



OPEN

Assessment and modeling using machine learning of resistance to scald (*Rhynchosporium commune*) in two specific barley genetic resources subsets

Houda Hiddar^{1,2}, Sajid Rehman², Berhane Lakew³, Ramesh Pal Singh Verma^{2,4}, Muamar Al-Jaboobi², Adil Moulakat^{1,2}, Zakaria Kehel², Abdelkarim Filali-Maltouf¹, Michael Baum² & Ahmed Amri²✉

Barley production worldwide is limited by several abiotic and biotic stresses and breeding of highly productive and adapted varieties is key to overcome these challenges. Leaf scald, caused by *Rhynchosporium commune* is a major disease of barley that requires the identification of novel sources of resistance. In this study two subsets of genebank accessions were used: one extracted from the Reference set developed within the Generation Challenge Program (GCP) with 191 accessions, and the other with 101 accessions selected using the filtering approach of the Focused Identification of Germplasm Strategy (FIGS). These subsets were evaluated for resistance to scald at the seedling stage under controlled conditions using two Moroccan isolates, and at the adult plant stage in Ethiopia and Morocco. The results showed that both GCP and FIGS subsets were able to identify sources of resistance to leaf scald at both plant growth stages. In addition, the test of independence and goodness of fit showed that FIGS filtering approach was able to capture higher percentages of resistant accessions compared to GCP subset at the seedling stage against two Moroccan scald isolates, and at the adult plant stage against four field populations of Morocco and Ethiopia, with the exception of Holetta nursery 2017. Furthermore, four machine learning models were tuned on training sets to predict scald reactions on the test sets based on diverse metrics (accuracy, specificity, and Kappa). All models efficiently identified resistant accessions with specificities higher than 0.88 but showed different performances between isolates at the seedling and to field populations at the adult plant stage. The findings of our study will help in fine-tuning FIGS approach using machine learning for the selection of best-bet subsets for resistance to scald disease from the large number of genebank accessions.

Barley (*Hordeum vulgare* L.) remains one of the most important cereal crops since its domestication in the “Fertile Crescent” 10,000 years ago¹. It ranks fourth among cereals in terms of production and acreage and is grown mainly in Russia, Australia, Europe, United Kingdom, Canada, North America, and Asia². Barley is a major crop in North Africa and West Asia, and it is grown on an average ~3 million ha in the dry areas of Morocco and Ethiopia³.

Barley is a multiple purpose crop used mainly as feed (green forage, straw, and grain), food, malt, and brewing. Due to its health benefits, it has become an important ingredient of the modern human diet in many developing as well as developed countries⁴. Particularly in Japan, more than 90% of barley is used as human food⁵. In Morocco, Ethiopia, India and China, barley grain is still widely consumed⁶. Recently, barley grain is being used in renewable energy for the production of biofuels⁷.

¹Laboratory of Microbiology and Molecular Biology, Faculty of Sciences, BioBio Research Center, University Mohammed V in Rabat, Rabat, Morocco. ²International Center for Agricultural Research in the Dry Areas (ICARDA), Biodiversity and Crop Improvement, 10010 Rabat, Morocco. ³Berhane Lakew, Ethiopian Institute of Agricultural Research (EIAR), Addis Ababa, Ethiopia. ⁴ICAR-Indian Institute of Wheat and Barley Research, Karnal, Haryana 132001, India. ✉email: a.amri@cgiar.org

Barley has a great potential to adapt to a wide range of environmental conditions, ranging from desert, to high latitudes and altitudes⁸. Around the globe barley can grow in the spring and winter seasons⁹. However, barley productivity and quality are affected by several biotic factors including viruses, fungi, and bacteria¹⁰. Among these phytopathogens, *Rhynchosporium commune* (formerly *R. secalis*), is one of the most serious and destructive fungi causing a foliar disease known as scald or leaf blotch. This disease is found in all continents, but is more frequent and devastating under cool and semi-humid barley growing regions¹¹. Yield losses due to scald can vary from 10 to 70%^{11–14}. Furthermore, Home-Grown Cereals Authority in the United Kingdom reported that despite the use of chemical fungicides, *R. commune* causes annual yield losses worth £4.8 million⁴.

Owing to its high genetic diversity and spontaneous mutations, *R. commune* populations can evolve rapidly¹⁵, and has the potential to quickly adapt and overcome deployed host plant resistance^{16,17}, making its control more challenging. Linde et al.¹⁸ found 76% of the global genetic diversity of the pathogen within a single barley field. *R. commune* is an imperfect fungus, and its perfect stage has not been reported. Pathogens with mixed reproduction system pose serious threats to longevity of resistance genes. A sexual reproduction can be predicted based on equal frequency of two mating types in the population. In Turkey, Çelik Oğuz et al.¹⁹ reported 35% of the isolates with MAT1-1 and 65% of the isolates with MAT1-2 which hints asexual reproduction. In addition, other mating type studies indicate asexual reproduction in Iran, and Syria, while in Switzerland, Australia, Ethiopia, South Africa, Scandinavia and California, sexual reproduction has been predicted^{20,21}. For developing control strategy, genetic diversity and its distribution among populations may be helpful. Furthermore, *R. commune* populations can adapt to abiotic selection pressure, comprising the variation of temperature and global warming²². Another challenge complicating the control of scald disease is its ability to survive, colonize and sporulate asymptotically on the resistant varieties, which can be a source of inoculum for the next cropping season infection²³. Many studies showed that polymerase chain reaction (PCR) can detect symptomless presence of *Rhynchosporium* in the seeds²⁴, which allows the transmission of the pathogen²⁵.

Virulence changes in pathogen populations affect the disease resistance status of the host plant genotypes. It is necessary to understand the evolutionary forces which impact pathogen biology in terms of emergence of individual genotypes with altered virulence. Conversely, plant genetic resources should be screened with newly emerged virulent pathotypes for effective management of the diseases. Several factors such as large population size, frequency-dependent selection, spontaneous mutation, gene flow, sexual reproduction and asexual recombination could explain the high genetic diversity of *R. commune* populations^{19,26}. Different studies have reported 2–72 pathotypes based on responses of diverse isolates on selected barley genotypes^{27–34}. Azamparsa and Karakaya³³ grouped 52 scald isolates into 30 pathotypes based on the screening on 17 differential barley genotypes. None of the pathotype was virulent on all differentials and two cultivars, Jet and Abyssinia, were the most resistant with a mean disease score of 0.2–0.3 on a scale of 0–4. Therefore, a continuous search for new effective resistance genes and their deployment through breeding is needed to cope with the emerging new virulent pathotypes of leaf scald pathogen.

The use of fungicides and cultural practices are not efficient in sustainable management of leaf scald disease, and resistance to multiple fungicides has been reported by many researchers^{35,36}. *R. commune* has developed insensitivity to demethylation inhibitor fungicides e.g., methyl benzimidazole carbamates²⁶, and to azole-based fungicides through mutation in multiple loci³⁶. Genetic resistance remains the most economical and environmentally friendly way to control scald disease³⁷. Thus genetic resources, such as barley landraces and wild barley accessions, held in the gene banks are key to crop improvement efforts, as they allow continuous supply of traits sought by the breeders, including genes for resistance/tolerance to major abiotic and biotic stresses³⁸. Wild barley (*Hordeum spontaneum*) offers great potential as source of resistance compared to landraces. The response of 198 landraces and 104 *H. spontaneum* accession from Turkey revealed that only 0.5% (1) of the landraces were resistant to six isolates of scald, but 26% (27) of *H. spontaneum* accessions were resistant to all isolates tested³⁹. Similarly, Rehman et al.⁴⁰ reported 58% (66) *H. spontaneum* accessions from ICARDA genebank resistant to scald under field conditions in Morocco. Silvar et al.⁴¹ screened 159 lines issued from a core collection of Spanish barley landraces and found 26% being resistant to scald. Yitbarek et al.⁴² found sources of resistance to scald in the Ethiopian germplasm with a higher rate of resistance in accessions originated from higher altitudes. Furthermore, van Leur et al.⁴³ found 13% lines being resistant to scald from a set of 100 lines derived from landraces collected from different regions of Syria and Jordan. They observed more variation within landraces and among collection sites, with more prevalence of resistant accessions from cooler environments. Furthermore, Düşünceli et al.⁴⁴ reported 8 out of 683 barley genotypes resistant at the seedling as well as at the adult plant stage.

