bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Towards Improved Molecular Identification Tools in Fine Fescue (*Festuca* L., Poaceae) Turfgrasses: Nuclear Genome Size, Ploidy, and Chloroplast Genome Sequencing

Yinjie Qiu[1,*], Cory D. Hirsch[2], Ya Yang[3] and Eric Watkins[1,*]

[1] Department of Horticultural Science, University of Minnesota, St. Paul, MN, 55108, USA; qiuxx221@umn.edu (Y.Q.); ewatkins@umn.edu (E.W.)

2 Department of Plant Pathology, University of Minnesota, St. Paul, MN, 55108, USA; cdhirsch@umn.edu (C. H.)

3 Department of Plant and Microbial Biology, University of Minnesota, St. Paul, MN, 55108, USA; yayang@umn.edu (Y.Y.)

* Correspondence: qiuxx221@umn.edu (Y.Q.); ewatkins@umn.edu (E.W.); Tel.: +1- 612-624-7496 (E.W.)

Abstract: Fine fescues (*Festuca* L., Poaceae) are turfgrass species that perform well in low-input environments. Based on morphological characteristics, the most commonly-utilized fine fescues are divided into five taxa: three are subspecies within *F. rubra* L. and the remaining two are treated as species within the *F. ovina* L. complex. Morphologically, these five taxa are very similar, both identification and classification of fine fescues remain challenging. In an effort to develop identification methods for fescues, we used flow cytometry to estimate genome size, ploidy level, and sequenced the chloroplast genome of all five taxa. Fine fescue chloroplast genome sizes ranged from 133,331 to 133,841 bp and contained 113 to 114 genes. Phylogenetic relationship reconstruction using whole chloroplast genome sequences agreed with previous work based on morphology. Comparative genomics suggested unique repeat signatures for each fine fescue taxon that could potentially be used for marker development for taxon identification.

Keywords: Fine fescue, chloroplast genome, phylogeny, comparative genomics

---

1. Introduction

With ca. 450 species, Fescues (*Festuca* L., Poaceae) is a large and diverse genus of perennial grasses (Clayton and Renvoize 1986). Fescue species are distributed mostly in temperate zones of both the northern and southern hemispheres, but most commonly found in the northern hemisphere (Jenkin 1959). Several of the fescue species have been commonly used as turfgrass. Based on both leaf morphology and nuclear ITS sequences, fescue species can be divided into two groups: broad-leaved fescues and fine-leaved fescues (Torrecilla and Catalán 2002). Broad-leaved fescues commonly used as turfgrass include tall fescue (*Festuca arundinacea* Schreb.) and meadow fescue (*Festuca pratensis* Huds.). Fine-leaved fescues are a group of cool-season grasses that include five commonly used taxa called fine fescues. Fine fescues include hard fescue (*Festuca brevipila* Tracey, 2n=6x=42), sheep fescue (*Festuca ovina* L., 2n=4x=28), strong creeping red fescue (*Festuca rubra* ssp. *rubra* 2n=8x=56), slender creeping red fescue (*Festuca rubra* ssp. *litoralis* (G. Mey.) Auquier 2n=6x=42), and Chewings fescue (*Festuca rubra* ssp. *fallax* (Thuill.) Nyman 2n=6x=42) (Ruemmele et al. 1995). All five taxa share very fine and narrow leaves and have been used for forage, turf, and ornamental purposes. They are highly tolerant to shade and drought, prefer low pH (5.5-6.5) and low fertility soils (Beard 1972). Additionally, fine fescues grow well in the shade or sun, have reduced mowing requirements, and do not need additional fertilizer or supplemental irrigation (Ruemmele et al. 1995).

Based on morphological and cytological features, fine fescues are currently divided into two groups referred to as the *F. rubra* complex (includes *F. rubra* ssp. *litoralis*, *F. rubra* ssp.

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

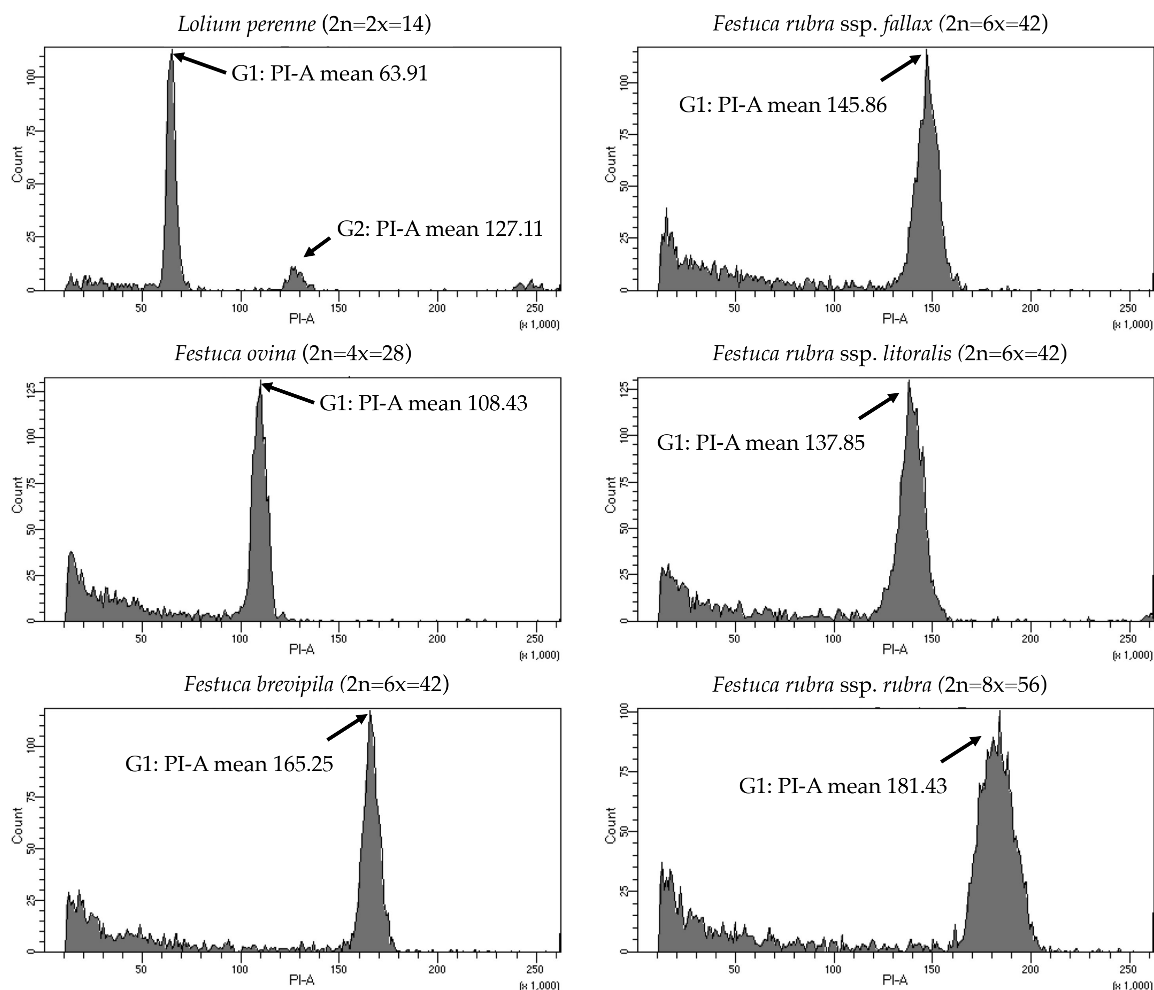
47 *rubra*, *F. rubra* ssp. *fallax*) and the non- rhizomatous *F. ovina* complex (includes *F. brevipila*
48 and *F. ovina*) (Ruemmele et al. 1995). While it is relatively easy to separate fine fescue taxa
49 into their proper complex based on the presence and absence of rhizome, it is challenging
50 to identify taxon within the same complex. In the *F. rubra* complex, both ssp. *litoralis* and
51 ssp. *rubra* are rhizomatous while ssp. *fallax* is non-rhizomatous. However, the separation of
52 ssp. *litoralis* from ssp. *rubra* using rhizome length is challenging. Taxon identification within
53 the *F. ovina* complex relies heavily on leaf characters; however, abundant morphological and
54 ecotype diversity within *F. ovina* makes taxa identification difficult (Piper 1906). This is
55 further complicated by inconsistent identification methods between different continents.
56 For example, in the United States, sheep fescue is described as having a bluish gray leaf color
57 and hard fescue leaf blade color is considered green (Beard 1972), while in Europe, it is the
58 opposite (Hubbard 1968). Because the ploidy level of the five taxa varies from tetraploid to
59 octoploid, beyond morphological classifications, laser flow cytometry has been used to
60 determine ploidy level of fine fescues and some other fescue species (Huff and Palazzo
61 1998). A wide range of DNA contents within each complex suggests that the evolutionary
62 history of each named species is complicated, and interspecific hybridization might interfere
63 with species determination using this approach. Plant breeders have been working to
64 improve fine fescues for turf use for several decades, with germplasm improvement efforts
65 focused on disease resistance, traffic tolerance, and ability to perform well under heat stress
66 (Casler 2003). Turfgrass breeders have utilized germplasm collections from old turf areas as
67 a source of germplasm (Bonos and Huff 2013); however, confirming the taxon identity in
68 these collections has been challenging. A combination of molecular markers and flow
69 cytometry could be a valuable tool for breeders to identify fine fescue germplasm (Hebert
70 et al. 2003).
71 Due to the complex polyploidy history of fine fescues, sequencing plastid genomes
72 provides a more cost-effective tool for taxon identification than the nuclear genome because
73 it is often maternally inherited, lacks of heterozygosity, is present in high copies and usable
74 even in partially degraded material (Bryan et al. 1999, Provan et al. 2001). Previous studies
75 have developed universal polymerase chain reaction (PCR) primers to amplify non-coding
76 polymorphic regions for DNA barcoding in plants for species identification (Baldwin et al.
77 1995, Demesure et al. 1995). However, the polymorphisms discovered from these regions
78 are often single nucleotide polymorphisms that are difficult to apply using PCR screening
79 methods. For these reasons, it would be helpful to assemble chloroplast genomes and
80 identify simple sequence repeat (SSR) and tandem repeats polymorphisms. Chloroplast
81 genome sequencing has been simplified due to improved sequencing technology. In
82 turfgrass species, high throughput sequencing has been used to assemble the chloroplast
83 genomes of perennial ryegrass (*Lolium perenne* cv. Cashel) (Diekmann et al. 2009), tall fescue
84 (*Lolium arundinacea* cv. Schreb) (Cahoon et al. 2010), diploid *Festuca ovina*, *Festuca pratensis*,
85 *Festuca altissima* (Hand et al. 2013), and bermudagrass (*Cynodon dactylon*) (Huang et al. 2017).
86 To date, there is limited molecular biology information on fine fescue taxon identification
87 and their phylogenetic position among other turfgrass species (Hand et al. 2013, Cheng et
88 al. 2016). In this study, we used flow cytometry to confirm the ploidy level of five fine fescue
89 cultivars, each representing one of the five commonly utilized fine fescue taxon. We then
90 reported the complete chloroplast genome sequences of these five taxa, carried out
91 comparative genomics and phylogenetic inference. Based on the genome sequence we
92 identified unique genome features among fine fescue taxa and predicted taxon specific SSR
93 and tandem repeat loci for molecular marker development.

