

1 **ForageGrassBase: Molecular resource for the forage grass *Festuca***
2 ***pratensis* Huds.**

3

4 Jeevan Karloss Antony Samy: jeevan.karloss@nmbu.no

5 Odd Arne Rognli: odd-arne.rognli@nmbu.no

6 Mallikarjuna Rao Kovi: mallikarjuna.rao.kovi@nmbu.no

7

8 **Institute address:**

9 Department of Plant Sciences (IPV), Faculty of Biosciences (BIOVIT), Norwegian University of Life
10 Sciences (NMBU), Ås, Akershus, 1432, Norway.

11

12 **Corresponding authors:** Mallikarjuna Rao Kovi

13 E. mail: mallikarjuna.rao.kovi@nmbu.no

14 Tel: +47-40566127

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32 **Abstract**

33 **Background**

34 Meadow fescue (*Festuca pratensis* Huds.) is one of the most important forage grasses in
35 temperate regions. *F. pratensis* is a diploid ($2n = 14$) outbreeding species that belongs to the
36 genus *Festuca*. Together with *Lolium*, they are the most important genera of forage grasses
37 in temperate regions. *F. pratensis* has good winter survival, with high quality dry matter
38 yields and persistency, and is suitable both for frequent-cutting conservation regimes and
39 for grazing. It is a significant component of species-rich permanent pastures in the
40 temperate regions, ensuring high forage yield under harsh climatic conditions where other
41 productive forage grass species are unable to grow. However, genomic resources for *F.*
42 *Pratensis* is not available so far.

43

44 **Results**

45 The draft genome sequences of two *F. pratensis* genotypes “HF7/2” and “B14/16” are
46 reported in this study. Here, using the draft genome, functional annotation datasets of two
47 *F. pratensis* cultivars, we have constructed the *F. pratensis* genome
48 database <http://foragegrass.org/>, the first open-access platform to provide comprehensive
49 genomic resources related to this forage grass species. The current version of this database
50 provides the most up-to-date draft genome sequence along with structural and functional
51 annotations for genes using Genome Browser (GBrowse). In addition, we have integrated
52 comparative genomic tracks for *F. pratensis* genomes by mapping *F. pratensis* genome to the
53 barley, rice, *Brachypodium* and maize genomes. We have integrated homologous search tool
54 BLAST also for the users to analyze their data. Combined, GBrowse, BLAST and downloadable
55 data gives an user friendly access to *F. pratensis* genomic resources. All data in the database
56 were manually curated.

57

58 **Conclusion**

59 To our knowledge, ForageGrassBase is the first genome database dedicated to forage
60 grasses. It provides valuable resources for a range of research fields related to *F.*
61 *pratensis* and other forage crop species, as well as for plant research communities in

62 general. The genome database can be accessed at <http://foragegrass.org>. In the near future,
63 we will expand the ForageGrassBase by adding genomic tools for other forage grass species,
64 as soon as their genomes become available.

65

66 **Keywords**

67 *Festuca pratensis*, meadow fescue, forage grass database, molecular data, online molecular
68 database, genomic data resources, Gbrowse

69

70

71 **Background**

72 Grasslands are covering very large portions of the earth's surface and they are important as
73 feed sources and pastures for livestock. In both developed and developing countries, many
74 millions of livestock farmers, ranchers and pastoralists depend on grasslands and conserved
75 products such as hay and silage from a range of fodder crops for their livelihoods. Among
76 several forage crops, meadow fescue (*Festuca pratensis* Huds.) is one of the most important
77 forage grass species in temperate regions of the world. It is a diploid ($2n = 14$) outbreeding
78 species that belongs to the genus *Festuca* (1)

79

80 Fescues in general have evolved superior adaptations to abiotic stresses, e.g., winter
81 survival in meadow fescue. The most abundant forage grass species in temperate regions,
82 perennial ryegrass (*Lolium perenne* L.), is known for its superior nutritive quality, rapid
83 establishment and growth but is lacking persistency under harsh environmental conditions.
84 The *Lolium-Festuca* species complex is unique since it is possible to combine *Lolium* and
85 *Festuca* genomes in interspecific hybrids (*Festulolium*) (2). Complementation of traits in
86 *Festulolium* hybrids is thus a very interesting strategy for developing novel germplasm and
87 cultivars with improved quality and persistency, which can contribute to a sustainable
88 forage production. Relatively modest genomic resources have been developed for meadow
89 fescue compared with other grass species like perennial ryegrass (*Lolium perenne*) (3).

