Genetic diversity of gliadin alleles in bread wheat (*Triticum aestivum* L.) from Northern Kazakhstan

Maral Utebayev [Corresp., 1, 2] , Svetlana Dashkevich [1] , Nina Bome [2] , Kulpash Bulatova [3] , Yuri Shavrukov [4]

[1] A.I.Baraev Research and Production Centre of Grain Farming, Shortandy, Kazakhstan

[2] Institute of Biology, University of Tyumen, Tyumen, Russia

[3] Kazakh Research Institute of Agriculture and Plant Growing, Almaty region, Kazakhstan

[4] College of Science and Engineering, School of Biological Sciences, Flinders University, Bedford Park, SA, Australia

Corresponding Author: Maral Utebayev
Email address: phytochem@yandex.ru

Background: Spring bread wheat (*Triticum aestivum* L.) represents the main cereal crop in Northern Kazakhstan. The quality of wheat grain and flour strongly depends on the structure of gluten, comprised of gliadin and glutenin proteins. Electrophoresis spectra of gliadins are not altered by environmental conditions or plant growth, are easily reproducible and very useful for wheat germplasm identification in addition to DNA markers. Genetic polymorphism of two *Gli* loci encoding gliadins can be used for selection of preferable genotypes of wheat with high grain quality.

Methods: Polyacrylamide gel electrophoresis was used to analyse genetic diversity of gliadins in a germplasm collection of spring bread wheat from Northern Kazakhstan.

Results: The highest frequencies of gliadin alleles were found as follows, in *Gli1*: - *A 1 f* (39.3%), - *B 1 e* (71.9%), and - *D 1 a* (41.0%); and in *Gli-2*: - *A 2 q* (17.8%), - *B 2 t* (13.5%), and - *D 2 q* (20.4%). The combination of these alleles in a single genotype may be associated with higher quality of grain as well as better adaptation to the dry environment of Northern Kazakhstan; preferable for wheat breeding in locations with similar conditions.

1 **Genetic diversity of gliadin alleles in bread wheat (*Triticum aestivum* L.) from**

2 **Northern Kazakhstan**

3

4 Maral Utebayev [1,2]*, Svetlana Dashkevich [1], Nina Bome [2], Kulpash Bulatova [3], Yuri Shavrukov [4]

5

6 [1] A.I. Barayev Research and Production Centre of Grain Farming, Shortandy, Kazakhstan;

7 [2] Institute of Biology, University of Tyumen, Tyumen, Russia;

8 [3] Kazakh Research Institute of Agriculture and Plant Growing, Almaty Region, Kazakhstan;

9 [4] College of Science and Engineering, School of Biological Sciences, Flinders University, SA,

10 Australia;

11

12 * Correspondence: Maral Utebayev. E-mail: phytochem@yandex.ru

13

14 **Abstract**

15 Spring bread wheat (*Triticum aestivum* L.) represents the main cereal crop in Northern

16 Kazakhstan. The quality of wheat grain and flour strongly depends on the structure of gluten,

17 comprised of gliadin and glutenin proteins. Electrophoresis spectra of gliadins are not altered by

18 environmental conditions or plant growth, are easily reproducible and very useful for wheat

19 germplasm identification in addition to DNA markers. Genetic polymorphism of two *Gli* loci

20 encoding gliadins can be used for selection of preferable genotypes of wheat with high grain

21 quality. Polyacrylamide gel electrophoresis was used to analyse genetic diversity of gliadins in a

22 germplasm collection of spring bread wheat from Northern Kazakhstan. The highest frequencies

23 of gliadin alleles were found as follows, in *Gli1*: *-A1f* (39.3%), *-B1e* (71.9%), and *-D1a* (41.0%);

24 and in *Gli-2*: *-A2q* (17.8%), *-B2t* (13.5%), and *-D2q* (20.4%). The combination of these alleles in

25 a single genotype may be associated with higher quality of grain as well as better adaptation to

26 the dry environment of Northern Kazakhstan; preferable for wheat breeding in locations with

27 similar conditions.

28

29   **Introduction**

30   Wheat flour remains one of main ingredients in quite a diverse range of foods for human

31   consumption and provides the major proteins gliadins and glutenins. In particular, glutenin can

32   make up at least 40% of the total protein in grain and flour (*Qi et al., 2006; Metakovsky et al.,*

33   *2018*). The genetic control of gliadin includes two major genes, *Gli-1* and *Gli-2*, mapped to the

34   short arms of chromosome groups 1 and 6, respectively, with corresponding homeologous genes,

35   *Gli-A1, -B1, -D1* and *Gli-A2, -B2, -D2* (*Metakovsky, Branlard & Graybosch, 2006; Metakovsky*

36   *et al., 2018*). Multiple alleles are typically found for both *Gli* loci. Each *Gli* allele encodes the

37   transcription of clusters of subunits, with several components of gliadin proteins showing linked

38   inheritance. Gliadin groups can differ in the number of components, their electrophoretic

39   mobility and molecular weight, and levels of expression (*Sozinov & Poperelya, 1980; Obukhova*

40   *& Shumny, 2016*). By its nature, gliadin is a complex protein with several components that can

41   be separated using polyacrylamide gel electrophoresis in aluminium-lactate buffer (pH=3.1)

42   (*Bushuk & Zillman, 1978*). The original protocol of gliadin electrophoresis has been since

43   modified (*Tkachuk & Metlish, 1980; Khan, Hamada & Patek, 1985; Metakovsky &*

44   *Novoselskaya, 1991*), and was used as the basis for the International standard procedure ISO

45   (*ISO 1993*). *Gli* alleles and their components have been widely studied and identified in

46   International wheat germplasm collections, resulting in published Catalogues. The genetic

47   polymorphism in the composition of *Gli* alleles in a given genotype was summarised as the

48   'Gliadin genetic formula' (GGF) in the Catalogues for bread wheat (*Metakovsky, 1991;*

49   *Metakovsky et al., 2018*) and for durum wheat (*Melnikova, Kudryavtseva & Kudryavtsev, 2012*).

50

51   As reported in many publications, wheat cultivars produced in each separate country often have

52   similar GGF despite the absence of any selection pressure based on gliadins (*Xynias, Kozub &*

53   *Sozinov, 2006; Aguiriano et al., 2008; Salavati et al., 2008; Melnikova et al., 2010;*

54   *Novoselskaya-Dragovich et al., 2011; Hailegiorgis et al., 2017*). A linkage between *Gli* alleles

55   and other genes or a group of genes encoding favourable traits can be preferable and beneficial

56   for wheat breeding (*Chebotar et al., 2012*). Therefore, a high frequency of *Gli* alleles can be used

57   as simple and convenient method based on protein marker analysis for wheat germplasm

58   identification and application in further breeding programs in the same environment.

59

60 Currently, molecular markers based on DNA analysis are widely used for genotyping and genetic

61 identification in various crops (*Shavrukov, 2016; Jatayev et al., 2017; Scheben, Batley &*

62 *Edwards, 2017; Burridge et al., 2018*). The application of molecular markers was successful in

63 the study of wheat genes controlling such traits as 1000-grain weight, protein and gluten content

64 (*Zhang et al., 2018*), grain hardness (*Nirmal et al., 2016*), flour production from grain milling

65 (*Nirmal et al., 2017*), and bread quality (*Henry, Furtado & Rangan, 2018*). Genome editing

66 using CRISPR/Cas9 technology represents a novel method in plants (*Khlestkina & Shumny,*

67 *2016; Liang et al., 2018; Borisjuk et al., 2019*), for production of wheat with low gluten content

68 (Sánchez-León et al. 2018), as required by people allergic to some components of gliadin in

69 traditional wheat cultivars (*Palosuo et al., 2001; Pastorello et al., 2007*).

70

71 Nevertheless, molecular markers are relatively expensive in the equipment and reagents required,

72 in typically well-established molecular laboratories. In contrast, biochemical markers based on

73 proteins such as enzymes and storage proteins offer an alternative method involving cheaper and

74 simpler protocols for crop breeding including wheat (*Shewry & Halford, 2001; Ghanti et al.,*

75 *2009; Al-Doss et al., 2010; Netsvetaev, Akinshina & Bondarenko, 2010; Hailegiorgis, Lee &*

76 *Yun, 2017*). Additionally, protein synthesis is encoded by genes, and modulation of gene

77 expression in response to changes in the environment directly results in different levels of the

78 corresponding proteins.

79

80 The aim of this study was to identify and analyse the genetic diversity of the *Gli* alleles in spring

81 bread wheat (*Triticum aestivum* L.) collection from Northern Kazakhstan, and to address the

82 question of which alleles of gliadins with highest frequencies are typical for modern wheat

83 produced and cultivated in the dry environment of this region.