ICARDA has the global mandate for the improvement of barley⁴⁵, and its genebank has one of the largest collections of barley worldwide totaling around 32,800 accessions, including more than 2500 accessions of wild *Hordeum* species⁴⁶. ICARDA distributes to collaborators the elite germplasm in the form of international nurseries and accessions of genebank based on requests and using the Standard Material Transfer Agreement (SMTA). The distribution of accessions was done in the past either using random sampling or using core collections with the latter attempting to include 90% of the diversity in 5–10% of the total holdings⁴⁷. The Generation Challenge Program (GCP) proposed the development of a reference set representing most of the genetic diversity in 10% of the core collection using molecular markers. ICARDA with the Australian and the Russian collaborators developed an approach named “Focused Identification of Germplasm Strategy (FIGS)”, which is either based on filtering or predictive modeling, to construct best-bet subsets for targeted traits⁴⁸. The predictive modeling approach allows to develop algorithms linking the sought plant traits to geographic, agro-climatic, and edaphic variables, which are used to construct manageable subsets for efficient evaluation and mining of genetic resources⁴⁹. The relationship between the trait and the environment was shown in case of barley by Endersen⁵⁰ and was tested for salinity tolerance⁵¹, and for net blotch resistance⁵². Furthermore, FIGS was successful in identifying sources of resistance to the Russian wheat aphid and Sunn pest in wheat which could not be identified using random

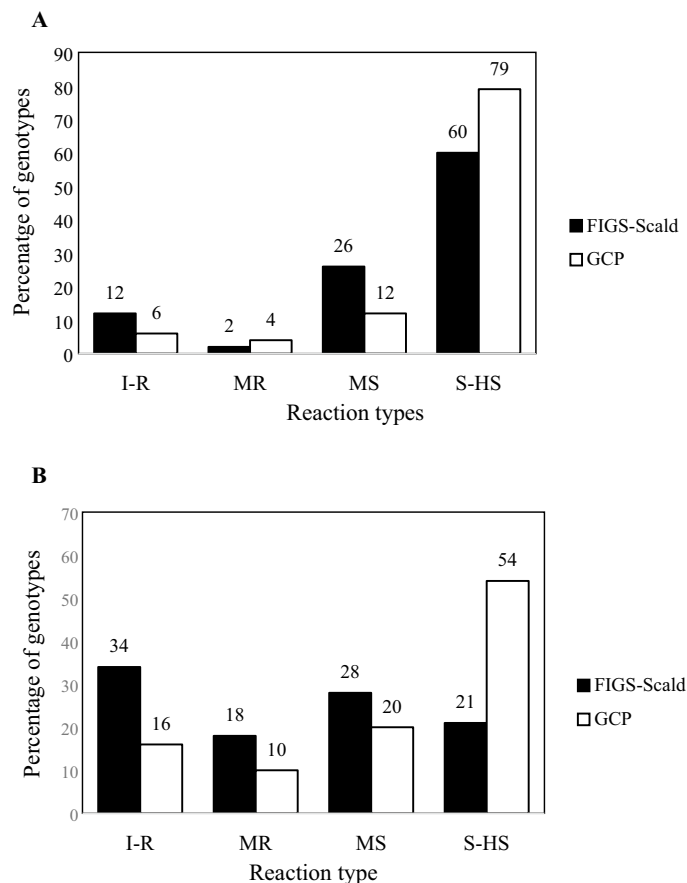


Figure 1. Seedling reaction to scald isolate 511 (A) and 1122 (B) of barley accessions of FIGS and GCP subsets under control conditions (in percent). Here *I-R* Immune and resistant, *MR* moderately resistant, *MS* moderately susceptible, *S-HS* susceptible and highly susceptible.

sampling of larger subsets^{53,54}. In addition, new alleles for resistance to powdery mildew in wheat⁵⁵ were identified from a targeted FIGS subset.

The present study aims at the identification of sources of resistance to scald in GCP and FIGS subsets, and to assess different predictive models for efficient mining of genetic resources for resistance to scald.

Results

Reaction of GCP and FIGS subsets to scald disease. Scald infections were successful under both controlled and field conditions as shown by the uniformity of the infection of the susceptible accessions as well as the checks. The average scores of susceptible (Tocada and Tissa) and resistant (ICARDA4 and Atlas 46) checks at the adult plant stage were 8 and 1, respectively. While at the seedling stage, average scores of resistant and susceptible checks for isolate SC-511 were 0 and 5, and for the isolate SC-1122 were 1 and 4, respectively.

The two subsets (FIGS and GCP) included sources of resistance to scald at both the seedling and adult plant stages, but their percentages differed based on isolates at the seedling stage and based on location and pathogen populations under field conditions. At the seedling stage, GCP subset showed 6% and 16% resistant accessions, and 79% and 54% susceptible accessions to scald isolates SC-511 and SC-1122, respectively. For FIGS subset, 12% and 34% of the accessions were resistant, while 60% and 21% were susceptible to SC-511 and SC-1122 isolates, respectively (Fig. 1A,B). Different percentages for the moderately resistant and moderately susceptible classes were found among the two subsets for both scald isolates tested. FIGS subset had 8% more accessions with MR and MS reactions compared with GCP for the SC-1122 isolate. For isolate SC-511, 4% of GCP and 2% of FIGS accessions were MR while 12% and 26% showed MS reaction, respectively. Isolate SC-511 showed less percentage of accessions with immune, resistant, and moderately resistant reaction and a higher percentage of susceptible accessions compared to isolate SC-1122 (Fig. 1).

At the adult plant stage, the percentage of accessions for scald disease severity differed based on the location, year, and subsets. The highest percentages of resistant accessions for both subsets (65% for GCP, and 82% for FIGS) were observed at Guich-Morocco 2018, and the lowest percentages at Holetta 2017 (25% for GCP, and 22% for FIGS) (Fig. 2). For Marchouch 2018, 37% of GCP and 51% of FIGS were resistant, while for Holetta-nursery 2018, these respective percentages were 46% and 53%, and 55% and 61% for Holetta-quarantine 2018. The highest percentages of susceptible (32% for GCP, and 44% FIGS), and highly susceptible accessions (10% for GCP,

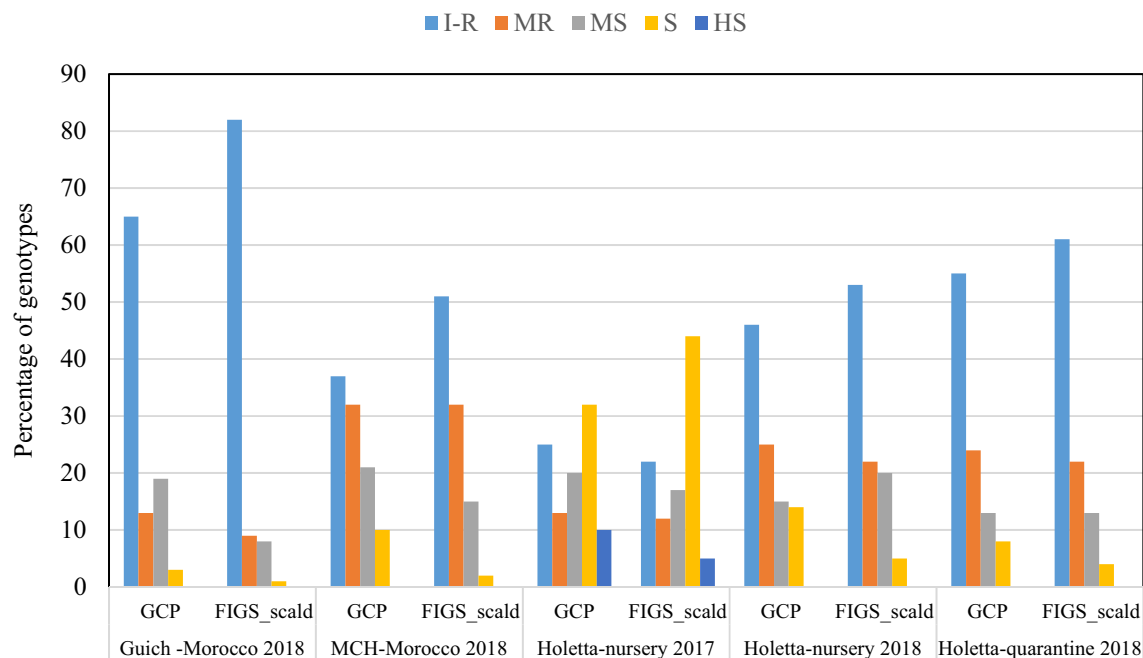


Figure 2. Adult plant reaction of GCP and FIGS barley subsets (in percent) to scald populations under field conditions. *I-R* Immune and resistant, *MR* moderately resistant, *MS* moderately susceptible, *S* susceptible, *HS* highly susceptible.

