


RESEARCH

Open Access



Variability in ITS1 and ITS2 sequences of historic herbaria and extant (fresh) *Phalaris* species (Poaceae)

Allison L. Graper¹, Andrzej K. Noyszewski^{1,2}, Neil O. Anderson^{1*}  and Alan G. Smith¹

Abstract

Background: *Phalaris* species (Poaceae) occupy diverse environments throughout all continents except Antarctica. *Phalaris arundinacea* is an important forage, ornamental, wetland restoration and biofuel crop grown globally as well as being a wetland invasive. The nuclear ribosomal internal transcribed spacer (ITS) region has been used for *Phalaris* barcoding as a DNA region with high nucleotide diversity for *Phalaris* species identification. Recent findings that *P. arundinacea* populations in Minnesota USA are most likely native and not European prompted this analysis to determine whether Eurasian vs. native North American *P. arundinacea* differed in ITS regions. Our objectives were to amplify and compare ITS regions (ITS1 and ITS2) of historic herbaria (1882–2001) and extant (fresh) *Phalaris* specimens; analyze ITS regions for species-specific polymorphisms (diagnostic SNPs) and compare ITS regions of historic *Phalaris* specimens with known, extant *Phalaris* species.

Results: We obtained complete ITS1 and ITS2 sequences from 31 *Phalaris* historic (herbaria samples, 1908 to 2001) and five extant (fresh) specimens. Herbaria *Phalaris* specimens did not produce new SNPs (single nucleotide polymorphisms) not present in extant specimens. Diagnostic SNPs were identified in 8/12 (66.6%) *Phalaris* species. This study demonstrates the use of herbaria tissue for barcoding as a means for improved species identification of *Phalaris* herbaria specimens. No significant correlation between specimen age and genomic DNA concentration was found. *Phalaris arundinacea* showed high SNP variation within its clade, with the North American being distinctly different than other USA and most Eurasian types, potentially allowing for future identification of specific SNPs to geographic origin.

Conclusions: While not as efficient as extant specimens to obtain DNA, *Phalaris* herbaria specimens can produce high quality ITS sequences to evaluate historic genetic resources and facilitate identification of new species-specific barcodes. No correlation between DNA concentration and age of historic samples (119 year range) occurred. Considerable polymorphism was exhibited in the *P. arundinacea* clade with several N. American accessions being distinct from Eurasian types. Further development of within species- and genus-specific barcodes could contribute to designing PCR primers for efficient and accurate identification of N. American *P. arundinacea*. Our finding of misidentified *Phalaris* species indicates the need to exercise stringent quality control measures on newly generated sequence data and to approach public sequence databases in a critical way.

Keywords: Herbarium specimens, DNA barcoding, PCR amplification, Diagnostic SNPs

Background

From the 1980s onwards, sequence-based molecular phylogenetic studies in plants relied primarily on plastid genome spacers and genes, particularly *rbcl*

*Correspondence: ander044@umn.edu

¹ Department of Horticultural Science, University of Minnesota, 1970 Folwell Avenue, Saint Paul, MN 55108, USA

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(ribulose-bisphosphate carboxylase in the chloroplast genome) [1]. Risks of using such uniparentally inherited sequences for phylogenetics necessitated the development and implementation of nuclear markers to reflect biparental trait inheritance [2]. The internal transcribed spacer (ITS) region of the nuclear ribosomal cistron, 18S–5.8S–26S, initially used by Baldwin's laboratory [3–5] quickly became the universally applied tool for molecular-based phylogenetic research for several reasons [6], not least of which included that 18S–26S rDNA arrays and their products are essential components of eukaryotic nucleolus organizing regions (NORs). Consequently, numerous plant families, genera and species have been analyzed for variance in ITS sequence differences, particularly for phylogenetic studies, species identification (barcoding) as well as ascertaining cultivar or genotype identities.

Such is the case with the genus *Phalaris* L. which is an important forage, ornamental, birdseed, wetland remediation/restoration and biofuel crop grown across the globe as well as being recognized as an invasive wetland species [7–9]. The genus *Phalaris* (Poaceae, grass family), classified in the Aveneae-Poeae section of the subfamily Pooideae, contains 20 species in the latest taxonomic treatises [10–15], although it previously included as many as 25 taxa. All *Phalaris* species in this monophyletic genus are cool-season grasses, with either annual or perennial life histories, from both the New and Old Worlds, varying in basic chromosome numbers of $x = 6$ and $x = 7$ and include a polyploid series from $2x$ to $8x$ [16–18]. Polyploid *Phalaris* exist in both Europe and North America, in fact polyploidy predominates throughout Europe and Asia. The species are widely adapted across the globe and are not restricted to one hemisphere [19]; six floret types distinguish the species as well as other diagnostic traits such as the presence of characteristic ligules [15]. The centers of origin and diversity are the Mediterranean Basin while a secondary center of diversity exists in western North America [7, 17–19]. The most cosmopolitan species is *P. arundinacea* L., circumpolar in distribution in the northern hemisphere. Most likely, a diploid ancestor of *P. arundinacea* came across on the land bridge of the Bering Strait into the present-day State of Alaska, USA [20] during the late Tertiary period [21].

Phalaris has a lengthy taxonomic history with the earliest references arising in the first Century CE (common era), with *P. canariensis* being described by Dioscorides with an accompanying drawing from the Byzantine period (~ 525 CE) [22], although [7] argued that the drawing did not clearly delineate this species. Two species were given scientific names in quadrinomial nomenclature early in the 1600s: *P. major semine albo* (cf. *P. canariensis* L.) and *P. major semine nigro* (cf. *P. minor* Retz.),

prior to the Linnaean era of binomial nomenclature [23, 24].

Of particular importance within the genus, two species are of primary commercial value as cultivated crops, i.e. *P. arundinacea* L. (reed canarygrass; grown for ornamental, forage, biofuel and remediation/restoration efforts) [7–9] and *P. canariensis* L. (canarygrass; grown for birdseed) [25], although additional species are also cultivated: *P. aquatica*, *P. minor* [17, 18] and *P. xdaviesii* [26]. Seed of both *P. arundinacea* and *P. canariensis* are commercially produced in Roseau, Minnesota USA [25], the state with the highest concentration and wetland surface area coverage (50–100% of wetland area) of native, yet invasive, *P. arundinacea* in the continental USA and in North America [27]. The species also is native to Eurasia [7, 28] and North America [29]. As a widely adaptable species, it is able to withstand a variety of conditions including frost, drought, partial shade and poorly drained soil [30]. While minor variation, i.e., plant height and biomass, exists due to genetic and environmental factors, both the native *P. arundinacea* North American types and those native to Eurasia are virtually indistinguishable for any morphological trait [31], since all possess ligules [7] and the same floret type, “Floret Type 4” [15], although the floret type cannot be used when collecting vegetative genotypes for analyses. The North American and Eurasian types (using both extant and historic or herbaria specimens) have been separated, however, using biochemical (allozymes) [28] and molecular markers, such as ISSRs (inter-simple sequence repeats) [32], AFLPs (amplified fragment polymorphisms) [33–35], SNPs (single nucleotide polymorphisms) [29, 36], as well as ITS regions [11, 15, 19, 20].

Herbarium specimens are museum-quality, pressed and dried plant samples deposited and preserved in global herbaria to serve in future research [37]. Plant specimens are typically mounted on acid-free paper with the primary goal of preserving specimen integrity and visual quality to allow for future systematic and taxonomic studies [37]. Herbaria contain an extensive collection of *Phalaris* species past germplasm (pre-1900), that were collected by early land surveyors during European settlement [38, 39]. Those herbarium specimens can serve as an informative resource for comparative genetics, genomics and systematics, to describe changes of *Phalaris* species germplasm over time. Less than optimal early specimen preservation methods can lead to various degrees of DNA degradation [29]. Estimation of the genetic distance between individuals and populations is based on single nucleotide polymorphism (SNP) and depends on the proper nucleotide base calling. Nucleotide misincorporations were found during amplification of ancient DNA and could cause false SNP recognition [40, 41]. False SNP

recognition in herbaria specimens can be highly problematic when proper relatedness needs to be determined in comparison to fresh tissue collection. Here we compared SNPs found in genomic DNA obtained from herbaria specimens (collection years from 1882 to 2001) and fresh tissue (June 17th, 2019) to determine if amplification of DNA obtained from herbaria specimens resulted in varying SNP differences in comparison to DNA obtained from fresh tissue.

Plastid *trnT-F*, *trnL-F* and the nuclear ribosomal ITS region (ITS1–5.8S–ITS2) were amplified/sequenced in the core tribe Aveneae (oats) for taxonomic reconstruction of the Aveneae-Poeae-Seslerieae complex in the Poaceae [11]. That study included the genera *Anthoxanthum*, *Hierochloa* and *Phalaris*, which reside in the Phalarideae (sub Panicoideae). Alignment or clustering vs. bootstraps of ITS showed *Phalaris canariensis* and *P. truncata* to be very strongly supported (100/100; posterior probability support or PPS/bootstrap support) whereas *P. coerulescens* scored 100/95 in relation to these two *Phalaris* species [11]. The use of plastid and nuclear ribosomal markers provided highly accurate taxonomic delineation of the associated genera and species in the Aveneae-Poeae-Seslerieae complex wherein *Phalaris* was classified as a “small, less-diversified satellite lineage” [11]. Chloroplast DNA (13 intergenic sequence regions) and AFLPs were subsequently used to distinguish among North American and European *Phalaris arundinacea* herbaria samples [9]. Chloroplast markers supported AFLP findings that North American races of *P. arundinacea* were distinctly different from European types (which had higher genetic diversity) plus the added finding of a separate Scandinavian chloroplast race from the rest of Europe. Subsequently, Voshell et al. [19] were the first to use ITS with plastid *trnT-F* in *Phalaris* to infer a phylogenetic tree for the genus as well as determine floret evolution and polyploidy relationships. The shortest ITS region length was 588 bases in *P. rotgesii* to the longest of 602 bases in *P. arundinacea* (both of which are closely related) and 142 of 169 variable characters were parsimony-informative [19]. ITS sequencing differentiated two main clades in the genus with an additional subclade, proving the utility and robustness of ITS in taxonomically distinguishing among species and within species’ genotypic differences as well as discerning ploidy and floret type variation. Subsequent studies using ITS in *Phalaris* [15, 20] were used to determine historic dispersal routes from the center of origin in the Mediterranean Basin into the Americas via the Bering land route. *Phalaris arundinacea* appeared in two ITS clades with one embedded in Europe while the other was distinctly North American; the mid-Miocene was identified as the epoch in which *Phalaris* species diversification occurred [20].