90

91 In order to develop better *Festulolium* hybrids, we have initiated sequencing of *F. pratensis*,
92 and combined with an efficient utilization of the close relationship with barley (*Hordeum*
93 *vulgare*), rice (*Oryza sativa*), *Brachypodium distachyon* and maize (*Zea mays*) through
94 comparative genomics approaches. High quality annotated *Festuca* genomes are now

95 available. As a first step, the genome sequences and genome annotations for two *F.*
96 *pratensis* genotypes are made available through ForageGrassBase (<http://foragegrass.org>).
97 ForageGrassBase was developed to make these substantial amounts of genomic data
98 accessible through visualizations and analytic tools in a common framework. Similar
99 resources for other forage grass species will be added to ForageGrassBase when they
100 become available.

101

102 **Construction and content**

103 Bootstrap (HTML, CSS), Javascript, PHP and Python were used to develop ForageGrassBase.
104 The Generic Genome Browser (GBrowse) [4] and BLAST [5] were also installed. R packages
105 are used for BLAST results visualizations.

106 *De novo* sequencing of the *Festuca pratensis* genomes were performed using Illumina mate
107 pair sequencing and assembly was performed by the SOAPDenovo assembler. Furthermore,
108 gene annotation was performed by in-house developed annotation pipelines and python
109 scripts.

110

111 **Utility and discussion**

112 **Genome browser (GBrowse)**

113 The generic genome browser (GBrowse) is simple and one of the most used genome
114 browsers for visualization of genomes. We installed GBrowse to visualize and share genomic
115 data of *F. pratensis* (Fig. 1). Currently, ForageGrassBase contains molecular data of two
116 *Festuca pratensis* genotypes; Festuca HF2/7, a Norwegian genotype originating from a
117 population selected for high frost tolerance and a Yugoslavian genotype, B14/1700, which is
118 used by our group to develop a mapping family for linkage map construction (6). Further, a
119 comparative genome analysis was performed against other grass species like barley,
120 Brachypodium, rice and maize. These comparative genomics tracks consisting of gene
121 names and chromosome positions were added to the genome browsers (Fig. 1). More data
122 and tracks will be added in the near future for other economically important forage grass
123 species like timothy (*Phleum pratense*) to expand the forage grass genomics resources in
124 ForageGrassBase .

125

126 **BLAST server**

127 We have installed a BLAST server to search for homologous regions in the *F. pratensis*
128 genome. Users having unknown sequences can use BLAST search to find the homologous
129 regions in *Festuca* and their corresponding homologous genes and their physical location in
130 *Brachypodium*, , barley, rice and maize (Fig. 2a). After the search, our algorithm chooses the
131 best hits and plots them in a unique way. BLAST results are connected to GBrowse, so the
132 users can view the homologous regions and nearby genes and other genomic features in all
133 these species (Fig. 2b).

134

135 **Future plans and integrations**

136 ForageGrassBase was developed based on high interest for the molecular data of *Festuca*
137 *pratensis*. Genetic variations and gene expression data will be added using Genetic variation
138 browser (GVBrowser) and Gene expression browser (GEBrowser) in the very near future.
139 Due to rapid developments and lower costs of high-throughput sequencing technologies,
140 we expect more forage grass genome sequence data to be available soon, and these
141 resources and new tools will be added under ForageGrassBase.

142

143 **Database access and feedback**

144 All the data used in developing this database are available through the 'Download' menu in
145 ForageGrassBase. Genome sequences and gene annotation files for the two *Festuca*
146 genotypes are available in "fasta" and "gff3" file formats to download and re-use. Users can
147 send their questions and comments through 'Contact form' under 'Contact' menu.

148

149 **Conclusions**

150 To the best of our knowledge, ForageGrassBase is the only online database to access,
151 visualize and download data for the forage grass species *Festuca pratensis* and its
152 homologous sequences/genes in rice, barley, *Brachypodium* and maize. Due to rapid
153 developments in high-throughput sequencing technology in recent years, we expect a huge
154 influx of data for forage grasses. Thus, the database ForageGrassBase will be updated by
155 adding new forage grass species genomic resources as soon as their genome sequences are
156 publicly available.

157

158 **List of abbreviations**

159 *F. pratensis*: *Festuca pratensis*

160 GBrowser: Genome browser
161 BLAST: Basic Local Alignment Search Tool
162

163

164 **Declarations**

165 **Acknowledgements**

166 We thank Torben Asp, Aarhus University, Dag Inge Våge, Teshome Dagne Mulugeta, Torfinn
167 Nome and the Orion computational facility at the Norwegian University of Life Sciences for
168 their support.