Locations	Sets	Reaction type						χ^2 (P-value)
		I	R	MR	MS	S	HS	
Guich-Morocco 2018	GCP	58	66	26	36	5	0	12.664 (P=0.013)
	FIGS_scald	48	35	9	8	1	0	
MCH-Morocco 2018	GCP	6	64	62	40	19	0	12.099 (P=0.017)
	FIGS_scald	8	44	32	15	2	0	
Holetta-nursery 2017	GCP	24	24	24	39	61	19	5.756 (P=0.331)
	FIGS_scald	9	13	12	17	44	5	
Holetta-nursery 2018	GCP	38	48	47	30	26	0	6.635 (P=0.156)
	FIGS_scald	26	27	22	20	5	0	
Holetta-quarantine 2018	GCP	42	61	46	25	15	0	5.231 (P=0.264)
	FIGS_scald	33	27	22	13	4	0	

Table 1. Test of independence of adult plant reaction of FIGS and GCP barley subsets to scald evaluated under field conditions in Ethiopia (2017–2018), and in Morocco (2018) for each disease reaction class. *I* Immune, *R* resistant, *MR* moderately resistant, *MS* moderately susceptible, *S* susceptible, *HS* highly susceptible.

and 5% for FIGS) were identified at Holetta-nursery 2017 for both subsets, while for all the other environments these classes of reaction ranged from 0 to 10%. When the disease scoring is grouped into three classes, (*I* + *R*, *MR* + *MS*, *S* + *HS*), the highest percentage of resistant accessions were found in the case of Guich Morocco-2018, and the lowest at Holetta-nursery 2017. For the moderately resistant and moderately susceptible classes, the GCP and FIGS subsets had similar percentages except at Guich-2018 where GCP (33%) had almost twice as much as FIGS (17%) (see Supplementary Fig. S1).

Comparison of GCP and FIGS subsets. The tests of independence showed that the reaction to scald is dependent on the subsets when tested in the field at Guich and Marchouch in Morocco with χ^2 P-values < 0.01 (Table 1). However, this dependency was not observed when the subsets were tested against the scald natural populations in Ethiopia. The same conclusions were confirmed when different classes of disease reactions were combined (see Supplementary Table S1). In case of dependency, FIGS appeared to deviate from theoretical distribution and had higher percentages of accessions with resistant and moderately resistant reactions and lower percentages of susceptible and highly susceptible accessions. At the seedling stage, the reaction types to both isolates were dependent on the subsets for all classes of reactions (P values < 0.01), except for the isolate SC-511 for the grouping (*I* + *R* + *MR*) and (*MS* + *S* + *HS*) with a p-value of 0.201 (Table 2).

Isolates	Set	Group of 6 classes	Group of 2 classes	Group of 3 classes
		χ^2 (P-value)	χ^2 (P-value)	χ^2 (P-value)
Test of independence				
SC-511	GCP	21.352 (<0.01)	1634 (P=0.201)	12.107 (<0.01)
	FIGS_scald			
SC-1122	GCP	35.140 (<0.01)	18,265 (<0.01)	30.290 (<0.01)
	FIGS_scald			
Test of goodness of fit				
SC-511	GCP	39.944 (<0.01)	2.851 (P=0.091)	22.143 (<0.01)
	FIGS_scald			
SC-1122	GCP	60.501 (<0.01)	18.265 (<0.001)	46.698 (<0.01)
	FIGS_scald			

Table 2. χ^2 tests of independence and of goodness of fit of different groups of disease reaction classes to two scald isolates (511, 1122) under controlled conditions using GCP and FIGS subsets of barley. Group of 6-classes includes reaction types I, R, MR, MS, S, HS, group of 2-classes include I + R + MR and MS + S + HS, group of three classes include I + R, MR + MS, and S + HS. FIGS Focused Identification of Germplasm Strategy subset, GCP Generation Challenge Program reference set.

Best model	SC-511 HS	SC-511 AS	SC-1122 HS	SC-1122 AS
	BCART	BCART	KNN	SVM
Sensitivity	0.20	0.11	0.25	0.25
Specificity	0.96	1.00	0.84	0.88
Precision	0.25	1.00	0.36	0.54
Accuracy	0.91	0.89	0.68	0.64
Kappa	0.17	0.18	0.10	0.14
Accuracy lower	0.82	0.80	0.57	0.53
Accuracy upper	0.96	0.95	0.79	0.75
Accuracy null	0.93	0.88	0.74	0.63
Mcnemar (P-value)	1.00	0.01	0.31	0.01

Table 3. Modeling performance from the best machine learning algorithm for seedling data for high (HS) and average (AS) disease scores for scald disease in barley set. BCART Bagged Carts, KNN K-nearest neighbors, SVM support vector machine.

For the test of goodness of fit, highly significant differences between FIGS and GCP subsets were declared based on a comparison of different classes of disease reactions at Guich and Marchouch (2018) in Morocco with P values < 0.01. These differences were significant (P-value = 0.033) in case of the trial in Holetta-nursery 2018, but not in Holetta-nursery 2017 (see Supplementary Table S2). When the differences were significant, FIGS subset showed higher percentages of resistant classes and lower percentages of susceptible classes than GCP subset at the adult plant stage. When the tests were done at the seedling stage using both isolates, the differences were highly significant between FIGS and GCP subsets for all groups of disease reaction classes except for the isolate SC-511 for the grouping (I + R + MR) and (MS + S + HS) with a p-value of 0.091 (Table 2).

Prediction of reaction to scald using machine learning approach. In our present study the performance of predictive models varied depending on the environment at the adult plant stage, and on the isolates at the seedling stage. At the seedling stage, the two modeled classes were not balanced for both high and averages scores (Table 3). The isolate SC-1122 showed a better balance than isolate SC-511 between resistant and susceptible classes. High score exhibited a higher but non-significant accuracy than the average score for both isolates. However, models for average accuracy were significantly higher than Null accuracy with 0.89 and 0.64 for isolate SC-511 and SC-1122, respectively. Models for both isolates, and both scores were not able to identify efficiently resistant accessions due to low sensitivity with maximum value of 0.25.

At the adult plant stage for the combined field data, only two models showed significant and accurate results, BCART and RF with an accuracy of 0.73 and 0.72, respectively (see Supplementary Table S3). Specificity was very low-to-medium for all the four assessed models, and the Null accuracy was higher than 0.5 and reached to 0.67. All models efficiently identified resistant accessions with specificities higher than 0.88. Modeling of field data per location/environment (Table 4) showed different patterns. Accuracy from modeling Marchouch was significantly higher than Null model with an accuracy of 0.81 and Kappa of 0.53.

Best model	GUICH	MCH	Holetta17	Holetta18
	BCART	RF	SVM	KNN
Sensitivity	0.94	0.95	0.41	0.73
Specificity	0.23	0.53	0.77	0.25
Precision	0.86	0.81	0.52	0.63
Accuracy	0.82	0.81	0.63	0.55
Kappa	0.21	0.53	0.19	-0.02
Accuracy lower	0.72	0.75	0.51	0.43
Accuracy upper	0.90	0.87	0.74	0.67
Accuracy null	0.83	0.69	0.62	0.63
Accuracy (P-value)	0.69	0.00	0.46	0.94

Table 4. Performance metrics for the best machine learning models for scald reaction of barley at adult plant stage in different environments. *BCART* Bagged Carts, *RF* Random forest, *SVM* support vector machine, *KNN* K-nearest neighbors, *MCH* Marchouch station.

