


RESEARCH ARTICLE

Open Access



Genetic variation for tolerance to the downy mildew pathogen *Peronospora variabilis* in genetic resources of quinoa (*Chenopodium quinoa*)

Carla Colque-Little^{1†}, Miguel Correa Abondano^{2†}, Ole Sogaard Lund¹, Daniel Buchvaldt Amby¹, Hans-Peter Piepho³, Christian Andreasen¹, Sandra Schmöckel³ and Karl Schmid^{2*} 

Abstract

Background: Quinoa (*Chenopodium quinoa* Willd.) is an ancient grain crop that is tolerant to abiotic stress and has favorable nutritional properties. Downy mildew is the main disease of quinoa and is caused by infections of the biotrophic oomycete *Peronospora variabilis* Gaüm. Since the disease causes major yield losses, identifying sources of downy mildew tolerance in genetic resources and understanding its genetic basis are important goals in quinoa breeding.

Results: We infected 132 South American genotypes, three Danish cultivars and the weedy relative *C. album* with a single isolate of *P. variabilis* under greenhouse conditions and observed a large variation in disease traits like severity of infection, which ranged from 5 to 83%. Linear mixed models revealed a significant effect of genotypes on disease traits with high heritabilities (0.72 to 0.81). Factors like altitude at site of origin or seed saponin content did not correlate with mildew tolerance, but stomatal width was weakly correlated with severity of infection. Despite the strong genotypic effects on mildew tolerance, genome-wide association mapping with 88 genotypes failed to identify significant marker-trait associations indicating a polygenic architecture of mildew tolerance.

Conclusions: The strong genetic effects on mildew tolerance allow to identify genetic resources, which are valuable sources of resistance in future quinoa breeding.

Keywords: *Chenopodium quinoa*, *Chenopodium album*, *Peronospora variabilis*, Downy mildew, Phenotyping, Linear mixed models, Quantitative resistance, Plant genetic resources

* Correspondence: karl.schmid@uni-hohenheim.de

[†]Carla Colque-Little and Miguel Correa Abondano contributed equally to this work.

²Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, Fruwirthstrasse 21, D-70599 Stuttgart, Germany
Full list of author information is available at the end of the article



© The Author(s). 2021 **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 Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Quinoa (*Chenopodium quinoa* Willd.) is a grain crop that was domesticated in South America and cultivated from Chile to Southern Colombia for thousands of years [1]. After the arrival of the Spanish, it was replaced by European crops in many regions [2]. More recently, quinoa has experienced renewed interest as alternative grain crop worldwide and became an important export commodity for countries like Bolivia, which exported approximately 40,000 tons in 2014 [3]. The interest in quinoa results from its nutritional properties and tolerance to abiotic stresses such as high salinity, drought, and frost [2, 4]. The increasing demand for quinoa and successful cultivation outside its native range led to multiple breeding programs aimed at improving yield, resistance and adaptation to novel cultivation regions or climate change [5–7]. Susceptibility to plant diseases are one important biotic factor that limits crop yield. Downy mildew, the most important disease of quinoa, is caused by the biotrophic oomycete *Peronospora variabilis* Gaüm, previously known as *Peronospora farinosa* f.sp. *chenopodii* [8]. It causes severe yield losses of up to 30–50% in tolerant cultivars, and an almost complete yield loss in susceptible cultivars under conditions of high humidity and absence of chemical control measures [9]. The disease is widely spread over continents where quinoa is cultivated and may have been spread by seeds that were contaminated with the pathogen [10–13]. *P. variabilis* also infects the closely related and widespread weed *C. album* (known as goosefoot, fat hen, or lambs quarter) [14, 15] which may act as a secondary host. *C. album* occurs throughout Europe and is frequently infected by downy mildew-causing pathogens that seem to be conspecific with *P. variabilis* infecting *C. quinoa*. Therefore, *C. album* may act as alternative host for *P. variabilis* and constitute a reservoir for this pathogen [8, 16]. Other weedy *Chenopodium* species such as *C. murale* (nettle leaf goosefoot), *C. ambrosioides* (Indian goosefoot) are also susceptible [17–19], but cross-infection with *C. quinoa* has not been reported so far. Since the disease is seedborne, tolerance to this pathogen is a critical trait in the development of new quinoa varieties [20].

Currently, very little is known about the physiological mechanisms involved in the *P. variabilis* - quinoa interaction, or about the genetic basis of downy mildew tolerance and the role of other phenotypic traits in disease susceptibility. Previous studies for quinoa tolerance using greenhouse experiments, seedlings, detached leaves and field scorings primarily focused on quantitative measures by scoring disease symptoms [9, 13, 21, 22]. Response to mildew infection utilizes visual scoring of disease severity, which is the proportion of leaf tissue with lesions caused by the pathogen [9]. Another measure is the extent of sporulation by the pathogen. It is measured

with a detached leaf assay and the identification of spore bodies on leaf surfaces [23]. Reliable and efficient scoring of tolerance to downy mildew is a key component in the development of improved quinoa varieties.

The objectives of the present study were to investigate the variation of quinoa genotypes from its native range in South America (Bolivia, Peru, Ecuador, and Chile) in their response to inoculation with the downy mildew pathogen. We investigated the robustness of phenotypic scoring under controlled conditions and characterized the relationship of the disease traits severity of infection, sporulation and incidence with other phenotypic traits. These traits included size and density of leaf stomata because *P. variabilis* enters leaf tissues through the stomata [8, 24], and seed saponin content because saponin extracts have antifungal properties [25, 26]. We estimated genetic variance components and heritability of the response to *P. variabilis* infection and conducted a genome-wide association study (GWAS) with whole genome sequences of a subset of accessions to identify genomic regions with putative tolerance genes.

Results

High variation in mildew tolerance

In total, 132 genotypes (5 control varieties, 21 cultivars and 106 accessions) were successfully grown, inoculated with mildew, phenotyped and scored in three independent greenhouse experiments. Severity of infection ranged from 5.0% (*Chenopodium album*) to 83.0% (Accession G9) with a mean of 46.2%, whereas sporulation ranged from 0.2% (Variety Puno) to 83.6% (Cultivar CV21) with a mean of 42.6%. Incidence of infection showed a smaller range among genotypes from 36.8% (Accession G41) to 92.0% (Accession G92) with a mean of 71.6%.

Analysis of mildew tolerance with linear mixed models (LMM)

The severity of infection and sporulation measurements are expressed as proportions. We therefore fitted the LMM in Eq. (1) with both the raw data and data that have been transformed with logit and angular functions, which are frequently used with proportions. Our goal was to assess the effect of data transformation and inclusion of control varieties on estimates of variance components, heritabilities, and genotype means. In addition, we evaluated the effects of a randomized block design with replicated control varieties and unreplicated genotypes per block by using dummy variables to remove replicated checks from the model in order to verify that estimates of variance components are not influenced by control varieties. The combination of these parameters resulted in six LMMs that were fitted to the traits severity and sporulation (Table 1). Genotypes were fitted as fixed effects in all models to estimate genotype means

Table 1 Linear mixed models used to analyze quantitative response variables severity of infection and sporulation

Trait	Data transformation	Control variety included	Variance type p -value
Severity	No transformation	Yes	1.72×10^{-10}
		No	1.92×10^{-04}
	Arcsine root	Yes	4.68×10^{-07}
		No	0.0014
Sporulation	No transformation	Yes	2.44×10^{-08}
		No	9.77×10^{-06}
	Arcsine root	Yes	1.45×10^{-05}
		No	0.0390
Logit	Yes	0.0049	
	No	0.199	
Logit	Yes	0.0019	
	No	0.0259	

Type of error variance refers to error variance structure between experiments. p -value of a REML ratio test comparing a null model with homogeneous variances of the error with a model with a heterogeneous variance structure

and to test for a genotype effect on disease traits. A REML ratio test showed that a heterogeneous error variance structure for the experiments provided a better model fit ($p < 0.05$) except for a single model (Table 1).

Untransformed data for severity of infection did not strongly deviate from normality in a histogram of residuals and a QQ-plot (Supplementary Fig. S2A and B). On the other hand, a residual vs. fitted plot shows increasing variance along the x -axis and indicates heterogeneity of variances (Supplementary Fig. S2C). One source of variation is the experiment (Supplementary Fig. S2D), which is consistent with the results of the REML ratio tests in Table 1. A Wald test for fixed effects of genotypes on severity of infection was highly significant for all six fitted models and tests without control varieties have considerably lower p -values (Table 2).

Estimation of variance components allows to model sources of variation and to account for the structure of an experimental design [27]. For the trait severity of infection, proportions of variance components were highly similar among models. We obtained the highest estimates and confidence intervals for between experiments variance (σ_E^2) and genotype by experiment interaction variance ($\sigma_{G:E}^2$) components. Estimates of variance components for experiment 1 to 3 (σ_{e1}^2 , σ_{e2}^2 , σ_{e3}^2) and variance of blocks nested within experiments ($\sigma_{E:B}^2$) were much lower (Supplementary Fig. S3 A-C, Supplementary Table

Table 2 Wald F -test for genotype fixed effects in a linear mixed model analysis (LMM)

Model	Data transformation	Control variety included	LMM analysis F -value	d.f.	p -value	Mean s.e.d.	σ_G^2	H^2
<i>Trait: Severity of infection</i>								
1	Untransformed	Yes	3.573	131	1.18×10^{-18}	0.12	0.02	0.72
2		No	23.95	130	3.56×10^{-93}	0.12	0.03	0.77
3	Arcsine root	Yes	3.79	131	3.35×10^{-20}	0.14	0.03	0.74
4		No	28.95	130	8.41×10^{-103}	0.14	0.03	0.78
5	Logit	Yes	3.75	131	6.30×10^{-20}	0.58	0.46	0.73
6		No	30.08	130	5.59×10^{-104}	0.59	0.61	0.77
<i>Trait: Sporulation</i>								
1	Untransformed	Yes	4.47	131	4.39×10^{-25}	0.16	0.04	0.78
2		No	8.37	130	1.55×10^{-44}	0.17	0.05	0.78
3	Arcsine root	Yes	4.89	131	7.30×10^{-28}	0.20	0.07	0.79
4		No	11.35	130	1.13×10^{-57}	0.20	0.08	0.80
5	Logit	Yes	4.96	131	2.47×10^{-28}	0.87	1.51	0.80
6		No	11.43	130	4.67×10^{-58}	0.91	1.85	0.82
<i>Trait: Incidence</i>								
		Yes	1.62	131	0.0003	0.55	0.10	0.4
		No	2.39	130	<0.0001	0.48	0.08	0.4

Mean standard error of the difference (s.e.d.), genetic variance (σ_G^2) and heritability (H^2) estimates from were obtained from a linear mixed model for the traits severity and sporulation, and from a generalized linear mixed model for the trait incidence

S2). Estimated variance components were highly similar in a comparison of models with and without control varieties because the 95% confidence intervals overlapped, but variances of models with controls varieties were smaller than without (Supplementary Fig. S3 A-C).

The results for the trait sporulation after infection are very similar to severity after infection. The Wald F-test for fixed effects was highly significant in every model fit with sporulation as response variable, which indicates that host genotypes differ in sporulation (Table 2). Removal of control varieties strongly reduced p -values. For sporulation, the largest variance component in every model was genotype by experiment interaction, $\sigma_{E:B}^2$, but in contrast to severity of infection, estimates of variance between experiments (σ_E^2) and variance of blocks within experiments ($\sigma_{E:B}^2$) were the lowest among all variance components in all models (Supplementary Fig. S3D-F). Taken together these analyses provide strong evidence for an effect of genotypes on severity of infection and sporulation that is robust with respect to the data transformation and the effects of a blocked design.