The recent discovery that, based on SNPs, all historic herbaria and extant riparian and cultivated populations of *P. arundinacea* in the State of Minnesota are most likely North American natives [29, 36] provided impetus for the present study to examine specific Minnesota / North American native *P. arundinacea* genotypes for their identity with reported ITS for the species as well as for comparison with additional *Phalaris*. The objectives of this study were to: 1) amplify and compare ITS regions (ITS1 and ITS2) of historic herbaria and extant (fresh) *Phalaris* specimens; 2) analyze ITS regions for species-specific polymorphisms (diagnostic SNPs) and 3) compare ITS regions of historic *Phalaris* specimens with known extant *Phalaris* species. Associated null hypotheses tested, respectively, were: 1) H_0 = There is no differences in ITS region between herbaria and fresh *Phalaris* specimens due to herbarium sample age; 2) H_0 = There is no ITS region polymorphism found within *Phalaris* species and 3) H_0 = There is no additional polymorphism between herbaria genotypes and currently known extant *Phalaris* species in GenBank, the National Center for Biotechnology Information (NCBI) [42].

Results

Historic specimen DNA degradation

Sampling of historic specimens showed variable levels of DNA degradation when compared to DNA obtained from extant *Phalaris* tissue (Fig. 1; Supplementary Fig. 1), similar to previous sampling [29]. Extant tissue DNA extracted from *P. aquatica* (PI 476288) and *P. arundinacea* (PI 241065) (lanes 1 and 2, respectively, Fig. 1) have a majority of higher molecular weight DNA fragments (<10 kb). Among representative historic samples *P. canariensis* (PI 619107; lane 3, Fig. 1), *P. paradoxa* (ISC-V-0021361; lane 5, Fig. 1), *P. canariensis* (PI 71229; lane 7, Fig. 1) and *P. minor* (PI 229774; lane 8, Fig. 1), all have highly degraded DNA fragments (<0.7 kb) when compared with other historic specimens, *P. brachystachys* (ISC-V-0021035; lane 4, Fig. 1) and *P. coerulescens* (ISC-V-0021204; lane 6, Fig. 1) with less degraded DNA (2–8 kb).

A standard DNA purification kit (OPS Diagnostics Laboratory, Lebanon, NJ) provided high quality and quantity of genomic DNA that allowed amplification of all extant tissue and most of the historic specimens (41/52 or 78.8%; Table 1). For the majority of herbaria samples, the Optical Density (OD)_{260/280} values were within the expected range for high quality DNA (2.10 ± 0.79, mean ± standard deviation; Table 1) and the second measure of DNA purity (OD_{260/230}) had higher variation (1.64 ± 1.41; Table 1). Fresh specimens yielded high quality DNA, based on both the OD_{260/280} (1.93 ± 0.05) and OD_{260/230} (1.93 ± 0.24; Table 2)

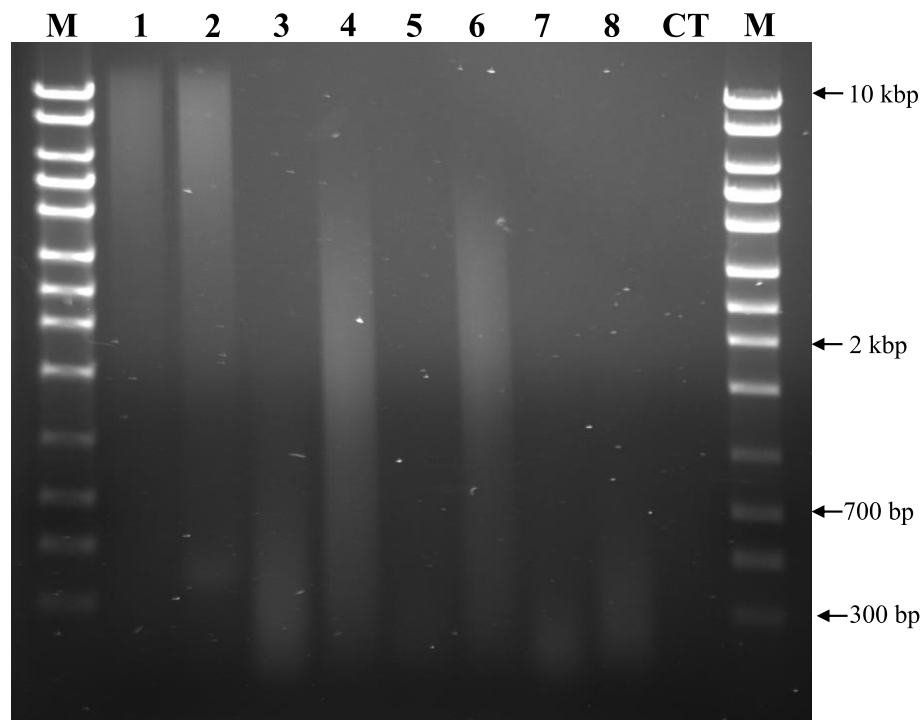


Fig. 1 Gel electrophoresis of genomic DNA (200 ng DNA/lane) providing a visualization of DNA degradation. Fresh tissue produced high molecular weight DNA with herbarium specimens producing variable range of DNA degradation, from samples that had DNA highly degraded (<0.7 kb; 3, 5, 7, 8) to less degraded (fragments in 2–8 kb; 4, 6). Genomic DNA from fresh tissue: (1) *P. aquatica* (PI 476288) and (2) *P. arundinacea* (PI 241065), both > 10 kb; from herbarium: (3) *P. canariensis* (619107) and (4) *P. brachystachys* (ISC-V-0021035); herbarium specimens that were PCR re-amplified (5) *P. paradoxa* (ISC-V-0021361) and (6) *P. coerulescens* (ISC-V-0021204). Herbarium specimens with unsuccessful PCR amplifications (7) *P. canariensis* (71229) and (8) *P. minor* (229774). ITS region amplification (fresh and herbarium samples) and re-amplification is presented on Fig. 3. M – DNA size marker; CT – no sample loaded. Full-length blots/gels are presented in Supplementary Fig. 1

markers. Both herbaria and fresh specimens have a wider range in $OD_{260/230}$ purity measures. However, for both purity measurements, herbaria specimens were less consistent than fresh specimens.

Concentration of DNA from historic *Phalaris* specimens varied widely, from 2.5 ng/ μ l (*P. paradoxa*, ISC-V-0021415; Table 1) to 87.7 ng/ μ l (*P. canariensis*, 619,107; Table 1), despite similar amounts of tissue used for extraction. No correlation between the age of the historic samples over a 119 year range and DNA concentration was found, based on the insignificant slope of the linear regression, $y = 0.095x + 14.7$ and a small but insignificant $r^2 = 0.0177$ (p -value = 0.4; Fig. 2a). Extant specimens exhibited much more consistent values with a range of 28.5 to 80.7 ng/ μ l (Table 2). No regression was fit to the concentration of fresh specimens (Fig. 2b) due to an inadequate sample size ($n = 7$). One of the oldest samples collected 29 July 1886 (*P. canariensis*; 71,229, lane 7, Fig. 1; Table 1) produced 41.9 ng/ μ l of DNA of relatively good quality. This specimen did not, however, yield a successful amplification or sequence of the ITS regions.

ITS amplification and sequencing of *Phalaris* specimens

Plant-specific ITS-P5 and ITS-U4 primers were efficient in PCR amplification of the ITS regions across a diverse set of *Phalaris* species, including a majority of *Phalaris* tissue extracted from herbarium specimens, producing complete sequences of both the ITS1 and ITS2 regions (31/52 or 59.6%, Table 1). The oldest specimen fully sequenced was collected in 1908: *P. californica* (ISC-V-0021043, MN811182.1; Table 1). In addition, most of the ITS amplifications were suitable for direct sequencing after clean up. Overall, 31/41 (75.6%) herbarium (+; Table 1) and 5/7 (71.4%) fresh specimens (+; Table 2) produced full sequences of the ITS region for 12 *Phalaris* species. An additional six sequences of fresh ($n = 2$; *P. arundinacea* PI 241065 and *P. aquatica* PI 476288) and herbaria samples ($n = 4$; *P. arundinacea* 71,166, *P. canariensis* 619,107 and 71,226 and *P. caroliniana* ISC-V-0021166) produced partial sequences of the ITS region but were not used for sequence analysis. High PCR amplification success for fresh specimens were evidenced across the sampling, i.e. *P. canariensis* (lane 1; Fig. 3a; Supplementary Fig. 2), *P. aquatica* (lane 2; Fig. 3a;

Table 1 *Phalaris* species historic herbaria specimen data from leaf sample collections at the Bell Museum Herbarium (MIN; University of Minnesota) and Ada Hayden Herbarium (ISC; Iowa State University), DNA concentration (ng/μl) of extracted samples, their corresponding DNA quality readings from the spectrophotometer (optical densities [OD] at two ratios: OD_{260/280} and OD_{260/230}), amplification success (+/−), herbaria accession codes and the National Center for Biotechnology Information (NCBI) GenBank accession numbers sequences

