Genotyping barley resources – a foundation for future breeding

Fluturë Novakazi[1] & Hrannar Smári Hilmarsson[2]

[1]Department of Plant Breeding,
Swedish University of Agricultural Sciences
Sundsvägen 10, 234 56 Alnarp Sweden
[2]Department of Cultivation and Food
Agricultural University of Iceland
311, Hvanneyri Iceland

Abstract

Barley was cultivated in Iceland from settlement until the 16th century and it was not until the end of the last century that cultivation increased. A breeding program for Icelandic barley started in the 1980s using foreign material as founders. The breeding program bred for earliness and has released eight cultivars. To dissect and investigate the origin of the Icelandic breeding population, we genotyped 65 Icelandic barley breeding lines, supplemented with 27 parental lines and compared them to potentially related genotypes. Additionally, the Icelandic panel (n=92) was analysed together with a core set (n=1000). From admixture analysis we inferred K = 2 and K = 4 ancestral populations in the data. The Icelandic genotypes had ancestry from Denmark, Finland, Norway, Sweden, Faroe Islands and Scotland. Out of the 65 Icelandic lines, 29 showed admixed ancestry. The linkage disequilibrium (LD) was calculated and decay ranged between 1.9 Mbp (4H) and 3.4 Mbp (3H), which is lower than the core set (~0.3 Mbp) and within Northern European material (~1 Mbp), but comparable with the Far East material (~2 Mbp). A PCA including the core set and the 92 genotypes based on 38,328 informative SNPs showed the first two components explained 9.09% and 7.34%. The Icelandic genotypes clustered with the Northern European genotypes, and showed considerable diversity along PC2. The Icelandic barley breeding population is substantially admixed with other Nordic barley. Chromosome-wise LD suggests that the Icelandic population is genetically diverse. Genetic diversity has been introduced by crossing with other Nordic material.

Samantekt

Bygg var ræktað á Íslandi frá landnámi fram að 16. öld en það var ekki fyrr en á 10. áratug síðustu aldar sem ræktun byggs jókst. Kynbótaverkefni fyrir bygg á Íslandi hófst á 9. áratugnum með því að nota erlendan efnivið sem stofnfeður. Verkefnið miðaði að því að kynbæta bygg fyrir flýti og hefur þegar skráð átta yrki til ræktunar. Til þess að skilgreina uppruna íslenska byggsins erfðagreindum við 65 íslenskar bygg arfgerðir og 27 arfgerðir af stofnefnivið með iSelect 50k einbasabreytileika örflögu og bárum saman við opinberlega aðgengileg gögn. Greiningarnar sýndu að fjöldi stofnhópa væru K=2 og K=4. Íslenska byggið átti upprunatengsl við danskt, norskt, sænskt, færeyskt og skoskt bygg. Af 65 íslenskum bygg arfgerðum sýndu 29 arfgerðir samblandaðan uppruna. Greining á tengslaójafnvægi (TÓ) og hnignun þess reyndist vera 1,9 Mbp og 3,4Mbp (3H) lægra en allt opna gagnasafnið (~0,3Mbp) og Norður Evrópska hópinn í opna gagnasafninu (~1Mbp) en nær byggi frá Austurlöndum fjær. Höfuðþáttagreining á íslenska efniviðinum og opna gagnasafninu byggt á 39.328 einbasabreytileikum sýndi að fyrstu tveir höfuðþættirnir útskýrðu 9,09% og 7,34% breytileikans. Íslenski efniviðurinn hópaðist með norður-evrópskum einstaklingum og sýndi talsverðan breytleika á höfuðþætti tvö. Íslenska byggið er verulega samblandað við annað Norrænt bygg. TÓ reiknað innan litninga bendir til þess að íslenska byggið sé erfðafræðilega fjölbreytt. Erfðafjölbreytileiki stofnsins á uppruna sinn frá víxlunum við annað Norrænt bygg.

