (U.S. Fish and Wildlife Service, 2007). Plant specific herbicides can target and kill *Phragmites*. Burning and cutting or mowing *Phragmites* is another way to remove the plant, however, this method should be used with caution due to the underground rhizomes. If rhizomes are left after burning or cutting a patch, the plant will easily regrow. Using any of these methods to remove *Phragmites* should begin with proper identification of the strain. The native and introduced strains need to be identified prior to removal in order to promote conservation of the native plant.

When focusing on eliminating only the introduced *Phragmites*, there are three well researched experimental methods. In one method, biological control, an introduced insect species (such as a European moth) is used to naturally feed on the plants, but this poses the adverse effect of the insect feeding on native plants as well (Kiviat *et al.,* 2019). A second method, microbial control, involves targeting the symbiotic relationships between endophytes and the *Phragmites* plants. Differences between microbe communities would be sources of regulation; however, the two strains' communities have been thoroughly researched to show no significant

differences (Bickford *et al.,* 2018). Therefore, the method of gene silencing, or silencing genes for essential functions, holds promise due to genetic differences between strains.

Gene silencing is a concept where genes are silenced so essential functions are disrupted in the plant. Potentially, this control method is specific to the species, and could target only introduced *Phragmites*. In research conducted by Onylio *et al*. (2018), gene silencing was used to control and lower the pathogenicity of a harmful fungus. The same technique has been used in virus induced gene-silencing to target plant genes and the production of proteins (Gouil *et al.,* 2015). In the case of *Phragmites*, similar research methods will be applied to find varying genes of interest. Research initially begins with identification of essential SNPs in central genes in order to develop this form of control. When the SNPs of interest are found, similar induced gene silencing methods like those utilized in the Gouil *et al*. 2015 study could be applied to silence the vital genes of introduced *Phragmites*.

When identifying genetic differences between native and introduced *Phragmites*, one method is to analyze the single nucleotide polymorphisms, or SNPs in chloroplast DNA. SNPs are a genetic variation where only one nucleotide in the genetic code is different (Genetics Home Reference, 2020). These differences in the genetic code are due to mutations, which can sometimes dramatically alter DNA. The SNPs used to identify native and introduced *Phragmites* are often point mutations, where one nucleotide is substituted for another. Depending on the base substitution, the SNP may or may not have an effect on the organism (Genetics Home Reference, 2020). Yet, these subtle differences in gene sequences show evidence of where the native and introduced strains diverged and their evolution. The genomic sequence of one strain typically has one nucleotide while the other strain has a different nucleotide at SNPs of interest. In order to identify SNPs in *Phragmites* DNA, mitochondrial DNA is isolated, PCR amplified, primed, and

the DNA area of interest is sequenced using overlapping portions of DNA chosen by specific start and end primers. The SNPs are found by comparing the sequenced DNA to a reference DNA sequence.

Like SNPs, microsatellites are a type of genetic variation that holds value in genotyping by using another form of DNA identification technology. Microsatellites, or simple sequence tandem repeats (SSTRs), are used in plant population studies because they are nuclear DNA sequences that can be used to estimate gene flow to help analyze evolution of mutations (Vieira *et al*., 2016). While the SNPs in *Phragmites* identification uses chloroplast DNA, microsatellites offer the use of identification using nuclear DNA. To identify microsatellites in DNA, DNA libraries have been constructed by fragmenting DNA, ligating to adaptors and inserting the DNA into bacterial vectors (Vieira *et al.,* 2016). Due to this course of amplification then resultant sequencing, only nuclear DNA is used rather than mitochondrial DNA. Microsatellites are more useful in cases of identifying plant DNA sequences because nuclear DNA includes DNA from both parent plants whereas mitochondrial DNA only includes DNA from the maternal plant.

Microsatellites and nuclear DNA identification are particularly useful due to the phenomenon of *Phragmites* hybridization- or the instance where the native and introduced strains produce a genetic hybrid by interbreeding between the two populations. In these occurrences, the maternal and paternal plants are different strains- one is introduced, and one is a native plant. Therefore, if SNPs were used to identify hybrids, the use of mitochondrial DNA would only give indication of which plant strain was the maternal plant. Gene flow between native and introduced *Phragmites* has been researched using microsatellite technology, with results showing hybridization occurs very rarely (Saltonstall, 2003). This is due in part to the fact that introduced *Phragmites* will outcompete native *Phragmites* and there is no opportunity for

interbreeding. Therefore, most native or introduced samples will be genetically pure, with no hybridization, with particular purity in the North American native strain (Saltonstall, 2003).

Although the instances of hybridization between native and introduced *Phragmites* are rare, it has occurred. This phenomenon must be carefully watched and considered due to the possible consequences of hybridization. Gene flow between two populations can create a new, hybrid population with much higher fitness than that of the parents (Todesco *et al.*, 2016). If hybrid offspring of *Phragmites* grew in thicker patches with higher fitness, both the native and the introduced strains could eventually go extinct. Hybrids hold the potential to create an unprecedented, more invasive strain of *Phragmites*, posing a larger risk to the already dwindling native *Phragmites* population. Extinction of parental strains in this way is known as genetic swamping, and in one study, in 69 of 143 cases of hybridization, genetic swamping caused extinction (Todesco *et al.*, 2016). Due to these risks, while identifying different strains of *Phragmites*, it is critical to use nuclear DNA to find, record, and potentially eliminate hybrid plants.

Current research on *Phragmites* hybridization started in an experiment. Plants were manually cross-pollinated between the strains and microsatellite markers were analyzed to prove hybridization (Meyerson *et al*., 2010). The results of their study showed indication of an aggressive hybrid that would further decline native *Phragmites* populations. Natural hybrids have been identified in the field at locations such as Seneca Falls, NY (Saltonstall, 2014). Additionally, hybrids have been found in the Las Vegas Wash watershed (Saltonstall *et al.,* 2016). In these natural, documented cases of hybridization, it was noted that disturbances, flooding, and the spread of *Phragmites* by rhizomes may increase the rare instances of natural hybridization.

All strains of *Phragmites* share evolutionarily conserved genes for essential proteins. For example, the high-affinity nitrate transporter gene (*NRT2*) is present in native, introduced, and hybrid strains with few genetic differences when analyzed according to Genbank accession no. *AB096061* sequence (Araki *et al.,* 2005). The *NRT2* encoded transporter protein is essential for nitrate uptake by the roots, and many mutations would render this protein nonfunctional (Araki *et al.,* 2005). However, some SNPs and single base heterozygosities exist between the strains that do not effect *NRT2's* function. Instead, these SNPs in the *AB096061* gene are indicative of evolutionary changes as the native and introduced strains diverged. At a few promising locations in the *AB096061* gene, all native samples had one SNP while all introduced samples had a different base pair. However, this information was limited to *Phragmites* sampled across a geographic area limited to the United States Midwest, United States Northeast, and Eastern portions of Canada.