<i>Phalaris</i> Species	Collection Date	DNA Conc. (ng/μl)	OD _{20/280}	OD _{260/230}	Amplification Success	Herbaria Accession Codes	GenBank Accession Number
<i>P. angusta</i> Nees ex Trin.	16 May 1936	15.8	1.94	1.44	+	ISC-V-0020926	MN811167.1
	16 April 1960	13.8	2.00	1.50	+	ISC-V-0020921	MN811165.1
	3 June 1964	14.5	1.97	1.25	+	ISC-V-0020922	MN811168.1
	12 April 1970	29.2	1.98	1.89	+	ISC-V-0020920	MN811166.1
	3 April 1982	6.1	2.48	1.27	+	ISC-V-0020919	MN811169.1
<i>P. aquatica</i> L.	8 September 1935	34.6	1.95	1.79	+	ISC-V-0021399	MN811171.1
	31 August 1959	9.1	2.05	1.13	+	ISC-V-0020927	MN811170.1
	8 July 1962	13.9	1.88	0.81	+	ISC-V-0021398	MN811172.1
<i>P. arundinacea</i> L.	30 June 1913	70.8	2.00	1.82	+	71,166	n/a
	28 May 1955	6.9	1.5	0.54	+	532,148	MN811175.1
	27 August 1983	40.7	1.93	2.42	+	753,216	MN811176.1
	26 June 2001	6.3	2.07	2.59	+	484,712	MN811200.1
<i>P. brachystachys</i> Link	8 April 1959	10.2	2.12	2.43	+	ISC-V-0021036	MN811180.1
	8 April 1959	10.5	2.37	1.61	+	ISC-V-0021039	MN811179.1
	8 April 1959	63.3	1.91	2.41	+	ISC-V-0021035	MN811178.1
	10 August 1962	7.8	2.21	3.90	+	ISC-V-0021037	MN811181.1
<i>P. californica</i> Hook. & Arn.	4 May 1903	16.2	2.03	2.11	+	ISC-V-0021040	n/a
	24 May 1907	10.6	1.47	0.66	−	ISC-V-0021044	n/a
	18 June 1908	19.9	1.86	1.04	+	ISC-V-0021043	MN811182.1
	1958	10.1	2.16	2.52	−	ISC-V-0021041	n/a
	ND*	8.3	1.78	1.79	−	ISC-V-0021042	n/a
<i>P. canariensis</i> L.	August 1884	26.1	1.91	2.62	+	71,226	n/a
	29 July 1886	41.9	1.87	2.39	−	71,229	n/a
	12 July 1938	73.1	1.95	2.12	+	367,474	MN811185.1
	July 1969	87.7	1.88	2.09	+	619,107	n/a
<i>P. caroliniana</i> Walter	30 May 1940	17.8	2.01	2.07	+	ISC-V-0021097	MN811188.1
	28 May 1964	2.8	2.53	1.74	+	ISC-V-0021081	MN811187.1
	6 April 1973	7.8	2.28	2.47	+	ISC-V-0021166	n/a
	3 May 1982	25.9	1.97	2.17	+	ISC-V-0021080	MN811186.1
<i>P. coeruleascens</i> Desf.	24 November 1958	23.8	2.05	1.61	+	ISC-V-0021199	MN811190.1
	15 December 1958	5.0	2.60	1.40	+	ISC-V-0021198	n/a
	8 April 1959	22.9	2.06	1.93	+	ISC-V-0021204	n/a
	8 April 1959	6.6	2.19	1.45	+	ISC-V-0021203	MN811189.1
<i>P. lemmonii</i> Vasey	1888	6.6	1.61	−4.95	−	ISC-V-0021333	n/a
	18 April 1946	26.2	2.02	1.89	+	ISC-V-0021334	n/a
	15 June 1955	42.9	1.79	0.79	+	ISC-V-0021329	MN811192.1
	7 May 1956	9.0	2.11	0.96	+	ISC-V-0021336	MN811191.1
	23 May 1980	8.8	1.80	0.63	+	ISC-V-0021328	n/a
<i>P. minor</i> Retz.	May 1882	12.6	1.92	0.88	−	ISC-V-0021344	n/a
	April 1901	36.3	1.81	2.07	−	229,774	n/a
	12 April 1918	17.0	1.80	0.72	−	ISC-V-0021338	n/a
	24 November 1958	12.7	2.11	1.74	−	ISC-V-0021342	n/a
	8 April 1959	15.3	2.12	1.70	+	ISC-V-0021341	MN811193.1

Table 1 (continued)

<i>Phalaris</i> Species	Collection Date	DNA Conc. (ng/μl)	OD _{20/280}	OD _{260/230}	Amplification Success	Herbaria Accession Codes	GenBank Accession Number
<i>P. paradoxa</i> L.	13 May 1914	19.6	2.03	2.20	+	ISC-V-0021360	n/a
	4 June 1915	10.9	1.71	0.61	-	ISC-V-0021363	n/a
	2 June 1934	71.8	1.68	0.75	+	ISC-V-0021361	MN811195.1
	3 May 1940	3.8	1.90	7.05	-	ISC-V-0021362	n/a
	8 April 1959	2.5	7.42	0.27	+	ISC-V-0021415	MN811194.1
<i>P. truncata</i> Guss. ex Bertol.	11 November 1958	19.7	1.88	2.68	+	ISC-V-0021377	MN811198.1
	24 November 1958	7.7	2.25	2.03	+	ISC-V-0021384	MN811199.1
	15 December 1958	30.5	2.06	1.81	+	ISC-V-0021395	MN811197.1
	24 December 1958	2.9	2.66	0.38	+	ISC-V-0021373	MN811196.1

ND* denotes no reported date for herbarium specimen collection

n/a denotes specimens with no GenBank sequence publication

+ denotes a successful PCR amplification (see Materials and Methods section)

- denotes an unsuccessful PCR amplification (see Materials and Methods section)

Table 2 Extant (fresh) *Phalaris* species specimens (USDA - GRIN) used for ITS sequencing, DNA concentration (ng/μl) of extracted samples, their corresponding DNA quality readings from the spectrophotometer (optical densities [OD] at two ratios: OD_{260/280} and OD_{260/230}), amplification success (+/-), GRIN accession numbers and the published National Center for Biotechnology Information (NCBI) GenBank accession numbers

<i>Phalaris</i> Species	DNA Concentration (ng/ul)	OD _{260/280}	OD _{260/230}	Amplification Success	GRIN Accession Number	GenBank Accession Number
<i>P. aquatica</i>	60.7	1.97	2.14	+	PI 476287	MN811177.1
	80.7	1.97	2.06	+	PI 476288	n/a
	59.9	1.96	1.93	+	PI 303825	MN811173.1
<i>P. arundinacea</i>	52.2	1.95	2.13	+	PI 241065	n/a
	28.5	1.84	1.47	+	PI 422030 (<i>loreed</i>)	MN811174.1
<i>P. canariensis</i>	53.6	1.89	1.99	+	PI 578800	MN811183.1
	57.8	1.97	1.79	+	PI 578798	MN811184.1

n/a denotes specimens with no GenBank sequence publication

+ denotes a successful PCR amplification

- denotes an unsuccessful PCR amplification

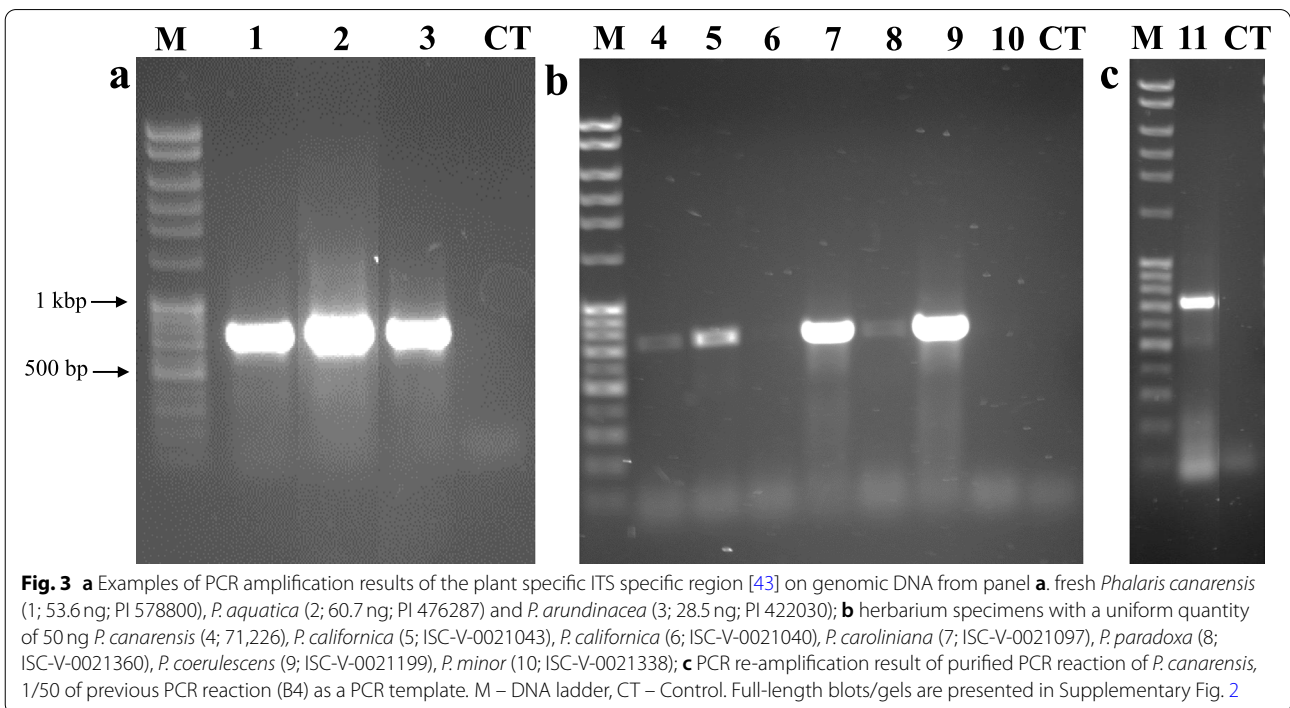
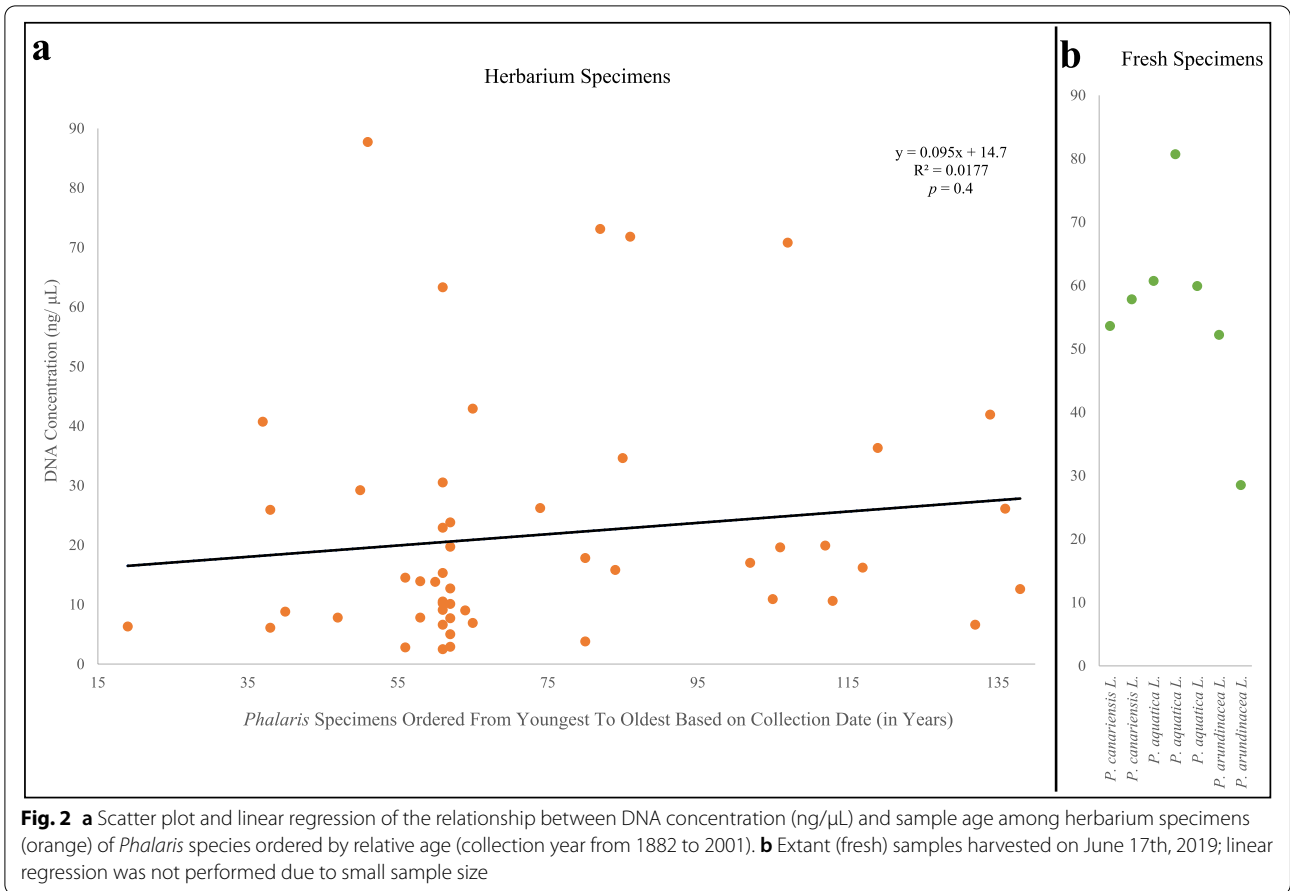
Supplementary Fig. 2) and *P. arundinacea* (lane 3; Fig. 3a; Supplementary Fig. 2).