169

170 **Funding**

171 This project has received financial support from the Research Council of Norway, project
172 numbers: 199664/I10 (VARCLIM), 255428 (GenSelTim) and 208481 (ELIXIR.NO). The funding
173 source had no role in study design, data collection and interpretation and in writing the
174 manuscript.

175

176 **Availability and requirements**

177 ForageGrassBase can be accessed at <https://foragegrass.org/>

178

179 **Availability of data and materials**

180 This work does not contain additional data.

181

182 **Authors' contributions**

183 MRK, JKAS and OAR conceived the idea of developing ForageGrassBase. OAR, MRK provided
184 the genome sequences and annotation files. JKAS developed ForageGrassBase with inputs
185 from MRK and OAR. JKAS and MRK wrote the manuscript and included comments from
186 OAR. All authors read and approved the final manuscript.

187

188 **Competing interests**

189 The authors declare that they have no competing interests.

190

191 **Consent to publish**

192 Not applicable.

193

194 **Ethics approval and consent to participate**

195 Not applicable

196

197

198 **Figure legends**

199 **Fig. 1:** ForageGrassBase genome browser contains *F. pratensis* genome annotation and their
200 orthologous regions in Brachypodium, barley, maize and rice.

201 **Fig. 2:** *F. pratensis* cultivars genome browsers with genome annotations and BLAST tool (a)
202 BLAST tool implemented to search for homologous regions in the reference genomes
203 available. (b) BLAST results page shows the homologous regions.

204

205 **References**

- 206 1. Rognli OA, Saha MC, Bhamidimarri S, van der Heijden S. Fescues. In 'B. Boller et al.
207 (eds.), Fodder Crops and Amenity Grasses, Handbook of Plant Breeding 5, 2010; DOI
208 10.1007/978-1-4419-0760-8_11, Springer Science+Business Media, pp 261-292.
- 209 2. Ghesquière M, Humphreys MW, Zwierzykowski Z. Festulolium. B. Boller et al. (eds.),
210 Fodder Crops and Amenity Grasses, Handbook of Plant Breeding 5, 2010; DOI
211 10.1007/978-1-4419-0760-8_12, Springer Science+Business Media, LLC 2010, pp
212 293-316.
- 213 3. Byrne SL, Nagy I, Pfeifer M, Armstead I, Swain S, Studer B, Mayer K, Campbell JD,
214 Czaban A, Hentrup S, et al. A synteny-based draft genome sequence of the forage
215 grass *Lolium perenne*. *Plant J.* 2015; 84(4):816–26.
- 216 4. Donlin MJ. Using the Generic Genome Browser (GBrowse). *Curr Protoc*
217 *Bioinformatics* 2009, Chapter 9:Unit 9 9.
- 218 5. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search
219 tool. *J Mol Biol* 1990, 215(3):403-410.
- 220 6. Alm V, Fang C, Busso CS, Devos KM, Vollan K, Grieg Z, Rognli OA. A linkage map of
221 meadow fescue (*Festuca pratensis* Huds.) and comparative mapping with other
222 Poaceae species. *Theor Appl Genet* 2003, 108 (1):25-40.

223

224

ForageGrassBase

Molecular Data Resource for Forage Grasses

File Help

HF 2 6A: 11 kbp from 51503358:1..11,000

Browser [Select Tracks](#) [Snapshots](#) [Custom Tracks](#) [Preferences](#)

Search

Landmark or Region:

51503358:1..11,000

Search

Download Sequence File Configure... Go

Save Snapshot Load Snapshot

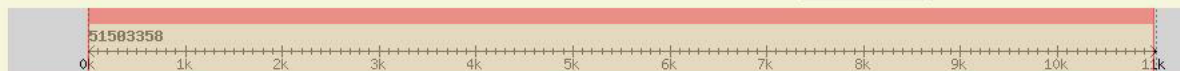
Examples: [AMP1](#), [51755314](#).