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

94 **2. Results**

95 *2.1 Species Ploidy Level Confirmation*

96 We used flow cytometry to estimate the ploidy levels of five fine fescue taxa by
97 measuring the DNA content in each nucleus. DNA content was reflected by the flow
98 cytometry mean PI-A value. Overall, fine fescue taxa had mean PI-A values roughly from
99 110 to 180 (**Figure 1 and Figure S1**). *Festuca rubra* ssp. *rubra* cv. Navigator II (2n=8x=56) had
100 the highest mean PI-A value (181.434, %rCV 4.4). *Festuca rubra* ssp. *litoralis* cv. Shoreline
101 (2n=6x=42) and *F. rubra* ssp. *fallax* cv. Treazure II (2n=6x=42) had similar mean PI-A values
102 of 137.852, %rCV 3.7 and 145.864, %rCV 3.5, respectively. *Festuca brevipila* cv. Beacon
103 (2n=6x=42) had a mean PI-A of 165.25, %rCV 1.9, while *F. ovina* cv. Quatro (2n=4x=28) had a
104 mean PI-A of 108.43, %rCV 2.9. Standard reference *L. perenne* cv. Artic Green (2n=2x=14) had
105 a G1 phase mean PI-A of 63.91, %rCV 3.0. USDA *F. ovina* PI 230246 (2n=2x=14) had a G1
106 mean PI-A of 52.73 (histogram not shown). The estimated genome size of USDA PI 230246
107 was 4.67 pg/2C. Estimated ploidy level of *F. brevipila* cv. Beacon was 6.3, *F. ovina* cv. Quatro
108 was 4.11, *F. rubra* ssp. *rubra* cv. Navigator II was 6.9, *F. rubra* ssp. *litoralis* cv. Shoreline was
109 5.2, and *F. rubra* ssp. *fallax* cv. Treazure II was 5.5 (**Table 1**). All newly estimated ploidy
110 levels roughly correspond to previously reported ploidy levels based on chromosome
111 counts.



112

113 **Figure 1.** Flow cytometry results for the five fine fescue taxa. *Lolium perenne* (2n=2x=14) was used as
114 the reference. Flow cytometry was able to separate *F. rubra* ssp. *rubra* from the other two subspecies

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

115 in the *F. rubra* complex. The mean PI-A values of *F. rubra* ssp. *fallax* and *F. rubra* ssp. *litoralis* were
116 similar (145.86 to 137.85).

117 **Table 1.** Summary of flow cytometry statistics, genome size estimation, and ploidy level estimation
118 of fine fescue species. *Lolium perenne* 2C DNA content was used to calculate fine fescue and USDA *F.*
119 *ovina* PI 230246 genome size, calculated PI 230246 DNA content was used as reference to infer fine
120 fescue ploidy level

Species name	Chromosome count	Cultivar name	Mean PI-A	%rCV *	Estimated Genome Size (pg/Nuclei)	Estimated Ploidy Level
F. brevipila	2n=6x=42	Beacon	165.3	1.9	14.6	6.3
F. ovina	2n=4x=28	Quatro	108.4	2.9	9.6	4.1
F. ovina PI 230246	2n=2x=14	NA	52.7	3.1	4.7	1.7
F. rubra ssp. rubra	2n=8x=56	Navigator II	181.4	4.4	16.1	6.9
F. rubra ssp. litoralis	2n=6x=42	Shoreline	137.9	3.7	12.2	5.2
F. rubra ssp. fallax	2n=6x=42	Treazure II	145.9	3.5	12.9	5.5
L. perenne	2n=2x=14	Artic Green	63.9	3.0	5.7	2.0

121 * %rCV: Quality of laser alignment. Low %rCV suggests high resolution sensitivity.

122 *2.2 Plastid Genome Assembly and Annotation of Five Fescue Taxa*

123 A total of 47,843,878 reads were produced from the five fine fescue taxa. After Illumina
124 adaptor removal, we obtained 47,837,438 reads. The assembled chloroplast genomes ranged
125 from 133,331 to 133,841 bp. The large single copy (LSC) and small single copy (SSC) regions
126 were similar in size between the sequenced fine fescue accessions (78 kb and 12 kb,
127 respectively). *Festuca ovina* and *F. brevipila* in the *F. ovina* complex had exactly the same size
128 inverted repeat (IR) region (42,476 bp). In the *F. rubra* complex, *F. rubra* ssp. *rubra* and *F.*
129 *rubra* ssp. *litoralis* had the same IR size (21,235 bp). Species in the *F. rubra* complex had a
130 larger chloroplast genome size compared to species in the *F. ovina* complex. All chloroplast
131 genomes shared similar GC content (38.4%) (**Figure 2**, **Table 2**). The fine fescue chloroplast
132 genomes encoded for 113-114 genes, including 37 transfer RNAs (tRNA), 4 ribosomal RNAs
133 (rRNA), and 72 protein-coding genes (**Table 2**). Genome structures were similar among all
134 five fine fescue taxa sequenced, except that the pseudogene *accD* was annotated in all three
135 subspecies of *F. rubra,* but not in the *F. ovina* complex (**Table S1**).

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

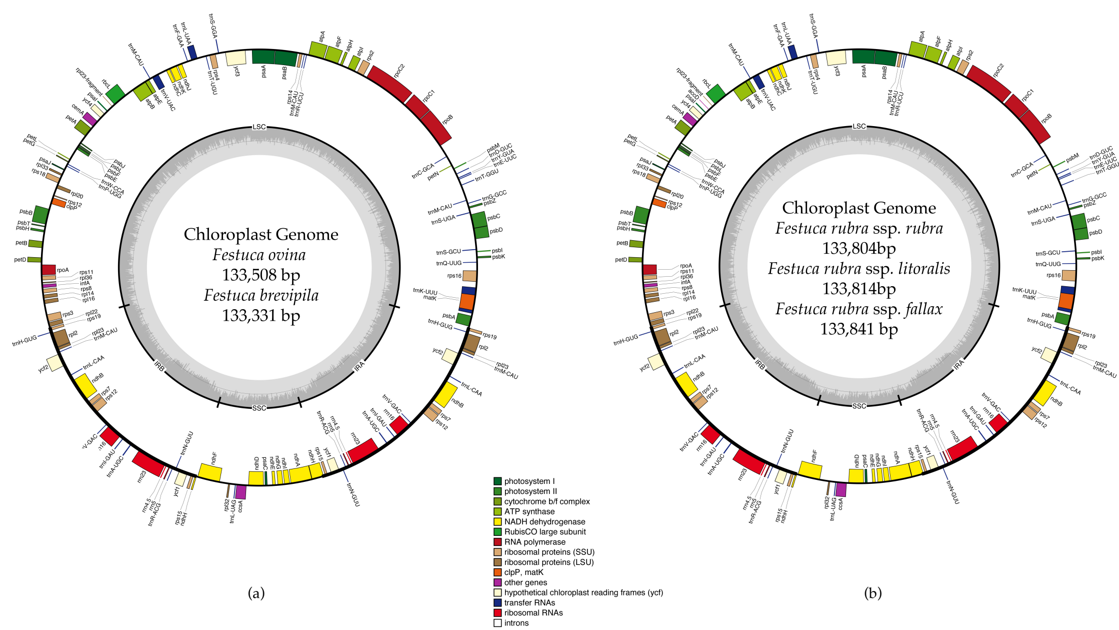


Figure 2. Whole chloroplast genome structure of *F. ovina* complex (a) and *F. rubra* complex (b). Genes inside the circle are transcribed clockwise, genes outside are transcribed counter-clockwise. Genes belong to different functional groups are color coded. GC content is represented by the dark gray inner circle, the light gray corresponded to the AT content. IRA(B), inverted repeat A(B); LSC, large single copy region; SSC, small single copy region.