84

85 **Materials and Methods**

86 ***Wheat germplasm and geographic locations***

87 A seed collection of 70 bread wheat cultivars was provided by the A.I. Barayev Research and

88 Production Centre of Grain Farming, Shortandy, Kazakhstan. The studied wheat accessions were

89 bred and produced at different times by Breeding Institutions (Karabalyk Agricultural Breeding

90   Station and Pavlodar Research Institute of Agriculture) in Northern Kazakhstan, as listed in

91   **Supplementary material 1**. Additional data for various wheat germplasms from Kazakhstan and

92   neighbouring regions, used for comparison of the results obtained for genetic diversity of *Gli*

93   alleles in wheats, were retrieved from papers published earlier (**Supplementary material 2)**. In

94   the map (**Figure 1)**, Northern Kazakhstan and two nearby regions in Russia with wheat Breeding

95   Research Organisations – Saratov (European part) and Omsk (Siberia) are indicated by ovals.

96   Briefly, Northern Kazakhstan is located at latitude 51°-55° N and longitude 61°-79° E, with a

97   territory of about 565K km^2 comprised largely of steppe or low-hilled forest. The strong

98   continental climate is characterized by a cold and long winter with high winds, but a hot and

99   short summer season. Average winter/summer temperatures are about -18°C and 20°C but

100   extreme levels of -45°C and 41°C, respectively can also be reached.

101

102   **[Insert Figure 1 here]**

103

104   ***Electrophoresis and identification of* Gli *alleles***

105   Polyacrylamide gel electrophoresis was carried out following a method published earlier

106   (*Metakovsky & Novoselskaya 1991*). Gliadins were extracted from individually milled seeds by

107   adding 150 µL of 70% ethanol. Acrylamide polymerization was initiated by 50 µL of 3 % H_2O_2

108   in 45 mL of gel solution. A vertical gel tank, model VE-20 (Helicon, Russia) was used and the

109   gels were 17.8 × 17.5 × 1 (mm) in size. Electrophoresis was conducted at optimal temperature

110   below 20°C, at 520 V for 4 h. 10 % trichloroacetic acid supplied with 0.05% of Coomassie

111   Brilliant Blue R-250 in ethanol (Sigma-Aldrich, USA) was used for gel fixation and staining.

112   The identification of gliadin components was conducted using the Protein Catalogue

113   (*Metakovsky, 1991*). Genes that encoded gliadins were identified in accordance to the Gene

114   Catalogue developed by (*McIntosh at al., 2008*) for *Gli-1* (*-A1*, *-B1*, and *-D1*) and for *Gli-2* (*-A2*,

115   *-B2*, and *-D2*). Alleles of the *Gli* locus were designated as additional Latin letters and total GGFs

116   were used as recommended for wheat cv. Chinese Spring with the following in full GGF: *Gli-*

117   *A1a*, *Gli-B1a*, *Gli-D1a*, *Gli-A2a*, *Gli-B2a*, *Gli-D2a*; and abbreviated GGF: *a, a, a, a, a, a*.

118

119   ***Computer and statistical analysis***

120   Intra-population diversity ($\mu \pm S_\mu$) and frequency of rare alleles ($h \pm S_h$) were calculated following

121   the method published by Zhivotovsky (*Zhivotovsky, 1991*), while genetic diversity (*H*) was

122   calculated by the formula described by Nei, where p_i is the frequency of alleles (*Nei, 1973*):

123                    $$H = 1 - \sum p_i^2$$

124   Phylogenetic tree construction and clustering analysis among the studied wheat genotypes was

125   carried out using the computer program software Statistica 6.0. (Statsoft, USA) following

126   instructions for Ward's method with Manhattan distances and applied for GGFs.   The 'Data

127   Standardization' option was applied to transform the allele identifications in letters into numbers

128   suitable for the computer program software

129   (http://documentation.statsoft.com/STATISTICAHelp.aspx?path=Cluster/ClusterAnalysis/Exam

130   ples/Example1JoiningTreeClustering).

131

132   **Results**

133   **Gli *allele diversity***

134   The alleles of loci *Gli-1* and *Gli-2* identified in the wheat germplasm collection (70 accessions)

135   and their GGFs are presented in **Supplementary material 1.** Most of the studied wheats were

136   monomorphic (76%) while the remaining 24% accessions were polymorphic. Grains of such

137   polymorphic wheats consisted of a mixture of genotypes, with variable alleles in one or more *Gli*

138   loci. For example, several biotypes of gliadins were present in polymorphic cv. Lutescence 65

139   with various spectra of gliadin components in three zones, α-, β- and γ, but identical in ω- zone

140   (**Figure 2, lanes 1-3**). Plants of cv. Byrlestik were monomorphic and represent the single type of

141   gliadin spectrum (**Figure 2, lanes 5-7**). In general, the inter-cultivar polymorphic alleles of *Gli*

142   encode the biosynthesis of gliadin components located in all four zones (α-, β-, γ- and ω- zones)

143   of the gliadin spectrum on the polyacrylamide gel electrophoregram (**Figure 2**).

144

145   **[Insert Figure 2 here]**

146

147   At the *Gli-1* locus, the highest frequencies were found in alleles *Gli-A1f* (38.7%), *-B1e* (62.1%),

148   and *-D1a* (33.6%). In contrast, the level of highest frequency of alleles was smaller at the *Gli-2*

149   locus and comprising *Gli-A2b* (17.14%), *-B2t* (12.9%), and *-D2q* (23.6%). Therefore, the GGF of

150   the majority of wheats bred and cultivated in Northern Kazakhstan is: *f, e, a, b, t, q,* based on

151   highest frequencies of the alleles. In total, results of gliadin electrophoresis revealed six and eight

152   alleles in *Gli-B1* and *Gli-D1* loci, respectively, 14 alleles in each of three loci, *Gli-A1*, *Gli-A2*

153   and *Gli-D2*, and 17 alleles in *Gli-B2* locus (**Figure 3**).

154

155   **[Insert Figure 3 here]**

156

157   Levels of genetic diversity (H), intra-population diversity (μ) and frequencies of rare alleles (h)

158   in 70 wheat accessions from Northern Kazakhstan were calculated based on allele frequencies in

159   *Gli* loci from **Supplementary material 1** and are presented in **Table 1A.** For comparison, four

160   other studies of wheat from Northern Kazakhstan (**Supplementary material 2**) with partial

161   overlap in the accessions studied were joined together with the current study, with the combined

162   results for 139 wheat accessions in total from Northern Kazakhstan presented in **Table 1B**.

163

164   **[Insert Table 1 here]**

165

166   Most of the results presented in **Table 1A** and **1B** are very similar, indicating for a representable

167   subset of 70 wheat accessions for Northern Kazakhstan. For example, genetic diversity, H, was

168   highest in loci *Gli-B2* (0.92/0.93) and *Gli-A2* (0.89/0.90), while smallest H=0.53/0.45 were

169   calculated for *Gli-B1* in both parts of Table 1. The same trend has been found for intra-

170   population diversity μ=15.13/17.37 and 12.04/14.44 for alleles of loci *Gli-B2* and *Gli-A2*,

171   respectively, with maximal number of the identified alleles (17 and 14 alleles, respectively). In

172   contrast, the locus *Gli-B1* had the smallest value of μ=3.65/4.33 with only six identified alleles

173   as the smallest number in this study and with highest frequency of the *Gli-B1e* allele (**Table 1**,

174   **Figure 3**).

175

176   The structure of intra-population diversity can be characterised by the frequencies of rare alleles

177   (h). A population can be estimated as 'balanced' if values of h are less than 0.3 and as small as

178   possible (*Zhivotovsky, 1980*). Therefore, the most balanced for intra-population diversity was

179   found for locus *Gli-B2* (h=0.11/0.13), while locus *Gli-B1* had the highest value for h (0.39/0.52)

180   due to the highest frequency of a single allele, *Gli-B1e*.

181

182   The highest frequencies of each gliadin allele in the combined group of 139 wheat accessions

183   were accounted as: *Gli-A1f* (39.3%), *-B1e* (71.9%), *-D1a* (41.0%), *-A2q* (17.8%), *-B2t* (13.5%),

184   and *-D2q* (20.4%). The GGF in the analysis of 139 wheat accession was as follows: *f, e, a, q, t,*

185   *q*, and almost identical to those identified in the current study, with only a single difference for

186   *Gli-A2-q* or *-b*. Therefore, the most typical GGF in wheat accessions from Northern Kazakhstan

187   can be identified as: *f, e, a, q+b, t, q*.

188

189   ***Comparative phylogenetic analysis of the biodiversity of gliadin-coding loci in bread wheat***

190   ***from Northern Kazakhstan and other origins***

191   A gliadin dendrogram (**Figure 4**) was established based on a cluster analysis of our combined

192   current and previously published results of allele variation in the *Gli* loci and GGF, in wheat

193   originating from Northern Kazakhstan (**Supplementary materials 1 and 2**) and other publicly

194   available data for wheat from other countries (**Table 2**).