Since the SNPs in the *AB096061* gene may be indicative of a smaller sample size, this study focused one expanding the sampled region to include native and introduced *Phragmites* samples from the Western United States. A hybrid plant was also studied for comparison. All samples will be genotyped according to native or invasive identification protocols and amplicons will be produced and isolated from the *AB096061* gene for sequencing. Sequencing results may verify or contradict the SNPs found in the gene to date.

## Methodology

### Sample Collection:

For the following protocols, samples were first obtained from various locations across the United States with a focus on the Western states Utah, Nevada, and California. Each sender was asked to cut a small part of a *Phragmites* leaf off a growing plant. Samples, some from

previously published *Phragmites* studies, were of green tissue and packaged individually in envelopes with information on the sample collection site and sent to Oakland University for analysis. Eleven samples were sent for the project: two native and two introduced samples each from Utah and California and one native, one introduced, and one hybrid from Nevada (Lambert *et al*., 2016).

Mixer Mill DNA Purification:

DNA purification protocol was based on Machery-Nagel Nucleospin Plant II DNA purification kit instructions with project modifications. Between 40 and 50 mg of green leaf tissue was used and placed in Mixer Mill tubes with a steel Mixer Mill ball. Each sample was shaken in the Mixer Mill machine at maximum frequency for 30 minutes or until completely powdered.

1000 µl of lysis buffer (PL1) and 10 µl of RNAse were added to the tubes and agitated until mixed. The tubes were incubated in a 65°C water bath for twenty minutes with periodic inversions to suspend the sample. Next, the samples were divided, and homogenate collected with the kit's Filter Tubes (centrifuge for four minutes at 11,000 x g). After collecting homogenate, 225 µl of binding buffer (Buffer PC) was added and mixed to every 200 µl homogenate. In 600 µl additions, the mixture was centrifuged (one minute at 11,000 x g) through the kit's Nucleospin Plant II column and the flow through discarded. The top of the Nucleospin Plant II column was washed once with 400 µl of wash buffer 1 (Buffer PW1) and then with 700 µl of wash buffer 2 (Buffer PW2); both centrifuged for one minute at 11,000 x g and the flow through discarded. Lastly, the columns were placed in a new 1.5 mL microcentrifuge tube,100 µl of elution buffer (Buffer PE) was added, and the assemblies were incubated at 65°C for five

minutes. The tubes were centrifuged (one minute at 11,000 x g) and the flow through saved and stored in the freezer.

The final flow through contained the DNA, and concentrations were confirmed and recorded using Nanodrop analysis.

Genotyping *Phragmites* (Differentiate Native and Introduced):

Polymerase chain reaction restriction fragment length polymorphism (PCR- RFLP) assay was used to differentiate native and introduced *Phragmites* strains according to a published identification protocol (Saltonstall, 2003). All samples were genotyped regardless of if the strain type was known prior to analysis. This assay utilized non-coding chloroplast DNA for genotyping according to the *rbcL* and *trnL* gene loci. With the *trnL* loci, native stains allows one digest (RsaI digest) to cut the DNA while the gene sequence of the introduced strain will not be cut by the same digest, allowing the difference to be seen during gel electrophoresis. When using a different digest (HhaI digest) on the *rbcL* loci, the introduced will be cut while the native is not.

Genotyping *Phragmites* (Differentiate Hybrids):

One sample sent to the project was a known hybrid used in previous research studies. This was verified by genotyping the hybrid sample and a known native and introduced sample by an indel assay in the *AB096061* gene. First, DNA of the *AB096061* gene is amplified using polymerase chain reaction (PCR). PCR reactions contained 3 µl DNA, 10 µl 2X Green Master Mix, 3 µl Mili-Q water, 2 µl of AB096061F primer (10 µM), and 2 µl of AB096061R primer (10 µM) in a total volume of 20 µl. The thermal cycling profile (‘PHRAG64’) was 2 min. at 94° followed by 30 cycles of 94° for 45 sec., 64° for 45 sec., 72° for one min., followed by a final extension of 72° for 2 min.

Following PCR, the products went restriction digest. 8 µl of PCR product are digested with 2 µl of restriction enzyme *XapI* with 13 µl of Milli-Q water and 2 µl of 10X Fast Digest Green Buffer. The digest is incubated in a 37°C water bath for 20 minutes. 20 µl of restriction fragments underwent gel electrophoresis in a 1.0 % agarose gel using SYBR Safe stain at 150 V for 30 min. and finally visualized under UV light.

For result interpretation, no cut producing only one band of about 1700 bp indicated the sample has the *AB096061* indel occurs in introduced samples. A cut sample, producing two bands, indicates the sample does not have the AB096061 indel and occurs in native samples. Hybrids will show three bands (an uncut band and two cut bands) all at the same intensity due to their heterozygosity.

Production of Amplicons of *AB096061*:

Amplicons of the *AB096061* gene were produced for use in sequencing. PCR began with a preparation of 200 ng of sample DNA, 34-x µl DNA of Milli-Q water, 50 µl 2X Green Master Mix, 8 µl of *AB096061-F* primer (10 µM), and 8 µl of *AB09606-R* primer (10 µM) (primer sequences given in Table 1) for a total volume of 100 µl. This preparation is divided equally into three 33 µl reactions. The thermal cycling profile PHRAG64 as previously described under 'Genotyping *Phragmites* (Differentiate Hybrids)' was used.

Following PCR, the three 33 µl reaction products were pooled back together and purified according to GeneJET Kit instructions. The purified products were then quantified by an assay of DNA concentration using the Qubit DNA Broad Range Assay Kit. Lastly, the amplicons are analyzed for quality control to verify fragment purity. According to Qubit DNA concentrations, 25 ng, 2 µl of loading dye, and a loadable volume of TE buffer are loaded and electrophoresed in

a 1.0 % agarose gel using SYBR Safe stain at 150 V for 30 min. Amplicons producing only one band indicate purity required for sequencing.

Sequencing:

Each sample was prepared for sequencing according to GeneWiz instructions. The six *AB096061* primers used in the reactions were *AB096061-F, 374F, 967F, 533R, 1035R*, and *AB096061-R* (See Table 1 for Primer Sequences).