Discussion

The good development of scald under natural infection in Holetta-Ethiopia and under artificial inoculation with the use of overhead sprinklers irrigation system in Morocco allowed reliable screening of FIGS and GCP subsets. Over the two cropping seasons, scald appeared as a permanent threat to barley in Ethiopia while in Morocco its natural development was more sporadic in time and space. The surveys conducted in North African countries during 1990–1994 through UNDP-ICARDA project on cereals and food legume diseases showed that scald is a major disease of barley in Tunisia, but in Morocco, infections were observed only once out of five years (unpublished data). Epidemics were observed in 1998 in Eritrea, Ethiopia, Turkey, Tunisia, and Morocco⁵⁶. During 2018 and 2019, the visits of farmers' fields in many regions of Morocco showed a low incidence of scald with fields showing medium severity including in mountainous and semi-arid regions (unpublished data). All released varieties in Morocco are moderately to highly susceptible to scald⁵⁷, and host plant resistance was chosen as the most-appropriate approach for integrated management of the disease.

Our results showed that resistance was found in both FIGS and GCP subsets at the seedling and adult plant stages. The reaction to the disease differed between isolates, seasons, field locations, and between subsets at both stages. The differences in reactions could be attributed to virulence differences among the isolates at the seedling stage with isolate SC-511 being more aggressive than isolate SC-1122, and to differences in virulence spectra among scald field populations at the adult plant stage (Fig. 1). A similar finding was reported by Albustan et al.⁵⁸, where they found 31% resistant genotypes at the adult plant stage and 19% at the seedling stage.

Although sources of resistance to scald were identified in both subsets at both stages, in general FIGS subset presented higher percentages of resistant accessions (54%) at both stages compared to GCP subset (45%). Likewise, 26% accessions in a Spanish barley core collection⁴¹, and 13% landraces from a set of lines derived from Syrian and Jordanian landraces⁴³ were resistant to scald. Our study is the first attempt to compare FIGS to another subset, and the results are supporting the efficiency of this approach in mining genebank collections. Several other studies have shown the limitation of the core collection approach for identifying and searching for rare and adaptive alleles^{59–61}. In addition, the core collection and the GCP reference set derived from it are targeting higher overall diversity of the accessions, while FIGS is targeting the diversity for resistance to scald. However, more research is needed to assess the efficiency of both GCP and FIGS of finding different effective genes for adaptive traits such as resistance to scald. Furthermore, genotyping and association mapping studies could decipher if resistance is conditioned by new resistance genes or by new allelic variants of the existing resistance genes.

The high likelihood of containing the targeted alleles in FIGS-subset is most probably due to the co-evolutionary process that occurs over time between the host and the pathogen, and which might allow the maintenance of a high number of effective avirulence genes in the pathogen populations. Several studies demonstrated that the environment has a great influence on the development of the disease as well as the genetic structure of both the host and pathogen through processes of natural selection, gene flow, and spatial/geographic differentiation^{62,63}. The filtering approach of FIGS was based on environmental conditions best suited for the natural development of the disease and this approach might not be applicable to the favorable areas where improved varieties are used at a large scale, and which has affected the diversity of the pathogen populations by continuous selection of more virulent races. In addition, our results support to some extent the theoretical basis of FIGS approach and presents further evidence that the geographical distribution of the disease resistance in nature is not random, and it is influenced by environmental conditions and the host diversity. However, our FIGS modeling approach using machine learning did not show a strong relationship between resistance to scald and environmental condition for all isolates and field populations. This could be due to the limited size of the sets, and the unbalanced distribution of reaction to scald. In addition, this could also be a result of the variability within a collection site⁴³, or a limited correlation between disease resistance and environmental conditions as Silvar et al.⁴¹ reported that resistance was found across the country. However, more resistance was found in winter types in Spain and at higher altitudes of Ethiopia⁴². This geographic structure of disease resistance distribution was also evidenced for stem rust in wheat⁶⁴, and powdery mildew in barley⁶⁵.

Several previous studies using FIGS approach in different crops supported the evidence of identification of desirable expression traits using eco-geographical data^{53–55}. FIGS approach was successfully used to identify source of resistance to stem rust and powdery mildew in wheat^{55,66,67}, and net-blotch in barley⁵². The superiority of FIGS filtering approach was not confirmed at Holetta-nursery 2017 in Ethiopia, showing the need for refining the sub-setting process through consideration of the virulence spectra of the pathogen populations. The filtering approach can also be refined or improved further by the search of the best predictive models that can explain the relationship between the resistance and agro-climatic conditions using machine learning.

Using single isolates at the seedling stage, and the use of natural and artificial infection under the field conditions, the results showed that the reactions of both FIGS and GCP subsets to *R. commune* population are highly variable. The instability of responses could be attributed to the variable nature of the *R. commune* pathogen in terms of pathogenicity, genetic makeup of barley genotypes grown in a region, and response to environmental factors. The major resistance gene will exert selection pressure on virulence allele frequency of the pathogen isolates which will eventually break down the deployed resistance within a short span⁶⁸. Selection is one of the evolutionary factors which can be controlled to some extent by human interventions to slow down the evolution of genetic variants with altered pathogenicity. Mert and Karakaya⁶⁹ found 7 barley cultivars resistant to all five *R. commune* isolates from Turkey. In another study, Azamparsa et al.⁷⁰ treated a barley cultivar (Tokak 157/37) with gamma irradiation and tested 25 advanced mutant lines with three virulent isolates of scald from Turkey. Though Tokak 157/37 was highly susceptible to three *R. commune* isolates, two mutant lines were resistant to at least two isolates tested. This approach is suitable to introduce resistance in popular barley cultivars which are difficult to replace. Similar findings of high variability of the pathogenicity of the *R. commune* were also reported across the world^{26,33,71}. All studies suggest complex interaction of scald isolates with barley genotypes and determination of pathotype diversity is essential to develop effective screening protocols.

The observed variability in virulence in *Rhynchosporium* is probably due to the high genetic variation of the pathogen population as reported by several studies^{72,73}. The use of molecular markers showed limited genetic diversity in scald isolates from West Asia and North Africa^{74,75}, however, high genetic variation was reported from Australia, California, Finland, and Norway *R. commune* populations^{76,77}. But a clear correlation between molecular markers and pathogenicity of scald populations is lacking which indicate the involvement of other molecular mechanisms which drive the evolution of virulence in *R. commune*. In general, virulence genes are located on chromosomal regions rich in transposable elements which are more prone to recombination/mutation providing an evolutionary advantage to the pathogen to come up with novel variations in its effector repertoire to evade host defenses⁷⁸. The study of the evolution of necrosis-inducing protein (NIP1) effector from 146 strains of four *Rhynchosporium* spp. revealed its presence in all species, but presence/absence polymorphism in addition to copy number variation in *R. commune* revealed a strong selection pressure on this effector protein⁷⁹. The deployment of major resistance genes in rotation along with prudent use of fungicides can reduce the selection of virulent isolates and enhance the longevity of deployed resistant cultivars⁸⁰.

The Focused Identification of Germplasm Strategy (FIGS) based on filtering agro-climatic conditions favoring the development of the disease has shown its efficiency in identifying a higher frequency of resistant accessions compared to the Reference Set of the Generation Challenge Program (GCP). However, more refinement of this approach is needed to take into consideration the specter of virulence of the pathogen and the development of best-bet models linking resistance to environmental conditions using machine learning.

Methods

Germplasm used. A total of 292 genotypes from the ICARDA genebank were evaluated for scald resistance including 101 accessions selected using FIGS approach (FIGS subset), and 191 accessions from the Generation Challenge Program reference set (GCP).

The filtering approach was used to develop FIGS subset based on the following parameters:

- Count number of days where the average daily temperature is between 8 and 12 °C, 10 days before the onset of the growing period and up to 15% into the vegetative phase.
- Remove sites with zero count from step 1.
- Sum daily rain for 10 days before the onset of the growing period up to 10% into the vegetative phase.
- Normalize both variables (steps 1 and 2) to range 0–1 for each site.
- Add variables to create index 1.
- Rank based on index 1 and remove the bottom 25 percent of sites.