195

196   **[Insert Figure 4 here]**

197

198   **[Insert Table 2 here]**

199

200   Two major Clades (designated as A and B) were found in GGF, with strong separation of the

201   analysed accessions. Wheat genotypes from Australia, America and Western Europe form Clade

202   A, while the more diverse Clade B includes accessions with *Gli* alleles mostly distributed in

203   Eastern Europe and Asia, with the exception of the UK. As expected, all wheat cultivars from

204   Northern Kazakhstan had GGF most closely related to Russian wheats, particularly those

205   developed in the two big Breeding Research Institutes in Saratov and Omsk, in the European and

206   Siberian part of Russia, respectively. These regions are very close to Northern Kazakhstan

207   geographically (**Figure 1**), and also have a long history of exchange of wheat germplasms within

208   the former Soviet Union.

209

210   **Discussion**

211   The presented study is an important part of the breeding program for seed quality in wheat, to

212   illustrate breeder selections for wheat genotypes with various combinations of gliadin alleles.

213   The received results can be used as the basis of a breeding strategy for wheat genotype selection

214   with preferred GGF and favourable combinations of *Gli* alleles. In the current study, genetic

215   origin, gliadin characteristics and the value of breeding for alleles in each of gliadin-coding loci

216   in 70 wheat accessions will be discussed in separate sub-sections.

217

218   *Locus Gli-A1*

219   Fifteen *Gli-A1* alleles were identified in the current study in wheat cultivars from Northern

220   Kazakhstan, out of 29 alleles published in the recent Catalogue of gliadin-coding genes

221   (*Metakovsky et al., 2018*). The highest frequency (0.39) was found in genotypes with allele *f*.

222   The wide-spread occurrence of the allele *f* in 27 wheat accessions from Northern Kazakhstan out

223   of the 70 studied seems to be related to introgression of the following high grain quality

224   cultivars: Cesium 111 (*f*), Albidum 24 (*f*) and Saratovskaya 29 (*j*+*f*) in the early stages of the

225   wheat breeding process in Kazakhstan (*Metakovsky et al., 2006*).

226

227   The possible origin and spread of other *Gli-A1* alleles, *i*, *o* and *b*, is likely also related to wheat

228   cultivars from Russia. For example, a series of cultivars entitled Omskaya 20, 22 and 23,

229   originating from the forest-steppe zone of South-Western Siberia, with allele *Gli-A1o*, seems to

230   be used in the exchange breeding process (*Metakovsky et al., 2006*). This conclusion is similar to

231   those in our previous published results using a different set of wheat cultivars from Northern

232   Kazakhstan (*Utebayev et al., 2016*).

233

234   *Locus Gli-B1*

235   Very limited genetic diversity was found in the *Gli-B1* locus, where the single allele *e* showed

236   the absolute highest frequency at 62% (**Figure 3**). It is important to note that this allele is quite

237   widely distributed, especially in southern regions of European Russia (*Novoselskaya-Dragovich*

238   *et al., 2003*) as well as in South-Eastern and South-Western Siberia, which are close and directly

239   neighbouring to Northern Kazakhstan, respectively (*Nikolaev, Pukhal'sky & Upelniek, 2009*).

240   The occurrence and quite frequent distribution of the allele *Gli-B1e* may be directly related to the

241   actively-used popular Russian drought tolerant cultivars with elite grain quality from the Saratov

242   region: Albidum 43 (*f, e, a, q, o, a*), Lutescence 62 (*j, e, a, q, o, a*), and Saratovskaya 29 (*j+f, e,*

243   *a, q+s, q+s, e*) (*Metakovsky et al., 2006*). However, these cultivars had some disadvantages,

244   particularly a sensitivity to a range of diseases (*Morgounov, Rosseeva & Koyshibayev, 2007*).

245

246   The second allele, *Gli-B1b,* with three-fold less frequency (29%) has a much wider distribution

247   among wheat cultivars from Scandinavian countries to Australia (*Metakovsky et al., 2018*) and

248   may therefore indicate the wide adaptability of genotypes with this allele. It is very likely that the

249   *Gli-B1b* allele is originated from historic and classical winter wheat cultivars bred in the former

250   Soviet Union, Besostaya 1 (*b, b, b, b, b, b*) and Mironovskaya 808 (*f, b, g, n, m, e*) (*Metakovsky*

251   *et al., 2006*).

252

253   ***Locus Gli-D1***

254   There were eight alleles identified in this locus in the studied wheat accessions, which is exactly

255   half of all that were published in the recent Catalogue of gliadin-coding loci (*Metakovsky et al.,*

256   *2018*). Three alleles showed the highest range of frequencies: *Gli-D1a*, 0.34; *-b*, 0.31; and *-f*,

257   0.16 (**Figure 3**). However, the spectrum of genetic diversity in the present study slightly differed

258   from our paper published earlier with another set of bread wheat accessions from Northern

259   Kazakhstan, where only four *Gli-D1* alleles were identified with the following frequencies: allele

260   *a*, 44.2%; and each of alleles *f* and *i*, 23.3%, respectively (*Utebayev et al., 2016*).

261

262   Similar to those indicated for other alleles above, Russian wheat cultivars were widely used in

263   the initial breeding programs in Northern Kazakhstan. Therefore, it is very likely that the most

264   commonly distributed allele, *a*, is originated from one or several cultivars, particularly Albidum

265   43, Lutescence 62, or Saratovskaya 29 (*Chernakov & Metakovsky 1994; Nikolaev, Pukhal'sky &*

266   *Upelniek, 2009*). Additionally, this allele, *Gli-D1a*, had quite high frequencies among wheat

267   cultivars in Southern Kazakhstan, with a very different environment, but the origin of the allele *a*

268   from the former Soviet Union wheat germplasm genepool is not in doubt (*Absattarova, 2002*).

269   This statement is in complete consensus with data for GGF in Kazakh wheats published in a

270   recent review (*Metakovsky et al., 2018*). The comparison of world-wide distribution of *Gli* allele

271   *a* among wheat genotypes bred and grown in Croatia, Finland and Spain (*Sontag-Strohm ,1997;*

272 *Metakovsky et al., 2018*), indicated for a possible association between allele *Gli-D1a* with

273 adaptability of wheat plants to various environments.

274

275 It is important to note that two *Gli-D1* alleles, *a* and *f*, encode the synthesis of almost identical

276 spectra of gliadin components. The only additional gliadin component present with smaller size

277 in the γ-zone of protein electrophoresis was recorded in wheat genotypes *Gli-D1* with allele *a* but

278 not with allele *f*. Therefore, it is hypothesised that wheat genotypes *Gli-D1a* and *-f* can have very

279 similar gliadin gene nucleotide sequences (*Chebotar et al. 2012*).

280

281 The moderately distributed allele *b* is also very likely to have originated from foreign wheat

282 accessions introgressed earlier in the Kazakh breeding program. However, it is interesting that

283 the *Gli-D1b* allele originates from a very different genepool of winter wheat, rather than spring

284 wheat. This statement is based on published data showing a quite high distribution of the allele *b*

285 among winter wheat, but not in spring wheat, in the former Soviet Union (*Kozub et al., 2009;*

286 *Novoselskaya-Dragovich et al., 2015*). Therefore, we can speculate that the possible

287 introgression of the *Gli-D1b* allele from winter wheat can indicate for the wide adaptability of

288 wheat genotypes, regardless of their responses to cold and vernalisation.

289

290 ***Locus Gli-A2***

291 The *Gli-2* gene is much more diverse in wheat, where the smallest number of alleles were

292 recorded in *Gli-A2* and accounted for 14 (**Figure 3**) of the 39 registered in the recent Catalogue

293 of *Gli* alleles (*Metakovsky et al., 2018*). The most commonly distributed alleles among the

294 studied wheat cultivars from Northern Kazakhstan were: *Gli-A2b* (17.1%), *-f* (12.1%), and *-q*

295 (15.0%). The first allele *b* was very typical for wheat cultivars from very diverse geographical

296 regions and had similarities to wheats from the UK, Eastern Europe and the Krasnodar region in

297 the southern part of Russia (*Metakovsky et al., 2018*). Winter wheat germplasm accessions also

298 had about 22% of the allele *Gli-A2b* (*Novoselskaya-Dragovich et al., 2015*), and this allele is

299 particularly spread among wheat cultivars with high tolerance to cold temperatures (*Markarova,*

300 *2015*). This leads us to the conclusion that the *Gli-A2b* allele may be associated with genotypes

301 with high adaptability to unfavourable conditions for plant growth.