Table 2. Characteristics of fine fescue chloroplast genomes.

	***F. brevipila* cv. Beacon**	***F. ovina* cv. Quatro**	***F. rubra* ssp. *rubra* cv. Navigator II**	***F. rubra* ssp. *litoralis* cv. Shoreline**	***F. rubra* ssp. *fallax* cv. Treazure II**
NCBI GenBank ID	MN309822	MN309824	MN309825	MN309823	MN309826
Total Genome Size (bp)	133,331	133,508	133,804	133,814	133,841
Large Single Copy (bp)	78,462	78,632	78,888	78,909	78,882
Small Single Copy (bp)	12,393	12,400	12,446	12,435	12,451
Inverted Repeat (bp)	42,476	42,476	42,470	42,470	42,508
Ratio of LSC (%)	58.85	58.9	58.96	58.97	58.94
Ratio of SSC (%)	9.29	9.29	9.3	9.29	9.3
Ratio of IRs (%)	31.86	31.82	31.74	31.74	31.76
GC content (%)	38.4	38.4	38.4	38.4	38.4

2.3 Chloroplast Genome IR Expansion and Contraction

Contraction and expansion of the IR regions resulted in the size variation of chloroplast genomes. We examined the four junctions in the chloroplast genomes, LSC/IRa, LSC/IRb, SSC/IRa, and SSC/IRb of the fine fescue and the model turfgrass species *L. perenne*. Although the chloroplast genome of fine fescue taxa were highly similar, some structural variations

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

were still found in the IR/LSC and IR/SSC boundary. Similar to *L. perenne*, fine fescue taxa chloroplast genes *rpl22-rps19, rps19-psbA* were located in the junction of IR and LSC; *rps15-ndhF* and *ndhH-rps15* were located in the junction of IR/SSC. In the *F. ovina* complex, the *rps19* gene was 37 bp into the LSC/IRb boundary while in the *F. rubra* complex and *L. perenne*, the *rps19* gene was 36 bp into the LSC/IRb boundary (**Figure 3**). The *rsp15* gene was 308 bp from the IRb/SSC boundary in *F. ovina* complex, 307 bp in *F. rubra* complex, and 302 bp in *L. perenne*. Both the *ndhH* and the pseudogene fragment of the *ndhH* (*ѱndhH)* genes spanned the junction of the IR/SSC. The *ѱndhH* gene crossed the IRb/SSC boundary with 32 bp into SSC in *F. brevipila* and *F. ovina*, 9 bp in *F. rubra* ssp. *rubra* and *F. rubra* ssp. *litoralis*, 10 bp in *F. rubra* ssp. *fallax,* and 7 bp in *L. perenne*. The *ndhF* gene was 88 bp from the IRb/SSC boundary in *F. brevipila* and *F. ovina*, 91 bp in *F. rubra* ssp. *rubra*, 84 bp in *F. rubra* ssp. *litoralis*, 77 bp in *F. rubra* ssp. *fallax*, and 102 bp in *L. perenne*. Finally, the *psbA* gene was 87 bp apart from the IRa/LSC boundary into the LSC in *L. perenne* and *F. ovina* complex taxa but 83 bp in the *F. rubra* complex taxa.

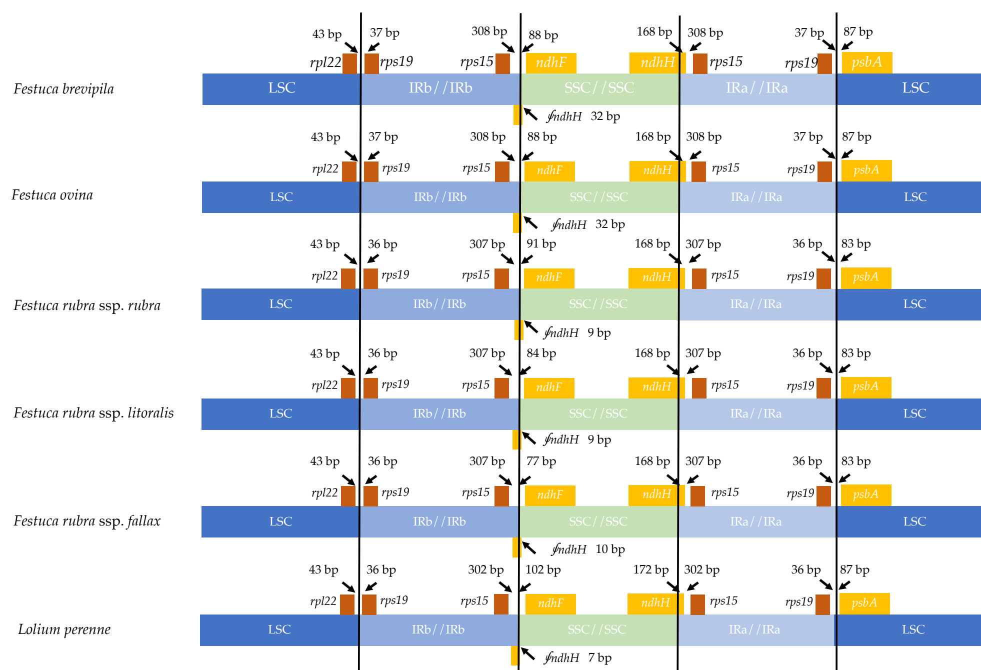


Figure 3. Comparison for border positions of LSC, SSC and IR regions among five fine fescues and *L. perenne*. Genes are denoted by boxes, and the gap between the genes and the boundaries are indicated by the number of bases unless the gene coincides with the boundary. Extensions of genes are also indicated above the boxes.

2.4 Whole Chloroplast Genome Comparison and Repetitive Element Identification

Genome-wide comparison among five fine fescue taxa showed high sequence similarity with most variations located in intergenic regions (**Figure 4**). To develop markers for species screening, we predicted a total of 217 SSR markers for fine fescue taxa sequenced (*F. brevipila* 39; *F. ovina* 45; *F. rubra* ssp. *rubra* 45; *F. rubra* ssp. *litoralis* 46; *F. rubra* ssp. *fallax* 42) that included 17 different repeat types for the fine fescue species (**Figure 5a, Table S2**). The most frequent repeat type was A/T repeats, followed by AT/AT. The pentamer AAATT/AATTT repeat was only presented in the rhizomatous *F. rubra* ssp. *litoralis* and *F. rubra* ssp. *rubra*, while ACCAT/ATGGT was only found in *F. ovina* complex species *F. brevipila* and *F. ovina*. Similar to SSR loci prediction, we also predicted long repeats for the fine fescue species and identified a total of 171 repeated elements ranging in size from 20 to 51 bp (**Figure 5b, Table**

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

S3). Complementary (C) matches were only identified in *F. brevipila* and *F. ovina*. *F. rubra* species had more palindromic (P) and reverse (R) matches. Number of forward (F) matches were similar between five taxa. Selected polymorphic regions were validated by PCR and gel electrophoresis assay (**Figure S2**).

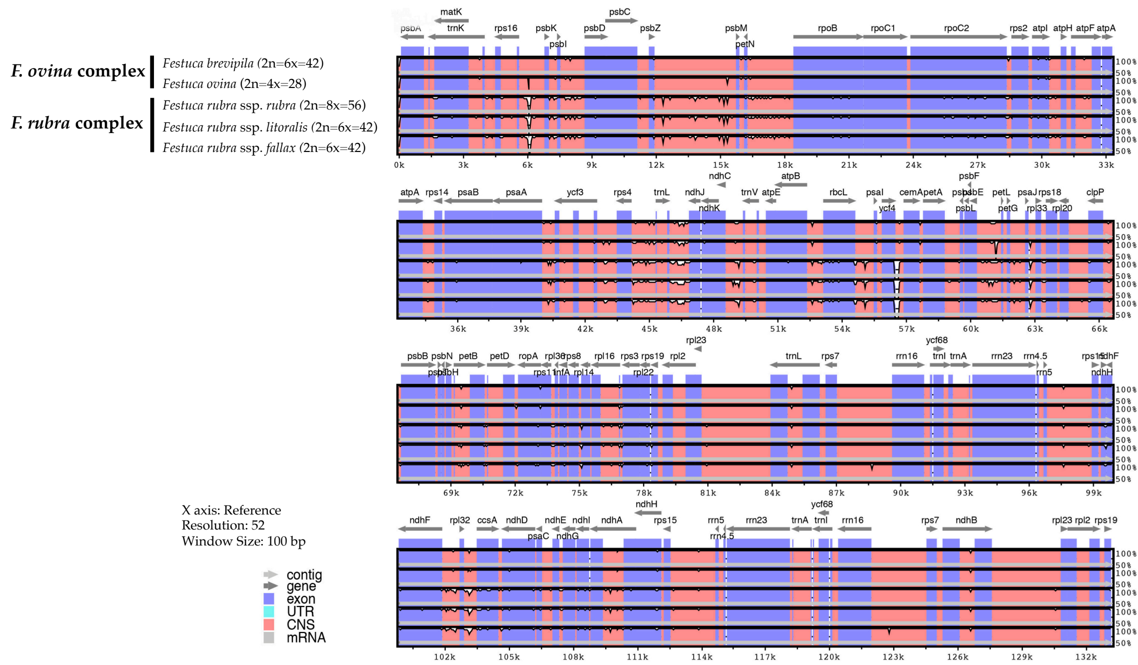


Figure 4. Sequence identity plot of fine fescues chloroplast genome sequences with *F. ovina* (2x) as the reference using mVISTA. A cut-off of 70% identify was used for the plots, and the percent of identity varies from 50% to 100% as noted on the y-axis. Most of the sequence variations between fine fescues were in intergenic regions. Taxa in the *F. ovina* complex, *F. brevipila* and *F. ovina* showed high sequence similarity. Similarly, subspecies within *F. rubra* complex also showed high sequence similarity.