For the remaining sites, the following was done:

- From 10% into the vegetative phase until onset of grain filling divide into 3 separate sub-phases of equal length.
- For each sub-phase count the number of days where the average daily temperature is between 15 and 20 °C.
- For each sub-phase determine the amount of precipitation.
- Remove sites if any of the variables = 0 (3 count variables and 3 precipitation variables).
- Normalize each variable for a range between 0 and 1.
- Add each variable and then add index 1 to create index 2.
- Rank sites using index 2 from largest to smallest.

Depending on the required subset size and the number of remaining candidate sites, proceed with the selection of accessions.

- If there are fewer candidate sites than the required set size, choose one accession in a cyclic manner working down the rank and then repeat starting at the top of the rank with sites that still have accessions after each pass. Alternatively, starting from the top ranked site, multiple accessions per site could be chosen until desired set size is achieved.
- If there are more sites than the desired set size, then one accession could be chosen randomly from each site starting at the top ranked site until the desired set size is reached. Alternatively, this approach could be taken after one candidate accession is donated by each country represented in the candidate site list.

Fungal isolates and inoculation. Scald infected barley leaves were collected during the disease surveys conducted from 2015 to 2018 in naturally infected barley fields from different agro-ecological zones of Morocco including Marchouch and Sidi Allal Tazi research stations. Infected barley leaves were dried in paper envelopes at room temperature for 4–5 days before their storage at 4 °C till further use. Air dried leaves were cut into 2 mm segments and soaked in sterile distilled water for 15 min followed by surface sterilization using 50% ethanol for 15 s and 10% sodium hypochlorite solution for 30 s, rinsed with sterile distilled water three times, and dried up within two layers of sterile Whatman filter paper. Then the leaf segments were incubated on Lima Bean agar (LMA) supplemented with Kanamycin and Streptomycin (50 mg per liter) at 14 °C in darkness for two weeks until the sporulating white or pink colonies were visible. A total of 50 single conidial isolates were prepared and stored at –80 °C freezer for long term storage. Based on race analysis, 12 pathotypes were revealed and two scald isolates with a wide virulence spectrum (Unpublished data), Marchouch (MCH-SC) and Sidi Allal Tazi (SAT-SC), were used for seedling screening in the greenhouse. While for adult plant screening, inoculum composed of a mixture of 15 pathotypes from different agro-ecological zones of Morocco was used.

Seedling screening (SRT). About 4–5 seeds of each accession were sown in peat moss supplemented with 14-14-14 NPK in a single cone of 3.8 cm diameter and 14 cm depth (Stuewe & Sons, Inc., Oregon, USA) in two replications. The seedlings were grown in a growth chamber at a photoperiod of 16 h light/8 h dark at 20 ± 1 °C. At the two leaves stage, inoculation was performed separately with two virulent isolates of barley leaf scald (SC-511, and SC-1122). The inoculum was prepared from 12 to 14 days old LBA plates by rubbing the agar surface gently followed by filtration with two layers of cheese cloth. Spore suspension of 5 × 10⁵ conidia/ml supplemented with surfactant Tween 20 (0.01%) was sprayed with a hand held sprayer till run off⁸¹. The inoculated seedlings were kept at 100% relative humidity in dark in the growth chamber for 72 h at 15 °C to enhance spore germination. Afterward, the seedlings were transferred to the greenhouse with a light/dark regime of 16/8 h at 16 ± 1 °C, respectively.

The disease evaluation at the seedling stage was carried after 16 days post-inoculation. Assessment of plant infection type and severity was done based on scoring scale (0–5) of Salamati and Tronsmo⁸² considering lesion morphology, size, chlorosis and necrosis. Where 0 = Immune (I), 1 = resistant (R), 2 = moderately resistant (MR), 3 = moderately susceptible (MS), 4 = susceptible (S), and 5 = highly susceptible (HS). For further analysis, the genotypes were categorized into either resistant (IR of 0, 1, 2) or susceptible (IR > 3, 4, 5) group.

Adult plant screening (APS). The adult plant screening was conducted in Ethiopia and Morocco. In Morocco, the trials were conducted at the INRA experimental stations of Marchouch (MCH; 33° 56' N, 6° 63' W) and at Guich (33° 58' N, 6° 51' W) during the cropping season of 2018. Each entry was seeded in 1 m length row with a row spacing of 0.5 m in an augmented design. Two susceptible (Tocada and Tissa) and resistant (ICARDA4 and Atlas 46) checks were planted after every 10 accessions and each block was surrounded by a perpendicular border row, composed of a mixture of scald susceptible genotypes (Tiddas, Aglou, Tissa, Adrar, Shepherd, Fleet, Baudin and Alester), to maintain high disease pressure. Artificial field inoculations were performed four times starting from Zadoks scale⁸³ GS30 at an interval of 10–12 days apart in the evenings with inoculum composed of spore suspension composed of a mixture of 50 *R. commune* isolates collected from different agro-ecological zones of Morocco using a knapsack sprayer. In addition, scald infected barley residues were spread uniformly in the field. Following inoculations, the covering of spreader rows with a plastic sheet overnight and the frequent use of a mist system with overhead sprinklers ensured good disease establishment and uniform spread of the disease. In Ethiopia, the experiment was conducted in Holetta experimental station (9° 00' N and 38° 30' E) during 2017 and 2018 cropping seasons using the same design as described above. There, the environmental conditions were favorable to allow a high natural infection level of scald for efficient screening of the germplasm. The disease severity was assessed at GS 73–75 using 0–9 scale⁸⁴. Genotypes were classified in the following categories: Resistant (0–2), moderately resistant (3–4), moderately susceptible (5–6), susceptible (7–8), and highly susceptible (9).

Data analysis. The statistical analysis was performed using R software⁸⁵. The disease assessment data from the seedling and adult plant stage were tested for statistical association between the sub-setting approaches using χ^2 tests of independence with significance level ($\alpha = 0.05$). The equation used to calculate chi-square is as follows:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}. \quad (1)$$

The equation for calculating expected values in a test of independence is as follows:

$$E_{ij} = \frac{\sum_{k=1}^c O_{ij} \sum_{k=1}^r O_{kj}}{N}, \quad (2)$$

where is the E_{ij} = expected value, $\sum_{k=1}^c O_{ij}$ is the Sum of the i th column, $\sum_{k=1}^r O_{kj}$ is the Sum of the k th row, N is the total number.

The expected values for the test of goodness were calculated using following equation:

$$E_i = np_i, \quad (3)$$

where E_i is the expected value, n is the total sample size, and p_i is the hypothesized proportion of observations in level i .

To find out the differences between FIGS and GCP subsets in terms of reaction to scald, the test of goodness of fit using χ^2 test at a significance level ($\alpha = 0.05$) was used where GCP was simulated to a random sample. Both tests were done using different groupings of reactions, all six classes (I, R, MR, MS, S and HS), three classes (I + R, MR + MS, S + HS) and two classes (I + R + MR and MS + S + HS).

Modeling of the reaction to scald disease. The machine learning approach using the obtained SRT and APS disease screening data was used to find out the best models that link adaptive traits, environments (and associated selection pressures) with genebank accessions. We used environmental data from WorldClim1 databases as predictors. WorldClim is an open access database providing global climatic layers describing past climatic profiles of collection sites and intended for spatial modeling or mapping.

It includes average, monthly minimum and maximum temperatures, precipitation, and bioclimatic variables⁸⁶.

The following machine learning algorithms were used: K-nearest neighbors (KNN)⁸⁷, Support Vector Machines (SVM)⁸⁸, Random Forest (RF)⁸⁹, and Bagged Carts (BCART)⁹⁰. Each machine learning model was tuned to select best tuning parameters using a training set (70% of the total set) and then the best model was selected between different machine learning models based on several metrics including accuracy, specificity, and Kappa. The modeling metrics were computed on the test set (30% of the total set). In this study, R language and caret library were used for machine learning⁹¹. Models were tuned for parameter's optimization and trained on 70% of the data and tested with 10 cross validation folds and 100 replications. Modeling was done for the two isolates for the seedling stage using average and high scores separately. For the adult plant reaction, modeling was done for the combined multi-locations data sets and for each location separately.

Statements of compliance. The corresponding author attests that: The plants were regenerated and handled using the standard operating procedures set by ICARDA genebank following the FAO genebank standards. All accessions used in this study were acquired from ICARDA genebank in accordance with the International Treaty on Plant Genetic Resources for Food and Agriculture.