302

303   The allele *Gli-A2f* was present in wheat cultivars originating from the Saratov region, Russia

304   (*Novoselskaya-Dragovich, Fisenko & Puhalskii, 2013*) and in some winter wheat cultivars

305   (*Novoselskaya-Dragovich et al., 2015*) but is known to show the highest frequencies in spring

306   wheat from Mexico and Portugal (*Metakovsky et al., 2018*).

307

308   The third allele, *Gli-A2q,* was very likely introgressed and spread widely in wheat cultivars in

309   Northern Kazakshtan from germplasm originating from the nearby Russian regions of Saratov

310   and Omsk (*Novoselskaya-Dragovich, Fisenko & Puhalskii, 2013*). For example, cv. Lutescence

311   62 was widely used for hybridisations in Kazakhstan with GGF (*j, e, a, q, o, a*) from the Saratov

312   Breeding Institute, and it was consequently bred during individual selection of plants of the

313   original historical cv. Poltavka (*f+j, e, a, q+k, o, a+e*) (*Rutz, 2005; Metakovsky et al., 2006*). The

314   influence of the wheat genepool originating from the Saratov region on the wheat breeding

315   program in Northern Kazakhstan was described in the genetic polymorphism of *Gli* alleles in

316   papers published a relatively long time ago (*Sozinov, Metakovsky & Koval, 1986; Metakovsky et*

317   *al., 1988*). However, among Kazakh wheat cultivars with elite quality of grain, only the allele

318   *Gli-A2q* had the highest frequency of distribution, indicating for a possible genetic association

319   with high grain quality (*Dobrotvorskaya et al., 2009*).

320

321   ***Locus Gli-B2***

322   Seventeen out of 45 *Gli* alleles described in recent Catalogues (*Metakovsky et al., 2018*) were

323   identified and analysed in the current study. The highest frequency was found for the allele *Gli-*

324   *B2t*, 12.8%, followed by 10.7% for alleles *-b* and *-g*, respectively. The origin of the first allele *t*

325   remains unclear because it was registered as a minor *Gli* allele in some modern wheat cultivars

326   from the Omsk Breeding Station, Russia (*Chernakov & Metakovsky, 1994*). We can propose that

327   the origin of the allele *Gli-B2t* is likely related to the old Russian cv. Cesium 111 used for

328   hybridisations with GGF (*f, m, i, j, t, i*) and published earlier (*Metakovsky et al., 2006;*

329   *Morgounov, Rosseeva & Koyshibayev, 2007*).

330

331   The occurrence and distribution of allele *Gli-B2b* is definitely related to the use and introgression

332   of wheat accessions from Eastern Europe and Russia, where this allele was exclusively present

333   (*Metakovsky et al., 2018*). In contrast, the *Gli* allele *g* very likely originates from one of the wide

334   geographically dispersed countries such as the Scandinavian group (*Metakovsky et al., 2018*), the

335   UK (*Chernakov & Metakovsky, 1994*), France (*Metakovsky & Branlard, 1998*), and China

336   (*Novoselskaya-Dragovich et al., 2011*).

337

338   ***Locus Gli-D2***

339   The sixth and last gliadin-coding locus, *Gli-D2*, was present with 14 alleles. The three most

340   widely distributed alleles were *q*, *b* and *a*, with corresponding percentage of frequencies: 23.5%,

341   17.8% and 11.4%, respectively. In the comparison with gliadin allele distributions, *Gli-D2**b*** was

342   originated from Russian wheat germplasm (*Metakovsky et al., 2018*). Both *q* and *a* alleles were

343   widely distributed in local wheats from Northern Kazakhstan, and regarding our previous study,

344   allele *Gli-D2**a*** was for the first time found in three Kazakh wheat cultivars, Milturum 45,

345   Tzelinogradka and Snegurka (*Utebayev et al., 2016*). These three cultivars were included in

346   wheat breeding in Northern Kazakhstan as genetic donors, and the first two of them (Milturum

347   45 and Tzelinogradka) were bred from original, old and polymorphic cv. Cesium 111 with GGF

348   − *f, m, i, j, t, <u>a</u>*+*e* (*Metakovsky et al., 2006*). It is more likely that modern Kazakh wheat

349   genotypes with allele *Gli-D2**a*** had a pedigree progenitor from one of the biotypes of cv. Cesium

350   111. Less likely, but still possible, is that the origin of the *a* allele is from other countries where it

351   was found, such as Croatia, Germany, France, Holland, Italy, Scandinavian countries, Spain or

352   the UK (*Metakovsky et al., 2018*), indicating for a possible wide interest for wheat breeding

353   programs.

354

355   ***Comparison of genetic diversity between* Gli-1 *and* Gli-2 *alleles***

356   In both our current and previous study (*Utebayev et al., 2016*), the three most popular and widely

357   distributed modern spring bread wheat cultivars from Northern Kazakhstan with elite grain

358   quality have the following GGF: Akmola 2 (***g, e, a, i, e, s***), Astana (***g**+**j, e, f**+**i, p, h, b***), and

359   Karabalykskaya 90 (***i**+**m**+**f, e, a**+**g, q**+**l, v, a***) (**Supplementary material 2**). These cultivars have

360   a similar composition of alleles in the gene *Gli-1,* with three homeologous loci (-*A1*, -*B1* and -

361   *D1*) to wheat cultivars with very high grain quality from the Russian Breeding Institutes, Saratov

362   and Omsk. Therefore, it was hypothesised that allele compositions in each of three loci of *Gli-1*

363   were directly related to grain and baked bread quality and its improvement (*Li et al. 2009;*

364   *Novoselskaya-Dragovich, Fisenko & Puhalskii, 2013*). In contrast, allele compositions in the

365   second gene *Gli-2* with three homeologous loci (*-A2*, *-B2* and *-D2*) located in chromosome group

366   6, were genetically associated with possible adaptation of plants to a dry environment

367   (*Novoselskaya-Dragovich, Fisenko & Puhalskii, 2013*).

368

369   Such a conclusion, made from the comparison between *Gli-1* and *Gli-2* genes, may explain how

370   non-pedigree related wheat cultivars from various geographic regions with a different climate

371   have very similar or identical compositions of *Gli-1* alleles. This is because one of the main

372   targets of wheat breeding is the production of wheat with elite quality of grain and baked bread,

373   where genetic diversity for allele composition in *Gli-1* is much smaller than in *Gli-2*: 14, 6 and 8

374   alleles for *Gli-A1*, *-B1* and *-D1*; and 14, 17 and 14 alleles for *Gli-A2*, *-B2* and *-D2*, respectively

375   (**Figure 3**). It is possible that a single perfect pedigree genotype with excellent grain quality was

376   used as a progenitor in many modern wheat cultivars, providing limited genetic variability in

377   allele composition of *Gli-1*. In contrast, the *Gli-2* gene, with much wider variability in allele

378   compositions, was more likely involved in plant adaptation to a dry environment. Because such

379   environments are quite variable in different countries and geographic regions, it may be reflected

380   in and explain the higher variability in allele diversity in *Gli-2*. The presented results reflect the

381   efforts of wheat breeders over many years of artificial selection based on phenotyping variability

382   in grain quality and tolerance to dry environments, as apparent in the results of genetic diversity

383   in both gliadin-coding genes based on gliadin analyses.

384

385   **Conclusions**

386   Genetic diversity in the alleles of gliadin-coding genes *Gli-1* and *Gli-2* was studied, and gliadin

387   genetic formulas were established following the results of gliadin electrophoresis in a set of 70

388   spring bread wheat cultivars from Northern Kazakhstan. The *Gli* alleles with highest frequencies

389   in the studied wheat material were identified as follows: *Gli-A1f* (39.3%), *-B1e* (71.9%), *-D1a*

390   (41.0%), *-A2q* (17.8%), *-B2t* (13.5%), and *-D2q* (20.4%). This allele combination of both *Gli*

391   genes was the most widely distributed in Northern Kazakhstan, and genotypes with such gliadin

392   formula can be used as prospective breeding material for elite grain quality, better adaptability to

393   the dry environment of the Northern Kazakhstan region and for wheat breeding under similar

394   conditions.

395

396   **Acknowledgments**

397   We thank Carly Schramm for critical comments and English editing.

398

399

400   **References**

401   **Absattarova AS. 2002**. Identification of Kazakhstan agroecotypes winter common wheat cultivars using

402   of gliadin components blocks. Dissertation of the Candidate of Biological Sciences. Kazakh

403   Research Institute of Agriculture and Plant Growing. Almalybak (in Russ.).

404   **Aguiriano E, Ruiz M, Fité R, Carrillo JM. 2008.** Genetic variation for glutenin and gliadins associated

405   with quality in durum wheat (*Triticum turgidum* L. ssp. *turgidum*) landraces from Spain. *Spanish*

406   *Journal of Agricultural Research* **6(4):**599-609 DOI 10.5424/sjar/2008064-353

407   **Al-Doss AA, Al-Hazmi AS, Dawabah AAM, Abdel-Mawgood AA, Al-Rehiayani SM, Solaiman A,**

408   **Moustafa K, Motawei M. 2010.** Impact of *Cre* and peroxidase genes of selected new wheat lines on

409   cereal cyst nematode (*Heterodera avenae* Woll) resistance. *Australian Journal of Crop Science*

410   **4(9):**737-743.