Data Source

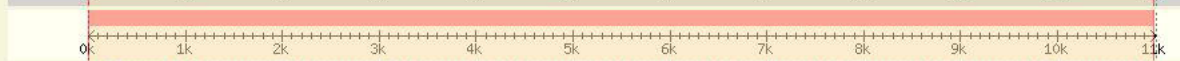
HF 2 6A

Scroll/Zoom: << < - Show 11 kbp + > >> Flip

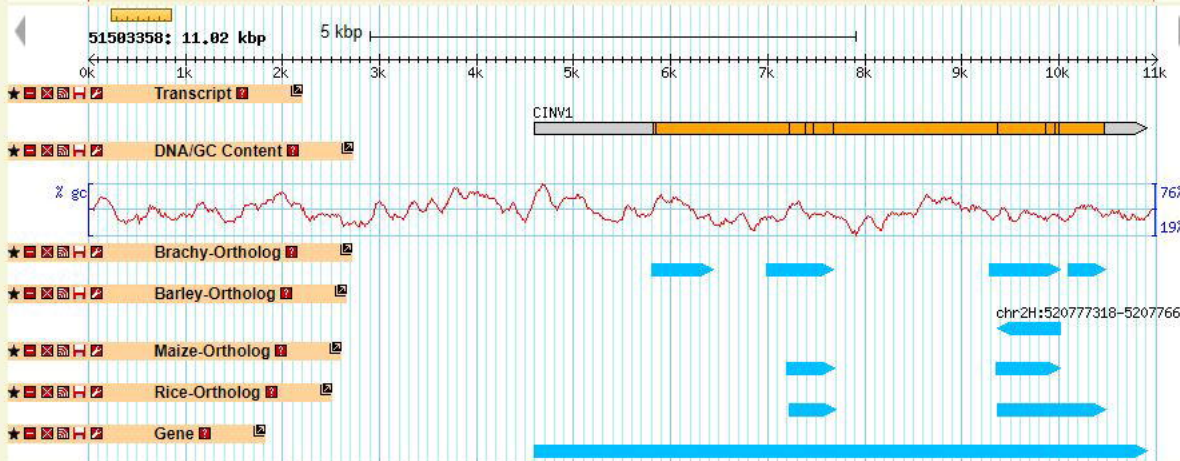
Overview



Region



Details



Select Tracks Clear highlighting

A

Or, upload file
 No file chosen

Choose database

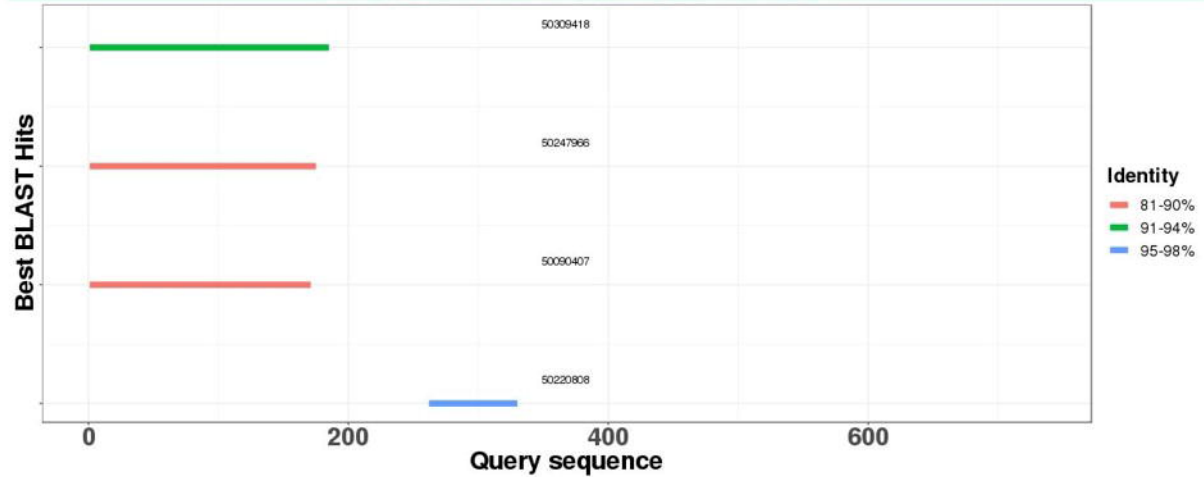
blastn
 megablast
 tblastn

Choose output format
 Table
 Alignment

E-value:

B

ForageGrassBase BLAST search results



Query	Subject	Identity	Length	Mismatches	Gaps	QStart	QEnd	SStart	SEnd	Score	E-Value	GBrowse View
ENA/AAW73220/AAW73220.1	50309418	93.51	185	12	0	1	185	2473	2289	276	3e-72	GBrowse View
ENA/AAW73220/AAW73220.1	50247966	89.71	175	18	0	1	175	4347	4521	224	1e-56	GBrowse View
ENA/AAW73220/AAW73220.1	50090407	87.79	172	19	2	1	171	8723	8553	200	2e-49	GBrowse View
ENA/AAW73220/AAW73220.1	50220808	97.10	69	2	0	262	330	6311	6243	117	2e-24	GBrowse View