Statements of consent. The corresponding author attests that: The seeds and plants were planted and evaluated with the consent and mutual understanding between ICARDA genebank and the Ethiopian Institute of Agricultural Research (EIAR). Standard operating procedures set by ICARDA genebank and EIAR were followed. We did not collect any barley leaves for this study. All the barley accessions were obtained from ICARDA genebanks using Standard Material Transfer Agreement.

Data availability

The data that support the findings of this study are stored in the ICARDA genebank database and can be made available upon request from the corresponding author.

Received: 30 March 2021; Accepted: 6 July 2021

Published online: 05 August 2021

References

1. Harwood, W. A. An introduction to barley: The crop and the model. In *Barley: Methods and Protocols* (ed Harwood, W. A.) 1–5 (Springer, 2019).
2. FAO GIEWS Country Brief on Morocco. <http://faostat.fao.org/>. Accessed Feb 2020 (2018).
3. FAO GIEWS Country Brief on Morocco. <http://faostat.fao.org/>. Accessed Feb 2020 (2016).
4. Ullrich, S. E. Significance, adaptation, production, and trade of barley. In *Barley: Production, Improvement and Uses* (ed. Ullrich, S. E) 3–13 (Wiley-Blackwell, 2010).
5. Aoki, E., Baba, T., Yamaguchi, O., Ito, S. & Moriwaki, J. Development of barley cultivars with resistance to scald (*Rhynchosporium secalis* (Oud.) Davis) in Japan. *Jpn. Agric. Res. Q. JARQ* **45**, 349–357. <https://doi.org/10.6090/jarq.45.349> (2011).
6. Ashman, H. & Beckley, J. Rediscovering barley. *Cereal Foods World* **51**, 2. <https://doi.org/10.1094/CFW-51-0038> (2006).
7. Griffey, C. *et al.* Grain composition of Virginia winter barley and implications for use in feed, food, and biofuels production. *J. Cereal Sci.* **51**, 41–49 (2010).
8. Dai, F. & Zhang, G. Domestication and improvement of cultivated barley. In *Exploration, Identification and Utilization of Barley Germplasm*, (ed. Zhang, L. C.) 1–26 (Elsevier, 2016).
9. Cuesta-Marcos, A. *et al.* Barley: Genetics and breeding. *Encycl. Food Grains* <https://doi.org/10.1016/B978-0-12-394437-5.00208-4> (2016).
10. Pessaraki, M. *Handbook of Plant and Crop Stress*. 4th edn. (CRC Press, 2020).
11. Zhan, J., Fitt, B. D. L., Pinnschmidt, H. O., Oxley, S. J. P. & Newton, A. C. Resistance, epidemiology and sustainable management of *Rhynchosporium secalis* populations on barley. *Plant. Pathol.* **57**, 1–14. <https://doi.org/10.1111/j.1365-3059.2007.01691.x> (2008).