Generalized linear mixed model (GLMM) analysis of incidence data

For the trait incidence of infection, we used a GLMM because it allows to fit non-normally distributed data like

discrete proportions and to include random effects. We used a *logit* link function and assumed homogeneous variances between experiments as indicated by the conditional Pearson residuals, i.e., there is no sign that the experiments are sources of variation that need to be accounted for (Supplementary Fig. S4). Using incidence as response variable in two GLMMs that differed by the inclusion and exclusion of control varieties, the test for fixed effects in both models was significant ($p < 0.001$). The genotype by experiment $\sigma_{G:E}^2$ variance component was larger in the model with control varieties (Fig. 1). Variance components reflecting the experimental design, σ_E^2 and $\sigma_{E:B}^2$, have the largest standard errors in both models. Additionally, in the GLMM without the control varieties the experiments σ_E^2 variance was zero, indicating that they are comparable between each other. The latter variance component and the residual error variance components were the largest regardless of the model used. In summary, like the other two disease traits, incidence of infection also shows a strong effect of genotypes on trait variation.

Correlations between traits

The similar variance component structures of the three disease traits (Fig. 1 and Supplementary Fig. S3) suggests that they are correlated. Adjusted means of the traits severity of

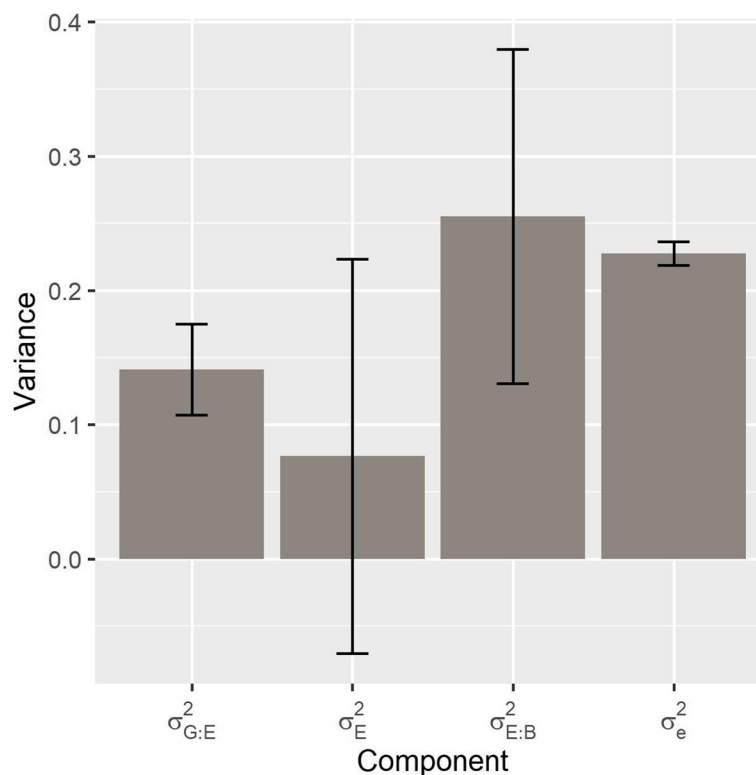


Fig. 1 Variance component estimates and their standard errors for incidence of infection in a GLMM with control varieties. $\sigma_{G:E}^2$, σ_E^2 , $\sigma_{E:B}^2$: Variance components for the genotype by experiment interaction, experiments and blocks nested within the replicates, σ_e^2 : Residual variance

infection and sporulation are highly correlated with control varieties ($R = 0.91$, $p < 0.001$; Fig. 2) and without control varieties included ($R = 0.9$, $p < 0.001$). The correlation of mildew incidence with both severity and sporulation was markedly lower ($R = 0.67$ and $R = 0.65$, respectively, $p < 0.001$).

Analysis of heritability

We also estimated heritability, \bar{H}^2 , of the three disease traits and evaluated the effect of data transformation on these estimates. The mean standard error of the difference (s.e.d.), which measures precision of pairwise comparisons in each model, and genetic variance, which is estimated when genotypes are fitted as random effect, are both components of \bar{H}^2 (Eq. (3)). Both parameters showed a small increase between models with and without control varieties, and we observed this difference with all data transformations (Table 2). Higher mean s.e.d values show that a removal of control varieties decreased the precision of pairwise comparisons. In consequence, \bar{H}^2 estimates in models without control varieties or with an arcsine root transformation resulted in marginally higher estimates than models with control varieties or with other transformations, respectively (Table 2). To summarize, data transformations and the exclusion of replicated control varieties have a little effect on the estimation of \bar{H}^2 , because heritability estimates remain within a narrow range from 0.72 to 0.78 for severity of infection, and a range from 0.78 to 0.82 for sporulation for all models analysed. For incidence of infection, estimated heritability was 0.40, the genetic variance was 0.10 and the mean s.e.d was 0.55 in the GLMM model with control varieties, and 0.08 and 0.48 in the model without control varieties.

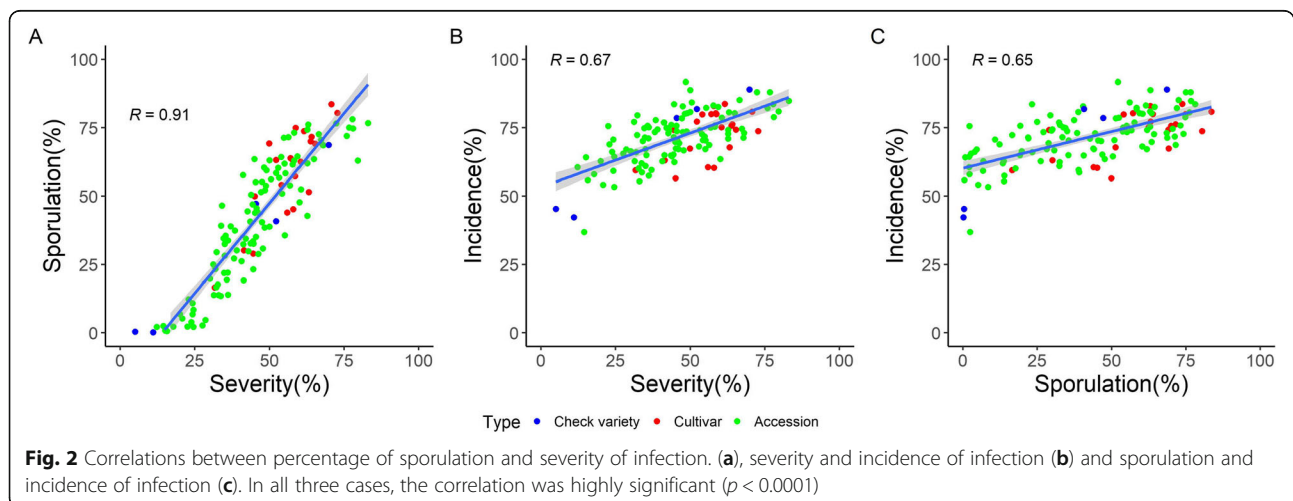
Ranking of genotypes by response to downy mildew infection

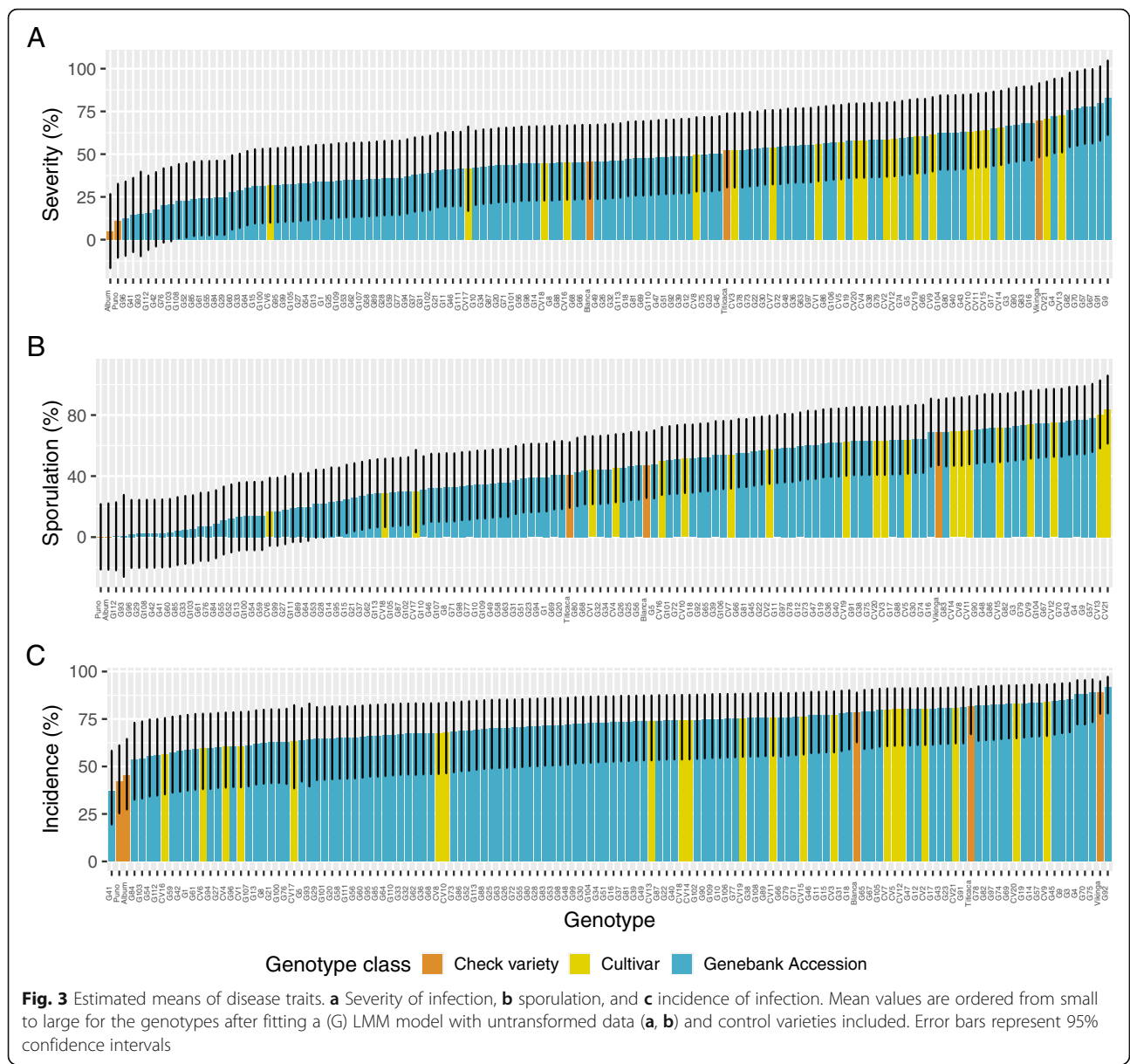
The strong effect of genotype on the three pathogen traits suggests that genebank accessions and varieties in our

sample are highly variable with respect to mildew tolerance. We therefore compared means in all three traits and identified substantial differences between genotypes. Adjusted mean values for severity of infection range from 5 to 83% in the LMM of untransformed data and with control varieties, and show a very similar range in the other models. Models without control varieties result in a smaller range for this trait because the two control varieties *C. album* and Puno had the lowest estimates for severity (Fig. 3a). We did not observe a strong effect of control varieties and the type of data transformation on the ordering of genotypes for severity of infection. Therefore, differences between genotypes for this trait are robust and allow to identify tolerant and susceptible genotypes. The control varieties *C. album*, Puno and genebank accessions G41, G42, G76, G93, G96 and G112 are most tolerant, whereas control variety Vikinga, cultivars CV13 and CV21, and accessions G4, G9, G57, G67, G82 and G91 are most susceptible (Fig. 3a).