The low PCR amplification success rate of herbaria specimens is most likely caused by DNA degradation that did not allow amplification of the ITS DNA, as found for samples *P. californica* (ISC-V-0021040), *P. paradoxa* (ISC-V-0021360) and *P. minor* (ISC-V-0021338; lanes 6, 8, 10, respectively; Fig. 3b; Supplementary Fig. 2). Overall, herbaria specimens that were collected before 1915 ± 28 (mean ± S.D.) years did not amplify the full ITS region, wherein most of the *Phalaris* herbaria samples collected later, in 1953 ± 23 years, were successful in ITS region amplification (Table 1). The 50 ng/μl DNA template was used as a standard DNA concentration for

PCR amplification of all *Phalaris* herbarium specimens (Fig. 3b). However, due to variable degradation of DNA from herbarium specimens, ITS amplification did not always allow use of those products for direct sequencing based on sequencing requirements (Fig. 3b, lane 8; University of Minnesota Genomics Center – Sanger Sequencing Classic). For some herbaria samples it was necessary to reamplify PCR products, which allowed for later product sequencing (Fig. 3c; Supplementary Fig. 2).

Polymorphism analysis of *Phalaris* species

Reconstruction of genetic distance from newly sequenced herbaria and fresh *Phalaris* specimens with combination of those available in GenBank showed proper specimen



classification and grouping within the same species (Fig. 4a; Supplementary Fig. 3). Generally, the sequences produced in this study matched with previous sequences of the same species. Shared SNPs from the matching sequences created unique clades by species in the distance tree (Fig. 4a). Some species (*P. brachystachys* and *P. canariensis*) are more closely related than others and the distance between clades suggests that the ITS region was insufficient to separate those two species into separate branches. Both species share nine unique SNPs that distinguish the two from the rest of *Phalaris* (Supplementary Fig. 3; Table 3).

A large *Phalaris* species sequence collection allowed us to identify species-specific SNPs that identified selected *Phalaris* species, based on single SNPs found within ITS1 and ITS2 (Table 3). Two additional species (*P. lemmonii* and *P. angusta*) lacked species-specific SNPs, based on their sequence alignment (Table 3).

Considerable polymorphism was exhibited in the *P. arundinacea* clade (Table 3; AR1, AR2, AR3, AR4, AR5 and AR6 diagnostic, species-specific, SNP sites in the ITS1 and ITS2 regions), particularly among Asian and those of N. American origin (Fig. 4b). We did not observe

sequence abnormalities between herbaria specimens and fresh tissue specimens of the same species despite suspected misidentified samples. Collection locations of *P. arundinacea* specimens reflect the relatedness within or distinctiveness among *P. arundinacea* genotypes (Fig. 4b). The *P. arundinacea* ITS sequences identified three clusters, the first with *P. arundinacea* originating in the United States which are N. American in origin [29, 36]: MN81175.1, MN811200.1, MN811176.1 (herbarium) and MN811174.1 (fresh) clustered tightly together (Fig. 4b), although they were grouped in between Korean accessions KF713257.1 and FJ766174.1. The second cluster of *P. arundinacea* was an adjacent clade, on one side of the N. American natives, which included one genotype from Korea (KF713257.1, Fig. 4b) and two others from the United States and the east coast USA State of Virginia (KF753779.1, JF951077.1). On the other side in closely related clades were four Eurasian genotypes from Taebaek-si and Yungyang-gun, S. Korea, specifically FJ766174.1, KF713256.1, KU883517.1 and KF13255.1 (Fig. 4b). A third, more distant cluster of two inter-related clades contained strictly Eurasian types from S. Korea ($n = 5$), China ($n = 2$) and Germany ($n = 1$) (Fig. 4b). The

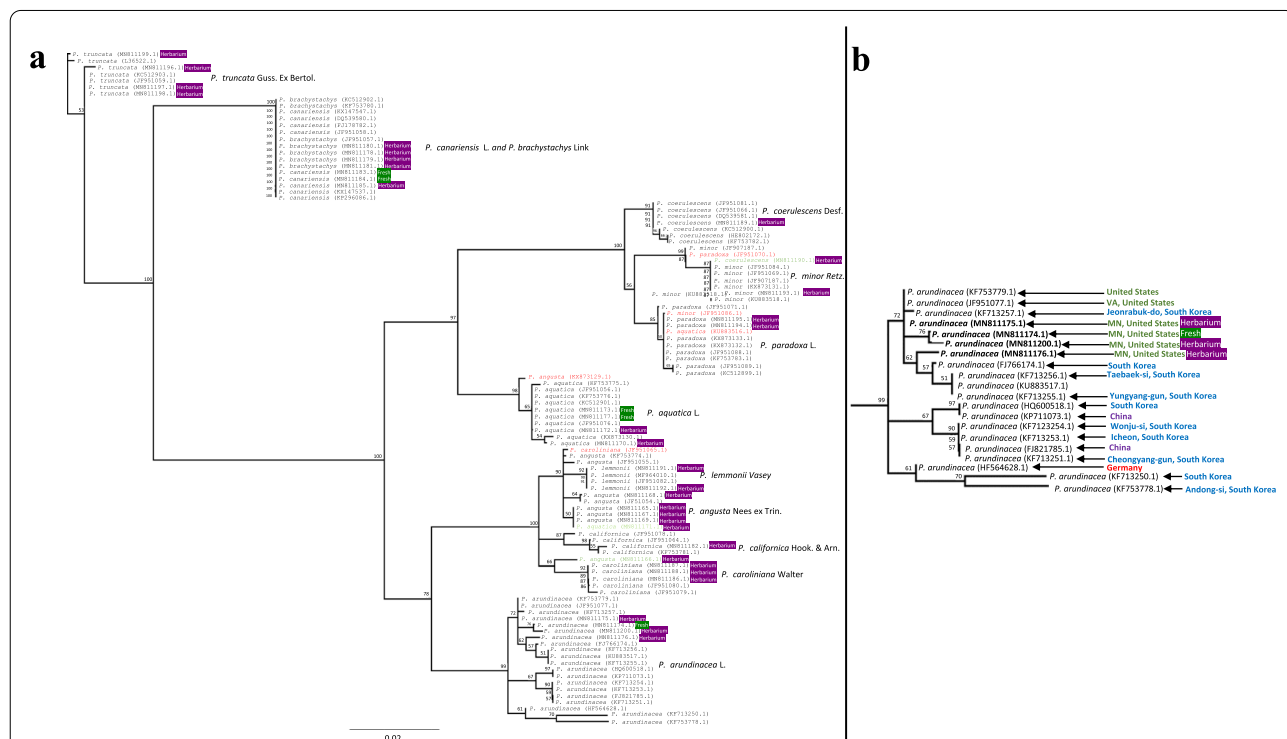


Table 3 Diagnostic, species-specific, single nucleotide polymorphism (SNP) sites in the ITS1 and ITS2 regions (with the species-specific letter code followed by the regions numbered 1–9) for *Phalaris* species, within each species-specific site, specifying the nucleotide substitution (A/C, A/G, A/T, C/A, C/G, C/T, G/A, G/C, T/A, T/C, T/G) and the relative position each site on the ITS1 and ITS2 multialignments (Supplementary Fig. 4). Two species (*P. lemmonii*, *P. angusta*) lack diagnostic SNPs, while two species (*P. brachystachys*, *P. canariensis*) are indistinguishable (N/A)

<i>P. truncata</i> Guss. ex Bertol.	T1 A/C ₂₃	T2 G/C ₁₅₅	T3 T/G ₁₆₆	T4 T/A or T/C ₂₄₅	T5 T/A or T/G ₃₈₁				
<i>P. brachystachys</i> Link & <i>P. canariensis</i> L.	B-C1 A/G ₅₂	B-C2 C/T ₆₅	B-C3 A/C ₈₈	B-C4 A/C ₁₆₂	B-C5 C/A ₂₄₆	B-C6 T/C ₃₁₅	B-C7 A/C ₃₃₉	B-C8 A/G ₃₄₆	B-C9 A/G ₃₄₉
<i>P. aquatica</i> L.	A1 T/G ₄₅	A2 A/T ₁₃₂	A3 A/T or A/C ₂₂₁	A4 C/T or C/A ₂₃₇	A5 A/C ₃₅₂				
<i>P. coerulescens</i> Desf	C1 DEL ₄₂₆	C2 T/C ₄₂₄	C3 A/T ₄₁₄						
<i>P. minor</i> Retz.	M1 T/C ₂₀₂	M2 T/G or T/A ₂₅₂	M3 T/C ₃₇₃						
<i>P. paradoxa</i> L	P1 C/G ₄₀₃								
<i>P. californica</i> Hook. & Arn.	CA1 G/A or G/C ₂₄₆	CA2 G/A ₄₁₆							
<i>P. caroliniana</i> Walter	CR1 A/T ₂₁₃	CR2 T/A ₃₈₅	CR3 A/G ₃₈₁						
<i>P. arundinacea</i> L.	AR1 T/G or T/C ₁₉₈	AR2 C/T ₂₁₅	AR3 A/G or A/T ₂₅₂	AR4 G/A ₂₈₁	AR5 A/C ₃₅₁	AR6 T/C ₃₃₇			
<i>P. lemmonii</i> Vasey	N/A								
<i>P. angusta</i> Nees ex Trin.	N/A								

relatedness, based predominantly on geographic distribution, illustrates high variation of the *P. arundinacea* samples within the ITS region along with the possibility to distinguish among *P. arundinacea* from different geographic locations on a continental scale (Fig. 4b).