Introduction

Spring barley (*Hordeum vulgare* L.) is an important cereal in Iceland cultivated on about 3000 hectares. Cereal production in Iceland was substantial until the 16th century, when it reduced and was discontinued (Júlíusson, 2018). In the beginning of the 20th century, cereals were reintroduced but only recently have become a new agricultural sector, in large part thanks to the breeding efforts of Jónatan Hermannsson in the last 30 years (Hilmarsson et al., 2017). The first cultivar released from the program was Skegla (2r, two-row) in 2002, but was replaced by Kria (2r) in 2004, which is still one of the most popular cultivars in Iceland today. In 2006 and 2008, Lómur and Skúmur, both 6r, were released but could not compete with foreign high-yielding 6r cultivars. In 2020, Smyrill (6r) was released in Iceland and is increasing its share on the seed market. Valur (6r) as well as Teista (2r) were registered as cultivars but it is not certain if Teista can outperform Kria in yield or if Valur has as strong a straw as Smyrill, which are important traits according to farmers (Hávarðardóttir, 2019).

In 2017, the former barley breeder of the Agricultural University of Iceland, Jónatan Hermannson, retired, leaving the barley breeding program orphaned. He left a selected collection of breeding lines from his barley crosses, but much of the genetic material was discarded during his final days in office. In the last few years, the viability of the rest of the collection has been threatened by flooding, mice pest, and inappropriate storage. To maintain the current Icelandic barley genetic diversity, we urgently needed to create a genetic database, propagate viable seeds and store fresh seeds under proper conditions.

The Icelandic barley germplasm has not been widely genotyped. It is important that the genetic material from the past decades of breeding are genotyped and that the population structure and genomic variability is analysed using all available seed from the Icelandic barley breeding program

The Icelandic breeding population has shown unique characteristics compared to other Nordic germplasm. According to the pedigree, Icelandic barley has Nordic ancestors but also Faroese, Scottish, Alaskan and even Central American. The trait heading date seems to have been the focus of the Icelandic breeding program. Barley selected for Icelandic conditions has shown to head earlier (Hilmarsson et al., 2017) and some genotypes are extremely early compared to other Nordic material (Göransson, et al., 2021) and that the extreme early lines are a week faster than the fastest barley lines on the market (Göransson, et al., 2021). Phenotypic data shows that Icelandic barley genotypes have been heading earlier while increasing in yield (Jökulsson, 2019). However, there has been a lack of genetic research on other traits in Icelandic barley.

Plant breeding is not a new science, although constantly evolving to increase the efficiency of plant breeding programs. A plant breeding program is a process that aims to develop and release new cultivars. To do so the program must generate genetic variability and select and test individuals to identify the best individuals for distribution. This selection can be done using phenotypic data, pedigree data, and genomic data. In recent years, breeding programs based on dense genomic information are being developed. Such programs have the potential to increase the annual genetic gain by a factor of 2-3 compared to classical breeding programs, which rely on phenotyping only (Tessema et al., 2020). To date, the Icelandic barley breeding program relied only on phenotypic data.

Genomic selection is a method that is based on using genotyping information and was introduced at the beginning of the century (Meuwissen et al., 2001). It has since revolutionized plant breeding in cereals such as wheat (Tessema et al., 2020). Genotyping and application of genomic selection give more accurate predictions of important traits, thus minimizing the risk of

releasing cultivars from the breeding program that do not perform well in farmers' fields as was the case for Skegla, Lómur, and Skúmur. Genomic selection in barley has increased genetic gain and lowered costs (Sallam & Smith, 2016) and has been applied in barley breeding programs in Finland (Haikka et al., 2020) and proved to be especially useful for traits that are difficult or expensive to screen for, such as malting quality (Schmidt et al., 2016) and nutrient value (Nielsen et al., 2016).

Aims

The main objective of this project is to produce and preserve genetic information of Icelandic barley resources and analyse the genetic diversity.

The objective of this project is a necessary first step to achieve other aims of the barley breeding program, which include:

1) Create an Icelandic barley gene bank database for our genetic barley resources by using recognized genotyping methods and assess population structure and genetic variability of the Icelandic barley breeding germplasm.
2) To perform genomic studies to identify novel quantitative trait loci associated with agronomic and quality traits and genome-wide association studies for traits of interest under Icelandic conditions.
3) Assess the feasibility of genomic prediction models under Icelandic conditions to determine a future implementation of genomic selection in a sustainable breeding program and compare the prediction accuracies based on phenotyping data, genotyping data and pedigree data.