**Table 1**

AB096061 *Primers for PCR and DNA Sequencing*

| Primer | Sequence | Use |
|---|---|---|
| *AB096061-F* | AAGACTCGAGAGGCCAGCTA | PCR, sequencing |
| *374F* | CGTCTTCTGCATGTCCCTCA | Sequencing |
| *967F* | CTACTACGACCACTTCGACCTA | Sequencing |
| *533R* | GCCAGAGAGAAGCCGATCAA | Sequencing |
| *1035R* | TTGGCCATTCCGAAGCAA | Sequencing |
| *AB096061-R* | ACGTGCGTCTTATACGTGCT | PCR, sequencing |

Sequence Assembly:

SnapGene Viewer program was used to analyze raw sequence data from GeneWiz. After sequencing, each sample had six traces. First, raw data was checked for heterozygosity (see 'Identifying Heterozygosity') and the appropriate base pairs were annotated. Raw base sequence data was rerecorded according to FASTA format. Then, using CAP3 online software, sequence contigs were assembled (http://doua.prabi.fr/software/cap3). The sequence contigs were then aligned and compared to a native *Phragmites* reference sequence (Figure 1) using Emboss Water online software (https://www.ebi.ac.uk/Tools/psa/emboss_water/). All polymorphisms (SNPs, insertions or deletions, and heterozygous bases) were recorded and compared to previous polymorphic data of the *AB096061* gene.

**Figure 1**

*Native Reference Sequence*

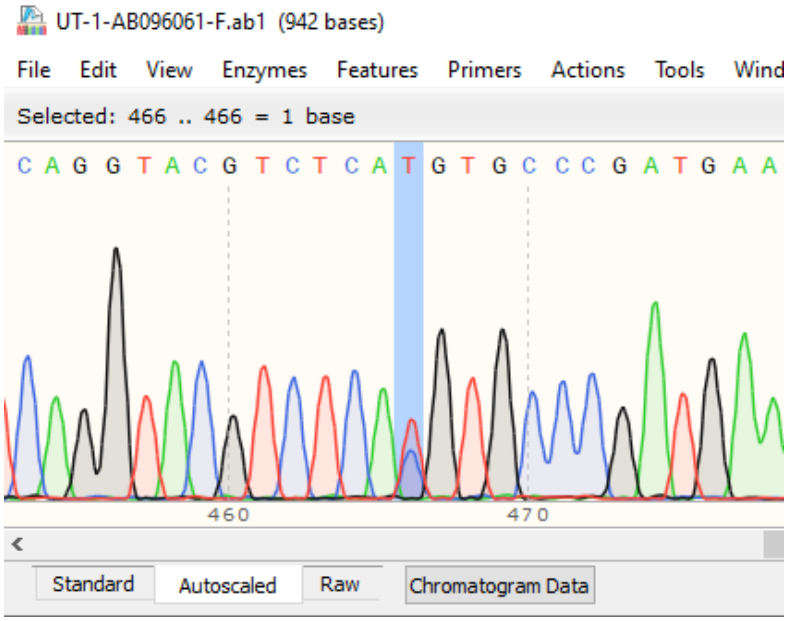
>Nat 11-001-AB096061 Contig
AGAGGCCAGCTATGGAGGTTGGCGCGCCGGGGAGCTCGCTGCACGGCGTCACGGGGCGCG
AGCCGGCGTTCGCCTTCTCGACGTCTGCGGTGCCCGACGATGTTGCAGCGAGCAAGTTCG
ACCTGCCGGTGGACTCGGAGCACAAGGCCAAGAGCATCCGGCTCTTCTCCTTCGCCAACC
CGCACATGCGCACCTTCCATCTCTCGTGGATCTCCTTCTTCACCTGCTTCGTCTCCACCT
TCGCCGCCGCGCCGCTCGTCCCCATCATCCGCGACAACCTCAACCTCACCAAGGCCGACA
TCGGCAACGCCGGCGTGGCCTCCGTCTCGGGATCCATCTTCTCCCGTCTCGCCATGGGCG
CCATCTGCGACCTCCTCGGCCCGCGCTACGGCTGCGCCTTCCTCATCATGCTCACGGCGC
CCACCGTCTTCTGCATGTCCCTCATCGACGACGCAGCCGGCTACATCGTCGTCAGGTACG
TCTCACGTGCCCGATGAAATTTATATGTACATTGCAACATAATTACTGAACGTCGTTGCC
GCCATGTTGCATGCATGCAGGTTCTTGATCGGCTTCTCTCTGGCCACGTTCGTGTCGTGC
CAGTACTGGATGAGCACCATGTTCAATAGCAAGATCATCGGGACGGTGAACGGGCTGGCG
GCCGGGTGGGGGAACATGGGCGGCGGCGCCACGCAGCTCCTCATGCCGCTCGTCTACGAC
GTCATCCGCAAGTGCGGCGCGACGCCCTTCACGGCGTGGCGCATCGCCTACTTCGTGCCG
GGGCTAATGCACGTGGTGATGGGCATCCTGGTGCTCACGCTGGGGCAGGACCTCCCCGAC
GGCAACCTCAGGAGCCTCCAGAAGAAGGGCGACGCCAACAAGGACAAGTTCTCCAAGGTG
CTCTGGTACGCCGTCACCAACTATCGCACCTGGATCTTCGTGCTGCTCTACGGCTACTGC
ATGGGCGTGGAGCTTACCACCGACAACGTCATCGCCGAGTACTACTACGACCACTTCGAC
CTAGACCTCCGCGTCGCCGGCATCATCGCCGCTTGCTTCGGAATGGCCAACATCGTGGCA
CGGCCCTTGGGCGGCATCCTCTCCGACATCGGCGCGCGCTACTGGGGCATGCGCGCGCGC
CTCTGGAACATCTGGATCCTCCAGACTGCTGGCGGCGCCTTCTGCCTCTGGCTCGGCCGC
GCCAGCACGCTTCCTGCCTCCATTACCGCCATGGTGCTCTTCTCCTTCTGCGCCCAGGCC
GCCTGCGGCGCCATTTTCGGCGTCACCCCTTTCGTCACCCGCCGCTCCCTCGGCATCATG
TCCGGGATGACGGGGGCTGGCGGCAACTTCGGCGCGGGGCTCACGCAGCTGCTCTTCTTC
ACCTCATCCAAGTACTCCACGGGCACGGGGCTGGAGTACATGGGCATCATGATCATGGCG
TGCACGCTTCCAGTGGTGTTCGTGCACTTCCCGCAGTGGGGATCCATGCTCCTCCCGCCC
AGCGCCGGCGCCGTCGAGGAGCACTACTACAACTCGGAGTGGAGTGAGGAGGAGAAGAGC
AAGGGGCTCCACAGCGCCAGCCTCAAGTTTTCAGAGAACTGCCGTTCAGAGCGCGGCAAC
CGCAACGTCATCCTCGCGGCACCAAACAGCACGCCCGA