411   **Borisjuk N, Kishchenko O, Eliby S, Schramm C, Anderson P, Jatayev S, Kurishbayev A,**

412   **Shavrukov Y. 2019.** Genetic modification for wheat improvement: from transgenesis to genome

413   editing. *BioMed Research International* **2019:** 6216304.

414   **Burridge AJ, Wilkinson PA, Winfield MO, Barker GL, Allen AM, Coghill JA, Waterfall C,**

415   **Edwards KJ. 2018.** Conversion of array-based single nucleotide polymorphic markers for use in

416   targeted genotyping by sequencing in hexaploid wheat (*Triticum aestivum*). *Plant Biotechnology*

417   *Journal* **16(4):**867-876 DOI 10.1111/pbi.12834

418   **Bushuk W, Zillman RR. 1978.** Wheat cultivar identification by gliadin electrophoregrams. I. Apparatus,

419   method and nomenclature. *Canadian Journal of Plant Science* **58:**505-515 DOI 10.4141/cjps78-076

420   **Chebotar SV, Blagodarova EM, Kurakina EA, Semenyuk IV, Polishchuk AM, Kozub NA, Sozinov**

421   **IA, Khokhlov AN, Ribalka AI, Sivolap YuM. 2012.** Genetic polymorphism of loci determining

422   bread making quality in Ukrainian wheat varieties. *Vavilov Journal of Genetics and Breeding*

423   **16(1):**87-98 (In Russ).

424   **Chernakov VM, Metakovsky EV. 1994.** Diversity of gliadin-coding locus allelic variants and evaluation

425   of genetic similarity of common wheat varieties from different breeding counters. *Genetika*

426   **30(4):**509-517 (in Russ).

427   **Dobrotvorskaya TV, Dragovich AY, Martynov SP, Pukhal'skii VA. 2009.** Genealogical and statistical

428   analyses of the inheritance of gliadin-coding alleles in a model set of common wheat *Triticum*

429   *aestivum* L. cultivars. *Russian Journal of Genetics* **45(6):**685-695 DOI 10.1134/S1022795409060088

430   **Ghanti SK, Sujata KG, Rao S, Udayakumar M, Kishor PBK. 2009.** Role of enzymes and

431   identification of stage-specific proteins in developing somatic embryos of chickpea (*Cicer arietinum*

432   L.). *In Vitro Cellular and Developmental Biology – Plant* **45:**667-672 DOI 10.1007/s11627-009-

433   9197-7

434   **Hailegiorgis D, Lee CA, Yun SJ. 2017.** Allelic variation at the gliadin coding loci of improved

435   Ethiopian durum wheat varieties. *Journal of Crop Science and Biotechnology* **20(4):**287-293 DOI

436   10.1007/s12892-017-0106-0

437   **Henry RJ, Furtado A, Rangan P. 2018.** Wheat seed transcriptome reveals genes controlling key traits

438   for human preference and crop adaptation. *Current Opinion in Plant Biology* **45:**231-236 DOI

439   10.1016/j.pbi.2018.05.002

440   **ISO 898:1993**, Wheat – Identification of varieties by electrophoresis.

441   **Jatayev S, Kurishbaev A, Zotova L, Khasanova G, Serikbay D, Zhubatkanov A, Botayeva M,**

442   **Zhumalin A, Turbekova A, Soole K, Langridge P, Shavrukov Y. 2017.** Advantages of

443   Amplifluor-like SNP markers over KASP in plant genotyping. *BMC Plant Biology* **17:**254 DOI

444   10.1186/s12870-017-1197-x

445   **Khan K, Hamada AS, Patek J. 1985.** Polyacrylamide gel electrophoresis for wheat variety

446   identification: effect of variables on gel properties. *Cereal Chemistry* **62(5):**310-313.

447 **Khlestkina EK, Shumny VK. 2016.** Prospects for application of breakthrough technologies in breeding:

448   The CRISPR/Cas9 system for plant genome editing. *Russian Journal of Genetics* **52(7):**676-687 DOI

449   10.1134/S102279541607005X

450 **Kozub NA, Sozinov IA, Sobko TA, Kolyuchii VT, Kuptsov SV, Sozinov AA. 2009.** Variation at

451   storage protein loci in winter common wheat cultivars of the central forest-steppe of Ukraine.

452   *Cytology and Genetics* **43(1):**55-62 DOI 10.3103/S0095452709010101

453 **Li Y, Song Y, Zhou R, Branlard G, Jia J. 2009.** Detection of QTLs for bread-making quality in wheat

454   using a recombinant inbred line population. *Plant Breeding* **128:**235–243 DOI 10.1111/j.1439-

455   0523.2008.01578.x

456 **Liang Z, Chen K, Zhang Y, Liu J, Yin K, Qiu JL, Gao C. 2018.** Genome editing of bread wheat using

457   biolistic delivery of CRISPR/Cas9 *in vitro* transcripts or ribonucleoproteins. *Nature Protocols*

458   **13(3):**413-430 DOI 10.1038/nprot.2017.145

459 **Markarova ZhR. 2015.** Frost hardiness and yield of soft winter wheat cultivars and accessions in the

460   Rostov region. *Nauchny Zhurnal Rossiiskogo NII Problem Melioratsii* **2(18):**137-145 (in Russ).

461 **McIntosh RA, Yamazaki Y, Devos KM, Dubkovsky J, Rogers J, Appels R. 2008.** MacGene 2003.

462   Catalogue of Gene Symbols for Wheat. In Proc. 10th Int. Wheat Genet. Symp. Paestum, Italy.

463   Available at http://wheat.pw.usda.gov/ggpages/wgc/2003. (Accessed: 07. 03. 2019).

464 **Melnikova NV, Ganeva GD, Popova ZG, Landjeva SP, Kudryavtsev AM. 2010.** Gliadins of

465   Bulgarian durum wheat (*Triticum durum* Desf.) landraces: genetic diversity and geographical

466   distribution. *Genetic Resources and Crop Evolution* **57(4):**587-595 DOI 10.1007/s10722-009-9497-0

467 **Melnikova NV, Kudryavtseva AV, Kudryavtsev AM. 2012.** Catalogue of alleles of gliadin-coding loci

468   in durum wheat (*Triticum durum* Desf.). *Biochimie* **94(2):**551-557 DOI 10.1016/j.biochi.2011.09.004

469 **Metakovsky E, Melnik V, Quijano MR, Upelniek V, Carrillo JM. 2018.** A catalog of gliadin alleles:

470   Polymorphism of 20th-century common wheat germplasms. *The Crop Journal* **6(6):**628-641 DOI

471   10.1016/j.cj.2018.02.003

472   **Metakovsky EV, Branlard G, Graybosch RA, Bekes F, Cavanagh CR, Wrigley CW, Bushuk W.**

473   **2006.** The gluten composition of wheat varieties and genotypes. Part I. Gliadin Composition Table.

474   Available at https://www.aaccnet.org/initiatives/definitions/Documents/GlutenFree/I_Gliadin.pdf.

475   (Accessed: 07. 03. 2019).

476   **Metakovsky EV, Branlard G. 1998.** Genetic diversity of French common wheat germplasm based on

477   gliadin alleles. *Theoretical and Applied Genetics* **96:**209-218 DOI 10.1007/s001220050729

478   **Metakovsky EV, Branlard GP, Graybosch RA. 2006.** Gliadins of common wheat: Polymorphism and

479   genetics. In: Wrigley C, Bekes F, Bushuk W, ed. Gliadin and Glutenin: the Unique Balance of Wheat

480   Quality. St. Paul American Association of Cereal Chemistry, 36.

481   **Metakovsky EV, Koval SF, Movchan VK, Sozinov AA. 1988.** Genetic formulas of gliadin in cultivars

482   of common wheat of Northern Kazakhstan. *Selekciya i Semenovodstvo* **1:**11-13 (in Russ).

483   **Metakovsky EV, Novoselskaya AY. 1991.** Gliadin allele identification in common wheat. 1.

484   Methodological aspects of the analysis of gliadin pattern by one-dimensional polyacrylamide - gel

485   electrophoresis. *Journal of Genetics and Breeding* **45:**317-324.