12. Jabbari, J. S. *Molecular Characterisation of Differentially Expressed Genes in the Interaction of Barley and Rhynchosporium secalis* PhD thesis, Adelaide University (2017).
13. Paulitz, T. C. & Steffenson, B. J. Biotic stress in barley: Disease problems and solutions. In *Barley: Production, Improvement and Uses* (ed. Ullrich, S. E.) 307–354 (Wiley-Blackwell, 2010).
14. Karakaya, A. *et al.* Current status of scald and net blotch diseases of barley in Turkey. In *IWBLD—1st International Workshop on Barley Leaf Diseases* (Salsomaggiore Terme, 2014).
15. Abang, M. *et al.* Pathogen evolution in response to host resistance genes: Evidence from fields experiments with *Rhynchosporium secalis* on barley. *Phytopathology* **96**, S2–S2 (2006).
16. Williams, K., Donnellan, S., Smyl, C., Scott, L. & Wallwork, H. Molecular variation in *Rhynchosporium secalis* isolates obtained from hotspots. *Australas. Plant Pathol.* **32**, 257–262. <https://doi.org/10.1071/AP03008> (2003).
17. Zaffarano, P. L., McDonald, B. A., Zala, M. & Linde, C. C. Global hierarchical gene diversity analysis suggests the fertile crescent is not the center of origin of the barley scald pathogen *Rhynchosporium secalis*. *Phytopathology* **96**, 941–950. <https://doi.org/10.1094/phyto-96-0941> (2006).
18. Linde, C. C., Zala, M., Ceccarelli, S. & McDonald, B. A. Further evidence for sexual reproduction in *Rhynchosporium secalis* based on distribution and frequency of mating-type alleles. *Fungal Genet. Biol.* **40**, 115–125. [https://doi.org/10.1016/S1087-1845\(03\)00110-5](https://doi.org/10.1016/S1087-1845(03)00110-5) (2003).
19. Çelik Oğuz, A., Ölmez, F., Karakaya, A. & Azamparsa, M. R. Genetic variation and mating type distribution of *Rhynchosporium commune* in Turkey. *Physiol. Mol. Plant Pathol.* **114**, 101614. <https://doi.org/10.1016/j.pmp.2021.101614> (2021).
20. Linde, C., Zala, M., Ceccarelli, F. & McDonald, B. Evidence for sexual reproduction in *Rhynchosporium secalis* based on distribution and frequency of mating-type alleles. *Fungal Genet. Biol. FG B* **40**, 115–125. [https://doi.org/10.1016/S1087-1845\(03\)00110-5](https://doi.org/10.1016/S1087-1845(03)00110-5) (2003).
21. Seifollahi, E., Sharifnabi, B., Javan-Nikkhah, M. & Linde, C. C. Low genetic diversity of *Rhynchosporium commune* in Iran, a secondary centre of barley origin. *Plant Pathol.* **67**, 1725–1734. <https://doi.org/10.1111/ppa.12886> (2018).
22. Stefansson, T. S., McDonald, B. A. & Willi, Y. Local adaptation and evolutionary potential along a temperature gradient in the fungal pathogen *Rhynchosporium commune*. *Evol. Appl.* **6**, 524–534. <https://doi.org/10.1111/eva.12039> (2013).
23. Atkins, S. D. *et al.* The epidemiological importance of asymptomatic infection of winter barley by *Rhynchosporium secalis* and its consequences for crop protection and breeding. In *The Dundee Conference. Crop Protection in Northern Britain 2010, Dundee, UK, 23–24 February 2010*. 81–86 (The Association for Crop Protection in Northern Britain, 2010).
24. Lee, H. K., Tewari, J. P. & Turkington, T. K. Quantification of seedborne infection by *Rhynchosporium secalis* in barley using competitive PCR. *Plant Pathol.* **51**, 217–224. <https://doi.org/10.1046/j.1365-3059.2002.00685.x> (2002).
25. Fountaine, J. M. *Epidemiological Studies of Rhynchosporium secalis (leaf blotch of barley)* PhD thesis, University of Reading, (2005).
26. Avrova, A. & Knogge, W. *Rhynchosporium commune*: A persistent threat to barley cultivation. *Mol. Plant Pathol.* **13**, 986–997. <https://doi.org/10.1111/j.1364-3703.2012.00811.x> (2012).
27. Fowler, A. M. & Owen, H. Studies on leaf blotch of barley (*Rhynchosporium secalis*). *Trans. Br. Mycol. Soc.* **56**, 137–152. [https://doi.org/10.1016/S0007-1536\(71\)80117-1](https://doi.org/10.1016/S0007-1536(71)80117-1) (1971).
28. Ali, S. M., Mayfield, A. H. & Clare, B. G. Pathogenicity of 203 isolates of *Rhynchosporium secalis* on 21 barley cultivars. *Physiol. Plant Pathol.* **9**, 135–143. [https://doi.org/10.1016/0048-4059\(76\)90033-3](https://doi.org/10.1016/0048-4059(76)90033-3) (1976).
29. Ceoloni, C. Race differentiation and search for sources of resistance to *Rhynchosporium secalis* in barley in Italy. *Euphytica* **29**, 547–553. <https://doi.org/10.1007/BF00023201> (1980).
30. Abbott, D. *et al.* The relationship between seedling infection types and field reactions to leaf scald in Clipper barley backcross lines. *Aust. J. Agric. Res.* **42**, 801–809. <https://doi.org/10.1071/AR9910801> (1991).
31. Tekauz, A. Pathogenic variation in *Rhynchosporium secalis* on barley in Canada. *Can. J. Plant Pathol.* **13**, 298–304. <https://doi.org/10.1080/07060669109500915> (1991).
32. Goodwin, S. B. The barley scald pathogen *Rhynchosporium secalis* is closely related to the discomycetes *Tapesia* and *Pyrenopeziza*. *Mycol. Res.* **106**, 645–654. <https://doi.org/10.1017/S0953756202006007> (2002).
33. Azamparsa, M. R. & Karakaya, A. Determination of the pathotypes of *Rhynchosporium commune* (Zaffarano, McDonald & Linde) in some regions of Turkey. *Plant Protect. Bull.* **60**, 5–14. <https://doi.org/10.16955/bitkorb.460235> (2020).
34. Çelik Oğuz, A. & Karakaya, A. Genetic diversity of barley foliar fungal pathogens. *Agronomy* **11**, 434 (2021).
35. Brunner, P. C., Stefansson, T. S., Fountaine, J., Richina, V. & McDonald, B. A. A global analysis of CYP51 Diversity and azole sensitivity in *Rhynchosporium commune*. *Phytopathology* **106**, 355–361. <https://doi.org/10.1094/phyto-07-15-0158-r> (2016).
36. Mohd-Assaad, N., McDonald, B. A. & Croll, D. Multilocus resistance evolution to azole fungicides in fungal plant pathogen populations. *Mol. Ecol.* **25**, 6124–6142. <https://doi.org/10.1111/mec.13916> (2016).
37. Oxley, S. J. P., Cooke, L. R., Black, L., Hunter, A. & Mercer, P. C. Management of *Rhynchosporium* in different barley varieties and cropping systems. Project Report PR315, <https://ahdb.org.uk/management-of-rhynchosporium-in-different-barley-varieties-and-cropping-systemsmanagement-of-rhynchosporium-in-different-barley-varieties-and-cropping-systems> Accessed Apr 2020 (2003).
38. Qualset, C. O. Sampling germplasm in a center of diversity: An example of disease resistance in Ethiopian barley. In *Crop Genetic Resources for Today and Tomorrow* (eds. Frankel, O. H. & Hawkes, J. G.) 81–98 (Cambridge University Press, 1975).
39. Azamparsa, M. R., Karakaya, A., Ergün, N., Duran, R. & Özbek, K. Identification of barley landraces and wild barley (*Hordeum spontaneum*) genotypes resistant to *Rhynchosporium commune*. *Tarım Bilimleri Dergisi* <https://doi.org/10.15832/ankutbd.441916> (2019).
40. Rehman, S. *et al.* Traits discovery in *Hordeum vulgare* sbsp. *spontaneum* accessions and in lines derived from interspecific crosses with wild *Hordeum* species for enhancing barley breeding efforts. *Crop Sci.* **61**, 219–233. <https://doi.org/10.1002/csc2.20360> (2021).
41. Silvar, C. *et al.* Screening the Spanish Barley Core Collection for disease resistance. *Plant Breed.* **129**, 45–52. <https://doi.org/10.1111/j.1439-0523.2009.01700.x> (2010).
42. Yitbarek, S. *et al.* Variation in Ethiopian barley landrace populations for resistance to barley leaf scald and netblotch. *Plant Breed.* **117**, 419–423. <https://doi.org/10.1111/j.1439-0523.1998.tb01966.x> (1998).
43. van Leur, J. A. G., Ceccarelli, S. & Grandó, S. Diversity for disease resistance in barley landraces from Syria and Jordan. *Plant Breed.* **103**, 324–335. <https://doi.org/10.1111/j.1439-0523.1989.tb00393.x> (1989).
44. Düşünceli, F. *et al.* Determination of the reactions of some barley cultivars and genotypes to scald under greenhouse and field conditions. *Tarım Bilimleri Dergisi* **14**(1), 46–50. https://doi.org/10.1501/Tarimbil_0000000526 (2008).
45. *International Center for Agricultural Research in the Dry Areas*. <https://www.icarda.org/> Accessed Feb 2020 (2018).
46. *Genesys PGR. Plant Genetic Resources*. <https://www.genesys-pgr.org/> Accessed Feb 2020 (2018).
47. Hintum, T., Theo, J.L., Brown, A., Spillane, C. & Hodgkin, T. Core collections of plant genetic resources. In *IPGRI Technical Bulletin No.3. International Plant Genetic Resources Institute, Rome, Italy*, vol. 3. [https://doi.org/10.1016/S0304-4238\(96\)00927-2](https://doi.org/10.1016/S0304-4238(96)00927-2) (1997).
48. Mackay, M. C. & Street, K. Focused identification of germplasm strategy—FIGS. In *Proceedings 11th Wheat Breeding Assembly*. 138–141 (Royal Australian Chemical Institute (RACI), 2004).
49. Azough, Z., Kehel, Z., Benomar, A., Bellafkih, M. & Amri, A. Predictive characterization of ICARDA genebank barley accessions using FIGS and machine learning. *Intell. Environ.* <https://doi.org/10.3233/AISE190031> (2019).
50. Endresen, D. T. F. Predictive association between trait data and ecogeographic data for Nordic barley landraces. *Crop Sci.* **50**, 2418–2430. <https://doi.org/10.2135/cropsci2010.03.0174> (2010).