The genotypes show a similar pattern with sporulation. The adjusted means of untransformed sporulation estimated including the control varieties ranged from 0.2% (Puno) to 83.6% (CV21). Transformation of the sporulation data had a small effect on the distribution of estimated means, independent of whether control varieties were included or excluded (Fig. 3b). The genotypes with the lowest values for sporulation are control varieties *C. album* and Puno, and genebank accessions G29, G41, G42, G93, G96, G106, G108, G112, whereas genotypes with highest values for sporulation are control variety Vikinga, cultivars CV12, CV13, and CV21, and accessions G4, G9, G43, G67, G70 and G104 (Fig. 3b).

The variation among genotypes for the trait incidence is lower than for the other two disease traits (Fig. 3c). Adjusted mean values were little affected by the inclusion or exclusion of control varieties. Genotypes with low incidence include control varieties Puno, *C. album*





and accession G41; while control variety Vikinga and genotypes G75 and G92 had the highest percentages of incidence (Fig. 3c).

Relationship of disease traits with genebank passport data and seed saponin content

The genotyped accessions included in the experiment were selected using information on mildew tolerance from passport data to obtain a set of accessions, which is polymorphic for this trait. The severity and incidence data recorded in the passport data of genebank accessions are highly incongruous with our results. For example, 35 of 106 accessions were recorded under the 0% severity category in the passport data but no accession was classified as such in our analysis; 16 accessions were

assigned to 0.1–25% group and 14 in this study; 26 accessions as 26–50% vs. 56 in our dataset. According to the passport data, 26 accessions are in the most susceptible category (75–100%), and only 6 accessions in the present study.

The only significant correlation of disease and stomatal traits was between severity and width of stomata ($r = 0.18, p = 0.041$). We also tested whether saponin content in seeds is correlated with disease susceptibility and carried out foam tests with seed harvested in two locations because saponin content varies between phenological stages and environments. 105 genotypes had seed available from both locations, Bolivia and Denmark, while 26 genotypes were harvested only in Denmark. Foam height measurements were not correlated between sources of

seed (Pearson's $r = 0.16$, $p = 0.11$). In a comparison of estimated average severity and sporulation of genotypes with and without saponins, we did not observe any systematic pattern based on the seed source or if accessions or cultivars were compared separately (t -test with $p > 0.05$).

Isolation of *P. variabilis* from *C. album* and cross-infection of *C. quinoa*

The isolate of *P. variabilis* used in this study was originally isolated from *C. album* and afterwards vegetatively propagated on *C. quinoa* (cv. Blanca and Vikinga). Spores harvested from these plants and inoculated onto *Chenopodium* spp. showed low disease severity (4%) and lowest sporulation (0.4%) to *C. album* compared to the *C. quinoa* genotypes (Fig. 3). This is the first time that cross infection from *C. album* has been reported. The BLAST comparison of the ITS DNA sequences used to validate the isolate showed a 100% match to an isolate obtained from *C. quinoa* cv. Atlas collected in 2001 (EU 113305).

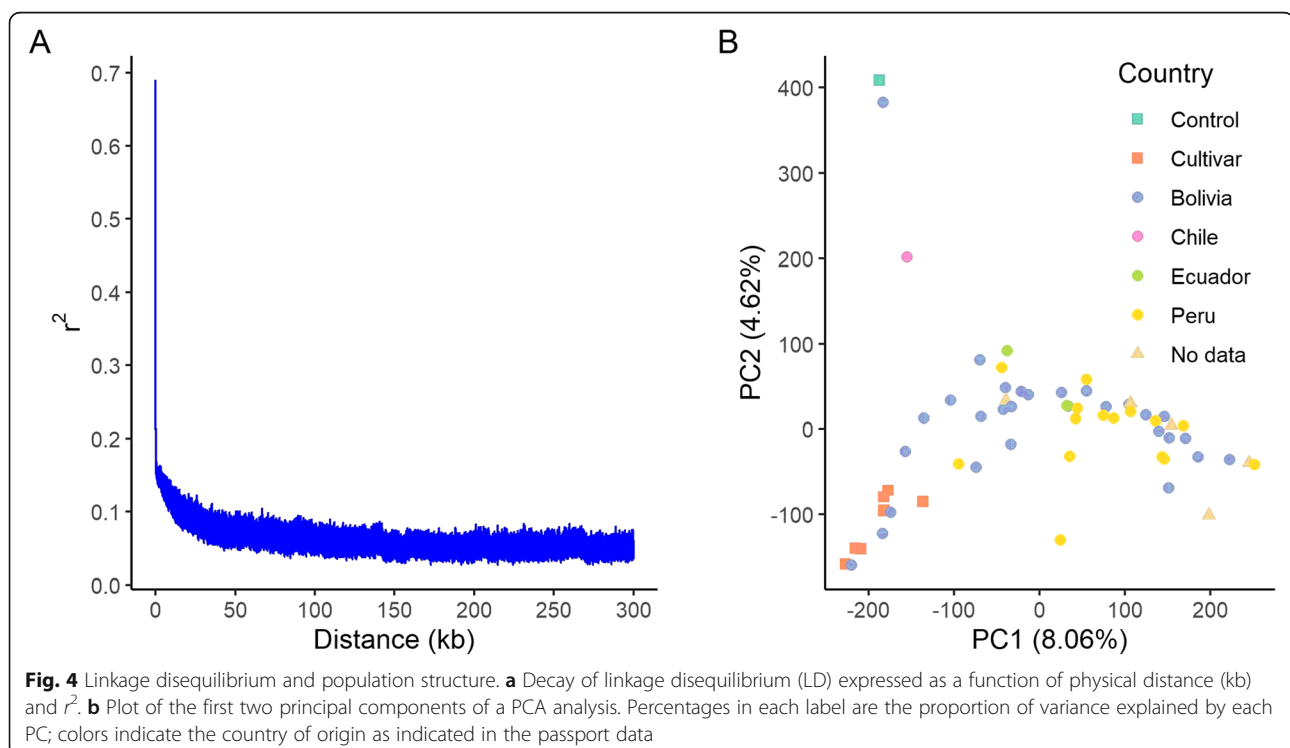
Whole genome sequencing of accessions

Given the highly significant genotypic effect on response to mildew infections, we sequenced a subset of 88 accessions and cultivars representing the range of phenotypic variation to conduct a GWAS. The sequence genotypes included two control varieties (Puno and Titicaca), 18 cultivars, and 68 genebank accessions of which 39

originated from Bolivia, 19 from Peru, two from Ecuador and 1 from Chile. The average severity of the sequenced genotypes ranged between 11% and 83%, with an average of 47%.

Sequencing of 88 samples produced 7.9×10^{11} bp in 2.6×10^9 read pairs with a length of 150 bp each. After mapping processed sequence reads to the quinoa reference genome version 1, sequence coverage per genotype ranged from 0.38 X (control variety Titicaca) to 9.17 X (Accession G37) with an average of 3.24 X. The proportion of mapped reads per sample ranged from 99.3% to 99.9%, with an average of 99.8%. In contrast to the nuclear genome, the chloroplast and mitochondrial genomes were overrepresented in our sample, with a coverage of 109 X and 32.6 X, respectively. Such a high coverage is expected because there are multiple copies of mitochondrial and plastid genomes per haploid nuclear genome.

Variant calling with GATK identified 18,017,831 biallelic SNPs across the genome. After filtering for minor allele frequency and sample missingness, and testing for departure from Hardy-Weinberg Equilibrium 4,131,562 variants remained. We imputed genotypes with coverage < 8 X, and sample and position missingness below < 0.7 , which produced 606,791 SNPs from 61 samples with an estimated accuracy of 96.9% based on 5,000 hidden markers. Fig. 4a shows that linkage disequilibrium, measured as r^2 , drops to 0.1 within 22-25Kb in the sequenced genotypes.



A principal component analysis (PCA) shows that all genotypes classified as cultivars cluster together (Fig. 4b, lower left corner). The control variety Puno and the genebank accession G71 (originating from Bolivia) are separated from the other genotypes. Accession G101 is also separated from the major group and originates from Chile, suggesting it is of the coastal ecotype. Accession G42, which appears to be separate from the main group (Fig. 4b, middle bottom), has very low severity and sporulation, which is comparable to the variety Puno. The remaining accessions originating from Bolivia, Peru and Ecuador are mixed and do not form distinct groups.

Association mapping for severity of downy mildew

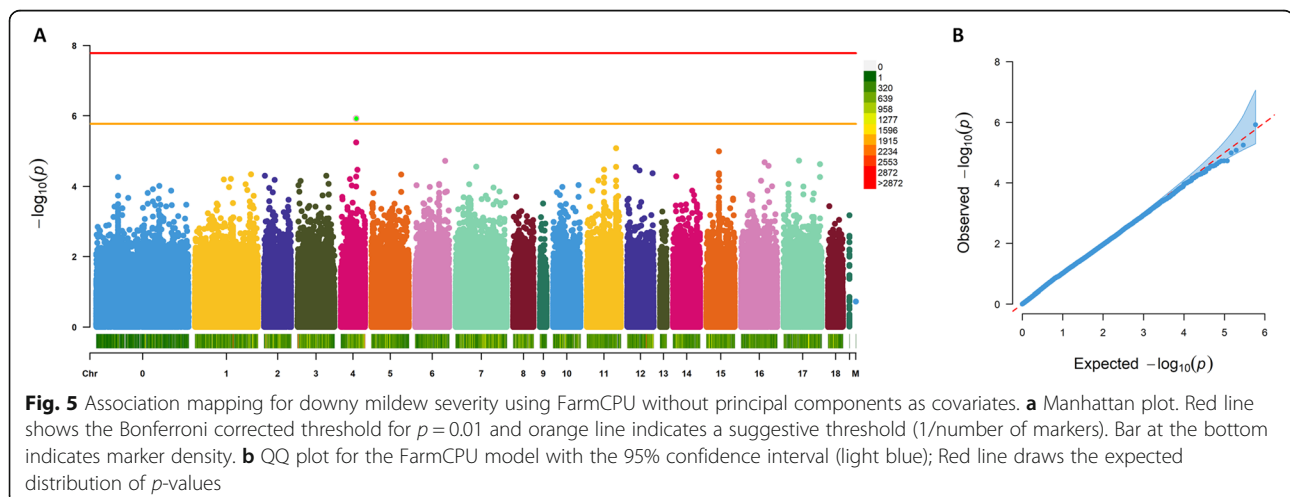
We carried out two different GWAS analyses to detect genomic regions associated with severity of downy mildew infection. Since severity and sporulation showed a high degree of correlation, we conducted the GWAS only with the first of the two traits. An analysis of 603,871 SNPs in 61 genotypes with FarmCPU did not uncover statistically significant associations with average severity when the model was fit with or without principal components (PCs) (Fig. 5a and Supplementary Fig. S5A, respectively). A single variant on chromosome 4 (S04_33782670) is located above a threshold (1.656×10^{-06}) in the model without correction for population structure using principal components (Fig. 5). The QQ plots showed no sign of inflation or deflation of p -values with respect to the theoretical expectation (Figs. 5b and Supplementary Fig. S5B, respectively) and therefore supports the absence of a significant association.

We also used a k -mer based approach because it allows the inclusion of additional genotypes with lower sequence coverage ($n = 88$) and is not biased to genomic regions included in a reference sequence. This analysis was based on an average of 570,741,731 k -mers of length

31 per sample. The control variety Titicaca (67,365,628) and genebank accessions G37 (814,316,239) were the genotypes with the lowest and highest numbers of k -mers counts, respectively. In total, 992,946.265 k -mers passed the filters and were included in the kinship matrix estimation and the subsequent GWAS. For the first stage of the GWAS, 880,137,481 k -mers were tested and 10,001 passed the first filter to be fit with GEMMA. The smallest p -value for any k -mer was 9.19×10^{-10} for a single k -mer. Therefore, this analysis also did not detect any significant associations with the trait severity of infection given a permutation-based 5% p -value threshold of 1.505×10^{-10} for the k -mer analysis.