Identifying Heterozygosity:

Heterozygous base pairs were identified according to the raw chromatogram sequence data in SnapGene Viewer. To identify these points, the entire sequence was analyzed for any chromatogram base positions where there are two peaks (Figure 2). These double peaks are classified as heterozygous bases. These sequencing data is then edited to include a letter indicating the type of heterozygosity at the sequence location: K for G or T, M for A or C, R for A or G, S for C or G, W for A or T, and Y for C or T.

**Figure 2**

*Heterozygosity in UT-1*




*Note*. Heterozygous base pair is highlighted. The two bases at this location are T and C, therefore the annotated base is edited to Y.

## Results

Sample Genotype Determination:

Eleven samples were sent to the project and given identification labels according to location as outlined in Table 2 and pictured in Figure 3. Each sample was genotyped as invasive or native as follows. Figure 4 and Figure 5 shows one of the electrophoresis results with bands showing how native and introduced haplotypes were distinguished.

In the case of sample UT-2, Figure 4 shows one uncut DNA band for the *RsaI* digest at the *trnL* gene locus. In Figure 5, UT-2 shows two bands for the *HhaI* digest at the *rbcL* locus, indicative of a cut in the DNA. These results indicated UT-2 is of the introduced haplotype.

**Table 2**

Phragmites *Samples and Haplotype*

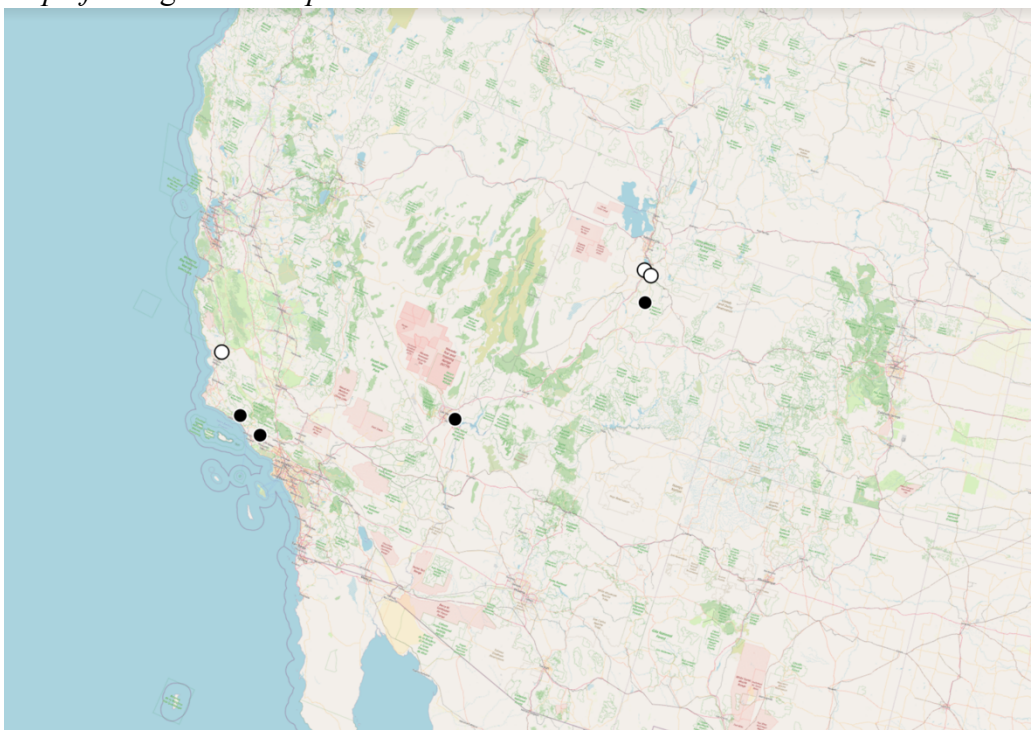
| Sample ID | Location | Haplotype |
|---|---|---|
| UT-1 | 40.142113, -111.802797 | Introduced |
| UT-2 | 40.082258, -111.602418 | Introduced |

| | | |
|---|---|---|
| UT-3 | 39.546634, -111.516889 | Native |
| UT-4 | Utah: exact location unknown | Native |
| CA-1 | 34.3553, -119.0064 | Native |
| CA-2 | 34.5403, -119.6200 | Introduced**** |
| CA-3 | 34.5339, -119.5639 | Native |
| CA-4 | 35.4993, -120.6520 | Introduced |
| NV-2 | 36.7938, -114.0828 | Native |
| NV-3 | 36.1220, -114.9045 | Native |
| NV-4 | 36.0932, -115.0152 | Native*** |

***NV-4 initial haplotype analysis indicates the sample as native, but this is a hybrid plant.
****CA-2 was of confirmed native strain from previous studies. Since this study found the sent sample to be contradictory and of introduced strain, the results from CA-2 were excluded from the study.