Materials and methods

The first step of the project was to disentangle the AUI seed storage at Keldnaholt and pair seed bags with pedigree and phenotype information. Of 92 genotyped individuals, 65 were Icelandic breeding lines and 27 individuals were considered as ancestor material of the Icelandic barley breeding population and are listed in Table 1.

Table 1. Names of barley genotypes and their ID number

ID		Barley name	ID		Barley name
IceBar	001	X 107-1	IceBar	047	X 134-2
IceBar	002	X 123-8	IceBar	048	X 124-11
IceBar	003	V 15-2	IceBar	049	X 251 F3
IceBar	004	X 172 8Ax Sj918011)	IceBar	050	V 6-1
IceBar	005	X 124-8	IceBar	051	ÁB-19
IceBar	006	VOH 2845	IceBar	052	X 97-11
IceBar	007	V 34-7-A	IceBar	053	X 33-11
IceBar	008	X 96-13	IceBar	054	X 106-9
IceBar	009	ÁB-4	IceBar	055	V 297-8
IceBar	010	X 2-36	IceBar	056	Klementína
IceBar	011	X 247 F3	IceBar	057	Olve
IceBar	012	X 135-4	IceBar	058	Bode

IceBar	013	X 123-7		IceBar	059	Porvenir C
IceBar	014	X 123-11		IceBar	060	Artturi
IceBar	015	X 257 F3		IceBar	061	Pohto
IceBar	016	X 251 F2		IceBar	062	Yrjar
IceBar	017	ÁB-3		IceBar	063	Alaska
IceBar	018	X 124-14		IceBar	064	Aapo
IceBar	019	X 90-3		IceBar	065	Etu
IceBar	020	X 124-6		IceBar	066	Agneta
IceBar	021	X 123-1		IceBar	067	Sj. 933247
IceBar	022	V 298-8		IceBar	068	Sj.D 865100
IceBar	023	X 124-10		IceBar	069	Sj. 933182
IceBar	024	V 76-4		IceBar	070	Sj. 922622
IceBar	025	X 248 F3		IceBar	071	Sj. 918011
IceBar	026	X 124-9		IceBar	072	Svå 84164
IceBar	027	ÁB-1		IceBar	073	340-7
IceBar	028	SIGUR-F		IceBar	074	X 25/F2
IceBar	029	SIGUR-C		IceBar	075	Filippa
IceBar	030	SIGUR		IceBar	076	Arve
IceBar	031	TAMPAR		IceBar	077	Nord
IceBar	032	054		IceBar	078	Nairn
IceBar	033	051		IceBar	079	Olsok
IceBar	034	SKOTLAND		IceBar	080	291-13
IceBar	035	X 101-14		IceBar	081	291-17
IceBar	036	X 121-10		IceBar	082	293-6
IceBar	037	V 298-10		IceBar	083	265-46
IceBar	038	X 120-2		IceBar	084	294-47
IceBar	039	V 2-5		IceBar	085	292-54
IceBar	040	X 124-5		IceBar	086	336-10
IceBar	041	V 34-7 B		IceBar	087	293-43
IceBar	042	X 96-14		IceBar	088	294-3
IceBar	043	ÁB-5		IceBar	089	293-45
IceBar	044	046		IceBar	090	333-42
IceBar	045	V 85-16 D		IceBar	091	294-17
IceBar	046	X 99-4		IceBar	092	342-6

Seeds of each individual were shipped to Trait Genetics where DNA was extracted and genotyped with a 50K iSelect SNP (single nucleotide polymorphism) array (Bayer et al., 2017) which has been shown to be a preferred option in the use of barley genetic resources (Darrier et al., 2019). The genotyped material was compared to available genotype data of barley individuals and two datasets were compiled, one with a total of 1093 individuals (the core set) and another with the Icelandic panel (n=92) together with 20 Nordic individuals (Nordic data set, n=112) from the core set. Monomorphic, unmapped (Morex v3) and SNPs with a call rate < 95% were removed from the analysis.