Discussion

The inoculation of quinoa varieties and genebank accessions with an isolate of the downy mildew pathogen *Peronospora variabilis* revealed substantial variation in the three infection related traits severity of infection, sporulation and incidence of mildew (Fig. 3). Using a mixed model approach we validated that estimates of genetic effects, variance components and heritabilities are robust with respect to data transformation and the inclusion or exclusion of control varieties between experiments. The differences in susceptibility to mildew infection have a strong genetic component as indicated by high genetic variance component estimates and high heritabilities, but we were not able to identify individual genomic regions strongly associated with mildew susceptibility in a GWAS. In previous studies, ecotype, environmental or physiological parameters like altitude of site of origin, seed saponin content, or size and density of stomata were postulated to be correlated with disease tolerance. We did not find any strong correlation of the three disease traits with any of these parameters except for width of stomata.



Large genetic effect of mildew tolerance

Mildew tolerance was scored in a single plant per block, in the case of accessions and cultivars, while control varieties had multiple plants because one person scored each leaf of all plants and the labor-intensive process has a time limitation for scoring window (max. 12 h). To evaluate the robustness of parameter estimates, we used linear mixed models and various combinations of data transformation and inclusion or exclusion of control varieties. Although the use of data transformation is under debate [28–30], it did not have a large effect on the distribution of residuals or on tests of fixed effects. This provides strong evidence for the robustness of our estimates because we phenotyped approximately the same number of plants per genotype, which reduces the effect of heteroscedasticity [31], and from the robustness of LMMs to heterogeneous error variances [32]. The effect of replicated control varieties in the three experiments on model fit and parameter estimation was minor, because their removal caused only a small increase in \bar{H}^2 estimates. This robustness results from a balanced experimental design, in which changes to differences are only small [33].

A limitation in fitting LMM and GLMM is that reliable estimation of the variance of a random effect requires at least five levels, i.e. locations, experiments, years, etc. [27]. The low number of groups in our experiments explains the large confidence intervals of the design variance components σ_E^2 and $\sigma_{E:B}^2$ (Supplementary Tables S2 and S3), which therefore are not reliable estimates of variation between experiments or blocks. We nevertheless modeled these effects as random to use inter-block information and to account for non-independence of the data, because genotypes were nested in blocks, which were nested within experiments [27, 34].

Previous work on quinoa yield traits found high estimates of genotype-by-environment variance components [35]. The genotype-by-experiment interaction variance ($\sigma_{G:E}^2$) in our study was also large in comparison to other variance components across traits and models, which reflects that even subtle differences in the environment can cause a genotype to respond differently to the disease. From a plant breeding perspective, the $\sigma_{G:E}^2$ variance component is important because it masks the genotypic component of phenotypic variance. A high observed $\sigma_{G:E}^2$ suggests that future studies of mildew tolerance using diverse quinoa genotypes should include multiple locations and years [36], or unbalanced designs with replicated control varieties that allow testing of many genotypes [37].

High correlations between disease-related traits

To identify different reactions of host plants, we measured three different disease-related traits. A comparison of severity of infection and sporulation is of interest

because pathogen populations typically harbor high levels of genetic variation for both virulence and fecundity [38, 39]. Some quinoa host genotypes might allow fast fructification of the pathogen while others may suppress its proliferation [40]. We found a very strong positive correlation between severity and sporulation and scoring of the former seems to be a good predictor of the latter. For this reason severity alone can be used to assess a panel of genotypes because it also showed a smaller error variance than either sporulation (Fig. S3D–F) or incidence (Fig. 1). The correlation coefficients of the disease traits in our study ($R = 0.9$) is larger than the coefficient reported in an evaluation of scoring methods for downy mildew in cucumber [41]. However, sporulation should be measured in highly resistant genotypes with <5% severity of infection, because in such a genetic background pathogen proliferation may be strongly impaired as observed in the wild relative *C. album*, control variety Puno and several gene bank accessions.

Analysis of heritability

Consistent with estimates of genetic variance, heritability estimates were moderately high and very similar between models with a range from 0.72 to 0.78 for severity and from 0.78 to 0.81 for sporulation. In comparison, estimated heritability of sporulation of downy mildew in grapevine (*Vitis vinifera*) was around 0.40 [42], and resistance to systemic infection by sorghum downy mildew in maize was in the range of 0.61–0.68 [43]. Although estimates of \bar{H}^2 from greenhouse experiments may differ from field trials [44], the high heritabilities for the disease traits indicate the selection for higher mildew tolerance is possible. In this respect our results are consistent with previous estimates of \bar{H}^2 for physiological, morphological and yield traits, which are also high (≈ 0.85) [45] and indicate that multiple traits of quinoa can be substantially improved by plant breeding.

Identification of accessions tolerant to downy mildew

Although our panel includes the main quinoa ecotypes [46], there was no correlation between elevation and severity of infection or sporulation. Since infection requires high humidity (Fig. 8), a lack of such a correlation reflects microclimatic variation of humidity in high altitudes. For example, the Bolivian highland (Altiplano) is more humid in the North than in the South [47]. In addition, ecotypes from the Andean Valley (2500–3500 masl), where humidity also tends to be higher, are among the most tolerant accessions [47–49]. These observations are consistent with our results because five accessions from the southern altiplano had the highest severity of infection (G4, G9, and G82 from Bolivia, and G67, and G99 from Peru).

We observed that a ranking of genotypes by their mildew tolerance is robust with respect to disease traits and analysis methods, which allows to identify accessions high tolerance for further investigations or utilization in breeding. Tolerant genotypes include the wild relative *C. album* and the Puno control variety, which was one of the genotypes with the lowest severity, sporulation, and incidence of downy mildew (Fig. 3a-c), as well as a set of Bolivian cultivars and genebank accessions. For example, cultivars Mañiqueña (CV21) and Phisankalla (CV10) perform well in dry areas like the southern Altiplano in Bolivia, but are susceptible to mildew in humid environments [3, 7, 48]. This phenotype was confirmed in our study because both cultivars are among the most susceptible under the humid conditions of our experiment (Fig. 3). It has been proposed that genotypes with good performance in dry and a high disease susceptibility in humid environments either have not been selected for disease tolerance during domestication or have an advantage in dry environments possibly because of the cost of resistance [47]. Current evidence is contradictory and does not establish such a relationship, because cultivar 'Rosa Blanca' (CV6) was developed in a dry region [3], but has a higher tolerance to the disease with an average severity of 32%, whereas the more susceptible cultivars 'Jach'a Grano' (CV15; 64% severity) and 'Aynoka' (CV20; 58% severity) originated from the same breeding program. Such a high variation may result from the interaction between genotypes and experiments (Fig. S3) and is consistent with similar GxE interactions anatomical and yield-related traits of *C. quinoa* [45, 50].

An important limitation of our study is the use of a single isolate only of *P. variabilis* for inoculation because it does not allow to test whether mildew tolerance is race-specific or reflects a quantitative resistance. Preliminary evidence supports the latter hypothesis, because cultivars Kurmi (CV16) and Mañiqueña Real (CV21) were inoculated with a Bolivian isolate of *P. variabilis* and classified as tolerant and susceptible, respectively, as evaluated by the disease progression of downy mildew *P. variabilis* [51]. Our results confirm the differences between these two cultivars with a different isolate because Kurmi (45.1% severity of infection) was less susceptible than Maniqueña (70.7%) (Fig. 3). Since Kurmi was developed for cultivation in the highlands (3600–3800 masl) and selected in field trials for downy mildew resistance [52], a high tolerance of Kurmi to two different isolates supports a quantitative disease tolerance. Our results also support previous work that found recombinant inbred lines from both Chilean and Peruvian origin segregating for quantitative mildew resistance in a $F_{2:6}$ population [23].

Comparison of passport data with our results is limited by missing information on scoring methods, phenotypical

stage of the plant, genetic constitution of accessions (e.g., extent of heterozygosity) and information about field trials in Bolivia. In addition, our data on severity of infection is based on one pathogenic isolate whereas the passport data is based on natural infections of local races whose virulence might differ from the Danish isolate used in this study [14]. Previous studies have shown that Andean isolates are genetically distinct when highly sensitive polymorphism methods of identification are used [14]. Based on the PCR sequencing data, our isolate showed complete sequence identity in the ITS region to specimen EU 113305 collected in Tåstrup, Denmark on *C. quinoa* cv. Atlas in 2001 [8].

Relationship of disease traits with stomatal traits and seed saponin content

P. variabilis enters host tissues through stomata, and its haustoria emerge from the stomatal pore to release spores [8, 24], which may explain the preference of pathogen for humid conditions because stomata are typically open under such conditions [53]. We therefore tested whether pathogen traits are correlated with stomatal traits and found that only severity of infection showed a weak correlation with stomatal width. This result should be taken with caution because our measurements were based on only a single leaf per genotype due to the high effort required to obtain the data. Future phenotyping should make use of automated image analysis methods to obtain larger data sets for stomatal traits.

Downy mildew remains dormant in the pericarp of the seed [54] and the saponin content of the seeds could influence the response of a genotype during the early stages of infection. However, we found no support for a relationship between saponin content of the seed and mildew severity. One explanation for the absence of a correlation may be that the foam test revealed GxE effects, because the scores for content of saponin differed between seed samples of the same genotypes obtained from plants cultivated in different locations. This is expected because the content of saponin is variable over time and depends on the water status of the plant [55, 56].

Therefore, our results do not support the hypothesis that mildew tolerance is substantially influenced by other traits such as stomata characteristics or saponin content.

Interpretation of GWAS for severity of downy mildew

We conducted a GWAS with severity of mildew infection and whole genome resequencing data to test whether the observed differences between genotypes are caused by few genomic regions. Both GWAS methods failed to detect significant associations of variants or *k*-mers with severity of infection. The power of GWAS

depends on the sample size of the association panel and on the genetic architecture of a trait of interest [57]. Our analysis was limited by a small sample size of 61 (Farm-CPU) and 88 (*k*-mer analysis) accessions, and possibly by the genetic architecture of severity of infection because distribution of phenotypic values suggests it is a polygenic quantitative trait (Fig. 3a). A polygenic response of quinoa to *P. variabilis* infections is supported by multiple studies that include greenhouse experiments and field trials [13, 23, 40, 58–61].

Our results provide a perspective for a more efficient resistance breeding in quinoa. The large variation found in mildew tolerance and high heritabilities of disease traits allows the development of QTL mapping populations by crossing genotypes from both ends of the distribution (e.g., Danish varieties Puno and Titicaca; Fig. 3). Previous work suggests that QTL mapping may identify major R genes that could be useful in quinoa breeding because mildew tolerance is modulated by incomplete gene effects [62], which depends on pathogen aggressivity, as observed on Ecuadorian material [21]. Furthermore, the segregation ratio of mildew severity in an F_2 mapping population derived from a cross of bitter and sweet (i.e., no seed saponins) genotypes suggested that mildew tolerance shows a dominant inheritance [63].

The analysis and utilization of genetic variation for mildew tolerance will be enhanced by more high throughput phenotyping methods. Our experimental setup was on targeted inoculations with a single isolate, which contributes to a robust and repeatable estimation of disease tolerance, but it is work intensive and limits the number of genotypes for genetic mapping. However, alternative approaches such as scorings of detached leaves or randomized selection of leaves on the field can be misleading because host genotype \times pathogen genotype \times environment (G \times G \times E) effects in the field are difficult to control. Furthermore, symptoms of pathogen infection are influenced by the position and age of leaf tissue [64] which results from induced resistance that occurs not only at the site of the initial infection but also in distal, uninfected parts [65]. Therefore, in addition to a controlled greenhouse experiment used in this study, multilocation field trials of segregating populations that use modern phenotyping technologies such as deep learning to score pathogen infections are a complementary approach in resistance breeding [66]. Both approaches in combination with genetic analysis will contribute to the development of improved quinoa varieties in both within and outside of the native cultivation range.