The composition of the *Phalaris* distance tree also indicates that there are GenBank sequences and herbaria specimens that may have been misidentified (Fig. 4a). In addition to proper classification and groupings, multiple ITS sequences found in GenBank did not follow the species branch assignment: *P. angusta* (KX873129.1) was classified with the *P. aquatica* clade; *P. paradoxa* (JF951070.1) was assigned to *P. minor* clade; *P. caroliniana* (JF951065.1) was assigned to *P. angusta* clade; *P. minor* (JF951086.1) and *P. aquatica* (KU883516.1) assigned to the *P. paradoxa* clade (Fig. 4a). Three of the sequences produced in this research did not group with a specific clade: *P. coerulescens* (MN811190.1; ISC-V-0021199), *P. aquatica* (MN811171.1; ISC-V-0021399) and *P. angusta* (MN811166.1; ISC-V-0020920), which clustered with *P. minor*, *P. angusta* and *P. caroliniana*, respectively. Both MN811189.1 and MN811190.1 (that did not group with a species clade) were *P. coerulescens* herbarium specimens grown from Turkish seed at Iowa State University but did not have the same SNPs, further suggesting possible misidentification. The misidentification is likely due to the similar morphological characteristics of the *Phalaris* genus and highlights the need for ITS

barcoding to distinguish among and between extant and historic *Phalaris* species. It also highlights the need to exercise stringent quality control measures on newly generated sequence data and to approach public sequence databases in a critical way.

Discussion

Historic specimen DNA degradation

Efficient DNA amplification (barcoding) requires the DNA template to be of high quality (measured by OD_{260/280} and OD_{260/230}), quantity (ng of DNA template/μl) and integrity (visualized on agarose gels). High DNA quality equates to the DNA being free from protein and chemical contaminants with a sufficient quantity of DNA template to start PCR amplification of targeted DNA regions [44]. A measure of DNA quality is the ability to be amplified. A second round of PCR can reduce inhibitors, as sometimes found in extracted DNA [45]. The potential effect of PCR inhibitors is reduced during the PCR reaction, in that isolated DNA from each reaction was diluted 1/25 in volume prepared for PCR in our experimentation. One of the most critical aspects when working with herbaria specimens is physical DNA degradation (fragmentation) that directly impacts amplification of selected DNA regions, such as the ITS regions.

In the case of historic *Phalaris* specimens [29], the additional factor of varying levels of genomic DNA degradation also affect DNA amplification [45]. Choi

[46] demonstrated that trends between age and purity were not generally significant in four evolutionarily, geographically and ecologically different plant lineages. Therefore, DNA purity of specimens is likely due to the post-harvest and storage techniques used. Historic samples that are preserved and stored in traditional herbaria cases maintain the plant tissue appearance and overall specimen condition to minimize tissue degradation and maintain the phenotype of the type specimens as illustrated in Supplementary Fig. 4 [38]. However, the process and conditions of tissue preservation (harvest and post-harvest tissue storage) may not preserve DNA integrity [29, 47], leading to degradation of DNA in herbarium specimens which reduces amplification of ITS regions for samples with significantly degraded DNA [48]. Other factors, such as sample age, are also inadequate assessors of DNA quality [49]. Nonetheless, despite obtaining ODs with high purity and DNA concentrations of 2.5 to 87.7 ng/ μ l (Table 1), only some of the *Phalaris* herbaria samples were extracted with high molecular weight DNA (Fig. 1) which was not as consistent as with fresh (extant) specimens, yielding 28.5 to 80.7 ng/ μ l (Table 2).

It is not sufficient to obtain a certain quantity of DNA template (i.e. 50 ng/PCR reaction), but evaluation of DNA degradation should also be taken into account as a major factor limiting proper PCR amplification [43, 50]. Another tactic to be considered is to amplify the genetic marker in a stepwise fashion when the DNA template is highly degraded [48]. Additionally, DNA reconstruction is a method that could be employed early on to increase DNA quality (OD_{260/280} ratios) and later PCR success and sequence length [51]. While an exhaustive study was not performed on reamplification options (nested PCR, whole genome amplification, or others) with the purpose of increasing the concentration of the PCR product to the required concentration for Sanger sequencing, it was found that a visible band was needed for sequencing. This method could be useful when working with old or precious tissue, such as herbarium, as a means to increase sequencing success.

The lack of consistency in herbaria specimens may be due to increased degradation and/or minor levels of contaminants, although the wide age range of 119 years across herbaria samples was not correlated with DNA concentration (Fig. 2a). When using the ITS region for barcoding herbaria specimens within the Juncaceae, a significantly negative association between age and sequencing success was found [47]. Therefore, the age relation to sequencing success may vary by species as well as storage treatments for herbarium specimens, rather than age – at least in the case of our *Phalaris* herbaria samples.

ITS amplification and sequencing of *Phalaris* specimens

Complete ITS1 and ITS2 sequences were obtained in 59.6% of the accessions across 12 species in the genus *Phalaris* (Tables 1, 2) and were within the expected range of 757 ± 140 bases [52]. The oldest herbarium specimen that was fully sequenced was *P. californica* collected in 1908 (Table 1). Six fresh and herbarium partial sequences of *P. arundinacea*, *P. aquatica*, *P. canariensis* and *P. caroliniana* (Tables 1, 2) were not used in the sequence analyses. The occurrence of partial sequences is not unusual, as previous researchers had difficulties with *P. peruviana* producing low quality sequences with missing data [15]. Design of a new set of ITS PCR primers that would be located farther from the beginning (and end) of the ITS region would most likely provide more high quality ITS sequences. Sanger sequencing tends to produce low quality sequences, just after the end of the sequencing primer position [53], that should be removed during sequence quality review.

Other factors can contribute to sequencing success such as age, preservation procedure, specimen color, the climate collected in and taxon related plant characteristics [53–55]. While Brewer et al. did not investigate the Poaceae family, they did determine that sequencing success can be related to taxon-specific traits in relation to the preservation technique used (herbaria vs. silica gel-dried specimens) [54]. The Poaceae family that the *Phalaris* genus belongs to was found to have a decreased sequence recovery of the ITS2 region with age and silica gel vs. herbaria preservation method [55]. This calls for the optimization of taxon-specific traits for species-focused studies such as this to achieve maximized sequence retrieval for barcoding.

Polymorphism analysis of *Phalaris* species

Multiple sequence alignment of the ITS region from historic herbaria and fresh *Phalaris* species showed that sequences and DNA polymorphisms found in newly sequenced *Phalaris* species resemble those previously reported [19, 56]. The similarity among the phylogenetic tree of the ITS region constructed by Voshell et al. [19] (fresh *Phalaris* tissue) and the distance tree constructed in this study (mainly herbaria *Phalaris* tissue) verifies the relatedness found among the *Phalaris* species and the genetic distinctions that can be extrapolated from SNPs in the ITS region with the use of herbarium tissue as old as 119 years.

Two species (*P. lemmonii*, *P. angusta*) lacked species-specific SNPs based on their sequence alignment (Table 3). Their relatedness may be expressed in flower morphology and may be due to geographic separation [19, 56]. In other instances, some *Phalaris* species, like *P. brachystachys* and *P. canariensis*, were too closely related

such that the ITS region was insufficient to separate them (Fig. 4a). To separate both species we recommend using other barcoding regions, such as *atpF*–*atpH*, *matK*, *rbcL*, *rpoB*, *rpoC1*, *psbK*–*psbI* and *trnH*–*psbA* to see if a SNP can be found [57].

When comparing the phylogenetic tree constructed by Voshell et al. [19] to our distance tree (Fig. 4) of predominantly herbarium based ITS regions, similar clade groupings can be identified. A separation of clades (*P. canariensis*, *P. brachystachys* and *P. truncata*) and remaining species was identified in our study (Fig. 4), similar to [19]. The trees from Fig. 4, Supplementary Fig. 3 and Voshell et al. [19] form a consistent cluster of the species *P. lemmonii*, *P. angusta*, *P. caroliniana*, *P. californica* and *P. arundinacea* interpreted to resemble lineages 1 and 2 of the *Phalaris* genus [19]. The divergence of lineage 3 including *P. coerulescens*, *P. minor*, *P. paradoxa* and *P. aquatica* was also found in this analysis [19]. The similarity between the ITS inferred distance in this study (mainly herbaria *Phalaris* tissue) and by Voshell et al. [19] (fresh *Phalaris* tissue) verifies the relatedness found among *Phalaris* species and the genetic distinctions that can be extrapolated from SNPs in the ITS region with the use of herbarium tissue as old as 112 years (MN811182.1). No false SNPs were produced in sequence alignment (Supplementary Fig. 3) and distance tree construction due to the age of tissue and relatedness [19], indicating the versatility of herbarium specimens in ITS barcoding (Fig. 4a). Nucleotide polymorphism exists among most *Phalaris* species that were shared among historic and extant tissue types.

The clustering of *P. arundinacea* ITS sequences into predominantly regional clades from the USA and Asia is a notable grouping (Fig. 4b). All genotypes from N. America were in a single clade, distinct from most other ones of Eurasian origin. Despite considerable polymorphisms in the species (AR1–AR6, Table 3), a tight and adjacent clustering of the four native N. American *P. arundinacea* accessions from Minnesota (extant and historic MN811175.1, MN811200.1, MN811176.1, MN811174.1; Fig. 4b) was unexpected. These were grouped among two accessions of Korean origin (KF713257.1 and FJ766174.1; Fig. 4b) which could be the original ancestors of *P. arundinacea* germplasm that came across the land bridge of the Bering Strait into the State of Alaska, USA [20] during the late Tertiary period [21]. Other genotypes of USA origin (either native N. American or of Eurasian ancestry) were separated in the clade from the four N. American (Minnesota) natives. While the separation of USA and Asia *P. arundinacea* was not 100% consistent, most of the Eurasian types were in distinct clades; the use of additional barcoding markers would most likely allow proper geographic association of *P. arundinacea*. The

relatedness of N. American vs. most Eurasian types illustrates the possibility to distinguish among *P. arundinacea* from different geographic locations on a continental scale (Fig. 4b). Future research is essential with our large germplasm base of native riparian, roadside, lakebed and cultivated N. American types in the Midwest (> 3000 genotypes) [29, 36] to potentially provide substantive confirmation of this distinction or a melding of all N. American types.

The high variation within the *P. arundinacea* and the detection of variants, forma and subspecies (Fig. 4b) further justifies the sequencing of more specimens within the ITS region and other barcoding regions [56–60] to better understand the relatedness and variation within this species and the genus as a whole. Herbaria curators at The Consortium of California Herbarium have recognized the *P. arundinacea* species to contain the subspecies *picta* and *typica*, four varieties (*arundinacea*, *colorata*, *geuina*, and *japonica*) and nine forma (*Arundinacea*, *coarcta*, *luteo-Picts*, *minor*, *pallens*, *pallida*, *Ramifera*, *Ramosa*, *variegata*; <http://www.cch2.org/portal/taxa/index.php?taxon=Phalaris+arundinacea&forms-submit=Search+Terms>). Further genetic structure analysis of *P. arundinacea* will provide the potential ability to genetically distinguish *P. arundinacea* by continents, even though it is morphologically indistinguishable. More exhaustive analysis of the ITS region may provide SNPs to discriminate the native vs. exotic status in N. American populations. Furthermore, the secondary structure of the ITS1 and ITS2 regions may provide a more accurate and detailed idea of the relatedness, sequence evolution and barcodes of the *Phalaris* species – especially when used in combination with sequence data [61].