The genetic variability within the germplasm was assessed with model-based clustering (STRUCTURE Software) to estimate individual ancestry of the 112 genotypes (Pritchard et al., 2000). The optimal K was assessed using the STRUCTURE HARVESTER stand-alone version in Python v2.7 with the EVANNO method (Evanno et al., 2005). Additionally, a principal component analysis (PCA) was conducted with the SMARTPCA function in the EIGENSOFT V5 software (Patterson et al., 2006). Linkage disequilibrium (LD) was estimated by calculating the squared allele frequency correlation r^2 between marker pairs and chromosome-wise LD decay was estimated by plotting the r^2 values against the physical distance with a second-degree smoothed loess curve.

Nei's genetic distance (Nei, 1972) was estimated with the function GENET.DIST in the R package HIERFSTAT and visualized as a heatmap with the package COMPLEXHEATMAP.

Results

After removing SNPs that were monomorphic, unmapped (Morex v3) and with a call rate < 95% 36 942 (n=112) and 38 328 (n=1093) SNPs were left for subsequent analyses.

The LD decay ranged between 1.9 Mbp (4H) and 3.4 Mbp (3H). The diversity within the Icelandic genotypes is therefore lower than within the core set (~0.3 Mbp) and lower than within Northern European material (~1 Mbp), but comparable with the Far East material (~2 Mbp).
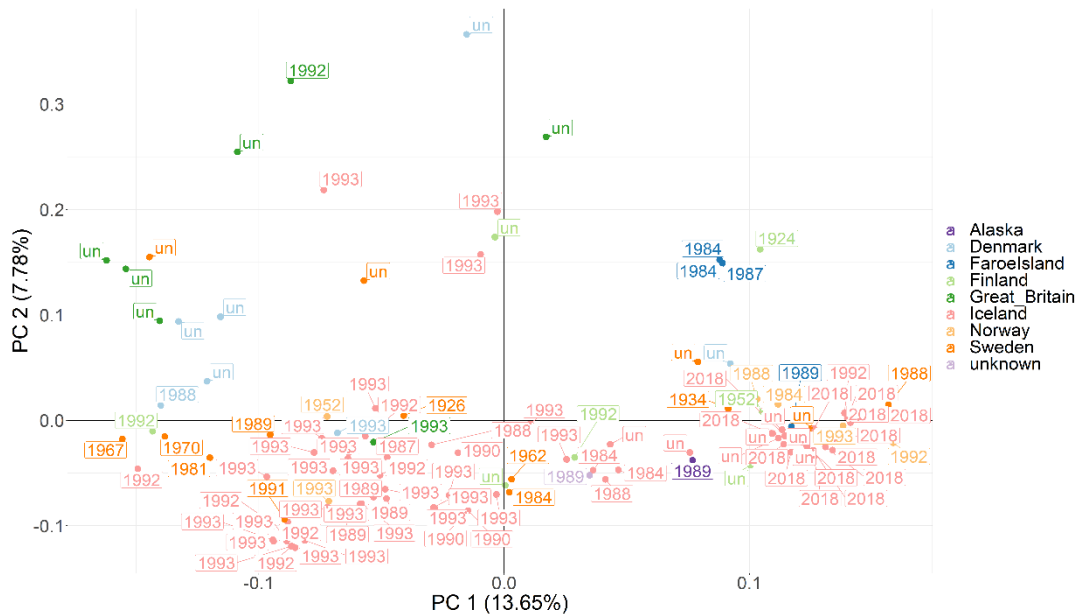
Icelandic and related Nordic material

For the Icelandic individuals and Nordic material from the core set the STRUCTURE HARVESTER results inferred K = 2, which corresponded with the row type. The results also showed a small peak at K = 4 (Figure 1). The Icelandic genotypes had ancestry from Denmark, Finland, Norway, Sweden, Faroe Islands, Great Britain and Alaska. Within the Icelandic breeding lines, 29 showed admixed ancestry, 15 belonged to subpopulation K = 1, 20 to subpopulation K = 3 and one line belonged to subpopulation K = 4. The Faroese landraces belonged almost exclusively to subpopulation K = 2.