Conclusion

Our study revealed a high level of variation of quinoa varieties and accessions to *Peronospora variabilis* infections. We have shown that cross-infection from *C. album*

to *C. quinoa* and vice-versa is feasible and this widely distributed weed is likely a reservoir for the pathogen and an alternate host for the *P. variabilis*, which has implications for quinoa cultivation in the presence of *C. album*. The substantial variation in mildew tolerance between genotypes has a strong genetic component is therefore is amenable to selection in breeding programs. However, inferences based on a single experiment - or a single location field trial - should be taken with care because a large genotype by experiment interaction was found, so future work on the resistance of *C. quinoa* to *P. variabilis* must take this into consideration during the design and planning phases.

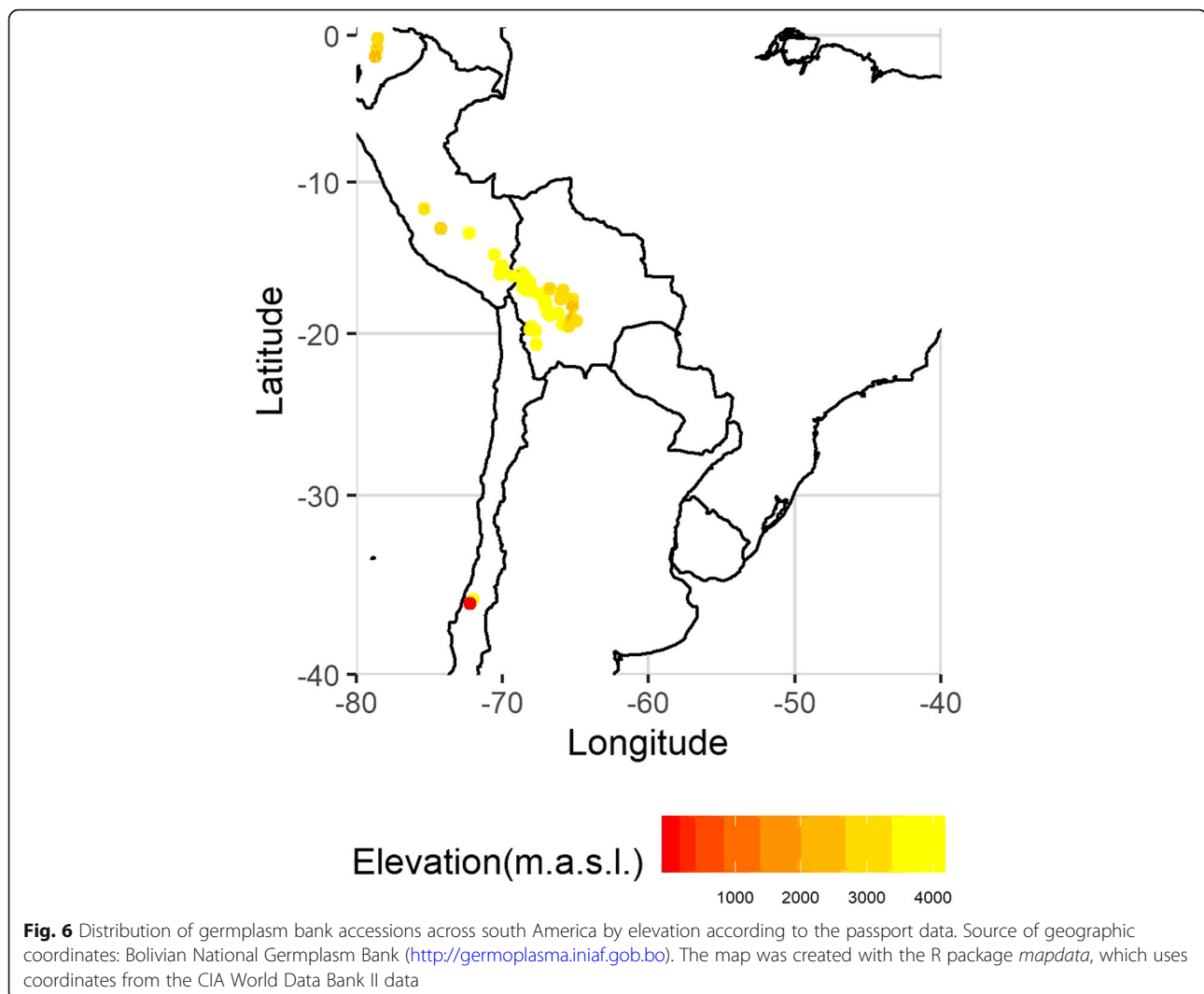
Methods

Plant material

The quinoa genotypes analyzed in this study consist of 106 accessions stored in the National Germplasm Bank of Bolivia. They include landraces collected in Bolivia (55 accessions), Peru (33), Ecuador (7) and Chile (4), in altitudes ranging from 2 m.a.s.l. to 4082 m.a.s.l. (Fig. 6). Seven accessions had no information about their origin. We also included 21 Bolivian cultivars, the Bolivian variety 'Blanca' ('Blanquita') and three Danish varieties 'Puno', 'Vikinga' and 'Titicaca'. The list of accessions and their passport data are provided in [Supplementary File 1](#) (Source: <http://germoplasma.iniaf.gob.bo>).

Genebank accessions were selected to represent both the geographic diversity of quinoa and variation in mildew tolerance, which was scored in field trials in La Paz (Bolivia) based on spontaneous infections by *P. variabilis*. Additional information on the genetic status of accessions, scoring method, phenological stage during scoring, and trial locations were not available from passport data.

Four quinoa varieties and the wild relative common goosefoot (*Chenopodium album* L.) were used as control varieties in greenhouse experiments. The four quinoa varieties include the cultivar 'Blanca', which is adapted to the Northern highlands and Inter-Andean valleys of Bolivia and partially resistant to downy mildew [3, 52]. The other three control varieties 'Titicaca', 'Puno' and 'Vikinga' were developed in the quinoa breeding program of the University of Copenhagen. Varieties Puno (KVL 37) and Titicaca were bred from Chilean and Peruvian landraces and selected for earliness and adaptation to European conditions [67, 68]. They showed different levels of downy mildew susceptibility in a field trial (S.-E. Jacobsen, personal communication), which was confirmed in a pilot experiment for this study (Supplementary Fig. S1). *Chenopodium album* is closely related to quinoa and a widely distributed weed. *C. album* seeds used in this study were collected in 2017 and 2018 at a former quinoa breeding field at the



experimental station of the Faculty of Science, University of Copenhagen (Højibakkegaard, Tåstrup).

Experimental research on these plants, including the collection of plant material, complied fully with institutional, national, and international guidelines. Greenhouse studies were conducted in accordance with local legislation. Permission was granted from INIAF-Bolivia to use the seed material of the Bolivian genebank accessions for the sole purpose of this research on the understanding all seeds will be destroyed on completion of the project. Seeds of the Danish varieties was provided as gift by the company QuinoaQuality Aps.

Peronospora variabilis isolate used for inoculation

Previous research recognized the role of alternate hosts on the evolution and spread of pathotypes [69] *P. variabilis* is a pathogen of both *C. quinoa* and *C. album* [8, 14, 15, 18, 20]. To obtain a defined isolate of *P. variabilis*, leaves from *C. album* with typical downy mildew sporulation were

collected in late September 2018 at a former quinoa breeding field on the research station Højibakkegaard. The isolate was inoculated for maintenance and propagation into two quinoa cultivars (Blanca and Vikinga) using a protocol by Danielsen and Ames [48]. We used these two cultivars because they differed in their latent period [40]. The latent period lasted 5 days in the Vikinga variety and 7–10 days in the Blanca variety. These differences allowed to maintain the pathogen on Blanca, and a quick propagation on the Vikinga variety.

We used DNA sequencing of the Internal Transcribed Spacer (ITS) region to confirm that the isolate was *P. variabilis*. A spore suspension (1×10^{11} spores/ml) produced from the maintained inoculum was filtered with a nylon filter with $20 \mu\text{m}$ pore size (Merck Millipore Ltd.) to capture *P. variabilis* spores, which were transferred to 1.5 ml microcentrifuge tubes containing glass beads ($425\text{--}600 \mu\text{m}$) and kept on ice. $200 \mu\text{l}$ lysis buffer (DNAeasy Plant Mini Kit, Qiagen) were added and

mycelia were pulverized with a sterile pestle. Further 200 μ l of lysis-buffer with 4 μ l of RNase were added. DNA was extracted with a DNeasy Plant Mini Kit (Qiagen) following manufacturer's instructions. Primers designed to amplify a 1150 bp fragment covering ITS-1 and ITS-2 region in members of the oomycete family Peronosporaceae, including species of *Peronospora*, *Pythium*, and *Phytophthora*, were amplified from genomic DNA by polymerase chain reaction (PCR) using Oomyc Fw-1: 5' cggaggatcattaccacac and Oomyc-Rv1: 5' cgcttattgatgcttaagtca as forward and reverse primers, respectively. PCR amplification was carried out with one cycle of 95C for 3 min; 35 cycles of 94C for 30 s, 55C for 30 s and 72C for 40 s, and one cycle of 72C for 3 min. Amplification products were purified using QIAquick PCR purification columns (Qiagen) and the DNA concentrations were determined on a NanoDrop Lite Spectrophotometer. DNA sequencing of the PCR amplified ITS was performed at Eurofins Genomics. DNA sequences were submitted to NCBI (accession MT895880) and compared against the NCBI nr database using BLASTN (<https://blast.ncbi.nlm.nih.gov>).

Identification of *P. variabilis* by microscopy and histopathology

The identity of the pathogen was also confirmed by visual analysis with microscopic and histopathology slides. Microscopic analysis was carried out with a sample of the *Peronospora* solution mounted on a glass slide. Histopathological slides were prepared from Blanca leaf pieces of ± 5 cm² collected 7 days after infection. Leaf samples were coated with nail polish on the abaxial side and dried for 24 h. The imprints were then removed and stained

with Lactophenol Aniline Blue [70]. Both microscopic and histopathological preparations were mounted on a Leica microscope MZ12.5 and photographed with a LEICA DFC420 camera under 40X magnification (Fig. 7a, b). Additionally, live infections were captured with a digital microscope (Dino-Lite, model AM4113/AD4113 Dino-Lite, Naarden, Holland) (Fig. 7c). After verification and calibration of the pathogen, the isolate was constantly propagated *in planta* on Vikinga and Blanca. The detailed protocol for the isolation and propagation of the pathogen is available in [Supplementary File 2](#).

Design of phenotypic characterization

Phenotypic data were collected in greenhouses of the University of Copenhagen between February and May 2019. The response of quinoa genotypes to downy mildew inoculation was evaluated in three sequential identical experiments that each included the complete set of 132 genotypes and the four control varieties with a randomized complete block design with four blocks each. Each experiment occupied a greenhouse allocated exclusively for the experiment to avoid infestations with insects or risk of cross contamination as well as the provision of biological control agents. Experiments started 2 weeks after the end of the prior one. Within blocks, accessions and cultivars were represented by a single plant while control varieties were represented by 2 to 5 plants.