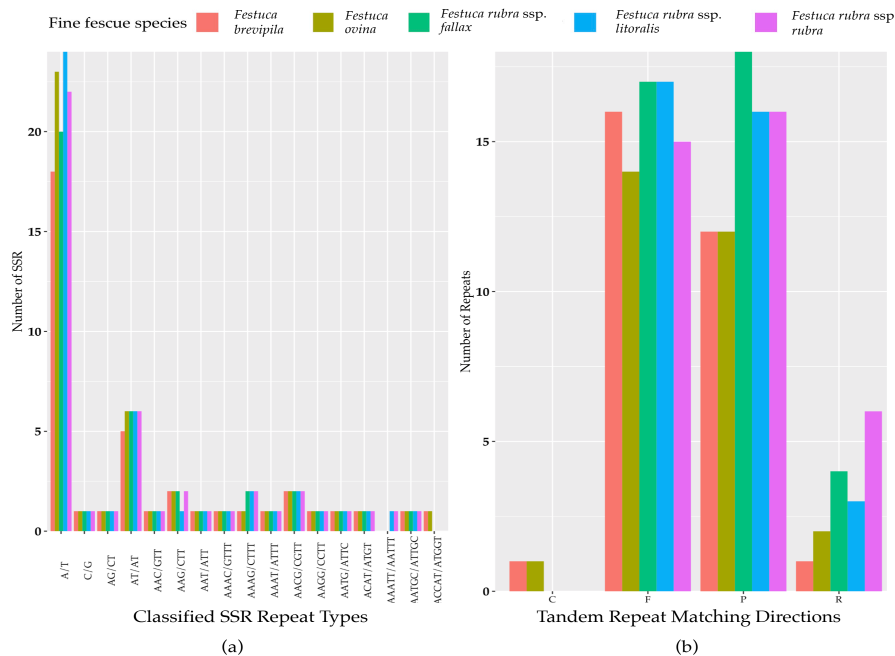


Figure 5. (a) SSR repeat type and numbers in the five fine fescue taxa sequenced. Single nucleotide repeat type has the highest frequency. No hexanucleotide repeats were identified in the fine fescue

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

chloroplast genomes sequenced. One penta-nucleotide repeat type (AAATT/AATTT) is unique to *F. rubra* ssp. *rubra* and *F. rubra* ssp. *litoralis*; One penta-nucleotide repeat type (ACCAT/ATGGT) is unique to *F. brevipila* and *F. ovina*. (b) Long repeats sequences in fine fescue chloroplast genomes. Complement (C) match was only identified in the *F. ovina* complex; Reverse (R) match has the most number variation in fine fescues.

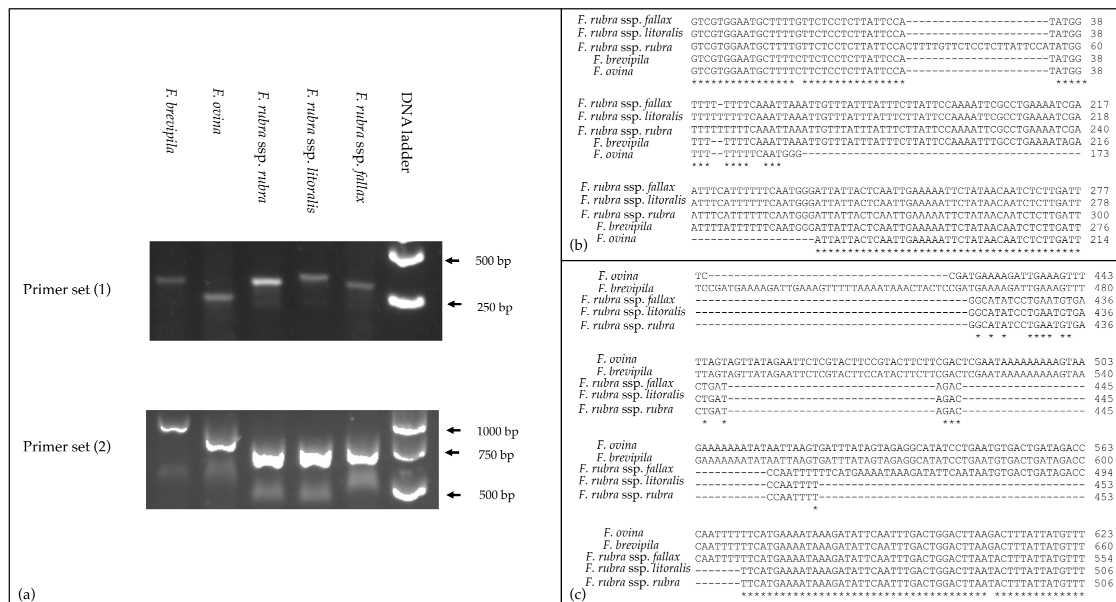


Figure S2. Examples of PCR validation of predicted repeat regions based on fine fescue chloroplast genomes. PCR primers were developed using Primer3 module(Untergasser, Cutcutache et al. 2012). Primers used for the PCR assays (1) Forward primer 5′-GTCGTGGAATGCTTTTGTTCTC-3′; Reverse primer 5′-AGTGGATTCATCAGATGATACA-3′; (2) Forward primer 5′-TTCCTCTTTTCATTG-CAAAGTGGT AT-3′; Reverse primer 5′-TACTCGGAGGTTCGAATCCTTCC-3′. PCR products were examined on 1% agarose gel and gel images showed fragment size separation between different taxa(a). Figure (b) and (c) showed partial sequence alignment of regions amplified by primer sets (1 and 2).

2.5 SNP and InDel Distribution in the Coding Sequence of Five Fine Fescue Species

To identify single nucleotide polymorphisms (SNPs, non-reference allele in this content), we used the diploid *F. ovina* chloroplast genome (JX871940.1) as the reference for the mapping and used the genome annotation file to identify genic and non-genic SNPs. The total genic and non-genic sequence of (JX871940.1) were 60,582 and 72,583 bp, respectively. We found SNP polymorphisms were over-present within intergenic regions in the *F. ovina* complex (~0.3 SNP/Kbp more), while were under-present in the *F. rubra* complex (~0.5 SNP/Kbp less). Most InDels were located in intergenic regions of the fine fescue species (**Table 3**). Between *F. ovina* and the *F. rubra* complex, *the ropC2* gene had the most SNPs (4 vs 31). The *rbcL* gene also has a high level of variation (1 vs 14.3). In addition, *rpoB*, *ccsA*, NADH dehydrogenase subunit genes (*ndhH*, *ndhF*, *ndhA*), and ATPase subunit genes (*atpA*, *atpB*, *aptF*) also showed variation between *F. ovina* and *F. rubra* complexes. Less SNP and InDel variation were found within each complex (**Table 3, Table S4 and S5**).

Table 3. Distribution of SNPs and InDels for the five fine fescue taxa sequenced in this study.

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

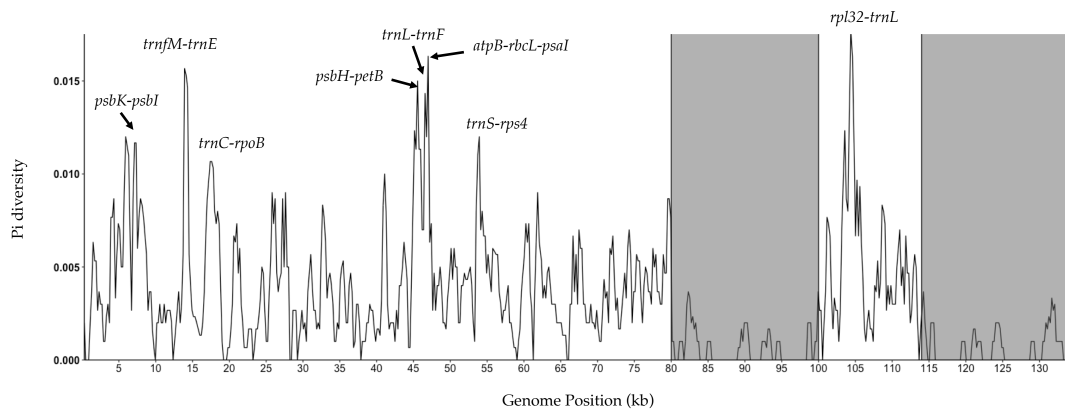
Species	***F. brevipila***	***F. ovina***	***F. rubra* ssp. *rubra***	***F. rubra* ssp. *litoralis***	***F. rubra* ssp. *fallax***
Total number of SNPs	98	134	638	615	624
SNPs in the coding region	35	52	306	301	300
SNPs in intergenic region	63	82	332	314	324
SNPs per Kbp in genic region	0.5777	0.8583	5.0510	4.9685	4.9520
SNPs per Kbp in non-genic region	0.8680	1.1297	4.5741	4.3261	4.4639
Total number of InDels	112	102	149	156	149
InDels in the coding region	22	17	27	26	27
InDels in intergenic region	90	85	122	130	122
Percentage of InDels in the intergenic region	80.36	83.33	81.88	83.33	81.88
Average sequencing depth	171.61	86.81	101.58	77.04	50.94

221 *2.6 Nucleotide Diversity Calculation*

222 A sliding window analysis successfully detected highly variable regions in the fine
223 fescue chloroplast genomes (**Figure 6, Table S6**). The average nucleotide diversity (Pi)
224 among fine fescue taxa was relatively low (0.00318). The IR region showed lower variability
225 than the LSC and SSC region. There were several divergent loci having a Pi value greater
226 than 0.01 (*psbK-psbI*, *trnfM-trnE*, *trnC-rpoB*, *psbH-petB*, *trnL-trnF*, *trnS-rps4*, *aptB-rbcL-psaI*,
227 and *rpl32-trnL*), but mostly within intergenic regions. The *rbcL-psaI* region contained a
228 highly variable *accD-like* region in some fine fescue taxa, so we looked at the structural
229 variation of 10 taxa in the *Festuca - Lolium* complex. We found taxa in broad-leaved fescue
230 and *F. rubra* complex had similar structure, while *F. ovina* (2x, 4x) and *F. brevipila* had a 276
231 bp deletion in the *rbcL-psaI* intergenic region (**Figure 7**).