Figure 1 Results of STRUCTURE analysis for 112 barley genotypes calculated on 36 942 SNPs.
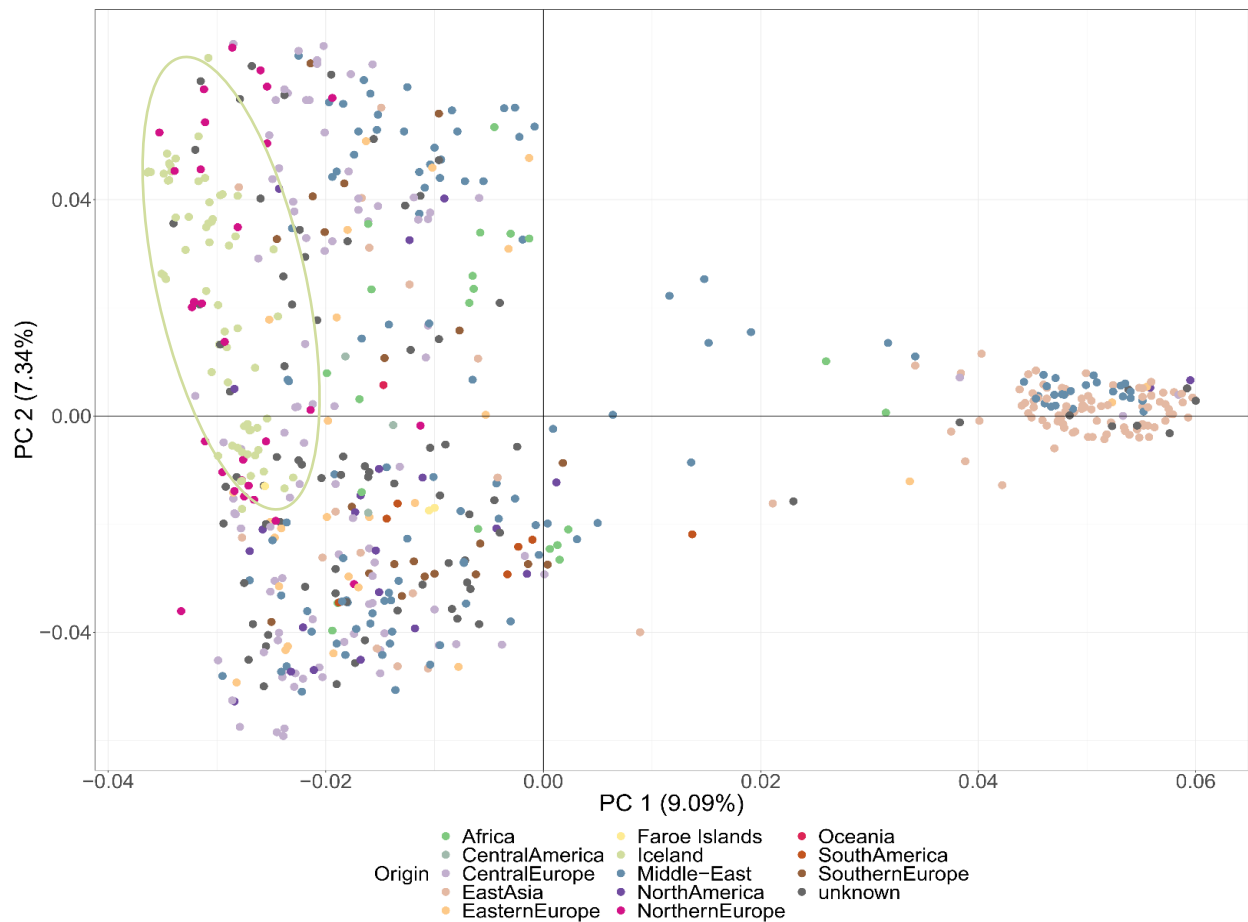
The first two principle components (PCs) explained 21.4% (Figure 2). The Icelandic breeding lines showed considerable diversity along the first PC, which explained 13.65%. This diversity within the Icelandic breeding lines is explained by the year the lines were first tested in the field, with the youngest lines forming a small cluster on the far right. This cluster corresponded to subpopulation K = 3 from the ancestry analysis.