**Figure 3**

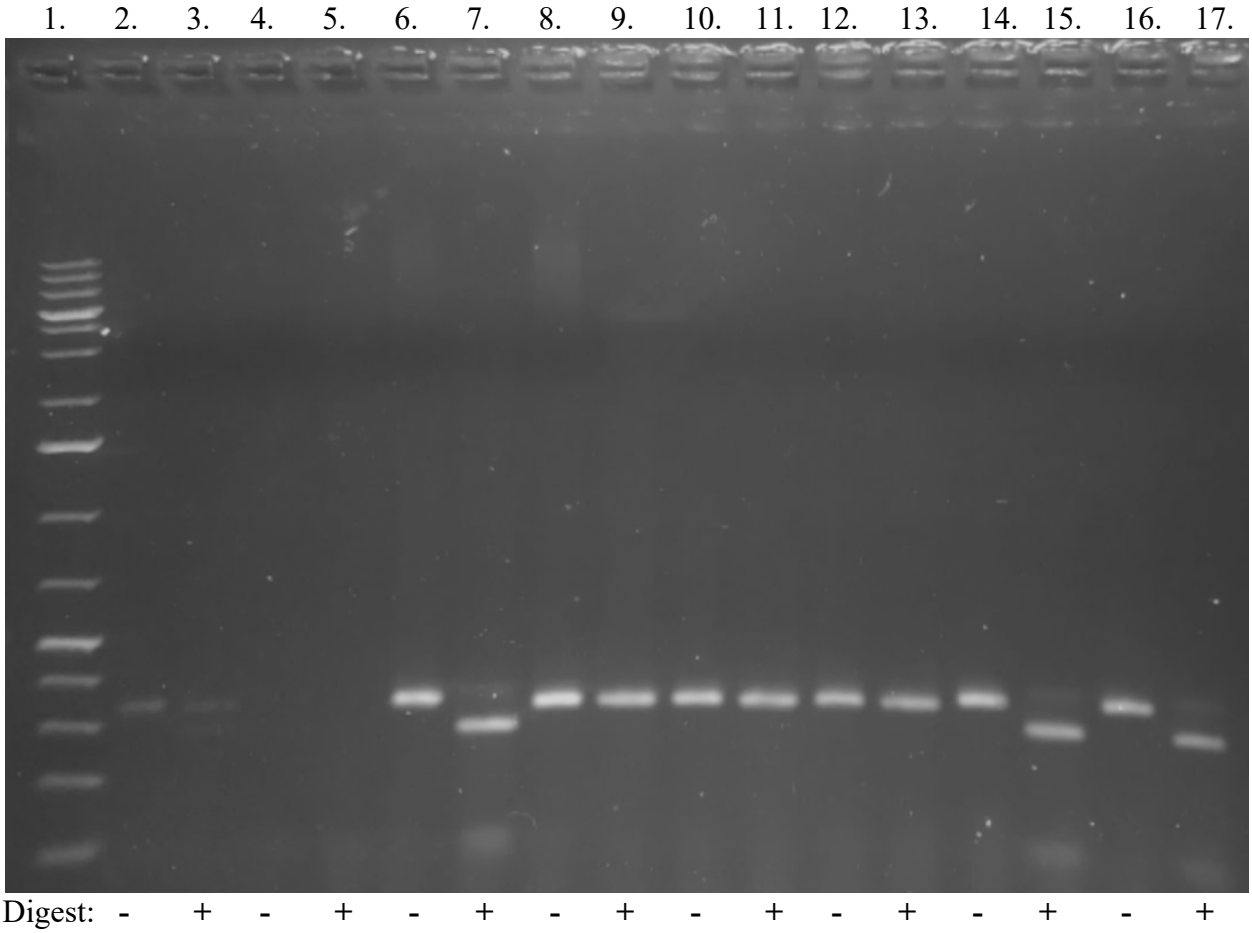
*Map of* Phragmites *Sample Locations*



*Note*. White dots label locations of introduced samples and black dots label locations of native samples. UT-4 is not included because sample location was not precise enough. CA-2 is not included because it was overall excluded from the study.

**Figure 4**

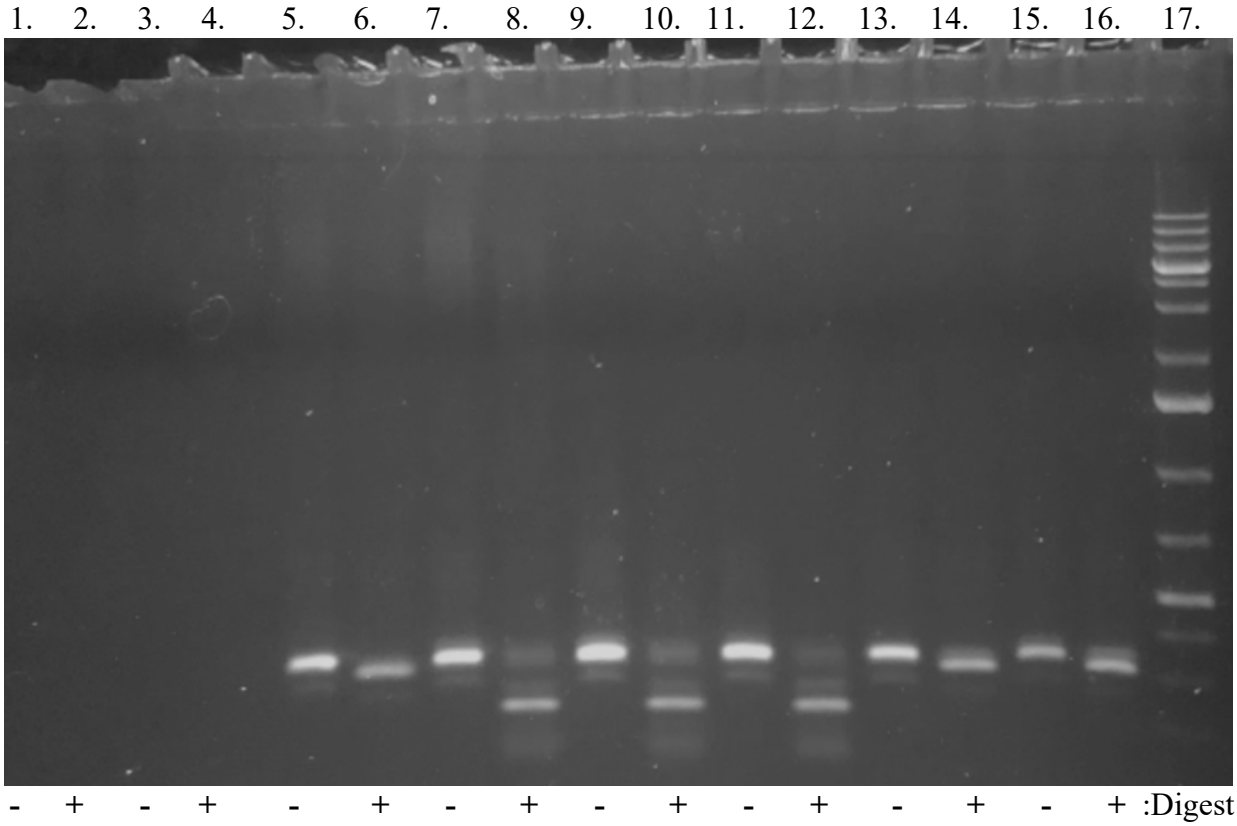
trnL *Locus Gel of Utah Samples With* RsaI *digest*




*Note*. Well contents as follows: 1. 1000 Kb DNA ladder, 2. and 3. preparation blank***, 4. and 5. no template control, 6. and 7. known native sample, 8. and 9. known introduced sample, 10. and 11. UT-1, 12. and 13. UT-2, 14. and 15. UT-3, and 16. and 17. UT-4.
*** Faint bands in wells 2. and 3. indicate some contamination of the blank sample prepared with all *Phragmites* samples during DNA extraction and purification.