486   **Metakovsky EV. 1991.** Gliadin allele identification in common wheat. 2 Catalogue of gliadin alleles in

487   common wheat. *Journal of Genetics and Breeding* **45:**325-344.

488   **Morgounov A, Rosseeva L, Koyshibayev M. 2007.** Leaf rust of spring wheat in Northern Kazakhstan

489   and Siberia: incidence, virulence, and breeding for resistance. *Australian Journal of Agricultural*

490   *Research* **58(9):**847-853 DOI 10.1071/AR07086

491   **Nei M. 1973.** Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy*

492   *of Sciences USA* **70(12):**3321-3323 DOI 10.1073/pnas.70.12.3321

493   **Netsvetaev VP, Akinshina OV, Bondarenko LS. 2010.** Genetic control of several α-amylase isozymes

494   in winter hexaploid wheat. *Russian Journal of Genetics* **48(3):**347-349 DOI

495   10.1134/S1022795412030106

496 **Nikolaev AA, Pukhal'sky VA, Upelniek VP. 2009.** Genetic diversity of local spring bread wheats

497 (*Triticum aestivum* L.) of West and East Siberia in gliadin genes. *Russian Journal of Genetics*

498 **45(2):**189-197 DOI 10.1134/S1022795409020094

499 **Nirmal RC, Furtado A, Rangan P, Henry RJ. 2017.** Fasciclin-like arabinogalactan protein gene

500 expression is associated with yield of flour in the milling of wheat. *Scientific Reports* **7:**12539. DOI

501 10.1038/s41598-017-12845-y

502 **Nirmal RC, Furtado A, Wrigley C, Henry RJ. 2016.** Influence of gene expression on hardness in

503 wheat. *PLoS One* **11(10):**e0164746 DOI 10.1371/journal.pone.0164746

504 **Novoselskaya-Dragovich AY, Bespalova LA, Shishkina AA, Melnik VA, Upelniek VP, Fisenko AV,**

505 **Dedova LV, Kudryavtsev AM. 2015.** Genetic diversity of common wheat varieties at the gliadin-

506 coding loci. *Russian Journal of Genetics* **51(3):**262–271 DOI 10.1134/S1022795415030102

507 **Novoselskaya-Dragovich AY, Fisenko AV, Puhalskii VA. 2013.** Genetic differentiation of common

508 wheat cultivars using multiple alleles of gliadin coding loci. *Russian Journal of Genetics* **49(5):**487–

509 496 DOI 10.1134/S1022795413020087

510 **Novoselskaya-Dragovich AY, Fisenko AV, Yankovsky NK, Kudryavtsev AM, Yang Q, Lu Z, Wang**

511 **D. 2011.** Genetic diversity of storage protein genes in common wheat (*Triticum aestivum* L.)

512 cultivars from China and its comparison with genetic diversity of cultivars from other countries.

513 *Genetic Resources and Crop Evolution* **58(4):**533-543 DOI 10.1007/s10722-010-9596-y

514 **Novoselskaya-Dragovich AY, Krupnov VA, Saifulin RA, Pukhalskiy VA. 2003.** Dynamics of genetic

515 variation at gliadin-coding loci in Saratov cultivars of common wheat *Triticum aestivum* L. over

516 eight decades of scientific breeding. *Russian Journal of Genetics* **39(10):**1130–1137 DOI

517 10.1023/A:1026170709964

518 **Obukhova LV, Shumny VK. 2016.** The inheritance of endosperm storage proteins by the line of the

519 Saratovskaya 29 cultivar of common wheat from its parental forms. *Russian Journal of Genetics*

520 **52(1):**49-55 DOI 10.1134/S1022795416010117

521 **Palosuo K, Varjonen E, Kekki OM, Klemola T, Kalkkinen N, Alenius H, Reunala T. 2001.** Wheat ω-

522 5 gliadin is a major allergen in children with immediate allergy to ingested wheat. *The Journal of*

523 *Allergy and Clinical Immunology* **108(4):**634-638 DOI 10.1067/mai.2001.118602

524 **Pastorello EA, Farioli L, Conti A, Pravettoni V, Bonomi S, Iametti S, Fortunato D, Scibilia J,**

525 **Bindslev-Jensen C, Ballmer-Weber B, Robino AM, Ortolani C. 2007**. Wheat IgE-mediated food

526 allergy in European patients: α-amylase inhibitors, lipid transfer proteins and low-molecular-weight

527 glutenins. *International Archives of Allergy and Immunology* **144(1):**10-22. DOI 10.1159/000102609

528 **Qi PF, Wei YM, Yue YW, Yan ZH, Zheng YL. 2006.** Biochemical and molecular characterization of

529 gliadins. *Molecular Biology* **40(5)**:713-723 DOI 10.1134/S0026893306050050

530 **Rutz RI. 2005.** Breeding Center of Siberian Research Institute of Agriculture is a leader of Siberian

531 breeding. *Vestnik VOGiS* **9(3)**:357-368 (in Russ.).

532 **Salavati A, Sameri H, Boushehri AAS, Yazdi-Samadi B. 2008.** Evaluation of genetic diversity in

533 Iranian landrace wheat *Triticum aestivum* L. by using gliadin alleles. *Asian Journal Plant Science*

534 **7(5)**:440-446 DOI 10.3923/ajps.2008.440.446

535 **Sánchez-León S, Gil-Humanes J, Ozuna CV, Giménez MJ, Sousa C, Voytas DF, Barro F. 2018.**

536 Low-gluten, non-transgenic wheat engineered with CRISPR/Cas9. *Plant Biotechnology Journal*

537 **16(4):**902-910 DOI 10.1111/pbi.12837

538 **Scheben A, Batley J, Edwards D. 2017.** Genotyping-by-sequencing approaches to characterize crop

539 genomes: choosing the right tool for the right application. *Plant Biotechnology Journal* **15(2):**149-

540 161 DOI 10.1111/pbi.12645

541 **Shavrukov Y. 2016.** Comparison of SNP and CAPS markers application in genetic research in wheat and

542 barley. *BMC Plant Biology* **16**:11 DOI 10.1186/s12870-015-0689-9

543 **Shewry PR, Halford NG. 2001.** Cereal seed storage proteins: structures, properties and role in grain

544 utilization. *Journal of Experimental Botany* **53(370)**:947-958 DOI 10.1093/jexbot/53.370.947

545   **Sontag-Strohm T. 1997.** Gliadin and glutenin subunit alleles on group 1 chromosomes in Finnish spring

546   wheats. *Acta Agriculturae Scandinavica, Section B - Soil and Plant Science* **47**:98-105 DOI

547   10.1080/09064719709362446

548   **Sozinov AA, Metakovsky EV, Koval SF. 1986.** Patterns of genotype formation in the breeding of the

549   wheat. *Vestnik Selskohozyaistvennoi Nauki* **3:**60-70 (in Russ.).

550   **Sozinov AA, Poperelya FA. 1980.** Genetic classification of prolamins and its use for plant breeding.

551   *Annales de Technologie Agricole* **29:**229-245.

552   **Tkachuk R, Metlish VJ. 1980.** Wheat cultivar identification by high voltage gel electrophoresis. *Annales*

553   *de Technologie Agricole* **29:**207-212.

554   **Utebayev M, Dashkevich S, Babkenov A, Shtefan G, Fahrudenova I, Bayahmetova S, Sharipova B,**

555   **Kaskarbayev Zh, Shavrukov Y. 2016.** Application of gliadin polymorphism for pedigree analysis

556   in common wheat (*Triticum aestivum* L.) from Northern Kazakhstan. *Acta Physiologiae Plantarum*

557   **38(8):**204 DOI 10.1007/s11738-016-2209-4

558   **Xynias IN, Kozub NO, Sozinov IA. 2006.** Seed storage protein composition of Hellenic bread wheat

559   cultivars. *Plant Breeding* **125(4)**:408-410 DOI 10.1111/j.1439-0523.2006.01242.x

560   **Zhang Y, Li D, Zhang D, Zhao X, Cao X, Dong L, Liu J, Chen K, Zhang H, Gao C, Wang D. 2018.**

561   Analysis of the functions of *TaGW2* homoeologs in wheat grain weight and protein content traits.

562   *The Plant Journal* **94(5):**857-866 DOI 10.1111/tpj.13903

563   **Zhivotovsky LA. 1980.** An index intrapopulation diversity. *Zhurnal Obshchei Biologii* **41(6):**828-836 (in

564   Russ).

565   **Zhivotovsky LA. 1991**. Population Biometrics. Moscow. Nauka RAN (in Russ).