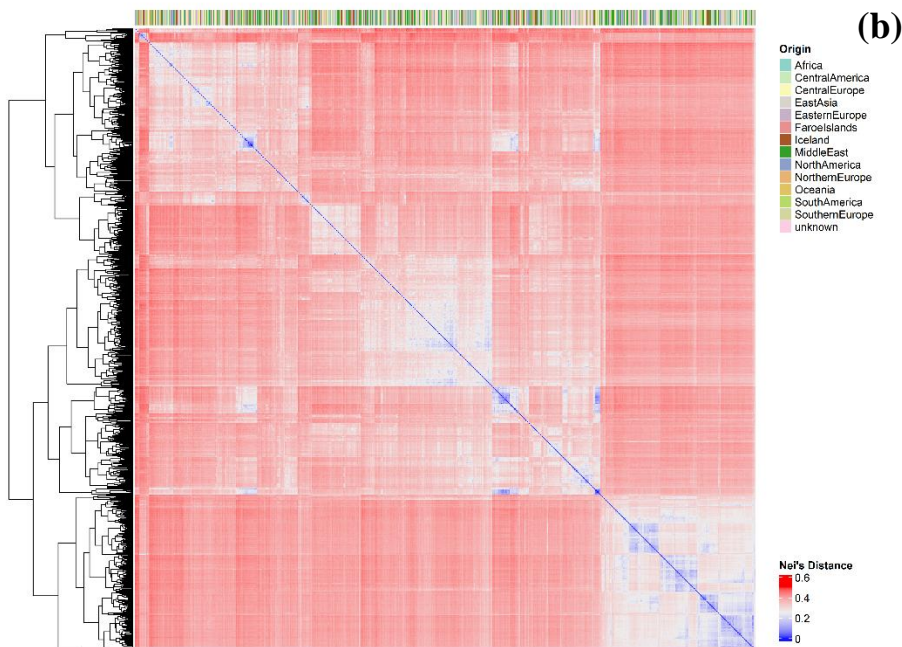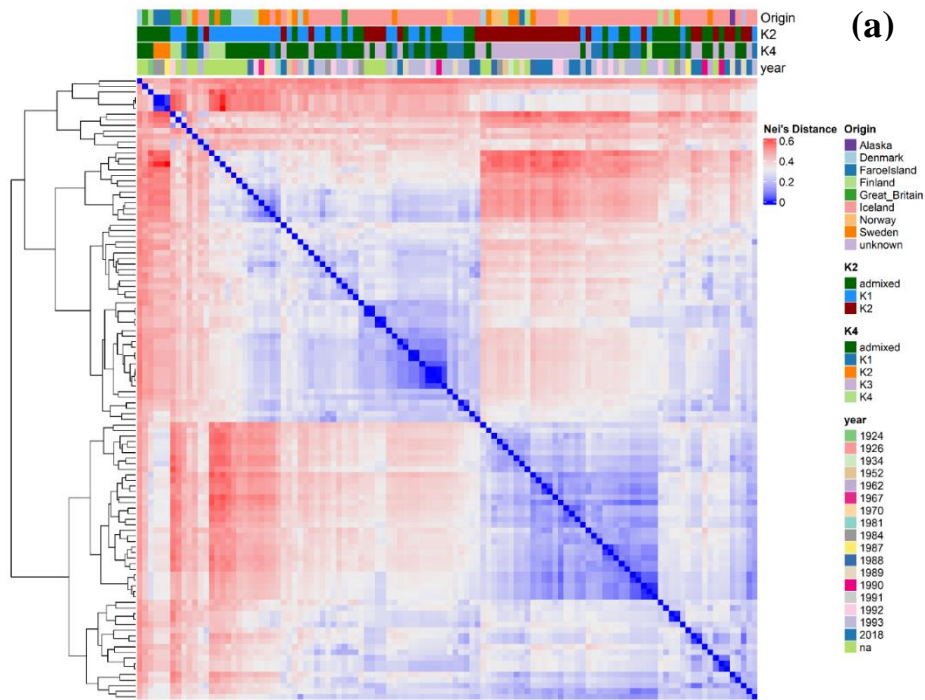**Figure 2** Principal component analysis for 112 barley genotypes calculated on 36 942 SNPs.