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

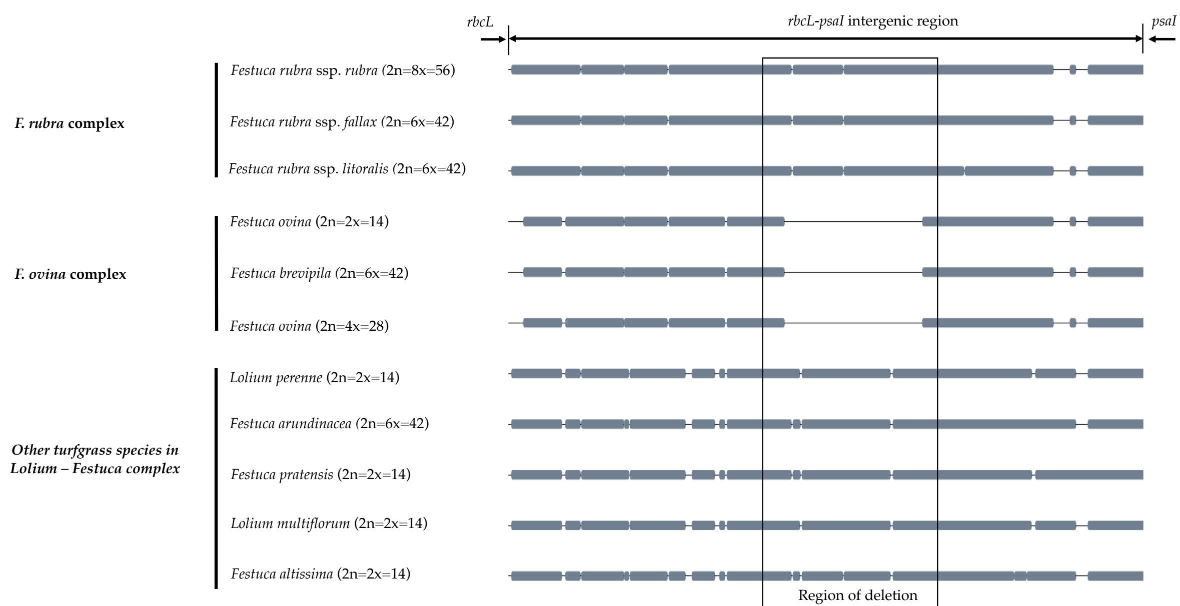
232



233

234 **Figure 6**. Sliding window analysis of fine fescue whole chloroplast genomes. Window size: 600 bp, step
235 size: 200 bp. X-axis: the position of the midpoint of a window (kb). Y-axis: nucleotide diversity of each
236 window. Inverted repeat regions are highlighted in grey. *rpl32-trnL* region has the most nucleotide
237 diversity followed by *psbH- petB-trnL-trnF-trnS-rps4* region.

238



239

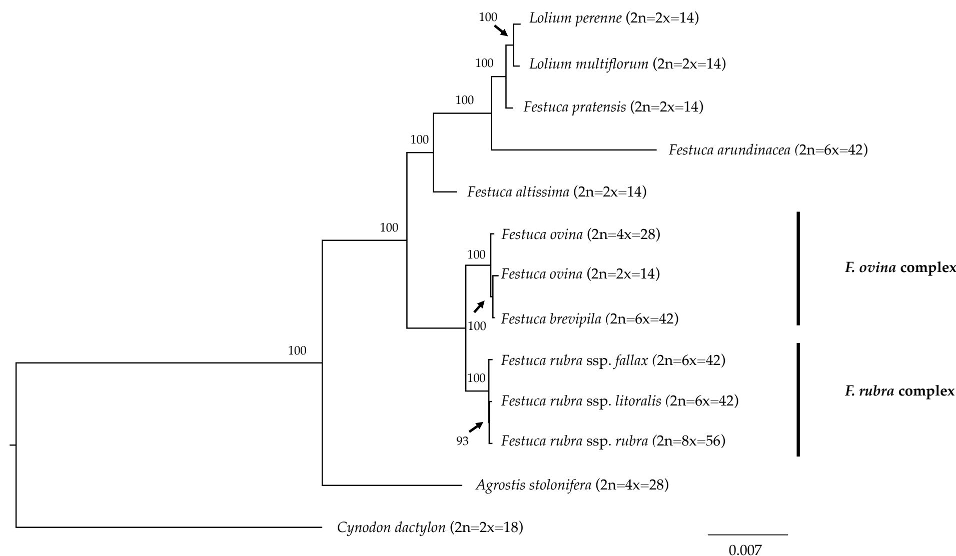
240 **Figure 7**. The alignment of *rbcL-psaI* intergenic sequence shows that the pseudogene *accD* is missing
241 in both *F. ovina* (2x, 4x) and *F. brevipila* but present in the *F. rubra* complex and other species examined
242 in this study. Species were ordered by complexes.

243 *2.7 Phylogenetic Reconstruction of Fine Fescue Species*

244 We reconstructed the phylogenetic relationships of taxa within the *Festuca* - *Lolium*
245 complex using the chloroplast genomes sequenced in our study and eight publicly available
246 complete chloroplast genomes including six taxa within the *Festuca-Lolium* complex (**Figure**
247 **8**). The dataset included 125,824 aligned characters, of which 3,923 were parsimony-
248 informative and 91.11% characters are constant. The five fine fescue taxa were split into two
249 clades ([ML]BS=100). In the *F. ovina* complex, two *F. ovina* accessions included in the

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

250 phylogenetic analysis, a diploid one from GenBank, and a tetraploid one newly sequenced
251 in this study are paraphyletic to *F. brevipila* ([ML]BS=100). All three subspecies of *F. rubra*
252 formed a strongly supported clade ([ML]BS=100). Together they are sisters to the *F. ovina*
253 complex ([ML]BS=100).



254

255 **Figure 8.** Maximum likelihood (ML) phylogram of the *Festuca - Lolium* complex with 1,000 bootstrap
256 replicates. Fine fescues were grouped into previous named complexes (*F. ovina* and *F. rubra*), sister to
257 broad leaved fescues in the *Festuca – Lolium* complex.

258 **3. Discussion**

259 In this study, we used flow cytometry to determine the ploidy level of five fine fescue
260 cultivars, assembled the chloroplast genomes for each, and identified structural variation
261 and mutation hotspots. We also identified candidate loci for marker development to
262 facilitate fine fescue species identification. Additionally, we reconstructed the phylogenetic
263 relationships of the *Festuca-Lolium* complex using plastid genome information generated in
264 this study along with other publicly-available plastid genomes.
265 While most crop plants are highly distinctive from their close relatives, *Festuca* is a
266 species-rich genus that contains species with highly similar morphology and different
267 ploidy level. Consequently, it is difficult for researchers to interpret species identity. In our
268 case, flow cytometry was able to successfully separate fine fescue taxa *F. brevipila* cv. Beacon,
269 *F. ovina* cv. Quatro and *F. rubra* ssp. *rubra* cv. Navigator II based on the estimated ploidy
270 levels. However, it is difficult to distinguish between *F. rubra* ssp. *litoralis* cv. Shoreline and
271 *F. rubra* ssp. *fallax* cv. Treazure II as they had similar PI-A values based on flow cytometry.
272 We noticed that the average mean PI-A of the diploid *L. perenne* (63.91) was higher than the
273 mean PI-A of diploid *F. ovina* (52.73), suggesting that *F. ovina* taxa have smaller genome size
274 than *L. perenne*. The ploidy estimation in the *F. ovina* complex are fairly consistent while the
275 estimations of genome sizes in the *F. rubra* complex are smaller than we expected, even
276 though these two complexes are closely related. Indeed, a similar finding was reported by
277 Huff et al (Huff and Palazzo 1998) who reported that *F. brevipila* has a larger genome size
278 than *F. rubra* ssp. *litoralis* and *F. rubra* ssp. *fallax*, both of which have the same ploidy level as
279 *F. brevipila*. The range of variation in DNA content within each complex suggest a
280 complicated evolutionary history in addition to polyploidization (Huff and Palazzo 1998).