Conclusions

Our study demonstrates that most *Phalaris* herbaria specimens can be used to produce high quality ITS sequences that are useful to evaluate past genetic resources. The methods used in this study are relatively simple to replicate and, most importantly, do not require specific protocol modifications. Overall, the use of a standard DNA isolation protocol, PCR amplification method and direct PCR product sequencing were sufficient to sequence the majority of the collected herbarium specimens. The lack of amplification of some specimens is most likely due to high DNA degradation. Furthermore, herbaria can provide the plant tissue and taxonomically identified specimens, facilitating the identification of new species-specific barcodes and evaluation of past genetic resources. The ITS region (ITS1 and ITS2) was sufficient to distinguish eight out of twelve *Phalaris* species in this study. In addition, separation of the genus and within species was possible, implying that

subsampling of multiple genotypes of this same species is necessary. The ITS region distinguished among most Eurasian vs. N. American accessions of *P. arundinacea*. A further development of within species and genus specific barcodes could contribute to designing PCR primers for efficient and accurate plant identification. Our finding of misidentified *Phalaris* species indicates the need for the DNA sequence database curation for proper specimen identification.

Methods

Germplasm collection

Extant (fresh) and historic herbaria samples of the *Phalaris* species were used for genomic DNA extraction. A total of $n=52$ historic specimens were collected from the Bell Museum Herbarium, University of Minnesota, St. Paul, MN (MIN; $n=9$; Table 1) and the Ada Hayden Herbarium, Iowa State University, Ames, IA (ISC; $n=43$; Table 1). Formal identification of each voucher specimen had already been accomplished by each herbaria curator. Destructive sampling of $\sim 2.5 \times 0.63$ cm of leaf tissue was performed on each herbarium specimen for leaf samples positioned in the back of the specimen to not decrease their visual integrity, similar to our previous methodology [29]. Permission for destructive tissue sampling was obtained from each herbarium curator in advance. Specimen notes were added to delineate the specimen-specific sampling. This herbarium collection represents a range of North American *Phalaris* herbarium specimens with collection dates ranging from 1882 (*P. minor* Retz.; ISC-V-0021344) to 2001 (*P. arundinacea* L.; 484,712; Table 1). Selection of the *P. arundinacea* herbaria samples were already tested for SNP genetic variation and provided sufficient and high quality nuclear DNA could be extracted, were determined to be most likely native North American genotypes [29, 36].

For additional sampling purposes, seeds were obtained from the U.S. Department of Agriculture Germplasm Resources Information Network (USDA GRIN; <https://www.ars-grin.gov/>). Formal identification of each seed specimen had already been accomplished by the USDA GRIN taxonomist; no voucher specimens were deposited since the seed is part of the national USDA GRIN system. Germinated seedlings were used as extant specimens of the *Phalaris* species. Fresh samples served as an amplification and SNP detection validation. Extant specimens included three *P. aquatica* (PI 476287; PI 476288; PI 303825), two *P. arundinacea* (PI 241065; PI 422030) and two *P. canariensis* (PI 578800; PI 578798; Table 2). Seeds were sown in 10 cm square pots filled with SunGro Professional Growing Mix (Sun Gro Horticulture; SKU:5105, Agawam, MA) and placed in a mist house

(greenhouse with an intermittent mist system). Once the seedlings germinated and true leaves were developed, plants were moved to the greenhouse for continued growth and leaf harvest. Environmental conditions in both greenhouses were $24.4 \pm 3.0/18.3 \pm 1.5$ °C day/night daily integral and a 16 h long day photoperiod (0600–2200 HR) lighting (400 W high pressure sodium high intensity discharge lamps, HPS-HID) at a minimum of $150 \mu\text{mol m}^{-2} \text{s}^{-1}$. Plants were fertilized twice daily, between 0700 and 0800 and 1600–1700, using a constant liquid feed (CLF) of 125 ppm N from water-soluble 20N–4.4P–16.6K (Scotts, Marysville, OH). Fungicide drenches were applied in monthly rotations. Leaf tissue of matured, fully-expanded leaf trips were harvested and stored at -20 °C in sealed plastic bags.

GenBank ITS resources

National Center for Biotechnology Information (NCBI) GenBank database searches were performed to collect existing ITS region sequences for comparative purposes from the twelve *Phalaris* species included in this study (Table 1). In this research, only *Phalaris* ITS sequences that contained full ITS1 and ITS2 regions were included and partial sequences were not compared. Partial sequences were not included because missing sequences could contain SNP polymorphism(s) that would contribute to less accurate distance analysis. Multiple ITS sequences from twelve *Phalaris* species were found ($n=68$, full sequences): *P. angusta* (KX873129.1, KF753774.1, JF951055.1, JF51054.1), *P. aquatica* (KU883516.1, KF753775.1, JF951056.1, KF753776.1, KC512901.1, JF951076.1, KX873130.1), *P. arundinacea* (KF753779.1, JF951077.1, KF713257.1, FJ766174.1, KF713256.1, KU883517.1, KF713255.1, HQ600518.1, KP711073.1, KF713254.1, KF713253.1, FJ821785.1, KF713251.1, HF564628.1, KF713250.1, KF753778.1), *P. brachystachys* (KC512902.1, KF753780.1, JF951057.1), *P. californica* (JF951078.1, JF951064.1, KF753781.1), *P. canariensis* (KX147547.1, DQ539580.1, FJ178782.1, JF951058.1, KX147537.1, KP296086.1), *P. caroliniana* (JF951065.1, JF951080.1, JF951079.1), *P. coerulescens* (JF951081.1, JF951066.1, DQ539581.1, KC512900.1, HE802172.1, KF753782.1), *P. lemmonii* (MF964010.1, JF951082.1), *P. minor* (JF907187.1, JF951084.1, JF951069.1, JF907187.1, KX873131.1, KU883518.1, JF951086.1), *P. paradoxa* (JF951070.1, JF951071.1, KX873133.1, KX873132.1, JF951088.1, KF753783.1, JF951089.1, KC512899.1) and *P. truncata* (L36522.1, KC512903.1, JF951059.1). The geographic source of the *P. arundinacea* specimens were recorded if information was available on GenBank or in the associated published work.

DNA extraction

Nuclear DNA was extracted from historic and extant *Phalaris* samples using Synergy 2.0 Plant DNA Extraction Kit (OPS Diagnostics Laboratory, Lebanon, NJ) with minimal adjustments to the protocol [29]. Tissue was loaded using scissors and forceps that were washed in soapy water, rinsed twice in distilled water and dried with paper towels. Extant samples, unlike historic ones, were kept cool on dry ice throughout the loading process. Samples were ground for a total of 15 min at 1500 rpm in homogenizer (Geno/Grinder; SPEX SamplePrep, Metuchen, NJ). Once purified, DNA was suspended in molecular grade water and stored at -20°C . DNA quality and quantity were checked with a spectrophotometer (NanoDrop 2000; Thermo Scientific, Waltham, MA). Quality guidelines were followed as recommended by Thermo Fisher Scientific as $\sim 1.8 \text{ O}_{260/280}$ and $1.8\text{--}2.2 \text{ O}_{260/230}$ considered as “pure” DNA (NanoDrop 2000/2000c Spectrophotometer V1.0 User Manual, 2009). Genomic DNA was visualized on 1% (w/v) agarose gels ($1 \times$ Tris-acetate-EDTA buffer) and stained with ethidium bromide, 6X loading dye (New England BioLabs; Ipswich MA) and a DNA ladder (FullRanger 1 kb; Norgen Biotek Corp, Thorold, ON, Canada). Only herbarium specimens with the highest quantity of DNA were used to visualize degradation (200 ng / lane on an agarose gel) due to the limited DNA.

To calculate correlation between herbarium tissue age and the concentration of genomic DNA obtained and to the PCR amplification success from herbarium tissue Microsoft Excel 365 (Microsoft, Redmond, WA) with Regression function was used to estimate *p*-value.

PCR amplification

Polymerase chain reaction (PCR) was performed using 10 μM of ITS-P5 and ITS-U4 primers [52]. PCR master mix (GoTaq Green Master Mix, M712; Promega, Madison, WI) plus 1 μL of DNA (with a minimum concentration $> 50 \text{ ng}/\mu\text{L}$) or DNA volume was adjusted to obtain 50 ng total DNA. The PCR protocol followed previous methodology [52]: 94°C for 4 min, then 34 cycles of 30 s at 94°C , 40 s at 55°C and 1 min at 72°C , finishing with 10 min at 72°C . PCR reactions were visualized using electrophoresis on a 1% (w/v) agarose gel ($1 \times$ Tris-acetate-EDTA buffer) with ethidium bromide and a DNA ladder (FullRanger 100 bases; Norgen Biotek Corp, Thorold, ON, Canada). Expected DNA amplification product of ITS-P5 and ITS-U4 primers was 757 ± 140 bases [52]. If a single amplification product was observed, amplified products were directly purified using PureLink™ Quick PCR Purification Kit (Thermo Fisher Scientific, Waltham, MA) and prepared for Sanger sequencing at the University of

Minnesota Genomics Center, following UMGC sample requirements (University of Minnesota Genomics Center – Sanger sequencing Classic, <http://genomics.umn.edu/sanger-sequencing-classic.php>). If the purified PCR reaction yielded less than the total DNA required for Sanger sequencing (25 ng/ μL), the PCR product was diluted in water (1/50) and 1 μL of diluted product was used and re-amplified following the same procedure (ITS-P5 and ITS-U4, primer set, PCR reaction composition, PCR program).

Sequence analysis

Sequencing results were edited and quality-checked using 4Peaks software (Nucleobytes, Alsmeer, Netherlands; <http://nucleobytes.com/4peaks/>) with additional manual sequence trimming. Sequence editing, alignments, annotations and manipulations were done using Geneious 11.1.5 software (Biomatters, Ltd., New Zealand; <https://www.geneious.com>) [58]. The *Arabidopsis thaliana* ($\times 52320.1$) ITS sequence was initially used to annotate full ITS regions of the newly obtained ($n = 36$) *Phalaris* DNA sequences (Tables 1, 2). Genetic distance among *Phalaris* species was inferred using the Neighbor-Joining method [59] and a bootstrap test was performed for each tree (with 100 replicates) among newly obtained and GenBank ITS sequence collection. A multiple sequence alignment was performed with use of MUSCLE alignment with version 3.8.425 on Geneious software [60]. Diagnostic, species-specific SNPs that differentiate among *Phalaris* species were determined based on multiple alignment of full-length ITS (ITS1 and ITS2 regions) sequences with exclusion of 5.8S ribosomal subunit. Sequences obtained in this study were deposited into the GenBank database. Accession numbers can be found in Tables 1 and 2.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-021-03284-z>.