Comparison of the Icelandic material to the core set

To put the Icelandic genotypes into international perspective, we conducted a PCA including 1000 genotypes from the barley core set in addition to the 92 genotypes based on 38,328 informative SNPs. The first two components explained 9.09% and 7.34% (Figure 3). The Icelandic genotypes clustered with the Northern European genotypes, and showed considerable diversity along the second PC.

Figure 3 Principal component analysis for 1093 barley genotypes calculated on 38.328 SNPs.

Based on Nei's genetic distance, two big clusters and four smaller clusters of lines were evident within the breeding population (Fig. 2a). Within the core set, two clusters were identified based on Nei's genetic distance (Fig. 2b), with the smaller cluster corresponding to the Middle-Eastern and East Asian material.

Figure 4 Nei's genetic distance for (a) 112 and (b) 1093 barley genotypes.

Discussion

The results for the admixture analysis at K=2 identify ancestry originated in Britain and Nordic countries, with Icelandic barley having higher British ancestry rather than Nordic. At K=4, STRUCTURE identifies ancestry which seems to belong to Swedish and Norwegian material, that is shared with Icelandic barley.

Faroese ancestry can be seen in the Icelandic material. Furthermore, the Icelandic barley breeding population is substantially admixed with all Nordic material (Figure 1). The STRUCTURE HARVESTER results for K=4 (Figure 1) are not convincing, but the row type explains the K=2. The results for the admixture for K=4 is supported by the pedigree information, where the Icelandic material has ancestry from Faroe Islands, Denmark and Scotland in the beginning but more Nordic relation in the later stages of the breeding program (Göransson et al., 2021). It could be the next step to compare the pedigree with the STRUCTURE results to defer the contribution of each founder or Nordic barleys to the Icelandic barley population similar to Thomasen et al (2013).

The PCA analysis for the Icelandic data and the core set showed the older Icelandic material to be furthest to the left in the upper corner, but the more modern Icelandic material is closer to the center of the other material (Figure 3) at least not showing variation on PC1 and somewhat decreasing through time. Hinting that Icelandic barley is not showing specific adaption to Iceland as a separate population but maybe signs of becoming genetically closer to modern varieties, which share traits like bigger kernels, shorter and stronger straw.

Nei's genetic distance analysis did not correspond with the admixture findings but does suggest four groups within the Icelandic material and its founders (Figure 4a)

Chromosome-wise LD suggests that the Icelandic population is genetically diverse. However, looking at the PCA in Figure 2, one could assume that the distances between the modern genotypes were shorter than the distances between the genotypes that were from the early stages of the breeding program. This could mean that the genetic diversity is being reduced or that the program had success in adapting barley to the Icelandic conditions. Furthermore, one could question whether the Icelandic material was developing closer to other Nordic barleys either through crossing or selection and development.

The data gathered here will have an important role in further research projects. Considerable amount of phenotypic data exists for the genotypes analyzed here and could be used together with more genotypic data to identify specific adaptation of barley to the extreme environment in Iceland.

This project will lay the foundation for a new barley breeding program, which will make sustainable use of Icelandic barley genetic resources.