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

281 When we cannot identify taxon based on the ploidy level, we need different approaches
282 to identify them. The presence and absence of rhizome formation could be taken into
283 consideration; for example, *F. rubra* ssp. *fallax* cv. Treazure II is a bunch type turfgrass, while
284 *F. rubra* ssp. *litoralis* cv. Shoreline forms short and slender rhizomes (Meyer and Funk 1989).
285 This method may not be effective because rhizome formation can be greatly affected by
286 environmental conditions (Yang et al. 2015, Ma and Huang 2016).
287 To further develop molecular tools to facilitate species identification, we carried out
288 chloroplast genome sequencing. We assembled the complete chloroplast genomes of five
289 low-input turfgrass fine fescues using Illumina sequencing. Overall, the chloroplast
290 genomes had high sequence and structure similarity among all five fine fescue taxa
291 sequenced, especially within each complex. All five chloroplast genomes share similar gene
292 content except for the three species in the *F. rubra* complex that have a pseudogene Acetyl-
293 coenzyme A carboxylase carboxyl transferase subunit (*accD*). The *accD* pseudogene is either
294 partially or completely absent in some monocots. Instead, a nuclear-encoded ACC enzyme
295 has been found to replace the plastic *accD* gene function in some angiosperm linage
296 (Rousseau-Gueutin et al. 2013). Indeed, even though the *accD* pseudogene is missing in the
297 *F. brevipila* chloroplast genome, the gene transcript was identified in a transcriptome
298 sequencing dataset (unpublished data), suggesting that this gene has been translocated to
299 nucleus genome. Previous studies have shown that broad-leaf fescues, *L. perenne*, *O. sativa*,
300 and *H. vulgare* all had the pseudogene *accD* gene, while it was absent in diploid *F. ovina*, *Z.*
301 *mays*, *S. bicolor*, *T. aestivum,* and *B. distachyon* (Hand et al. 2013). Broad-leaf and fine-leaf
302 fescues diverged around 9 Mya ago (Fjellheim et al. 2006), which raises an interesting
303 question about the mechanisms of the relocation of *accD* among closely related taxa in the
304 *Festuca-Lolium* complex and even within fine fescue species.
305 In plants, chloroplast genomes are generally considered "single copy" and lack
306 recombination due to maternal inheritance (Ebert and Peakall 2009). For this reason,
307 chloroplast genomes are convenient for developing genetic markers for distinguishing
308 species and subspecies. We have identified a number of repeat signatures that are unique to
309 a single species or species complex in fine fescue. For example, complement match is only
310 identified in *F. ovina* complex, and *F. rubra* complex has more reversed matches. We also
311 identified two SSR repeats unique to each of the two complexes. The first one consists of
312 AAATT/AATTT repeat units is unique to *F. rubra* ssp. *litoralis* and *F. rubra* ssp. *rubra*, and
313 the second one consists of ACCAT/ATGGT repeat units is unique to *F. brevipila* and *F. ovina*.
314 In cases like the identification of hexaploids *F. brevipila*, *F. rubra* ssp. *fallax*, and *Festuca rubra*
315 ssp. *litoralis*, it is critical to have these diagnostic repeats given all three taxa share similar
316 PI-A values from flow cytometry. Taxon-specific tandem repeats could also aid the SSR
317 repeats for species identification. We used chloroplast sequence developed candidate
318 primer sets to solve the problem. Primer set (1) provided a clear separation of *F. rubra* ssp.
319 *litoralis* cv. Shoreline and *F. rubra* ssp. *fallax* cv. Treazure II when flow cytometry was not
320 able to separate them. Primer set (2) provided clear separation of *F. brevipila* cv. Beacon and
321 *F. ovina* cv. Quatro, which provided an alternative method for *F. ovina* complex taxa
322 identification. By combining both flow cytometry and candidate primer sets developed in
323 this study, researchers will be able to identify fine fescue taxa within and between two
324 complexes.
325 Nucleotide diversity analysis suggested that several variable genome regions exist
326 among the five fine fescue taxa sequenced in this study. These variable regions included
327 previously known highly variable chloroplast regions such as *trnL-trnF* and *rpl32-trnL*
328 (Demesure et al. 1995, Dong et al. 2012). These regions have undergone rapid nucleotide

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

329 substitution and are potentially informative molecular markers for characterization of fine
330 fescue species.

331 Phylogeny inferred from DNA sequence is valuable for understanding the evolutionary
332 context of a species. The phylogenetic relationship of fine fescue using whole plastid
333 genome sequences agrees with previous classification based on genome size estimation and
334 morphology (Huff and Palazzo 1998, Cheng et al. 2016). The *F. ovina* complex includes *F.*
335 *ovina* and *F. brevipila* and the *F. rubra* complex includes *F. rubra* ssp. *rubra*, *F. rubra* ssp. *litoralis*
336 and *F. rubra* ssp. *fallax*, with the two rhizomatous subspecies (ssp. *rubra* and ssp. *literalis*)
337 being sister to each other. Within the *Festuca – Lolium* complex, fine fescues are monophyletic
338 and together sister to a clade consists of broad-leaved fescues and *Lolium*. In our analysis, *F.*
339 *brevipila* (6x) is nested within *F. ovina* and sister to the diploid *F. ovina*. It is likely that *F.*
340 *brevipila* arose from the hybridization between *F. ovina* (2x) and *F. ovina* (4x). Considering
341 the complex evolutionary history of this genus, further research using nuclear loci sequences
342 are needed to provide a more accurate phylogeny tree and validate this hypothesis.

343 The diversity of fine fescue provides valuable genetic diversity for breeding and
344 cultivar development. Breeding fine fescue cultivars for better disease resistance, heat
345 tolerance, and traffic tolerance could be achieved through screening wild accessions and by
346 introgressing desired alleles into elite cultivars. Some work has been done using *Festuca*
347 accessions in the USDA Germplasm Resources Information Network (GRIN)
348 (https://www.ars-grin.gov) to breed for improved forage production in fescue species
349 (Robbins et al. 2016). To date, there are 229 *F. ovina* and 486 *F. rubra* accessions in the USDA
350 GRIN. To integrate this germplasm into breeding programs, plant breeders and other
351 researchers need to confirm the ploidy level using flow cytometry and further identify them
352 using molecular markers. Resources developed in this study could provide the tools to
353 screen the germplasm accessions and refine the species identification so breeders can
354 efficiently use these materials for breeding and genetics improvement of fine fescue species.

355 ## 4. Materials and Methods

356 ### *Plant Material*

357 Seeds from the fine fescue cultivars were obtained from the 2014 National Turfgrass
358 Evaluation Program (www.ntep.org, USA) and planted in the Plant Growth Facility at the
359 University of Minnesota, St. Paul campus under 16 hours daylight (25 ℃) and 8 hours dark
360 (16 ℃) with weekly fertilization. Single genotypes of *F. brevipila* cv. Beacon, *F. rubra* ssp.
361 *litoralis* cv. Shoreline, *F. rubra* ssp. *rubra* cv. Navigator II, *F. rubra* ssp. *fallax* cv. Treazure II,
362 and *F. ovina* cv. Quatro were selected and used for chloroplast genome sequencing.

363 ### *Flow Cytometry*

364 To determine the ploidy level of the cultivars used for sequencing and compare them
365 to previous work (2n=4x=28: *F. ovina*; 2n=6x=42: *F. rubra* ssp. *litoralis*, *F. rubra* ssp. *fallax*, and
366 *F. brevipila*; 2n=8x=56: *F. rubra* ssp. *rubra*), flow cytometry was carried out using *Lolium*
367 *perenne* cv. Artic Green (2n=2x=14) as the reference. Samples were prepared using CyStain
368 PI Absolute P (Sysmex, product number 05-5022). Briefly, to prepare the staining solution
369 for each sample, 12 μL propidium iodide (PI) was added to 12 mL of Cystain UV Precise P
370 staining buffer with 6 μL RNase A. To prepare plant tissue, a total size of 0.5 cm x 0.5 cm
371 leaf sample of the selected fine fescue was excised into small pieces using a razor blade in
372 500 μL CyStain UV Precise P extraction buffer and passed through a 50 μm size filter
373 (Sysmex, product number 04-004-2327). The staining solution was added to the flow-

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

374 through to stain nuclei in each sample. Samples were stored on ice before loading the flow
375 cytometer. Flow cytometry was carried out using the BD LSRII H4760 (LSRII) instrument
376 with PI laser detector using 480V with 2,000 events at the University of Minnesota Flow
377 Cytometry Resource (UCRF). Data was visualized and analyzed on BD FACSDiva 8.0.1
378 software. To estimate the genome size, *L. perenne* DNA (5.66 pg/2C) was used as standard
379 (Arumuganathan, Tallury et al. 1999), USDA PI 230246 (2n=2x=14) was used as diploid fine
380 fescue relative (unpublished data). To infer fine fescues ploidy, estimation was done using
381 equations (1) and (2) (Doležel et al. 2007).
382
383 $$\text{Sample 2C DNA Content} = \text{Standard 2C DNA Content (pg DNA)} \times \frac{\text{(Sample G1 Peak Mean)}}{\text{(Standard G1 Peak Mean)}} \quad (1)$$
384
385 $$\text{Sample Ploidy} = \frac{2n \times \text{Sample pg/Nucleus}}{\text{Diploid Relative pg/Nucleus}} \quad (2)$$

386 *DNA Extraction and Sequencing*

387 To extract DNA for chloroplast genome sequencing, 1 g of fresh leaves were collected
388 from each genotype and DNA was extracted using the Wizard Genomic DNA Purification
389 Kit (Promega, USA) following manufacturer instructions. DNA quality was examined on
390 0.8% agarose gel and quantified via PicoGreen (Thermo Fisher, Catalog number: P11496).
391 Sequencing libraries were constructed by NovoGene, Inc. (Davis, CA) using Nextera XT
392 DNA Library Preparation Kit (Illumina) and sequenced in 150 bp paired-end mode, using
393 the HiSeq X Ten platform (Illumina Inc., San Diego, CA, USA) with an average of 10 million
394 reads per sample. All reads used in this study were deposited in the NCBI Sequence Read
395 Archive (SRA) under BioProject PRJNA512126.
396
397 *Chloroplast Genome Assembly and Annotation*