Prior to the experiment, the Bolivian gene bank accessions were self-pollinated once to increase homozygosity because the heterozygosity of genebank accessions was unknown. Seeds were produced in a greenhouse between February and August 2017 with a 12 h photoperiod, an

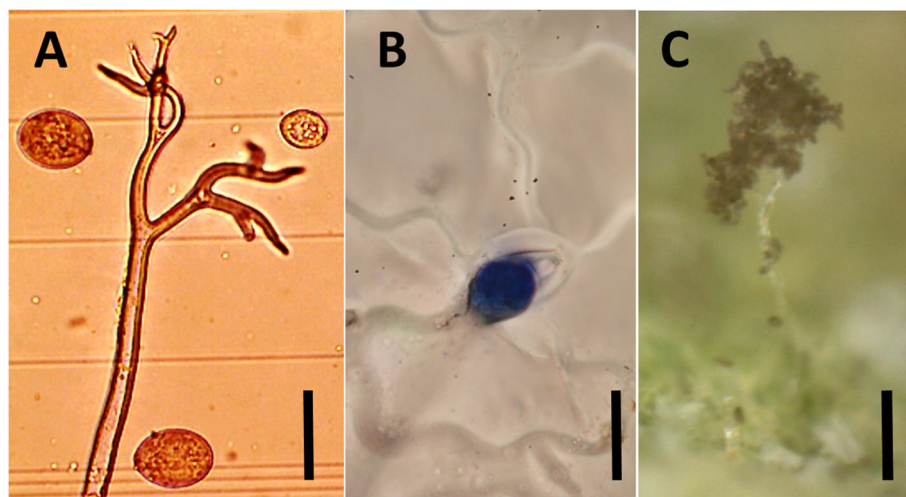


Fig. 7 Microscopic and histopathological imaging of *Peronospora variabilis*. **a** Sporangiophore and spores. **b** Spore and hyphae penetrating a stoma during infection (cultivar Blanca, 7 days after inoculation). **c** Branched sporangiophore bearing sporangia emerging from leaf surface. Scale bar is 20 μ m for (a), (b), and 100 μ m for (c)

average temperature of 24C during the day, 18C at night, and irrigated with fertilized water (NPK 14–3–23 + mg EC 1.9). After the day length surpassed the plants requirements, greenhouse curtains were used to maintain the photoperiod. Seeds were harvested, cleaned and stored at natural conditions until January 2018. Harvested seeds were sown in jiffy pots containing peat to assure the provision of plantlets for transplantation and grown for seven to ten days. To avoid infestation with flies, the compost was watered with a solution of gnatrol (10% v/v). Plantlets were then transplanted to 550 cm³ pots and grown for 3 weeks in greenhouses under the same conditions. The greenhouse management included biological control agents against common greenhouse-borne pests.

Three weeks after transplanting, plants were moved to different greenhouses and inoculated with a calibrated solution (1×10^5 spores/ml) of *P. variabilis* spores and Tween 20 (1%). The solution was sprayed comprehensively onto each plant with a pressure paint gun. 50 ml of solution were used for each block. Blocks with inoculated plants were covered with a plastic sheet 5 days after inoculation (Fig. 8a) for 24 h under complete darkness and a night-time temperature of 15C to create conditions that stimulate infection. After removal of the cover, plants were grown under greenhouse conditions. Once symptoms were observed (Fig. 8b), usually between 5 and 6 days after inoculation, plants were covered again for 24 h to promote sporulation.

The plant response to pathogen infection was measured with the three variables: severity, sporulation, and incidence. Scoring of severity and sporulation by visual analysis of the foliar area covered by lesions of chlorotic or other color on the adaxial leaf side and the area of diseased tissue with visible spores on the abaxial leaf side, respectively. Measurements were recorded as percentages for each leaf and then averaged per plant. Incidence was calculated as the proportion of leaves with symptoms. These parameters were measured by the same person to avoid operator bias and all plants had the same age when scored.

Phenotyping of stomata

To measure width, length, and density of stomata on leaf surfaces, a resin cast of the abaxial surface of one leaf from each genotype was made [71]. After drying, a layer of nail polish was added to the cast, left to dry, removed, mounted onto a glass slide and covered with a glass coverslip. Slides were mounted on a Leica MZ12.5 optical microscope and three fields were photographed using a Leica DFC420 digital microscope camera with 40X magnification. Width and length were measured using the Leica Application Suite software. Stomatal density was estimated as the number of stomata per unit of area and stomatal counts were obtained by using Stomata Counter, a web-based application, followed by manual curation of the data [72].



Fig. 8 Setup of inoculation experiment. **a** Block of the experiment with plastic sheet covers to provide high humidity. **b** Plants within a block during the darkness phase of infection. **c** Severity symptoms ranging from hypersensitive reactions causing pale yellowish spots (left) to high susceptibility with chlorotic lesions covering the whole leaf (right) **d** Leaf of tolerant Puno control variety showing chlorotic lesions but no sporulation. **e** Leaf of Cultivar 21 (Mañiqueña) with signs of sporulation

Analysis of phenotypic data with linear mixed models (LMM)

The following mixed model was used to estimate the mean severity and sporulation of the disease for each genotype in the panel, using ASREML-R package version 3.0 [73]:

$$y_{ijk} = \mu + \rho_i + \beta_{ij} + \alpha_k + (\rho\alpha_{ij}) + e_{ijk} \quad (1)$$

where y_{ijk} is the response (severity or sporulation) of the k -th genotype in the j -th block of the i -th experiment, μ is the general mean, ρ_i is the effect of the i -th experiment, β_{ij} is the effect of the j -th block nested within the i -th experiment, α_k is the genotype effect, $\rho\alpha_{ij}$ is the genotype-experiment interaction, and e_{ijk} is the residual error term. The effects for experiments, blocks within experiments and the genotype-experiment interaction were treated as random effects because experiments were considered as a random factor and hence all effects involving a random factor need to be modelled as random [74], whereas main effects of genotypes were treated as fixed. To avoid the influence of outliers on estimates of genetic variance, outliers were detected after fitting the model and removed from the dataset using the default method of the PLABSTAT package [75], as described in [76].

The residual error for the experiments was initially modeled as normally distributed and independent with a common variance component, where n is the total number of observations. In addition, we fitted a model such that the variance-covariance matrix of the vector of errors (sorted by experiments) was $R \sim N(0, \sigma_e^2 I_n)$, where n is the total number of observations. In addition, we fitted a model with independent variance components for each experiment [77]:

$$R = \bigoplus_{j=1}^s R_j = \begin{bmatrix} R_1 & 0 & 0 \\ 0 & R_2 & 0 \\ 0 & 0 & R_3 \end{bmatrix}$$

where $\bigoplus_{j=1}^s R_j$ is the direct sum of matrices and R_1 , R_2 and R_3 are variance-covariance structures for each experiment, each taking the form $R_j = \Sigma_e^2(j) * I_{nj}$, where n_j is the number of observations in the j -th experiment.

Analysis with generalized linear mixed models (GLMM)

We used a Generalized Linear Mixed Model (GLMM) to analyze the incidence of downy mildew in quinoa and fit the following model with the PROC GLIMMIX procedure of SAS software:

$$\text{logit}(\pi_{ijk}) = \eta_{ijk} = \mu + \rho_i + \beta_{ij} + \alpha_k + \rho\alpha_{ik} \quad (2)$$

where logit is the link function between the linear predictor and the observations (p_{ijk}), ρ_i is the effect of

the i -th experiment, β_{ij} is the effect of the j -th block nested within the i -th experiment, α_k is the genotype effect, and $\rho\alpha_{ik}$ is the genotype-experiment interaction. The model included a scale parameter account for over-dispersion of the data through the residual keyword in the RANDOM statement of PROC GLIMMIX in SAS version 9.0 [78].

Heritability estimation

The ad-hoc broad-sense heritability was estimated as:

$$\bar{H}^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\bar{v}}{2}} \quad (3)$$

where σ_g^2 is the genetic variance and \bar{v} is the mean variance of the difference of the adjusted means [33]. To estimate \bar{H}^2 , models were fit with genotypes as a random effect using the ASREML-R package for severity and sporulation and PROC GLIMMIX in SAS for incidence to obtain an estimate of σ_g^2 . The models with genotypes as fixed effect were used to estimate \bar{v} .

Model comparisons

To compare models with different error variance structures, the restricted likelihood ratio test implemented in the asremlPlus R package [79] was used to test if heterogeneous error variances improved the model. The effect of replicated control varieties on estimates of genetic variance was addressed by adding a dummy variable to the severity of infection, sporulation and incidence models [80]. Such a model was formulated as

$$y_{ijk} = \mu + \rho_{ij} + \beta_{ij} + Y\alpha_k + W(\rho\alpha_{ij}) + e_{ijk} \quad (4)$$

where Y and W are vectors with 0 for reference varieties and 1 for cultivars and accessions, α_k is the genetic random effect. The remaining effects are the same as in Eqs. (1) and (2). These models were compared using the mean standard error of the difference (s.e.d.) and heritability. The s.e.d.'s were calculated using the predictplus function of the asremlPlus R package [79].

The effect of transforming our severity of infection and sporulation scorings on heritability estimates was evaluated by repeating the steps outlined above with data transformed with the logit ($\log[\frac{p}{1-p}]$) and the angular, or arcsine root, transformation ($\sin^{-1}[\sqrt{p}]$), where p are the severity of infection or sporulation observations. Because the logit function is undefined at 0 or 1, the data at these limits was adjusted by adding and subtracting 0.025 from the original value. The fixed effect of genotypes was tested by using Wald's F-test as implemented in the ASREML R

package for LMMs and type II tests of fixed effects of the Proc GLIMMIX procedure of SAS.

Comparisons between means

The mean severity and sporulation for the downy mildew infection on each genotype, their confidence intervals and all pairwise comparisons were estimated with the *asremlPlus* R package [79] for severity and sporulation, and the PROC GLIMMIX procedure of SAS for incidence. Comparisons between the means were based on *t*-tests with a significance threshold $\alpha = 0.05$.

Correlation between traits

To identify any correlations between phenotypic traits and traits related to the tolerance of quinoa against mildew, i.e. severity, sporulation, and incidence, we used our data from measurements of stomata (width, length, and density). Pearson correlation coefficients were estimated for each pair of variables with a significance threshold $\alpha = 0.05$, using the R package *Hmisc* [81].

Relationship between saponin presence and downy mildew severity

Saponin content of seeds was assessed using the foam test [82], which consists of placing 0.5 g seeds with 5 ml distilled water in a test tube and shaking vigorously for 30 s. Foam height was recorded to the nearest 0.1 cm after shaking. To estimate the robustness of this saponin assay, two seed samples per gene bank accession were evaluated, one from plants grown in Bolivia and one from plants propagated at Højbakkegaard. All accessions with reads equal to zero (i.e., no foam was observed after shaking) were labeled as “no saponin” and all others were marked as “with saponin”. To test for a relationship between saponin presence or absence and downy mildew severity, we conducted a *t*-test using the adjusted means obtained from a LMM with heterogeneous variances between experiments using an untransformed data without the control varieties. This set of means was used because there was no indication from the previous analysis that fitting models with transformed data improved accuracy of the estimates.

Whole genome DNA sequencing

For DNA extraction, two plants per genotype were grown in a greenhouse of the Taastrup campus at the University of Copenhagen, and two healthy leaves from a single two-months old plant were collected and stored with silica gel for drying. DNA was extracted using the AX Gravity DNA extraction kit (A&A Biotechnology, Gdynia, Poland) following manufacturer’s instructions. Purity and quality of DNA were controlled by agarose gel electrophoresis and concentration determined with a Qubit instrument using SYBR green staining. DNA

sequencing libraries were constructed using the protocol of Baym et al. [83]. Whole-genome sequencing was done with short-read Illumina sequencing on an Illumina NovaSeq machine (Novogene).

Genome sequencing, variant calling and genotype imputation

Processing of the raw reads, mapping, and variant calling were done with a custom Snakemake pipeline [84]. Raw reads were trimmed with *Trim_galore* v 0.6.4 [85] (parameters `-q 30 -fastqc -paired`). Reads were then sorted and indexed with *SAMTOOLS* 1.10 [86] and deduplicated with the *MarkDuplicates* (parameter `REMOVE_DUPLICATES = TRUE`) tool of *PICARD* v2.21.9 [87]. The resulting FASTQ files were mapped against the quinoa reference genome version 1.0 [88] and the organellar genomes [89] using the *Burrows-Wheeler Aligner* v0.7.17 [90] with default parameters.