**Figure 5**

rbcL *Locus Gel of Utah Samples With* HhaI *Digest*




*Note*. Well contents as follows: 1. and 2. preparation blank, 3. and 4. no template control, 5. and 6. known native sample, 7. and 8. known introduced sample, 9. and 10. UT-1, 11. and 12. UT-2, 13. and 14. UT-3, 15. and 16. UT-4, and 17. 1000 Kb DNA ladder.

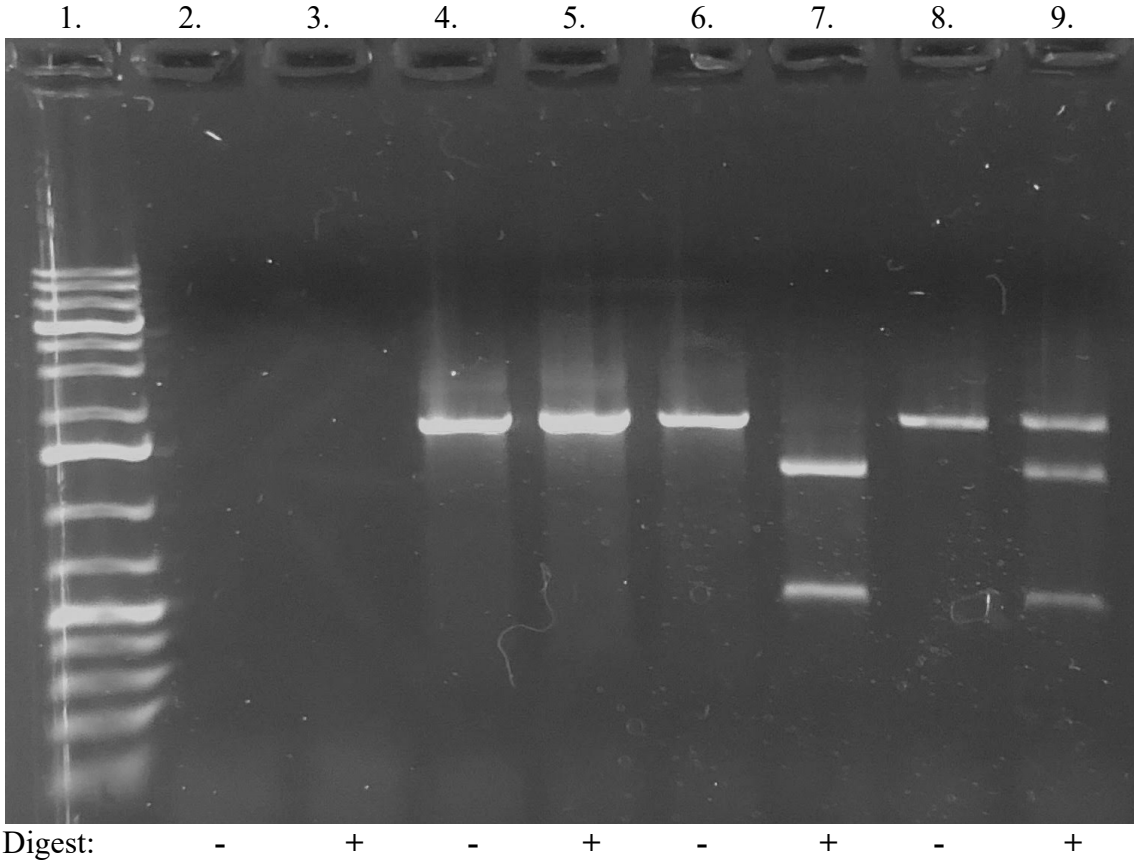
Hybrid Genotyping:

NV-4 was a confirmed hybrid sent from researchers with previous *Phragmites* work. The genotyping according to Saltonstall's 2003 RFLP assay only used chloroplast DNA, therefore the assay would not indicate hybridization. Hybridization determination requires the nuclear DNA from both parental strains. In the first RFLP assay of the *trnL* and *rbcL* loci, NV-4 indicated it was a native plant. Indel assay of the *AB096061* gene was required and conducted to confirm the suspected hybrid. Figure 6 shows the electrophoresis results of the indel assay,

where NV-4 digest created three distinct bands: two cut and one uncut. Results indicate NV-4 is of the hybrid strain. Since the first RFLP assay on chloroplast DNA indicated NV-4 as a native plant, it can be concluded that the maternal plant of NV-4 was of the native strain.

**Figure 6**

AB096061 *Indel Assay With* XapI *Digest*




*Note*. Well contents as follows: 1. 1000 Kb DNA ladder, 2. and 3. no template control, 4. and 5. UT-1 (invasive), 6. and 7. UT-4 (native), and 8. And 9. NV-4 (hybrid).

Sequence Assembly:

Raw sequence data from GeneWiz for the ten native and introduced samples indicated the following polymorphisms when compared to the native *Phragmites* reference sequence

(Table 3). Heterozygosity was common in introduced *Phragmites* samples, but not in native samples.

Two of the polymorphisms are deletions in the introduced strain. At reference position 499-502, all native *Phragmites* samples contain the sequence ATTT while the introduced samples have a four base pair deletion. Likewise, at reference position 531-533, all native samples contain the sequence CGT while introduced samples have a three base pair deletion.

The other seven polymorphisms of interest are the single nucleotide polymorphisms. The SNPs at each reference position are as follows, with the base in the native strain listed first followed by the base in the introduced: 101 T to C, 103 T to A, 975 T to G, 1287 C to T, 1452 A to C, 1508 G to A, and 1532 A to G.