398 Raw reads were trimmed of Illumina adaptor sequences using Trimmomatic (v. 0.32)
399 (Bolger et al. 2014). Chloroplast genomes were *de novo* assembled using NovoPlasty v. 2.0
400 (Dierckxsens et al. 2016). Briefly, *rbcL* gene sequence from diploid *F. ovina* (NCBI accession
401 number: JX871940) was extracted and used as the seed to initiate the assembly. NovoPlasty
402 assembler configuration was set as follows: *k-mer* size = 39; insert size = auto; insert range =
403 1.8; and insert range strict 1.3. Reads with quality score above 25 were used to complete the
404 guided assembly using *F. ovina* (NCBI accession number: JX871940) as the reference.
405 Assembled plastid genomes for each taxon were manually corrected by inspecting the
406 alignments of reads used in the assembly. The assembled chloroplast genomes were
407 deposited under BioProject PRJNA512126, GenBank accession numbers MN309822-
408 MN309826.
409 The assembled chloroplast genomes were annotated using the GeSeq pipeline (Tillich,
410 Lehwark et al. 2017) and corrected using DOGMA online interface
411 (https://dogma.ccbb.utexas.edu) (Wyman, Jansen et al. 2004). BLAT (a BLAST-like
412 alignment tool (Kent 2002) protein, tRNA, rRNA, and DNA search identity threshold was
413 set at 80% in the GeSeq pipeline using the default reference database with the generate
414 codon-based alignments option turned on. tRNAs were also predicted via tRNAscan-SE
415 v2.0 and ARAGORN v 1.2.38 using the bacterial/plant chloroplast genetic code (Lowe and
416 Eddy 1997, Laslett and Canback 2004). The final annotation was manually inspected and
417 corrected using results from both pipelines. The circular chloroplast map was drawn by the
418 OrganellarGenomeDRAW tool (OGDRAW) (Lohse et al. 2007).

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

419 *Nucleotide Polymorphism of Fine Fescue Species*

420 To identify genes with the most single nucleotide polymorphism, quality trimmed
421 sequencing reads of the five fine fescues were mapped to the diploid *Festuca ovina*
422 chloroplast genome (NCBI accession number: JX871940) using BWA v.0.7.17 (Li and Durbin
423 2009). SNPs and short indels were identified using bcftools v.1.9 with the setting "mpileup
424 -Ou" and called via bcftools using the -mv function (Quinlan and Hall 2010). Raw SNPs
425 were filtered using bcftools filter -s option to filter out SNPs with low quality (Phred score
426 cutoff 20, coverage cutoff 20). The subsequent number of SNPs per gene and InDel number
427 per gene was calculated using a custom perl script SNP_vcf_from_gene_gff.pl
428 (https://github.com/qiuxx221/fine-fescue-).
429 To identify simple sequence repeat (SSR) markers for plant identification,
430 MIcroSAtellite identification tool (MISA v 1.0) was used with a threshold of 10, 5, 4, 3, 3, 3
431 repeat units for mono-, di-, tri-, tetra-, penta-, and hexanucleotide SSRs, respectively (Thiel
432 et al. 2003). The identification of repetitive sequences and structure of whole chloroplast
433 genome was done via REPuter program online server (https://bibiserv.cebitec.uni-
434 bielefeld.de/reputer) (Kurtz et al. 2001). Program configuration was set with minimal repeat
435 size set as 20 bp and with sequence identify above 90%. Data was visualized using ggplot2
436 in R (v 3.5.3). Finally, the sliding window analysis was performed using DnaSP (v 5) with a
437 window size of 600 bp, step size 200 bp to detected highly variable regions in the fine fescue
438 chloroplast genome (Librado and Rozas 2009).

439 *Comparative Chloroplast Genomics Analysis*

440 To compare fine fescue species chloroplast genome sequence variations, the five
441 complete chloroplast genomes were aligned and visualized using mVISTA, an online suite
442 of computation tools with LAGAN mode (Brudno et al. 2003, Frazer et al. 2004). The diploid
443 *Festuca ovina* (NCBI accession number: JX871940) chloroplast genome and annotation were
444 used as the template for the alignment.

445 *Phylogenetic Analysis of Fine Fescues and Related Festuca species*

446 To construct the phylogenetic tree of the fine fescues using the whole chloroplast
447 genome sequence, chloroplast genomes of 8 species were downloaded from GenBank. Of
448 the 8 downloaded genomes, perennial ryegrass (*Lolium perenne*, AM777385), Italian ryegrass
449 (*Lolium multiflorum*, JX871942), diploid *Festuca ovina* (JX871940), tall fescue (*Festuca*
450 *arundiancea*, FJ466687), meadow fescue (*Festuca pratensis*, JX871941), and wood fescue
451 (*Festuca altissima*, JX871939) were within the *Festuca- Lolium* complex. Turfgrass species
452 outside of *Festuca-Lolium* complex including creeping bentgrass (*Agrostis stolonifera* L.,
453 EF115543) and *Cynodon dactylon* (KY024482.1) were used as an outgroup. All chloroplast
454 genomes were aligned using the MAFFT program (v 7) (Katoh and Standley 2013);
455 alignments were inspected and manually adjusted. Maximum likelihood (ML) analyses was
456 performed using the RAxML program (v 8.2.12) under GTR+G model with 1,000 bootstrap
457 (Stamatakis 2006). The phylogenetic tree was visualized using FigTree (v 1.4.3)
458 (https://github.com/rambaut/figtree) (Rambaut 2012).

459 **5. Conclusions**

460 Five newly-sequenced complete chloroplast genomes of fine fescue taxa were reported
461 in this study. Chloroplast genome structure and gene contents are both conserved, with the
462 presence and absence of *accD* pseudogene being the biggest structural variation between the

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

463 *F. ovina* and the *F. rubra* complexes. We identified SSR repeats and long sequence repeats of
464 fine fescues and discovered several unique repeats for marker development. The
465 phylogenetic constructions of fine fescue species in the *Festuca - Lolium* complex suggested
466 a robust and consistent relationship compared to the previous identification using flow
467 cytometry. This information provided a reference for future fine fescue taxa identification.

468 **Supplementary Materials**:

469 1. **Figure S1**. Flow cytometry nuclei population distribution of *L. perenne*, fine fescues, and diploid USDA
470 PI accession. G1 populations were gated in red, G2 population was only gated in *L. perenne* in green.
471 2. **Figure S2.** Examples of PCR validation of predicted repeat regions based on fine fescue chloroplast
472 genomes.
473 3. **Table S1.** Fine fescue chloroplast genomes gene content by gene category.
474 4. **Table S2**. SSR loci types and number distributions of fine fescue species predicted using MISA
475 program.
476 5. **Table S3**. Tandem repeat loci and repeat types predicted using PEPuter program in fine fescue species.
477 6. **Table S4**. SNPs number per gene distribution for fine fescue species. *rpoC2* gene has the most SNPs
478 (31) in *F. rubra* complex comparing to *F. ovina* species.
479 7. **Table S5**. InDel number per gene distribution for fine fescue species. *ndhA* gene had the most InDels
480 *F. rubra* species. *atpI* gene had the most InDesl in *F. ovina* species.
481 8. **Table S6**. Sliding window analysis using DnaSP at window size of 600 bp, step size 200 bp to detect
482 highly variable regions in the fine fescue chloroplast genome

483

484 **Author Contributions:** Y. Q. performed the experiments, analyzed the data, and wrote the manuscript; C. H.
485 helped analyze data, wrote perl scripts; Y. Y. helped with phylogenetic analysis; E. W. secured funding for this
486 project, supervised this research, provided suggestions, and comments. All authors contributed to the revision
487 of the manuscript and approved the final version.

488 **Funding:** This research is funded by the National Institute of Food and Agriculture, U.S. Department of
489 Agriculture, Specialty Crop Research Initiative under award number 2017-51181-27222.

490 **Acknowledgments:** The authors would like to thank Minnesota Supercomputing Institute for the high-
491 performance computing clusters. We would also like to thank Jill Ekar, Jason Motl, and Therese Martin for
492 providing instruction on operating the flow cytometer.