Additional file 1: Supplementary Figure 1. Original uncropped image of gel electrophoresis of 200 ng of genomic DNA in Fig. 1. Where lanes 2 and 3 are fresh tissue *P. aquatica* (PI 476288) and *P. arundinacea* (PI 241065) respectively. Remaining lanes (4–9) are from herbarium tissue consisting of *P. cananensis* (619107), *P. brachystachys* (ISC-V-0021035), *P. paradoxa* (ISC-V-0021361), *P. coeruleascens* (ISC-V-0021204), *P. cananensis* (71229) and *P. minor* (229774). Lanes 1 and 11 are a DNA size marker (FullRanger DNA ladder 1 kb) while lane 10 is a control lane with no sample loaded. Dotted white line over image illustrate where the image was cropped to form Fig. 1.

Additional file 2: Supplementary Figure 2. Original uncropped image of gel electrophoresis of PCR amplification examples of the plant specific ITS region in Fig. 3 [51] where panel labels (a, b, c) correspond. Panel a consist of fresh *P. cananensis* in lane 2 (53.6 ng; PI 578800), *P.* in lane 3 (57.8 ng; PI 578798), *P. aquatica* in lane 4 (60.7 ng; PI 476287), *P. aquatica* in lane 5 (80.7 ng; PI 476288), *P. aquatica* in lane 6 (59.9 ng; PI 303825), *P.*

arundinacea in lane 7 (52.2 ng; PI 241065), *P. arundinacea* in lane 8 (28.5 ng/μl; PI 422030), *P. canariensis* in lane 11 (20 ng; PI 578800), *P. canariensis* in lane 12 (20 ng; PI 578798), *P. aquatica* in lane 13 (20 ng; PI 476287), *P. aquatica* in lane 14 (20 ng; PI 476288), *P. aquatica* in lane 15 (20 ng; PI 303825), *P. arundinacea* in lane 16 (20 ng; PI 241065) and *P. arundinacea* in lane 17 (20 ng; PI 422030). Panel b consists of herbarium specimens of uniform 50 ng quantity where lane 21 is *P. canariensis* (71226), lane 22 is *P. californica* (ISC-V-0021043), lane 23 is *P. californica* (ISC-V-0021040), lane 24 is *P. caroliniana* (ISC-V-0021097), lane 25 is *P. paradoxa* (ISC-V-0021360), lane 26 is *P. coerulescens* (ISC-V-0021199) and lane 27 is *P. minor* (ISC-V-0021338). Panel c contains PCR re-amplification results of 1/50 dilutions of purified PCR reactions with *P. canariensis* in lane 31 (484712), *P. californica* in lane 32 (ISC-V-0021043), *P. californica* in lane 33 (ISC-V-0021040), *P. caroliniana* in lane 34 (ISC-V-0021081), *P. truncata* in lane 35 (ISC-V-0021373), *P. truncata* in lane 36 (ISC-V-0021384) and *P. paradoxa* in lane 37 (ISC-V-0021361). Lanes 9, 18, 28 and 38 are control lanes with no sample loaded. Lanes 1, 10, 19, 20, 29, 30 and 39 are a DNA size marker (FullRanger DNA ladder 100 bases). Dotted white line over image illustrate where the image was cropped to form Fig. 3.

Additional file 3: Supplementary Figure 3. Sequence alignment used for the distance tree in Fig. 4 with indication of specific single nucleotide polymorphism for *Phalaris* species.

Additional file 4: Supplementary Figure 4. Contrast between fresh *P. arundinacea* (PI 422030; MN811174.1) in panel a with a historic herbarium specimen of *P. arundinacea* (753,216; MN811176.1) from the University of Minnesota Herbarium in panel b.

Acknowledgements

We sincerely thank Drs. Timothy Whitfeld and Deborah Lewis for their time and work spent in preserving and providing the availability of the herbaria specimens. Furthermore, we thank the University of Minnesota Bell Museum Herbarium and Iowa State University Ada Hayden Herbarium for the institutional lending services provided.

Authors' contributions

NOA directed the entire research project. NOA, AKN and AGS designed the research. ALG and AKN performed the experimentation, AKN and ALG analyzed the data. ALG, NOA, AKN wrote the complete manuscript with equal contributions; AGS edited the manuscript along with NOA, AKN and ALG. All authors read and approved the final draft of the manuscript.

Funding

Funding for this research was from a University of Minnesota Undergraduate Research Opportunities Program Grant which paid for Allison Graper's salary, limited lab supplies and the Minnesota Agricultural Experiment Station, MAES21-0045, which funded sequencing and greenhouse charges.

Availability of data and materials

The DNA sequences produced are deposited in the NCBI GenBank data (<https://www.ncbi.nlm.nih.gov/nuccore/>). Sequence lists can be found in the Materials and Methods section (Tables 1 and 2) and are listed in the format below of, "***Phalaris* Species**: herbarium accession code or GRIN accession number, GenBank accession number.". Specimens without a GenBank accession number are labeled with "n/a".

***P. angusta* Nees ex Trin:** ISC-V-0020926, MN811167.1; ISC-V-0020921, MN8111655.1; ISC-V-0020922, MN811168.1; ISC-V-0020920, MN811166.1; ISC-V-0020919, MN811169.1.

***P. aquatica* L.:** ISC-V-0021399, MN811171.1; ISC-V-0020927, MN811170.1; ISC-V-0021398, MN811172.1; PI 476287, MN811177.1; PI 476288, n/a; PI 303825, MN811173.1.

***P. arundinacea* L.:** 71166, n/a; 532148, MN811175.1; 753216, MN811176.1; 484712, MN811200.1; PI 241065, n/a; PI 422030, MN811174.1.

***P. brachystachys* Link:** ISC-V-0021036, MN811180.1; ISC-V-0021039, MN811179.1; ISC-V-0021035, MN811178.1; ISC-V-0021037, MN811181.1.

***P. californica* Hook. & Arn.:** ISC-V-0021040, n/a; ISC-V-0021044, n/a; ISC-V-0021043, MN811182.1; ISC-V-0021041, n/a; ISC-V-0021042, n/a.

***P. canariensis* L.:** 71226, n/a; 71229, n/a; 367474, MN811185.1; 619107, n/a; PI 578800, MN811183.1; PI 578798, MN811184.1.

***P. caroliniana* Walter:** ISC-V-0021097, MN811188.1; ISC-V-0021081, MN811187.1; ISC-V-0021166, n/a; ISC-V-0021080, MN811186.1.

***P. coerulescens* Desf.:** ISC-V-0021199, MN811190.1; ISC-V-0021198, n/a; ISC-V-0021204, n/a; ISC-V-0021203, MN811189.1.

***P. lemmonii* Vasey:** ISC-V-0021333, n/a; ISC-V-0021334, n/a; ISC-V-0021329, MN811192.1; ISC-V-0021336, MN811191.1; ISC-V-0021328, n/a.

***P. minor* Retz.:** ISC-V-0021344, n/a; 229774, n/a; ISC-V-0021338, n/a; ISC-V-0021342, n/a; ISC-V-0021341, MN811193.1.

***P. paradoxa* L.:** ISC-V-0021360, n/a; ISC-V-0021363, n/a; ISC-V-0021361, MN811195.1; ISC-V-0021362, n/a; ISC-V-0021415, MN811194.1.

***P. truncata* Guss. ex Bertol.:** ISC-V-0021377, MN811198.1; ISC-V-0021384, MN811199.1; ISC-V-0021395, MN811197.1; ISC-V-0021373, MN811196.1.

Declarations

Ethics approval and consent to participate

All *Phalaris* plant material was obtained from either herbaria leaf or seed samples. Dried leaf samples were hand sampled from herbaria voucher specimens archived at the Bell Museum Herbarium (MIN; University of Minnesota) and Ada Hayden Herbarium (ISC; Iowa State University). The accession numbers of each herbarium voucher specimen are listed in Table 1. Approval for destructive leaf sampling was provided by the respective curators of each herbarium. Seed samples were obtained from the U.S. Department of Agriculture Germplasm Resources Information Network with permission granted by this agency for conducting genetic (molecular) research. Experimental research on these plant materials (either cultivated or wild) comply with relevant institutional, national and international guidelines and legislation.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Horticultural Science, University of Minnesota, 1970 Folwell Avenue, Saint Paul, MN 55108, USA. ²Present Address: Quality Control Analyst III, Aldevron, ND, Fargo, USA.

Received: 1 December 2020 Accepted: 18 October 2021

Published online: 06 November 2021

References

- Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD, et al. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Ann Mo Bot Gard*. 1993;80:528–80.
- Rieseberg LH, Wendel JF. Introgression and its consequences in plants. In: Harrison R, editor. Hybrid zones and the evolutionary process. Oxford: Oxford University Press; 2000. p. 70–109.
- Baldwin BG. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. *Mol Phylogenet Evol*. 1992;1:3–16.
- Baldwin BG. Molecular phylogenetics of *Calycadenia* (Compositae) based on ITS sequences of nuclear ribosomal DNA: chromosomal and morphological evolution reexamined. *Am J Bot*. 1993;80:222–38.
- Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Ann Mo Bot Gard*. 1995;82:247–77.
- Alvarez I, Wendel JF. Ribosomal ITS sequences and plant phylogenetic inference. *Mol Phylogenet Evol*. 2003;29:417–34. [https://doi.org/10.1016/S1055-7903\(03\)00208-2](https://doi.org/10.1016/S1055-7903(03)00208-2).
- Anderson DE. Taxonomy and distribution of the genus *Phalaris*. *Iowa State Collect J Sci*. 1961;36:1–96.
- Casler MD, Cherney JH, Brummer EC. Biomass yield of naturalized populations and cultivars of reed canary grass. *BioEnergy Res*. 2009;2:165–73. <https://doi.org/10.1007/s12155-009-9043-0>.