566

567 **Tables**

568 **Table 1.** Genetic diversity (H), intra-population diversity (μ) and frequencies of rare alleles (h)

569 in 70 (**A**) and 139 combined (**B**) wheat accessions from Northern Kazakhstan

570

571 **Table 2.** GGF of wheat cultivars from various countries compared to the GGF of Northern

572 Kazakhstan wheats

573

574

575 **Figure Legends**

576

577 **Figure 1**. A map of Kazakhstan and nearby regions of Russia. The red oval shows Northern

578 Kazakhstan, while the Russian regions, Saratov (European part) and Omsk (Siberia), are shown

579 in blue and black, respectively. The map was taken from the web-site:

580 http://theconversation.com/russias-borders-moscows-long-alliance-with-kazakhstan-is-strong-

581 but-not-unbreakable-36457.

582

583 **Figure 2.** Electrophoregram of the gliadin spectrum of polymorphic cv. Lutescence 65 (Lanes 1-

584 3) in comparison to cv. Besostaya 1 (Lane 4, used as a Standard) and monomorphic cv. Byrlestik

585 (Lanes 5-7). Sub-fractions α, β, and γ with polymorphic bands in Lanes 1-3 are indicated by

586 arrows.

587

588 **Figure 3.** Allele frequencies in *Gli* loci identified in the studied collection of 70 accessions of

589 spring bread wheat from Northern Kazakhstan.

590

591 **Figure 4.** Gliadin dendrogram showing the allele diversity in *Gli* loci of bread wheat from
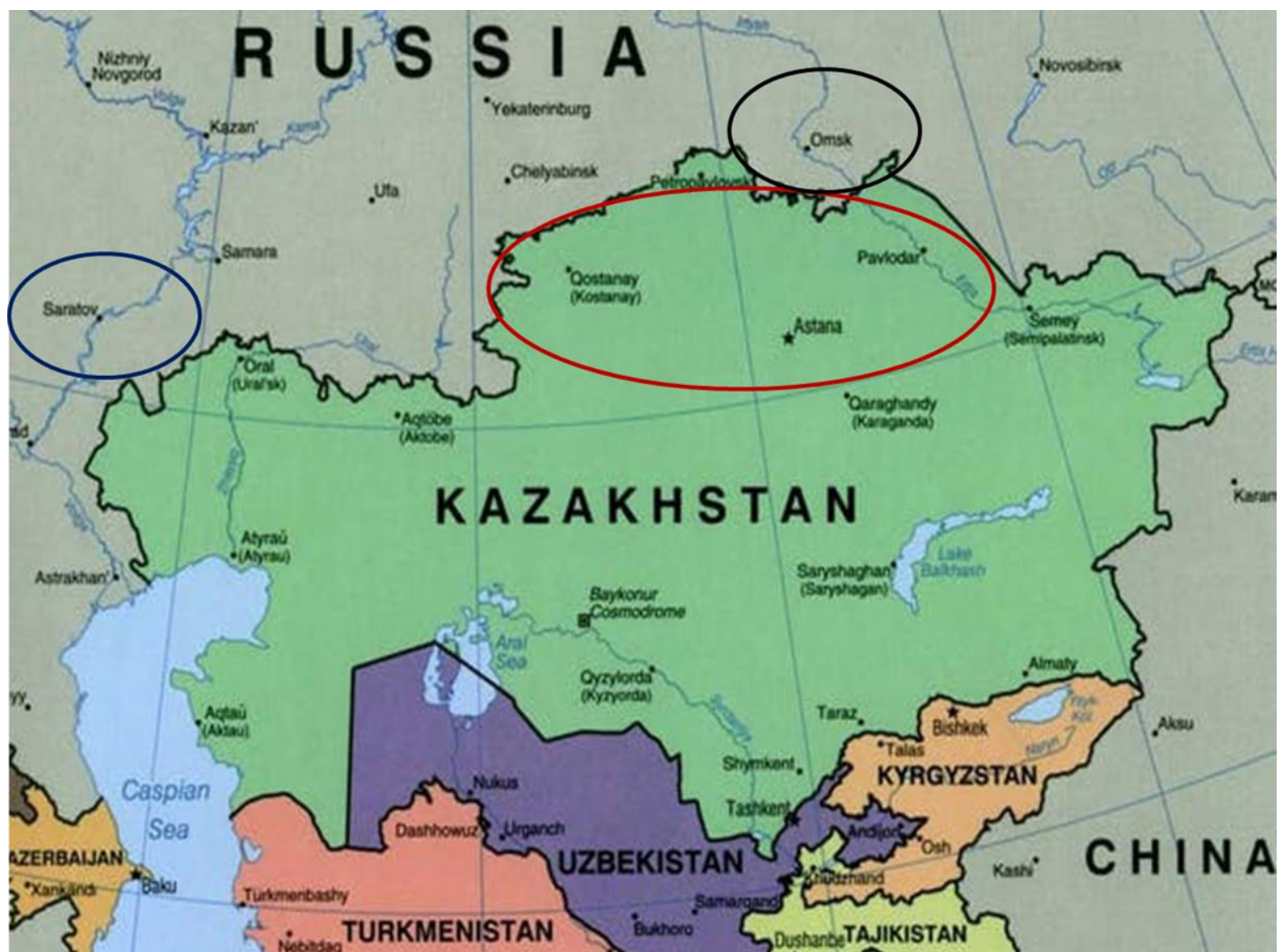
592 Northern Kazakhstan and other countries.

593

Figure 1

A map of Kazakhstan

Figure 1. A map of Kazakhstan and nearby regions of Russia. The red oval shows Northern Kazakhstan, while the Russian regions, Saratov (European part) and Omsk (Siberia), are shown in blue and black, respectively. The map was taken from the web-site: http://theconversation.com/russias-borders-moscows-long-alliance-with-kazakhstan-is-strong-but-not-unbreakable-36457
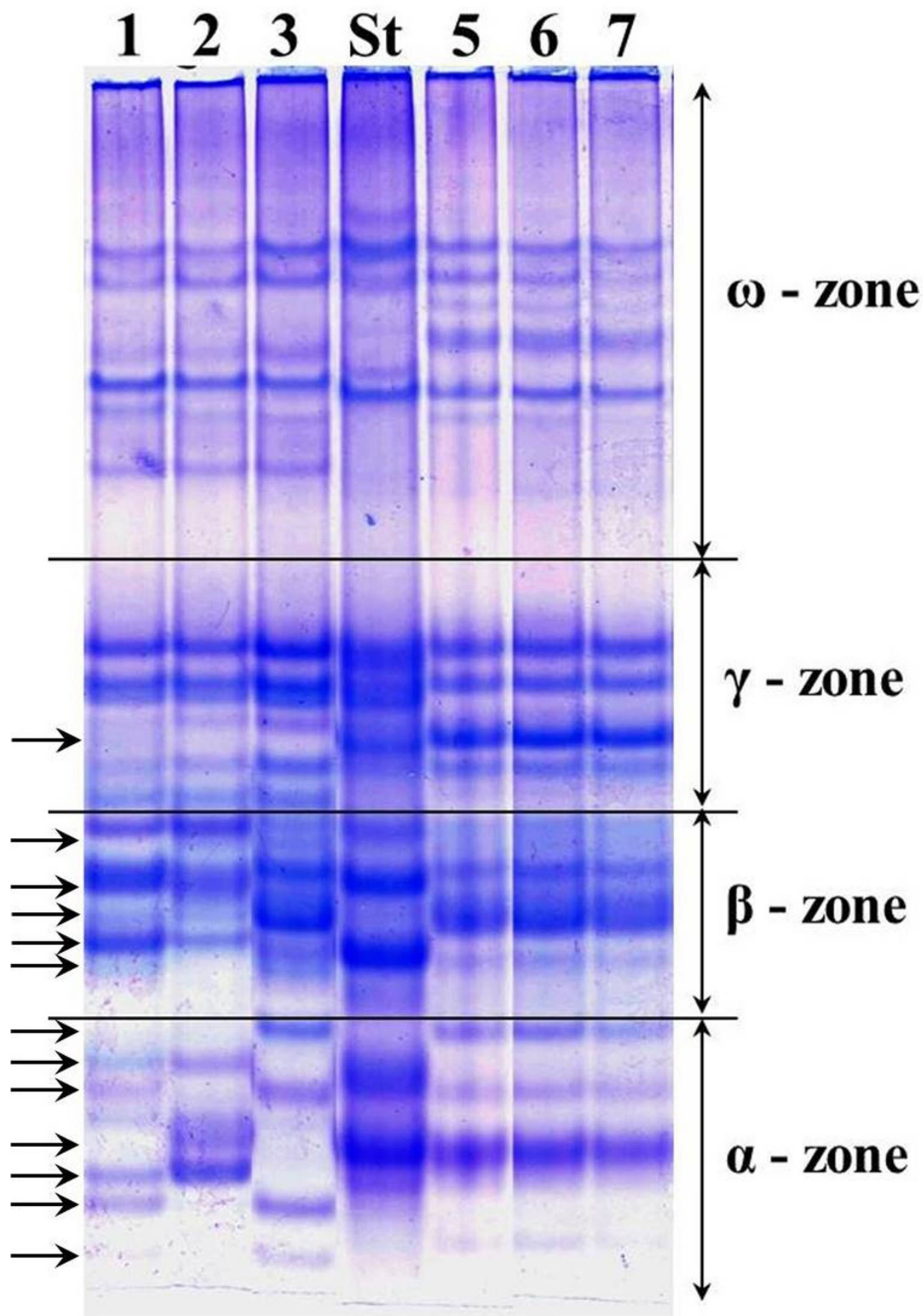
Figure 2

Electrophoregram of the gliadin spectrum

Figure 2. Electrophoregram of the gliadin spectrum of polymorphic cv. Lutescence 65 (Lanes 1-3) in comparison to cv. Bezostaya 1 (Lane 4, used as a Standard) and monomorphic cv. Byrlestik (Lanes 5-7). Subfractions α , β, γ and ω with polymorphic bands are indicated.