493 This manuscript has been released as a pre-print at bioRxiv, (Qiu, et al. 2019)

494 **Conflicts of Interest:** The authors declare no conflict of interest.
495
496

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

497 **References**

498 Arumuganathan, K., S. Tallury, M. Fraser, A. Bruneau and R. Qu (1999). "Nuclear DNA content of thirteen
499 turfgrass species by flow cytometry." Crop science **39**(5): 1518-1521.
500 Baldwin, B. G., M. J. Sanderson, J. M. Porter, M. F. Wojciechowski, C. S. Campbell and M. J. Donoghue (1995).
501 "The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny."
502 Annals of the Missouri botanical garden: 247-277.
503 Beard, J. B. (1972). "Turfgrass: Science and culture."
504 Bolger, A. M., M. Lohse and B. Usadel (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data."
505 Bioinformatics **30**(15): 2114-2120.
506 Bonos, S. A. and D. R. Huff (2013). "Cool-season grasses: Biology and breeding." Turfgrass: Biology, use, and
507 management(turfgrassbiolog): 591-660.
508 Brudno, M., C. B. Do, G. M. Cooper, M. F. Kim, E. Davydov, E. D. Green, A. Sidow, S. Batzoglou and N. C. S.
509 Program (2003). "LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic
510 DNA." Genome research **13**(4): 721-731.
511 Bryan, G., J. McNicoll, G. Ramsay, R. Meyer and W. De Jong (1999). "Polymorphic simple sequence repeat
512 markers in chloroplast genomes of Solanaceous plants." Theoretical and Applied Genetics **99**(5): 859-867.
513 Cahoon, A. B., R. M. Sharpe, C. Mysayphonh, E. J. Thompson, A. D. Ward and A. Lin (2010). "The complete
514 chloroplast genome of tall fescue (Lolium arundinaceum; Poaceae) and comparison of whole plastomes
515 from the family Poaceae." American Journal of botany **97**(1): 49-58.
516 Casler, M. D. (2003). Turfgrass biology, genetics, and breeding, John Wiley & Sons.
517 Cheng, T., C. Xu, L. Lei, C. Li, Y. Zhang and S. Zhou (2016). "Barcoding the kingdom Plantae: new PCR primers
518 for ITS regions of plants with improved universality and specificity." Molecular Ecology Resources **16**(1):
519 138-149.
520 Clayton, W. D. and S. A. Renvoize (1986). "Genera graminum. Grasses of the world." Genera graminum.
521 Grasses of the World. **13**.
522 Demesure, B., N. Sodzi and R. Petit (1995). "A set of universal primers for amplification of polymorphic non-
523 coding regions of mitochondrial and chloroplast DNA in plants." Molecular ecology **4**(1): 129-134.
524 Diekmann, K., T. R. Hodkinson, K. H. Wolfe, R. van den Bekerom, P. J. Dix and S. Barth (2009). "Complete
525 chloroplast genome sequence of a major allogamous forage species, perennial ryegrass (Lolium perenne
526 L.)." DNA research **16**(3): 165-176.
527 Dierckxsens, N., P. Mardulyn and G. Smits (2016). "NOVOPlasty: de novo assembly of organelle genomes from
528 whole genome data." Nucleic acids research **45**(4): e18-e18.
529 Doležel, J., J. Greilhuber and J. Suda (2007). "Estimation of nuclear DNA content in plants using flow
530 cytometry." Nature protocols **2**(9): 2233.
531 Dong, W., J. Liu, J. Yu, L. Wang and S. Zhou (2012). "Highly variable chloroplast markers for evaluating plant
532 phylogeny at low taxonomic levels and for DNA barcoding." PloS one **7**(4): e35071.
533 Ebert, D. and R. Peakall (2009). "Chloroplast simple sequence repeats (cpSSRs): technical resources and
534 recommendations for expanding cpSSR discovery and applications to a wide array of plant species."
535 Molecular Ecology Resources **9**(3): 673-690.
536 Fjellheim, S., O. A. Rognli, K. Fosnes and C. Brochmann (2006). "Phylogeographical history of the widespread
537 meadow fescue (Festuca pratensis Huds.) inferred from chloroplast DNA sequences." Journal of
538 Biogeography **33**(8): 1470-1478.
539 Frazer, K. A., L. Pachter, A. Poliakov, E. M. Rubin and I. Dubchak (2004). "VISTA: computational tools for
540 comparative genomics." Nucleic acids research **32**(suppl_2): W273-W279.

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

541 Hand, M. L., G. C. Spangenberg, J. W. Forster and N. O. Cogan (2013). "Plastome sequence determination and
542 comparative analysis for members of the Lolium-Festuca grass species complex." G3: Genes, Genomes,
543 Genetics: g3. 112.005264.
544 Hebert, P. D., A. Cywinska, S. L. Ball and J. R. Dewaard (2003). "Biological identifications through DNA
545 barcodes." Proceedings of the Royal Society of London. Series B: Biological Sciences **270**(1512): 313-321.
546 Huang, Y.-Y., S.-T. Cho, M. Haryono and C.-H. Kuo (2017). "Complete chloroplast genome sequence of
547 common bermudagrass (Cynodon dactylon (L.) Pers.) and comparative analysis within the family
548 Poaceae." PloS one **12**(6): e0179055.
549 Hubbard, C. E. (1968). "Grasses. A guide to their structure, identification, uses, and distribution in the British
550 Isles." Grasses. A guide to their structure, identification, uses, and distribution in the British Isles.
551 Huff, D. R. and A. J. Palazzo (1998). "Fine fescue species determination by laser flow cytometry." Crop science
552 **38**(2): 445-450.
553 Jenkin, T. J. (1959). "Fescue Species (Festuca L.)." In: Roemer, T. & W. Rudorf. Handbuch der
554 Pflanzenzüchtung(4): 418–434.
555 Katoh, K. and D. M. Standley (2013). "MAFFT multiple sequence alignment software version 7: improvements
556 in performance and usability." Molecular biology and evolution **30**(4): 772-780.
557 Kent, W. J. (2002). "BLAT—the BLAST-like alignment tool." Genome research **12**(4): 656-664.
558 Kurtz, S., J. V. Choudhuri, E. Ohlebusch, C. Schleiermacher, J. Stoye and R. Giegerich (2001). "REPuter: the
559 manifold applications of repeat analysis on a genomic scale." Nucleic acids research **29**(22): 4633-4642.
560 Laslett, D. and B. Canback (2004). "ARAGORN, a program to detect tRNA genes and tmRNA genes in
561 nucleotide sequences." Nucleic acids research **32**(1): 11-16.
562 Li, H. and R. Durbin (2009). "Fast and accurate short read alignment with Burrows–Wheeler transform."
563 bioinformatics **25**(14): 1754-1760.
564 Librado, P. and J. Rozas (2009). "DnaSP v5: a software for comprehensive analysis of DNA polymorphism
565 data." Bioinformatics **25**(11): 1451-1452.
566 Lohse, M., O. Drechsel and R. Bock (2007). "OrganellarGenomeDRAW (OGDRAW): a tool for the easy
567 generation of high-quality custom graphical maps of plastid and mitochondrial genomes." Current genetics
568 **52**(5-6): 267-274.
569 Lowe, T. M. and S. R. Eddy (1997). "tRNAscan-SE: a program for improved detection of transfer RNA genes in
570 genomic sequence." Nucleic acids research **25**(5): 955-964.
571 Ma, X. and B. Huang (2016). "Gibberellin-stimulation of rhizome elongation and differential GA-responsive
572 proteomic changes in two grass species." Frontiers in plant science **7**: 905.
573 Meyer, W. A. and C. R. Funk (1989). "Progress and Benefits to Humanity from Breeding Cool-Season Grasses
574 for Turf 1." Contributions from breeding forage and turf grasses(contributionsfr): 31-48.
575 Piper, C. V. (1906). North American species of Festuca, US Government Printing Office.
576 Provan, J., W. Powell and P. M. Hollingsworth (2001). "Chloroplast microsatellites: new tools for studies in
577 plant ecology and evolution." Trends in ecology & evolution **16**(3): 142-147.
578 Qiu, Y., C. D. Hirsch, Y. Yang and E. Watkins (2019). "Chloroplast Genome Sequencing and Comparative
579 Analysis for Fine Fescue (Festuca L., Poaceae) Turfgrasses." bioRxiv: 708149.
580 Quinlan, A. R. and I. M. Hall (2010). "BEDTools: a flexible suite of utilities for comparing genomic features."
581 Bioinformatics **26**(6): 841-842.
582 Rambaut, A. (2012). FigTree v1. 4.

bioRxiv preprint doi: https://doi.org/10.1101/708149; this version posted October 20, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

583 Robbins, M. D., J. E. Staub and B. S. Bushman (2016). "Development of fine-leaved Festuca grass populations
584 identifies genetic resources having improved forage production with potential for wildfire control in the
585 western United States." Euphytica **209**(2): 377-393.
586 Rousseau-Gueutin, M., X. Huang, E. Higginson, M. Ayliffe, A. Day and J. N. Timmis (2013). "Potential
587 functional replacement of the plastidic acetyl-CoA carboxylase subunit (accD) gene by recent transfers to
588 the nucleus in some angiosperm lineages." Plant physiology **161**(4): 1918-1929.
589 Ruemmele, B., L. Brilman and D. Huff (1995). "Fine fescue germplasm diversity and vulnerability." Crop
590 science **35**(2): 313-316.
591 Stamatakis, A. (2006). "RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of
592 taxa and mixed models." Bioinformatics **22**(21): 2688-2690.
593 Thiel, T., W. Michalek, R. Varshney and A. Graner (2003). "Exploiting EST databases for the development and
594 characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.)." Theoretical and applied
595 genetics **106**(3): 411-422.
596 Tillich, M., P. Lehwark, T. Pellizzer, E. S. Ulbricht-Jones, A. Fischer, R. Bock and S. Greiner (2017). "GeSeq–
597 versatile and accurate annotation of organelle genomes." Nucleic acids research **45**(W1): W6-W11.
598 Torrecilla, P. and P. Catalán (2002). "Phylogeny of broad-leaved and fine-leaved Festuca lineages (Poaceae)
599 based on nuclear ITS sequences." Systematic Botany **27**(2): 241-252.
600 Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm and S. G. Rozen (2012). "Primer3—
601 new capabilities and interfaces." Nucleic acids research **40**(15): e115-e115.
602 Wyman, S. K., R. K. Jansen and J. L. Boore (2004). "Automatic annotation of organellar genomes with
603 DOGMA." Bioinformatics **20**(17): 3252-3255.
604 Yang, M., L. Zhu, C. Pan, L. Xu, Y. Liu, W. Ke and P. Yang (2015). "Transcriptomic analysis of the regulation of
605 rhizome formation in temperate and tropical lotus (Nelumbo nucifera)." Scientific reports **5**: 13059.

606