9. Casler MD, Phillips MM, Krohn AL. DNA polymorphisms reveal geographic races of reed canarygrass. 2009. <https://pubag.nal.usda.gov/catalog/37445>. Accessed 1 Oct 2019.
10. Doering E, Albrecht J, Hilu KW, Roeser M. Phylogenetic relationships in the Aveneae/Poaceae complex (Pooideae, Poaceae). Kew Bull. 2007;62:407–24.
11. Quintanar A, Castroviejo S, Catalan P. Phylogeny of the tribe Aveneae (Poideae, Poaceae) inferred from plastid TRNT-F and nuclear ITS sequences. Am J Bot. 2007;94:1554–69.
12. Schneider J, Doering E, Hilu KW, Roeser M. Phylogenetic structure of the grass subfamily Pooideae based on comparison of plastid matK gene-3' trnK exon and nuclear ITS sequences. Taxon. 2009;58:405–24.
13. Saarela JM, Wysocki WP, Barrett CF, Soreng RJ, Davis JL, Clark LG, et al. Plastid phylogenetics of the cool-season grass subfamily: clarification of relationships among early-diverging tribes. AoB Plants. 2015;7:plv046. <https://doi.org/10.1093/aobpla/plv046>.
14. Soreng RJ, Peterson PM, Romaschenko K, Davidse G, Zuloaga FO, Judziewicz EJ, et al. A worldwide classification of the Poaceae (Gramineae). J Syst Evol. 2015;53:117–37. <https://doi.org/10.1111/jse.12150>.
15. Voshell SM, Baldini RM, Hilu KW. Infrageneric treatment of *Phalaris* (canary grasses, Poaceae) based on molecular phylogenetics and floret structure. Aust Syst Bot. 2015;28:355–67.
16. Jenkin TJ, Sethi BL. *Phalaris arundinacea*, Ph. Tuberosa, their F1 hybrids and hybrid derivatives. J Genet. 1932;26:1–38.
17. Baldini RM. The genus *Phalaris* L. (Gramineae) in Italy. Webbia. 1993;47:1–53. <https://doi.org/10.1080/00837792.1993.10670531>.
18. Baldini RM. Revision of the genus *Phalaris* L. (Gramineae). Webbia. 1995;49:265–329.
19. Voshell S, Baldini R, Kumar R, Tatalovich N, Hilu K. Canary grasses (*Phalaris*, Poaceae): molecular phylogenetics, polyploidy and floret evolution. Taxon. 2011;60:1306–16. <https://doi.org/10.1002/tax.605007>.
20. Voshell SM, Hilu KW. Canary grasses (*Phalaris*, Poaceae): biogeography, molecular dating and the role of floret structure in dispersal. Mol Ecol. 2014;23:212–24.
21. Abbott RJ, Smith LC, Milne RI, Crawford RMM, Wolff K, Balfour J. Molecular analysis of plant migration and refugia in the Arctic. Science. 2000;289:1343–6. <https://doi.org/10.1126/science.289.5483.1343>.
22. Matthioli PA. Petri Petri Andreae Matthioli medici senensis commentarii. In: Libros sex pedacii dioscoridis anazarbei, de materia medica, adjectis quam plurimum plantarum & animalium imaginibus, eodem autore; 1554.
23. Bauhin C. Pinax Theatri botanici. Basileae: Ludovici Regis; 1623.
24. Baldini RM, Jarvis CE. Typification of some Linnaean names in *Phalaris* (Gramineae). Taxon. 1991;40:475–85.
25. Anderson NO. Throwing out the bathwater but keeping the baby: Lessons learned from purple loosestrife and reed canarygrass. HortTech. 2019. <https://doi.org/10.21273/HORTTECH04307-19>.
26. Blake ST. A synthetic new species of *Phalaris* (Gramineae). Proc R Soc Qld. 1956;67:27–9.
27. Galatowitsch SM, Anderson NO, Ascher PD. Invasiveness in wetland plants in temperate North America. Wetlands. 1999;19:733–55. <https://doi.org/10.1007/BF03161781>.
28. Lavergne S, Molofsky J. Reed canary grass (*Phalaris arundinacea*) as a biological model in the study of plant invasions. Crit Re Plant Sci. 2004;23:415–29. <https://doi.org/10.1080/07352680490505934>.
29. Noyszewski AK, Anderson NO, Smith AG, Kilian A, Dalbotten D, Ito E, et al. Challenges of establishing native versus exotic status of herbarium specimens. HortTech. 2019;29:549–53. <https://doi.org/10.21273/HORTTECH04313-19>.
30. Nelson MF, Anderson NO. Variation among genotypes and source habitats in growth and fecundity of the wetland invasive plant *Phalaris arundinacea* L. Wetlands. 2015;35:1175–84. <https://doi.org/10.1007/s13157-015-0704-9>.
31. Akiyama Y, Kimura K, Kubota A, Fujimori M, Yamada-Akiyama H, Takahara Y, et al. Comparison of genome size in reed canarygrass (*Phalaris arundinacea* L.) exotic and putative native Japanese genotypes by flow cytometry. Jpn Agric Res Q. 2015;49:345–50. <https://doi.org/10.6090/jarq.49.345>.
32. Nelson MF, Anderson NO, Casler MD, Jakubowski AR. Population genetic structure of N. American and European *Phalaris arundinacea* L. as inferred from inter-simple sequence repeat markers. Biol Inv. 2014;16:353–63. <https://doi.org/10.1007/s10530-013-0525-9>.
33. Jakubowski A, Jackson R, Casler M. The history of reed canarygrass in North America: persistence of natives among invading Eurasian populations. Crop Sci. 2014;54:210. <https://doi.org/10.2135/cropsci2013.05.0342>.
34. Jakubowski AR, Casler MD, Jackson RD. Genetic evidence suggests a widespread distribution of native north American populations of reed canarygrass. Biol Inv. 2013;15:261–8. <https://doi.org/10.1007/s10530-012-0300-3>.
35. Anderson NO, Jocienė L, Krokaitė E, Rekašius T, Paulauskas A, Kupčinskienė E. Genetic diversity of *Phalaris arundinacea* populations in relation to river regulation in the Merkys basin, Lithuania. River Res Appl. 2018;34:300–9. <https://doi.org/10.1002/rra.3259>.
36. Noyszewski AK, Anderson NO, Smith AG, Kilian A, Dalbotten D, Ito E, et al. Riparian populations of Minnesota reed canarygrass (*Phalaris arundinacea*) are most likely native, based on SNPs (DARtseqLD). Wetl Ecol Mgt. 2020; [In Review].
37. Besnard G, Bianconi ME, Hackel J, Manzi S, Vorontsova MS, Christin P-A. Herbarium genomics retraces the origins of C4-specific carbonic anhydrase in Andropogoneae (Poaceae). Bot Lett. 2018;165:419–33. <https://doi.org/10.1080/23818107.2018.1469429>.
38. Lavoie C, Dufresne C, Delisle F. The spread of reed canary grass (*Phalaris arundinacea*) in Québec: a spatio-temporal perspective. Écosci. 2005;12(3):355–75.
39. Merigliano MF, Lesica P. The native status of reed canarygrass (*Phalaris arundinacea* L.) in the inland northwest, USA. Nat Areas J. 1998;18:223–30.
40. Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. PLoS One. 2012;7(3):e34131. <https://doi.org/10.1371/journal.pone.0034131>.
41. Stiller M, Green RE, Ronan M, Simons JF, Du L, He W, et al. Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. PNAS. 2006;135:78–84. <https://doi.org/10.1073/pnas.0605327103>.
42. Sayers EW, Cavanaugh M, Clark K, Ostell J, Pruitt KD, Karsch-Mizrachi I. GenBank. Nucleic Acids Res. 2020;48(D1):D84–6. <https://doi.org/10.1093/nar/gkz956>.
43. Wandeler P, Smith S, Morin PA, Pettifor RA, Funk SM. Patterns of nuclear DNA degeneration over time—a case study in historic teeth samples. Mol Ecol. 2003;12:1087–93. <https://doi.org/10.1046/j.1365-294x.2003.01807.x>.
44. Kneibelsberger T, Stöger I. DNA extraction, preservation, and amplification. In: Kress WJ, Erickson DL, editors. DNA barcodes: methods and protocols; methods in molecular biology. Totowa: Humana Press; 2012. p. 311–38. https://doi.org/10.1007/978-1-61779-591-6_14.
45. Savolainen MK, Loikkanen J, Naarala J. Amplification of glutamate-induced oxidative stress. Toxicol Lett. 1995;82–83:399–405. [https://doi.org/10.1016/0378-4274\(95\)03490-0](https://doi.org/10.1016/0378-4274(95)03490-0).
46. Ribeiro RA, Lovato MB. Comparative analysis of different DNA extraction protocols in fresh and herbarium specimens of the genus *Dalbergia*. Genet Mol Res GR. 2007;6:173–87.
47. Staats M, Cuenca A, Richardson JE, Ginkel RV, Petersen G, Seberg O, et al. DNA damage in plant herbarium tissue. PLoS One. 2011;6:e28448. <https://doi.org/10.1371/journal.pone.0028448>.
48. Do D, Závěská DL. Herbarium tale: the utility of dry specimens for DNA barcoding Juncaceae. Plant Syst Evol. 2018;304:281–94. <https://doi.org/10.1007/s00606-017-1476-x>.
49. Erkens RHJ, Cross H, Maas JW, Hoenselaar K, Chatrou LW. Assessment of age and greenness of herbarium specimens as predictors for successful extraction and amplification of DNA. Blumea Biodiv Evol Biogeo Plants. 2008;53(2):407–28 (22). <https://doi.org/10.3767/000651908X608052>.
50. Miller CR, Joyce P, Waits LP. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics. 2002;160:357–66.
51. Xu C, Dong W, Shi S, Cheng T, Li C, Liu Y, et al. Accelerating plant DNA barcode reference library construction using herbarium specimens: improved experimental techniques. Mol Ecol Resour. 2015;15:1366–74. <https://doi.org/10.1111/1755-0998.12413>.
52. Cheng T, Xu C, Lei L, Li C, Zhang Y, Zhou S. Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity. Mol Ecol Resour. 2016;16:138–49. <https://doi.org/10.1111/1755-0998.12438>.
53. Metzker ML. Emerging technologies in DNA sequencing. Genome Res. 2005;15:1767–76. <https://doi.org/10.1101/gr.3770505>.

54. Brewer GE, Clarkson JJ, Maurin O, Zuntini AR, Barber V, Bellot S, et al. Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. *Front Pl Sci*. 2019;10:1102. <https://doi.org/10.3389/fpls.2019.01102>.
55. Kuzmina ML, Thomas WA, Braukmann AJ, Fazekas SW, Graham SL, Roodrigues A, et al. Using herbarium-derived DNAs to assemble a large-scale DNA barcode library for the vascular plants of Canada. *Appl Pt Sci*. 2017;5(12):1700079. <https://doi.org/10.3732/apps.1700079>.
56. Winterfeld G, Becher H, Voshell S, Hillu K, Röser M. Karyotype evolution in *Phalaris* (Poaceae): the role of reductional dysploidy, polyploidy and chromosome alteration in a wide-spread and diverse genus. *PLoS One*. 2018;13. <https://doi.org/10.1371/journal.pone.0192869>.
57. CBOL Plant Working Group, Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, et al. A DNA barcode for land plants. *PNAS*. 2009;106:12794–7. <https://doi.org/10.1073/pnas.0905845106>.
58. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinforma Oxf Engl*. 2012;28:1647–9. <https://doi.org/10.1093/bioinformatics/bts199>.
59. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol*. 1987;4:406–25. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>.
60. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32:1792–7. <https://doi.org/10.1093/nar/gkh340>.
61. Wolf M. ITS so much more. *Trends Gen*. 2015;31(4):174–5. <https://doi.org/10.1016/j.tig.2015.02.005>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