Table 1(on next page)

Geneticdiversity (H), intra-population diversity (μ) and frequencies of rare alleles(h) in 70 (**A**) and 139 combined (**B**) wheat accessions from NorthernKazakhstan

1   **Table 1.** Genetic diversity (H), intra-population diversity (μ) and frequencies of rare alleles (h)

2   in 70 (**A**) and 139 combined (**B**) wheat accessions from Northern Kazakhstan

3

Diversity estimates	Gliadin-coding *Gli* loci					
	A1	*B1*	*D1*	*A2*	*B2*	*D2*
A. 70 wheat accessions from Northern Kazakhstan (Supplementary material 1)						
H	0.81	0.53	0.76	0.89	0.92	0.87
$\mu \pm S_\mu$	10.43±0.73	3.65±0.35	6.00±0.41	12.04±0.58	15.13±0.64	11.56±0.63
$h \pm S_h$	0.25±0.05	0.39±0.06	0.25±0.05	0.14±0.04	0.11±0.04	0.17±0.05
B. 139 wheat accessions from Northern Kazakhstan (Supplementary materials 1 and 2)						
H	0.80	0.45	0.75	0.90	0.93	0.89
$\mu \pm S_\mu$	12.32±0.71	4.33±0.18	6.78±0.40	14.44±0.61	17.37±0.57	13.88±0.56
$h \pm S_h$	0.32±0.04	0.52±0.04	0.32±0.04	0.20±0.03	0.13±0.03	0.18±0.03

4

Table 2(on next page)

GGF of wheat cultivars fromvarious countries compared to the GGF of Northern Kazakhstan wheats
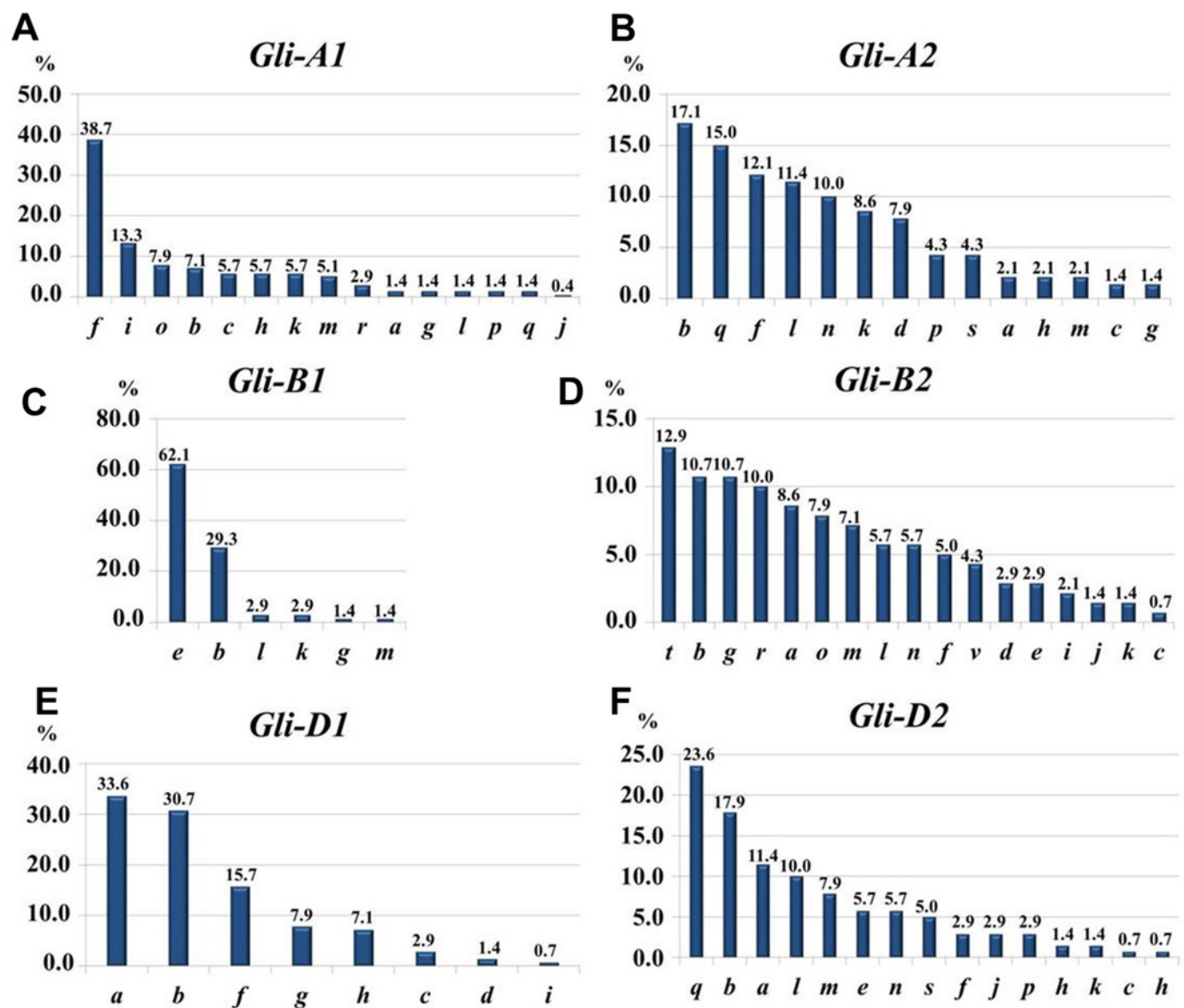
1   **Table 2.** GGF of wheat cultivars from various countries compared to the GGF of Northern

2   Kazakhstan wheats

3

Countries/regions	Gliadin coding loci *Gli*						References
	A1	*B1*	*D1*	*A2*	*B2*	*D2*	
Australia	*g*	*b*	*f*	*c*	*c*	*w*	*Metakovsky et al.,*
Canada	*m*	*d*	*j + a*	*m*	*c*	*h + m*	*2018*
France	*af + c*	*b +f*	*b*	*m*	*o + c*	*h*	
Italy	*a*	*g*	*k*	*g + o*	*o*	*a*	
Mexico	*o + a*	*d + b*	*b + a*	*f*	*c*	*m + j*	
Portugal	*a*	*c + l*	*b*	*f*	*c*	*c + j*	
Russia (Saratov)	*f*	*e*	*a*	*q*	*s*	*e*	
Bulgaria	*b + a*	*b*	*b*	*b + g*	*b*	*b*	
Croatia	*a + b*	*e*	*a*	*e*	*e*	*m + a*	
Serbia	*a + b*	*b + l*	*b*	*b + g*	*b*	*b + a*	
Ukraine	*b*	*b*	*g*	*f + b*	*b*	*e*	
Russia (Omsk)	*f*	*e*	*a*	*q*	*b*	*a*	*Novoselskaya-Dragovich, Fisenko & Puhalskii, 2013*
Iran	*f*	*f*	*b*	*g + l*	*o*	*a + n + h*	*Salavati et al., 2008*
China	*o + a*	*l + e*	*a + f*	*g +f*	*I + h*	*b + a*	*Novoselskaya-Dragovich et al., 2011*
UK	*f*	*f*	*b*	*l*	*g*	*a*	*Chernakov & Metakovsky, 1994*
Northern Kazakhstan	*f*	*e*	*a*	*q*	*t*	*q*	Current study

4

Figure 3

Allele frequencies

Figure 3. Allele frequencies in *Gli* loci identified in the studied collection of 70 accessions of spring bread wheat from Northern Kazakhstan.

Figure 4

Gliadin dendrogram

Figure 4. Gliadin dendrogram showing the allele diversity in *Gli* loci of bread wheat from Northern Kazakhstan and other countries.