Bibliography

Bayer, M. M., Rapazote-Flores, P., Ganal, M., Hedley, P. E., Macaulay, M., Plieske, J., Ramsay, L., Russell, J., Shaw, P. D., Thomas, W., & Waugh, R. (2017). Development and Evaluation of a Barley 50k iSelect SNP Array. *Frontiers in Plant Science*, *8*. https://doi.org/10.3389/fpls.2017.01792

Darrier, B., Russell, J., Milner, S. G., Hedley, P. E., Shaw, P. D., Macaulay, M., Ramsay, L. D., Halpin, C., Mascher, M., Fleury, D. L., Langridge, P., Stein, N., & Waugh, R. (2019). A Comparison of Mainstream Genotyping Platforms for the Evaluation and Use of Barley Genetic Resources. *Frontiers in Plant Science*, *10*. https://doi.org/10.3389/fpls.2019.00544

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Göransson, M., Hallsteinn Hallsson, J., Bengtsson, T., Bjørnstad, Å., & Lillemo, M. (2021). Specific adaptation for early maturity and height stability in Icelandic spring barley. *Crop Science*, *61*(4), 2306–2323. https://doi.org/10.1002/csc2.20459

Göransson, M., Sigurdardottir, T. H., Lillemo, M., Bengtsson, T., & Hallsson, J. H. (2021). The Winter-Type Allele of HvCEN Is Associated With Earliness Without Severe Yield Penalty in Icelandic Spring Barley (Hordeum vulgare L.). *Frontiers in Plant Science*, *12*. https://www.frontiersin.org/articles/10.3389/fpls.2021.720238

Haikka, H., Knürr, T., Manninen, O., Pietilä, L., Isolahti, M., Teperi, E., Mäntysaari, E. A., & Strandén, I. (2020). Genomic prediction of grain yield in commercial Finnish oat (Avena sativa) and barley (Hordeum vulgare) breeding programmes. *Plant Breeding*, *139*(3), 550–561. https://doi.org/10.1111/pbr.12807

Hávarðardóttir 1989-, G. L. (2019). *Kornrækt á Íslandi. Yrkjaval og reynsla bænda* [Thesis]. https://skemman.is/handle/1946/36071

Hilmarsson, H. S., Göransson, M., Lillemo, M., Kristjánsdóttir, Þ. A., Hermannsson, J., & Hallsson, J. (2017). *An overview of barley breeding and variety trials in Iceland in 1987-2014.* https://doi.org/10.16886/IAS.2017.02

Jökulsson, Í. (2019). *Tengsl skriðdags, uppskeru og þurrefnisprósentu í 2- og 6- raða byggi við íslenskar aðstæður*. Landbúnaðarháskóli Íslands.

Júlíusson, Á. D. (2018). *Af hverju strái: Saga af byggð, grasi og bændum 1300-1700*. Sagnfræðistofnun Háskóla Íslands.

Meuwissen, T. H., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829.

Nei, M. (1972). Genetic Distance between Populations. *The American Naturalist*, *106*(949), 283–292. https://doi.org/10.1086/282771

Nielsen, N. H., Jahoor, A., Jensen, J. D., Orabi, J., Cericola, F., Edriss, V., & Jensen, J. (2016). Genomic Prediction of Seed Quality Traits Using Advanced Barley Breeding Lines. *PLoS ONE*, *11*(10). https://doi.org/10.1371/journal.pone.0164494

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Sallam, A. H., & Smith, K. P. (2016). Genomic Selection Performs Similarly to Phenotypic Selection in Barley. *Crop Science*, *56*(6), 2871–2881. https://doi.org/10.2135/cropsci2015.09.0557

Schmidt, M., Kollers, S., Maasberg-Prelle, A., Großer, J., Schinkel, B., Tomerius, A., Graner, A., & Korzun, V. (2016). Prediction of malting quality traits in barley based on genome-wide marker data to assess the potential of genomic selection. *Theoretical and Applied Genetics*, *129*(2), 203–213. https://doi.org/10.1007/s00122-015-2639-1

Tessema, B. B., Liu, H., Sørensen, A. C., Andersen, J. R., & Jensen, J. (2020). Strategies Using Genomic Selection to Increase Genetic Gain in Breeding Programs for Wheat. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.578123

Thomasen, J. R., Sørensen, A. C., Su, G., Madsen, P., Lund, M. S., & Guldbrandtsen, B. (2013). The admixed population structure in Danish Jersey dairy cattle challenges accurate genomic predictions1. *Journal of Animal Science*, *91*(7), 3105–3112. https://doi.org/10.2527/jas.2012-5490