51. Hijmans, R. J., Jacobs, M., Bamberg, J. B. & Spooner, D. M. Frost tolerance in wild potato species: Assessing the predictivity of taxonomic, geographic, and ecological factors. *Euphytica* **130**, 47–59. <https://doi.org/10.1023/A:1022344327669> (2003).
52. Endresen, D. T. F., Street, K., Mackay, M., Bari, A. & De Pauw, E. Predictive association between biotic stress traits and eco-geographic data for wheat and barley Landraces. *Crop Sci.* **51**, 2036–2055. <https://doi.org/10.2135/cropsci2010.12.0717> (2011).
53. El Bouhssini, M. E., Street, K., Joubi, A., Ibrahim, Z. & Rihawi, F. Sources of wheat resistance to Sunn pest, *Eurygaster integriceps* Puton, in Syria. *Genet. Resour. Crop Evol.* **56**, 1065. <https://doi.org/10.1007/s10722-009-9427-1> (2009).
54. El Bouhssini, M. *et al.* Sources of resistance in bread wheat to Russian wheat aphid (*Diuraphis noxia*) in Syria identified using the Focused Identification of Germplasm Strategy (FIGS). *Plant Breed.* **130**, 96–97. <https://doi.org/10.1111/j.1439-0523.2010.01814.x> (2011).
55. Bhullar, N. K., Street, K., Mackay, M., Yahiaoui, N. & Keller, B. Unlocking wheat genetic resources for the molecular identification of previously undescribed functional alleles at the *Pm3* resistance locus. *Proc. Natl. Acad. Sci.* **106**, 9519–9524. <https://doi.org/10.1073/pnas.0904152106> (2009).
56. Yahyaoui, A. H. Occurrence of barley leaf blight diseases in central western Asia and North Africa. In *2nd International Workshop on Barley Leaf Blights, Aleppo (Syria)*, 7–11 April 2002 (ICARDA, 2004).
57. Saidi, S., Jilal, A., Amri, A., Grondo, S. & Ceccarelli, S. Amélioration génétique de l'orge au Maroc. *La création variétale à l'INRA Méthodologie, acquis Perspect* **2006/0228**, 99–137 (2005).
58. Albustan, S., Çetin, L. & Düşünceli, F. Orta Anadolu için geliştirilen bazı arpa nörselerinde arpa yaprak lekeli hastalığına (*Rhynchosporium secalis* (Oudem.) J.J. Davis) dayanıklı genotiplerin sera ve tarla koşullarında belirlenmesi. In *VIII. Fitopatoloji Kongresi Bildir*, Ankara (Türkiye): 30–33, 21–25 Eylül (1998).
59. Dwivedi, S. L. *et al.* Enhancing crop gene pools with beneficial traits using wild relatives. *Plant Breed. Rev.* **30**, 179–230 (2007).
60. Pessoa-Filho, M., Rangel, P. H. N. & Ferreira, M. E. Extracting samples of high diversity from thematic collections of large gene banks using a genetic-distance based approach. *BMC Plant Biol.* **10**, 127. <https://doi.org/10.1186/1471-2229-10-127> (2010).
61. Xu, Y. Plant genetic resources: Management, evaluation and enhancement. In *Mol. Plant Breed.* (ed. Xu, Y.) 151–194 (CABI, 2010).
62. Spieth, P. T. Environmental heterogeneity: A problem of contradictory selection pressures, gene flow, and local polymorphism. *Am. Nat.* **113**, 247–260. <https://doi.org/10.1086/283383> (1979).
63. Epperson, B. K. Spatial autocorrelation of genotypes under directional selection. *Genetics* **124**, 757–771 (1990).
64. Bonman, J. M., Bockelman, H. E., Jin, Y., Hijmans, R. J. & Gironella, A. I. N. Geographic distribution of stem rust resistance in wheat landraces. *Crop Sci.* **47**, 1955–1963. <https://doi.org/10.2135/cropsci2007.01.0028> (2007).
65. Paillard, S. *et al.* Evolution of resistance against powdery mildew in winter wheat populations conducted under dynamic management. II. Adult plant resistance. *Theor. Appl. Genet.* **101**, 457–462. <https://doi.org/10.1007/s001220051502> (2000).
66. Bhullar, N. K., Zhang, Z., Wicker, T. & Keller, B. Wheat gene bank accessions as a source of new alleles of the powdery mildew resistance gene *Pm3*: A large scale allele mining project. *BMC Plant Biol.* **10**, 88. <https://doi.org/10.1186/1471-2229-10-88> (2010).
67. Bari, A. *et al.* Focused identification of germplasm strategy (FIGS) detects wheat stem rust resistance linked to environmental variables. *Genet. Resour. Crop Evol.* **59**, 1465–1481. <https://doi.org/10.1007/s10722-011-9775-5> (2012).
68. Wolfe, M. S. & McDermott, J. M. Population genetics of plant pathogen interactions: The Example of the *Erysiphe graminis-Hordeum vulgare* Pathosystem. *Annu. Rev. Phytopathol.* **32**, 89–113. <https://doi.org/10.1146/annurev.py.32.090194.000513> (1994).
69. Mert, Z. & Karakaya, A. Assessment of the seedling reactions of Turkish barley cultivars to scald. *J. Phytopathol.* **152**, 190–192. <https://doi.org/10.1111/j.1439-0434.2004.00817.x> (2004).
70. Azamparsa, M. R. *et al.* Seedling response of two barley cultivars and gamma ray-induced advanced barley lines to *Rhynchosporium commune*. *Tarla Bitkileri Merkez Araştırma Enstitüsü Dergisi* **24**, 75–78 (2015).
71. Brown, J. S. Pathogenic variation among isolates of *Rhynchosporium secalis* from barley grass growing in South Eastern Australia. *Euphytica* **50**, 81–89. <https://doi.org/10.1007/BF00023164> (1990).
72. Linde, C. C., Zala, M. & McDonald, B. A. Molecular evidence for recent founder populations and human-mediated migration in the barley scald pathogen *Rhynchosporium secalis*. *Mol. Phylogenet. Evol.* **51**, 454–464. <https://doi.org/10.1016/j.ympev.2009.03.002> (2009).
73. McDonald, B. A. How can research on pathogen population biology suggest disease management strategies? The example of barley scald (*Rhynchosporium commune*). *Plant. Pathol.* **64**, 1005–1013. <https://doi.org/10.1111/ppa.12415> (2015).
74. Von Korff, M., Udupa, S. M., Yahyaoui, A. & Baum, M. Genetic variation among *Rhynchosporium secalis* Populations of West Asia and North Africa as revealed by RAPD and AFLP Analysis. *J. Phytopathol.* **152**, 106–113. <https://doi.org/10.1111/j.1439-0434.2003.00809.x> (2004).
75. Bouajila, A. *et al.* Genetic diversity of *Rhynchosporium secalis* in Tunisia as revealed by pathotype, AFLP, and microsatellite analyses. *Mycopathologia* **163**, 281–294. <https://doi.org/10.1007/s11046-007-9012-0> (2007).
76. McDonald, B. A., Zhan, J. & Burdon, J. J. Genetic structure of *Rhynchosporium secalis* in Australia. *Phytopathology* **89**, 639–645. <https://doi.org/10.1094/phyto.1999.89.8.639> (1999).
77. Salamati, S., Zhan, J., Burdon, J. J. & McDonald, B. A. The genetic structure of field populations of *Rhynchosporium secalis* from three continents suggests moderate gene flow and regular recombination. *Phytopathology* **90**, 901–908. <https://doi.org/10.1094/phyto.2000.90.8.901> (2000).
78. Grandaubert, J., Balesdent, M.-H. & Rouxel, T. Evolutionary and adaptive role of transposable elements in fungal genomes. In *Advances in Botanical Research* vol. 70 (ed. Martin, F. M.) 79–107 (Academic Press, 2014).
79. Mohd-Assaad, N., McDonald, B. A. & Croll, D. The emergence of the multi-species NIP1 effector in *Rhynchosporium* was accompanied by high rates of gene duplications and losses. *Environ. Microbiol.* **21**, 2677–2695. <https://doi.org/10.1111/1462-2920.14583> (2019).
80. Tucker, M. A. *et al.* Development of genetic SSR markers in *Blumeria graminis* f. sp. *hordei* and application to isolates from Australia. *Plant Pathol.* **64**, 337–343. <https://doi.org/10.1111/ppa.12258> (2015).
81. Mert, Z. & Karakaya, A. Determination of the suitable inoculum concentration for *Rhynchosporium secalis* seedling assays. *J. Phytopathol.* **151**, 699–701. <https://doi.org/10.1046/j.0931-1785.2003.00770.x> (2003).
82. Salamati, S. & Tronsmo, A. M. Pathogenicity of *Rhynchosporium secalis* isolates from Norway on 30 cultivars of barley. *Plant. Pathol.* **46**, 416–424. <https://doi.org/10.1046/j.1365-3059.1997.d01-20.x> (1997).
83. Zadoks, J. C., Chang, T. T. & Konzak, C. F. A decimal code for the growth stages of cereals. *Weed Res.* **14**, 415–421. <https://doi.org/10.1111/j.1365-3180.1974.tb01084.x> (1974).
84. Saari, E. & Prescott, J. M. A scale for appraising the foliar intensity of wheat diseases. *Plant Dis. Rep.* **59**, 377–380 (1975).
85. Team, R. C. R: *A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2018).
86. Fick, S. E. & Hijmans, R. J. WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315. <https://doi.org/10.1002/joc.5086> (2017).
87. Kotsiantis, S. Supervised machine learning: A review of classification techniques. *Informatica (Slovenia)* **31**, 249–268 (2007).
88. Hsu, C.-W., Chang, C.-C. & Lin, C.-J. *A Practical Guide to Support Vector Classification* (National Taiwan University, 2003).
89. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32. <https://doi.org/10.1023/A:1010933404324> (2001).
90. Kolcz, A. N-tuple network, CART, and bagging. *Neural Comput.* **12**, 293–304. <https://doi.org/10.1162/089976600300015790> (2000).
91. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **28**(26), 2008. <https://doi.org/10.18637/jss.v028.i05> (2008).

Acknowledgements

The authors would like to thank Mr. Othmane Bouhlal, Nouredine El-Haddad, Hasnae Choukri and Meriem Fardioui for their help.

Author contributions

A.A., H.H. formulated the research problem and designed the approaches. H.H., M.J., S.R., B.L. raised experiments in fields and recorded phenotyping data. H.H., Z.K. performed statistical analysis of data, A.A. and S.R. planned and supervised research activity. H.H. and A.A. wrote the initial draft of the paper, S.R. revised the manuscript. H.H. prepared the tables and the figures. All authors contributed to the final draft of the paper and approved the final version of the manuscript. All authors reviewed the article and approved the submitted version.

Funding

This study is supported financially through GIZ-attributed funding to ICARDA genebank, the CAIGE-GRDC-ICA00010 project and the CGIAR-Genebank Platform.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-94587-6>.

Correspondence and requests for materials should be addressed to A.A.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2021