**Table 3**

*Single Nucleotide Polymorphisms in Native and Introduced Samples*

| ←Introduced | | | | | Native→ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| UT-1 | UT-2 | CA-2 | CA-4 | Reference Position | UT-3 | UT-4 | CA-1 | CA-3 | NV-2 | NV-3 |
| C | C | C | C | **101** | T | T | T | T | T | T |
| A | A | A | A | **103** | T | T | T | T | T | T |
| C | C | C | C | 260 | C | C | C | C | C | C |
| C | C | C/T | C/T | 486 | C | C | C | C | C | C |
| ---- | ---- | ---- | ---- | **499-502** | ATTT | ATTT | ATTT | ATTT | ATTT | ATTT |
| C | G/C | G/C | G/C | 514 | G | G | G | G | G | G |
| --- | --- | --- | --- | **531-533** | CGT | CGT | CGT | CGT | CGT | CGT |
| C | C | T/C | T/C | 536 | T | T | T | T | T | T |
| C | C | T/C | T/C | 547 | T | T | T | T | T | T |
| T | T | T | T | 726 | C | C | C | C | C | C |
| G | G | C | C | 849 | C | C | C | C | C | C |
| G | G | G | G | **975** | T | T | T | T | T | T |
| C | C | C | C | 1179 | C | T | T | T | T | T |
| C | C | C/T | T | 1185 | C | C | C | C | C | C |
| T | T | T | T | **1287** | C | C | C | C | C | C |
| G | G | G | G | 1449 | T | G | G | A | G | G |
| C | C | C | C | **1452** | A | A | A | A | A | A |
| A | A | A | A | **1508** | G | G | G | G | G | G |
| G | G | G | G | **1532** | A | A | A | A | A | A |

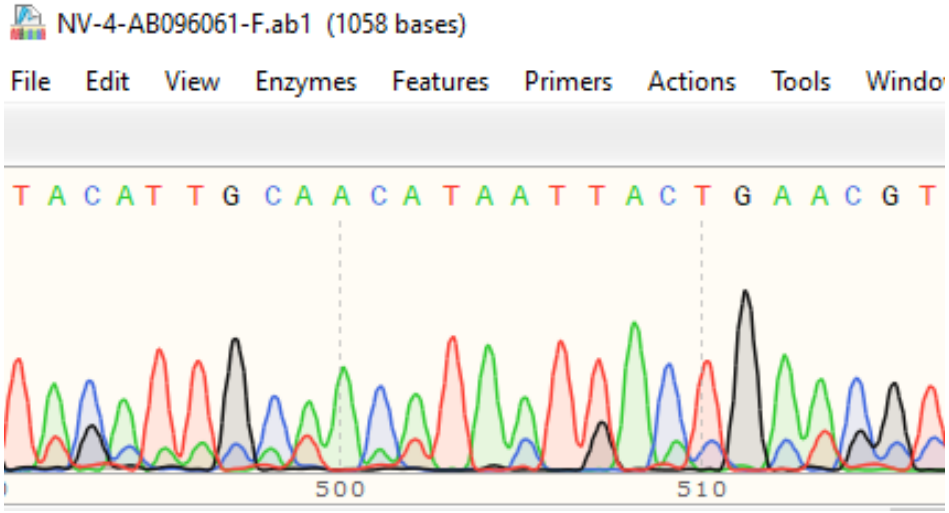
*Note*. Reference sequence found in Figure 1. Heterozygosity noted in UT-2 and CA-4. No heterozygosity noted in native samples. Although CA-2 was sequenced, as stated in the note under table 2, the CA-2 sample was considered invalid and results were not considered.

NV-4 (Hybrid) Sequencing Results:

Sequencing chromatograms showed multiple peaks (imitating heterozygous peaks) for nearly every sequenced base pair (Figure 6). This was due to the four base pair and three base pair deletions within the *AB096061* gene of its introduced sequence. Polymorphism identification as performed with native and introduced strains could not be conducted in the same manner.

**Figure 6**

*Chromatogram Data of Hybrid NV-4*




Therefore, to identify polymorphisms, a variant reference sequence preceding the reference position polymorphisms were used to identify location, and if the SNP was present, there would be a heterozygous chromatogram peak directly after the variant reference sequence. The nine reference positions of interest correspond to the same nine polymorphisms of interest that this research focused on. Table 4 shows this data and confirms whether the NV-4 hybrid contained the SNPs. By applying the reference sequences preceding the polymorphisms to hybrid data, it was found that all nine polymorphisms were present in the hybrid sample.

**Table 4**

NRT2 *Polymorphisms Reference Sequences and Polymorphism Appearance in NV-4*

| Reference Position | Variant Sequences and Polymorphisms [Native/Introduced] | Polymorphism Present in NV-4? |
|---|---|---|
| 101 | TGCCCGACGA[T/C]GTTGCAGCGA | Yes |
| 103 | CCCGACGATG[T/A]TGCAGCGAGC | Yes |
| 499-502 | GCCCGATGAA[ATTT/del]ATATGTACAT | Yes |
| 531-533 | AATTACTGAA[CGT/del]CGTTGCCGCC | Yes |
| 975 | GCGTGGAGCT[T/G]ACCACCGACA | Yes |
| 1287 | TCGGCGTCAC[C/T]CCTTTCGTCA | Yes |
| 1452 | GCACGCTGCC[A/C]GTGGTGTTCG | Yes |
| 1508 | CCCAGCGCCG[G/A]CGCCGTCGAG | Yes |
| 1532 | CACTACTACA[A/G]CTCGGAGTG | Yes |

*Note*. Reference sequence found in Figure 1.

## Discussion

Within the *AB096061* gene, there are nine polymorphisms of interest between the native and introduced *Phragmites* strains (reference positions are bolded within Table 3). At these points, every introduced sample has the same base and every native sample has the same base. These are the same nine locations in the gene where the polymorphisms continued to hold true for previously sampled locations. The hybrid sample, NV-4, also exhibited all nine polymorphisms, appearing as heterozygous locations at the polymorphism reference positions. This indicates the hybrid contains the polymorphisms from both the native and introduced *Phragmites* strains, as expected.

With respect to general heterozygosity in nonhybrid plants, the previous findings where heterozygotes in native samples were rare and heterozygotes in introduced samples rather common continued to hold true. None of the sequenced native samples showed any heterozygosity while three of the four introduced samples had heterozygous base pairs. The reference positions of heterozygosity in the samples were at 486, 514, 536, 547, and 1185. Again, continuous with previous findings, the points of heterozygosity continued to be at the same reference positions with the same base pairs at each location. Californian introduced samples showed the most heterozygosity and reference position 514 was the most common point

of heterozygosity. From the data, it is suggested that heterozygosity is more common in the Californian samples than in Utah samples, but even in this location, the points are not completely universal, and the data pool is most likely to small to fully draw this conclusion.