Variants (SNPs and indels) were called using *GATK* 3.8 [91] by using the *HaplotypeCaller* tool with a minimum per-base quality score of 20 and a minimum mapping quality score of 30. The GVCF files per sample were merged with the *GenotypeGVCFs* tool of *GATK* with default parameters. Missing data was imputed and filtered using *LinkImputeR* 1.2.3, which allows the user to define a series of filters and evaluate their effect on the accuracy and final number of imputed markers [92]. Thresholds for imputation were `depth = 8` (Number of reads including a position) and `missingness = 0.7` (Proportion of positions/samples with less than the threshold depth). Variant data were filtered to a minor allele frequency ≥ 0.05 and a deviation from Hardy-Weinberg equilibrium $p > 0.01$ using a likelihood ratio test [93]. Linkage disequilibrium was estimated with a pair-wise correlation coefficient between variants, r^2 , using the final VCF file as input for *PopLDdecay* [94] with default parameters.

Genome-wide association study (GWAS)

We used two methods for association mapping, *FarmCPU*, for use with sequence variants (SNPs) and a *k*-mer-based method. *FarmCPU* (Fixed and random model Circulating Probability Unification) uses SNP in a two-step iterative process with fixed and random effects models to improve computation times, reduce the confounding effects of structure and improve power to identify significant marker-trait associations in comparison to other methods [95, 96]. The model was run with and without the inclusion of the first three principal components with a *p*-value threshold of 0.01 (Bonferroni corrected) for both the inclusion of a marker during the first iteration of the model as well as the genome-wide significance threshold.

The *k*-mer based method by [97] identifies genotype-phenotype associations using sequencing reads instead of molecular variants to address the lack of a reference genome or account for structural variation. We implemented the method in a Snakemake pipeline using the following parameters: *k*-mer length of =31 nucleotides, minor allele count =3 minor allele frequency =0.05. This method requires a kinship matrix, which was estimated with a method used by EMMA (Efficient Mixed-Model Association) and consists of an identical-by-state (IBS) allele-sharing matrix under the assumption that every variant has a small random effect on the phenotype [98].

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-020-02804-7>.

Additional file 1: Figures S1 to S5 and Tables S1 to S3.

Supplementary Figures and Tables.

Additional file 2: Supplementary File 1. List of genebank accessions and passport data; Mildew infection raw data; saponin and stomatal measurements.

Additional file 3: Supplementary File 2. Detailed protocol for the isolation and maintenance of the downy mildew pathogen *Peronospora variabilis*.

Abbreviations

GLMM: Generalize linear mixed model; GWAS: Genome-wide association study; LD: Linkage disequilibrium; LMM: Linear mixed model; PCA: Principle components analysis; SNP: Single nucleotide polymorphism

Acknowledgments

We are grateful to INIAF (Bolivian National Institute for Innovation in Agriculture and Forestry) and to the greenhouse team. The authors wish to thank laboratory technician Lene Klem at University of Copenhagen for the PCR analysis on the pathogen. We thank Elisabeth Kokai-Kota for the construction of DNA sequencing libraries and to Mireia Vidal for advice on bioinformatics. The authors acknowledge support by the state of Baden-Württemberg through bwHPC. We express our appreciation to Sven-Eric Jacobsen for the provision of the Danish cultivars, support and knowledge.

Authors' contributions

CCL, MCA, HPP and KS designed the study. CCL conducted the field house experiment and phenotyping work. OSL designed the pilot study and the molecular identification of the pathogen isolate. CCL and MCA analysed the phenotyping data. DBA, CA and SS advised on experimental work and data analysis. MCA analysed sequencing data and carried out GWAS. CCL, MCA and KS wrote the manuscript. All authors read, revised and agreed on the manuscript.

Funding

This work was sponsored by fellowship to CCL by the Plurinational State of Bolivia through the program "Scholarships for the sovereignty for Science and Technology." Part of this work was funded through the F.W. Schnell Professorship of the Stifterverband to KS. Funding agencies had no role in the study design, data collection and analysis, or data interpretation. Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

All phenotypic data generated or analysed during this study are included in this published article in [Supplementary File 1](#) (Complete mildew infection dataset). The sequencing data for the identification of *Peronospora variabilis* are available from NCBI Genbank with ID MT895880, <https://www.ncbi.nlm.nih.gov/nucleotide/MT895880>. Raw whole genome sequencing read data are

available from the European Nucleotide Archive (ENA) with project number ID PRJEB39907, <https://www.ebi.ac.uk/ena/browser/view/PRJEB39907>.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests

Author details

¹Department of Plant and Environmental Sciences, University of Copenhagen, Højbakkegaard Allé 13, DK-2630 Taastrup, Denmark. ²Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, Fruwirthstrasse 21, D-70599 Stuttgart, Germany. ³Institute of Crop Science, University of Hohenheim, Fruwirthstrasse 21, D-70599 Stuttgart, Germany.

Received: 24 August 2020 Accepted: 16 December 2020

Published online: 14 January 2021

References

- Mujica A, Jacobsen S-E. La quinoa (*Chenopodium quinoa* Willd.) y sus parientes silvestres. In: Moraes RM, Øllgaard B, Kvist LP, Borchsenius F, Balslev H, editors. Botánica Económica de los Andes Centrales. La Paz: Universidad Mayor de San Andrés; 2007. p. 449–57.
- Gómez L, Aguilar E. Guía del Cultivo de Quinoa. Second. Lima: Universidad Nacional Agraria La Molina; 2016. <http://www.fao.org/documents/card/es/c/3a12f679-22a1-46a0-a91e-6853ca5bb5dd/>.
- Gandarillas A, Rojas W, Bonifacio A, Ojeda N. Quinoa in Bolivia: The PROINPA Foundation's Perspective. In: Bazile D, Bertero D, Nieto C, editors. State of the art report on quinoa around the world in 2013. Rome: FAO regional office for Latin America; the Caribbean; 2015. p. 344–61. <http://www.fao.org/3/a-i4042e.pdf>.
- Zurita- A, Fuentes F, Zamora P, Jacobsen S-E, Schwember AR. Breeding quinoa (*Chenopodium quinoa* Willd.): Potential and perspectives. Mol Breed. 2014;34:13–30.
- Bazile D, Bertero D, Nieto C. State of the art report on quinoa around the world in 2013: FAO; CIRAD; 2015. <http://www.fao.org/3/a-i4042e.pdf>.
- Bazile D, Pulvento C, Verniau A, Al-Nusairi MS, Ba D, Breidy J, et al. Worldwide evaluations of quinoa: preliminary results from post international year of quinoa FAO projects in nine countries. Front Plant Sci. 2016;7:1–8.
- Murphy KM, Matanguihan JB, Fuentes FF, Gómez-Pando LR, Jellen EN, Maughan PJ, et al. Quinoa breeding and genomics. In: Goldman I, editor. Plant Breeding Reviews. Hoboken: John Wiley & Sons, Inc.; 2018. p. 257–320.
- Choi Y-J, Danielsen S, Lübeck M, Hong S-B, Delhey R, Shin H-D. Morphological and molecular characterization of the causal agent of downy mildew on quinoa (*Chenopodium quinoa*). Mycopathologia. 2010; 169:403–12.
- Danielsen S, Munk L. Evaluation of disease assessment methods in quinoa for their ability to predict yield loss caused by downy mildew. Crop Prot. 2004;23:219–28.
- Danielsen S, Jacobsen S-E, Hockenhull J. First Report of Downy Mildew of Quinoa Caused by *Peronospora farinosa* f.sp. *Chenopodii* in Denmark. Plant Dis. 2002;86:1175.
- Testen AL, McKerny JM, Backman PA. First report of quinoa downy mildew caused by *Peronospora variabilis* in the United States. Plant Dis. 2012;96:146.
- Choi YJ, Choi IY, Kim JS, Shin HD. First report of quinoa downy mildew caused by *Peronospora variabilis* in Republic of Korea. Plant Dis. 2014;98:1003.
- Khalifa W, Thabet M. Variation in downy mildew (*Peronospora variabilis* Gäum) resistance of some quinoa (*Chenopodium quinoa* Willd) cultivars under Egyptian conditions. Middle East J Agric Res. 2018;7:671–82.
- Danielsen S, Lübeck M. Universally primed-PCR indicates geographical variation of *Peronospora farinosa* ex. *Chenopodium quinoa* J Basic Microbiol. 2010;50:104–9.
- Kara M, Soyulu EM, Uysal A, Kurt S, Choi Y-J, Soyulu S. Morphological and molecular characterization of downy mildew disease caused by