Beyond the conserved findings of universal SNPs, the sequence data on Western samples yielded a few unique results. For one, the SNP at reference position 1179 was previously considered uncommon. At this position, most samples (both native and introduced) have a C at this position. Only some native samples had the unique T SNP. However, in the *Phragmites* samples from the Western states, all native samples but one carried the T SNP. The SNP was not found in any introduced samples as well. This lends to the idea that the SNP at reference position 1179 may be more common than previously thought, and more common in some regional areas.

Additionally, one new SNP was found in the sample data. At reference position 1449, most samples contain either a G or T. CA-3's sequencing data showed a new SNP at this location: an A. This base substitution SNP has not been recorded before within the *AB096061* gene. Yet, this SNP is still at a previously recorded reference position. Perhaps this SNP is a regional mutation in California, but it is interesting that the other native California sample, CA-1, does not share this difference.

The most promising finding from the data is continued sampling from a larger area still upheld the same SNP patterns of the nine universal SNPs. There were no changes from previous data at these locations, continuing to suggest the SNPs as very old in strain divergence. As these SNPs continue to be present within the gene, they can eventually be used for native and introduced identification, studying the SNP effect on *NRT2* function, and even introduced strain control within the United States.

Additionally, sampling from a larger range still found the same points of heterozygosity. Western samples had the same SNP locations for all variations- there were no new SNP reference positions within the *AB096061* gene, signifying any variations between native and introduced samples diverged long before the strains spread across North America.

It is both interesting and promising the nine polymorphisms continue through a wider range of sampling. This gives a better picture of the evolutionary history and divergence of *Phragmites* strains. Future analysis of samples from more geographic areas may be required to confirm the SNP conclusions, including sampling in Southern and Central United States. Additionally, more sampling and studies with hybrid *Phragmites* plants should be added to the current repertoire of knowledge. There are many additional routes to take to continue to study the *AB096061* gene and the various strains of the *Phragmites* plant in order to understand more about its evolution and eventually methods for the introduced strain's control.

## Additional Notes

Special thank you to Lyle Bingham and Adam Lambert for their contribution of *Phragmites* samples to this project.

References

Araki, R., Mori, M., Mori, M., & Hasegawa, H. (2005). Genetic Differences in Nitrate Uptake in Two Clones of the Common Reed, *Phragmites australis. Breeding Science,* 55, 297-302.

Bickford, W. A., Goldberg, D. E., Kowalski, K. P., & Zak, D. R. (2018). Root endophytes and invasiveness: no difference between native and non-native *Phragmites* in the Great Lakes Region. *Ecosphere*, *9*(12). doi: 10.1002/ecs2.2526

Bourke, K. & Bickford, W. (2015). *Native vs Non-Native Phragmites*. Great Lakes Phragmites. https://www.greatlakesPhragmites.net/blog/native-vs-invasive-Phragmites/

Genetics Home Reference. (2020). *What Are Single Nucleotide Polymorphisms (SNPs)?*. Retrieved from: https://ghr.nlm.nih.gov/primer/genomicresearch/snp

Gouil, Q., Novák, O., & Baulcombe, D. (2015). SLTAB2 is the paramutated SULFUREA locus in tomato. *Journal of Experimental Botany*, *67*(9), 2655–2664. doi: 10.1101/033605

Kiviat, E., Meyerson, L. A., Mozdzer, T. J., Allen, W. J., Baldwin, A. H., Bhattarai, G. P., … Cronin, J. T. (2019). Evidence does not support the targeting of cryptic invaders at the subspecies level using classical biological control: the example of Phragmites. *Biological Invasions*, *21*(8), 2529–2541. doi: 10.1007/s10530-019-02014-9

Lambert, A., Saltonstall, K., Long, R., & Dudley, T. (2016). Biogeography of *Phragmites australis* lineages in the southwestern United States. *Biological Invasions*, *18*(9), 2597–2617. https://doi.org/10.1007/s10530-016-1164-8

Meyerson, L., Viola, D., & Brown, R. (2010). Hybridization of invasive *Phragmites australis* with a native subspecies in North America. *Biological Invasions*, 12, 103-111. doi: 10.1007/s10530-009-9434-3

Onyilo, F., Tusiime, G., Tripathi, J. N., Chen, L., Falk, B., Stergiopoulos, I., Tushemereirwe, W., Kubiriba, J. & Tripathi, L. (2018). Silencing of the Mitogen-Activated Protein Kinases (MAPK) Fus3 and Slt2 in *Pseudocercospora fijiensis* Reduces Growth and Virulence on Host Plants. *Frontiers in Plant Science*, *9*. doi: 10.3389/fpls.2018.00291

Saltonstall, K. (2003). A rapid method for identifying the origin of North American *Phragmites* populations using RFLP analysis. *Wetlands*, *23*(4), 1043–1047. https://doi.org/10.1672/0277-5212(2003)023[1043:ARMFIT]2.0.CO;2

Saltonstall, K. (2014). Hybrid Phragmites australis in North America. Great Lakes Phragmites. https://www.greatlakesPhragmites.net/blog/hybrid-Phragmites-australis-in-north-america/

Saltonstall, K., Lambert, A., & Rice, N. (2016). What happens in Vegas, better stay in Vegas: *Phragmites australis* hybrids in the Las Vegas Wash. *Biological Invasions,* 18, 2463-2474. doi: 10.1007/s10530-016-1167-5

Saltonstall, K. (2003). Microsatellite variation within and among North American lineages of *Phragmites australis*, *Molecular Ecology,* 12, 1689-1702. doi:10.1046/j.1365-294X.2003.01849.x

Saltonstall, K., Peterson, P., & Soreng, R. (2004). Recognition of *Phragmites australis* subsp. americanus (Poaceae: Arunidnoideae) in North America: evidence from morphological and genetic analyses. *SIDA, Contributions to Botany*, 21 (2), 683-692.

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., Heredia, S. M., Hahn, M. A., Caseys, C., Bock, D. G., & Rieseberg, L. H. (2016). Hybridization and extinction. *Evolutionary applications*, *9*(7), 892–908. https://doi.org/10.1111/eva.12367

U.S. Fish and Wildlife Service. (2007). *Phragmites: Questions and Answers.*

Retrieved from: https://www.fws.gov/gomcp/pdfs/Phragmitesqa_factsheet.pdf

Vieira, M. L., Santini, L., Diniz, A. L., & Munhoz, C. (2016). Microsatellite markers: what they mean and why they are so useful. *Genetics and molecular biology*, *39*(3), 312–328. https://doi.org/10.1590/1678-4685-GMB-2016-0027