- Peronospora variabilis* on *Chenopodium album* in Turkey. Aust Plant Dis Notes. 2020;15:10. <https://doi.org/10.1007/s13314-020-0381-2>.
16. Thines M, Choi Y-J. Evolution, diversity, and taxonomy of the peronosporaceae, with focus on the genus *peronospora*. Phytopathology®. 2016;106:6–18. <https://doi.org/10.1094/PHYTO-05-15-0127-RWW>.
 17. Verma S. C. and Chauhan, L. S. and Mathur, R. S. *Peronospora farinosa* (Fr.) Fr. On *Chenopodium murale* L.-a new record for India. Curr Sci. 1964;33:720–1.
 18. Aragón L, Gutiérrez W. Downy mildew on four *Chenopodium* species. Fitopatología. 1992;27:104–9.
 19. Baiswar P, Chandra S, Kumar R, Ngachan SV. *Peronospora variabilis* on *Chenopodium murale* in India. Aust Plant Dis Notes. 2010;5:45–7.
 20. Testen AL, Del M J-GM, Ochoa JB, Backman PA. Molecular detection of *Peronospora variabilis* in quinoa seed and phylogeny of the quinoa downy mildew pathogen in South America and the United States. Phytopathology. 2014;104:379–86.
 21. Ochoa J, Frinking HD, Jacobs T. Postulation of virulence groups and resistance factors in the quinoa/downy mildew pathosystem using material from Ecuador. Plant Pathol. 1999;48:425–30.
 22. Bonifacio A. *Chenopodium* sp.: Genetic resources, ethnobotany, and geographic distribution. Food Rev Int. 2003;19:1–7. <https://doi.org/10.1081/FRI-120018863>.
 23. Benhabib O, Boujartani N, Maughan PJ, Jacobsen SE, Jellen EN. Elevated genetic diversity in an F2:6 population of quinoa (*Chenopodium quinoa*) developed through an inter-ecotype cross. Front Plant Sci. 2016;7:1–9.
 24. Kitz L. Evaluation of Downy Mildew (*Peronospora farinosa* f.sp. *Chenopodii*) Resistance among Quinoa Genotypes and Investigation of *P. farinosa* Growth using Scanning Electron Microscopy. Thesis: Brigham Young University; 2008. <https://scholarsarchive.byu.edu/etd/1512>.
 25. Jacobsen S-E. The Worldwide Potential for Quinoa (*Chenopodium quinoa* Willd.). Food Rev Int. 2003;19:167–77.
 26. Tenorio R, Terrazas E, Alvarez MT, Vila JL, Mollinedo P. Concentrados de saponina de *Chenopodium quinoa* y de *Caiphora andina*: Alternativas como biocontroladores de hongos fitopatógenos. Rev Boliviana Quím. 2010;27:33–40.
 27. Harrison XA, Donaldson L, Correa-Cano ME, Evans J, Fisher DN, Goodwin CED, et al. A brief introduction to mixed effects modelling and multi-model inference in ecology. PeerJ. 2018;6:1–32.
 28. Milligan GW. The use of the arc-sine transformation in the analysis of variance. Educ Psychol Meas. 1987;47:563–73.
 29. O'Hara RB, Kotze DJ. Do not log-transform count data. Methods Ecol Evol. 2010;1:118–22.
 30. Warton DI, Hui FKC. The arcsine is asinine: the analysis of proportions in ecology. Ecology. 2011;92:3–10.
 31. Zimmerman DW. Two separate effects of variance heterogeneity on the validity and power of significance tests of location. Stat Methodol. 2006;3: 351–74.
 32. Jacquín-Gadda H, Sibillot S, Proust C, Molina J-M, Thiébaud R. Robustness of the linear mixed model to misspecified error distribution. Comput Stat Data Anal. 2007;51:5142–54.
 33. Piepho H-P, Möhring J. Computing heritability and selection response from unbalanced plant breeding trials. Genetics. 2007;177:1881–8.
 34. Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, et al. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol Evol. 2009;24:127–35.
 35. Bertero HD, La Vega AJ, Correa G, Jacobsen SE, Mujica A. Genotype and genotype-by-environment interaction effects for grain yield and grain size of quinoa (*Chenopodium quinoa* Willd.) as revealed by pattern analysis of international multi-environment trials. Field Crop Res. 2004; 89:299–318.
 36. Hayward MD, Romagosa I, Bosemark NO, Cerezo M. In: Hayward M, Bosemark NO, Romagosa I, editors. Plant breeding: Principles and Prospects: Springer Netherlands; 1993.
 37. Singh P, Bhatia D. Incomplete block designs for plant breeding experiments. Agric Res J. 2017;54:607.
 38. Sacristán S, García-arenal F. The evolution of virulence and pathogenicity in plant pathogen populations. Mol Plant Pathol. 2008;9:369–84.
 39. Barrett LG, Kniskern JM, Bodenhausen N, Zhang W, Bergelson J. Continuum of specificity and virulence in plant host–pathogen interactions: causes and consequences. New Phytol. 2009;183:513–29.
 40. Mhada M, Ezzahiri B, Benhabib O. Assessment of downy mildew resistance (*Peronospora farinosa*) in a quinoa (*Chenopodium quinoa* Willd.) germplasm. Int J Biol Med Res. 2015;6:4748–52.
 41. Pitrat M, editor. Cucurbitaceae 2008: proceedings of the IXth EUCARPIA meeting on genetics and breeding of cucurbitaceae, 21–24 may 2008, Avignon, France. Avignon: INRA; 2008.
 42. Divilov K, Barba P, Cadle-Davidson L, Reisch BI. Single and multiple phenotype QTL analyses of downy mildew resistance in interspecific grapevines. Theor Appl Genet. 2018;131:1133–43.
 43. Lohithaswa HC, Jyothi K, Sunil Kumar KR, Puttaramanaik HS. Identification and introgression of QTLs implicated in resistance to sorghum downy mildew (*Peronosclerospora sorghi* (Weston and Uppal) C. G. Shaw) in maize through marker-assisted selection. J Genet. 2015;94:741–8.
 44. Gardner KM, Latta RG. Heritable variation and genetic correlation of quantitative traits within and between ecotypes of *Avena barbata*. J Evol Biol. 2008;21:737–48.
 45. Santis G, D'Ambrosio T, Rinaldi M, Rascio A. Heritabilities of morphological and quality traits and interrelationships with yield in quinoa (*Chenopodium quinoa* Willd.) genotypes in the Mediterranean environment. J Cereal Sci. 2016;70:177–85. <https://doi.org/10.1016/j.jcs.2016.06.003>.
 46. Tapia M, Gandarillas H, Alandia S, Cardozo A, Mujica A. La quinua y la kañiwa: Cultivos Andinos. Bogotá: Centro Internacional de Investigaciones para el Desarrollo (CIID), Instituto Interamericano de Ciencias Agrícolas (IICA); 1979.
 47. Danielsen S, Bonifacio A, Ames T. Diseases of quinoa (*Chenopodium quinoa*). Food Rev Int. 2003;19:43–59.
 48. Danielsen S, Ames T. El mildiu (*peronospora farinosa*) de la quinua (*Chenopodium quinoa*) en la zona andina: Manual practico para el estudio de la enfermedad y el patogeno. Lima: Centro Internacional de la Papa (CIP); 2000. <http://cipotato.org/wp-content/uploads/2014/10/AN60198.pdf>.
 49. Gabriel J, Luna N, Vargas A, Magne J, Angulo A, La Torre J, et al. Quinoa de valle (*Chenopodium quinoa* Willd.): Fuente valiosa de resistencia genética al mildiu (*Peronospora farinosa* Willd.). J Selva Andina Res Soc. 2012;3:27–44.
 50. Al-Naggar A, El-Salam R, Badran A, El-Moghazi M. Heritability and Interrelationships for Agronomic, Physiological and Yield Traits of Quinoa (*Chenopodium quinoa* Willd.) under Elevated Water Stress. Arch Curr Res Int. 2017;10:1–5.
 51. Rollano-Peñaloza OM, Palma-Encinas V, Widell S, Rasmusson AG, Mollinedo P. The disease progression and molecular defense response in *Chenopodium quinoa* infected with *peronospora variabilis*, the causal agent of quinoa downy mildew. bioRxiv. 2019. <https://doi.org/10.1101/607465>.
 52. Bonifacio A. Quinoa breeding and modern variety development. In: Bazile D, Bertero D, Nieto C, editors. State of the art report on quinoa around the world in 2013. Rome: FAO regional office for Latin America; the Caribbean; 2015. p. 454–65.
 53. Lange OL, Lösch R, Schulze ED, Kappen L. Responses of stomata to changes in humidity. Planta. 1971;100:76–86.
 54. Danielsen S, Mercado VH, Ames T, Munk L. Seed transmission of downy mildew (*peronospora farinosa* f.sp. *Chenopodii*) in quinoa and effect of relative humidity on seedling infection. Seed Sci Technol. 2004;32:91–8.
 55. Solíz-Guerrero JB, De Rodríguez DJ, Rodríguez-García R, Angulo-Sánchez JL, Méndez-Padilla G. Quinoa saponins: Concentration and composition analysis. In: Janick J, Whipkey A, others, editors. Trends in new crops and new uses. Alexandria: ASHS Press; 2002. p. 110–114.
 56. Martínez EA, Veas E, Jorquera C, San Martín R, Jara P. Re-introduction of quinoa into arid Chile: cultivation of two lowland races under extremely low irrigation. J Agron Crop Sci. 2009;195:1–10.
 57. Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods. 2013;9:29.
 58. Gandarillas A, Saravia R, Plata G, Quispe R, Ortiz-Romero R. Principal Quinoa Pests and Diseases. In: Bazile D, Bertero D, Nieto C, editors. State of the art report on quinoa around the world in 2013. Rome, FAO regional office for Latin America; the Caribbean; 2015. p. 192–215. <http://www.fao.org/3/a-i4042e.pdf>.
 59. Kumar A, Bhargava A, Shukla S, Singh HB, Ohri D. Screening of exotic *Chenopodium quinoa* accessions for downy mildew resistance under mid-eastern conditions of India. Crop Prot. 2006;25:879–89.
 60. McElhinny E, Mazón N, Rivera MM, Peralta IE. Líneas promisorias de quinua con resistencia cuantitativa al mildiu en Ecuador. In: Danial DL, editor. Agro-biodiversidad y producción de semilla con el sector informal a través del mejoramiento participativo en la Zona Andina. PREDUZA; 2003. p. 40–47.
 61. Curti RN, La Vega AJ, Andrade AJ, Bramardi SJ, Bertero HD. Multi-environmental evaluation for grain yield and its physiological determinants of quinoa genotypes across Northwest Argentina. Field Crop Res. 2014;166:46–57.

62. Poland JA, Balint-Kurti PJ, Wisser RJ, Pratt RC, Nelson RJ. Shades of gray: the world of quantitative disease resistance. *Trends Plant Sci.* 2009;14:21–9.
63. Mastebroek HD, Loo R v. Breeding of quinoa—state of the art. In: Parente G, Frame J, editors. Abstracts/proceedings of COST 814 conference, crop development for cool and wet regions of Europe. Office of Official Publications of the European Communities; 2000. p. 491–6.
64. Calixtro M, Gómez-Pando L, Ibañez M. Evaluación de la resistencia de quinua al mildiú (*Peronospora variabilis*) y su transferencia por semillas en condiciones del valle del mantaro, junín—perú. In: Resúmenes de exposiciones del VI congreso mundial de la quinua y III simposio internacional de granos andinos, Perú 2017; 2017. p. 29.
65. Conrath U. Molecular aspects of defence priming. *Trends Plant Sci.* 2011;16:524–31.
66. Sperschneider J. Machine learning in plant–pathogen interactions: empowering biological predictions from field scale to genome scale. *New Phytol.* 2019;228:35.
67. Stikić R, Glamoclija D, Demin M, Vucelic-Radovic B, Jovanovic Z, Milojkovic-Opšenica D, et al. Agronomical and nutritional evaluation of quinoa seeds (*Chenopodium quinoa* Willd.) as an ingredient in bread formulations. *J Cereal Sci.* 2012;55:132–8. <https://doi.org/10.1016/j.jcs.2011.10.010>.
68. Sun Y, Liu F, Bendevis M, Shabala S, Jacobsen S-E. Sensitivity of two quinoa (*Chenopodium quinoa* Willd.) varieties to progressive drought stress. *J Agron Crop Sci.* 2014;200:12–23. <https://doi.org/10.1111/jac.12042>.
69. Lewis CM, Persoons A, Bebbler DP, Kigathi RN, Maintz J, Findlay K, et al. Potential for re-emergence of wheat stem rust in the United Kingdom. *Commun Biol.* 2018;1. <https://doi.org/10.1038/s42003-018-0013-y>.
70. Silva M, Rijo L, Rodrigues C Jr. Differences in aggressiveness of two isolates of race III of *Hemileia vastatrix* on cultivar caturra of *Coffea arabica*. In: Proceedings of the 11th international scientific colloquium on coffee. Lomé: ASIC; 1985. p. 635–44.
71. Scarpeci TE, Zanon MI, Valle EM. Estimation of Stomatal aperture in *Arabidopsis thaliana* using silicone rubber imprints. *Bio Protoc.* 2017;7:e2347.
72. Fetter KC, Eberhardt S, Barclay RS, Wing S, Keller SR. StomataCounter: a neural network for automatic stomata identification and counting. *New Phytol.* 2019;223:1671–81.
73. Butler DG, Cullis BR, Gilmour AR, Gogel BJ. ASReml-R reference manual (version 3). Brisbane, Qld: The State of Queensland, Department of Primary Industries and Fisheries; 2009.
74. Piepho HP, Buchse A, Emrich K. A Hitchhiker's guide to mixed models for randomized experiments. *J Agron Crop Sci.* 2003;189:310–22.
75. Utz HF. PLABSTAT: a computer program for statistical analysis of plant breeding experiments. Stuttgart: Institute for Plant Breeding, Seed Science and Population Genetics, University of Hohenheim; 2001.
76. Bernal-Vasquez A-M, Utz H-F, Piepho H-P. Outlier detection methods for generalized lattices: a case study on the transition from ANOVA to REML. *Theor Appl Genet.* 2016;129:787–804.
77. Isik F, Holland J, Maltecca C. Chapter 3: variance modeling in ASReml. In: Genetic data analysis for plant and animal breeding. Cham: Springer; 2017. p. 87–106.
78. Stroup WW. Generalized linear mixed models: modern concepts, methods and applications. Boca Raton: CRC Press; 2013.
79. Brien C. asremlPlus: Augments the use of ASReml-R in fitting mixed models; 2019.
80. Piepho HP, Williams ER, Fleck M. A note on the analysis of designed experiments with complex treatment structure. *HortScience.* 2006;41:446–52.
81. Harrel FE, Dupont C. Hmisc: Harrel Miscellaneous; 2019.
82. Koziol MJ. Afrosimetric estimation of threshold saponin concentration for bitterness in quinoa (*Chenopodium quinoa* Willd.). *J Sci Food Agric.* 1991;54:211–9.
83. Baym SAL Michael AND Kryazhimskiy Inexpensive Multiplexed Library Preparation for Megabase-Sized Genomes *PLoS One* 2015;10:1–5. doi: <https://doi.org/10.1371/journal.pone.0128036>.
84. Köster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics.* 2012;28:2520–2.
85. Krueger F. Trim galore! 2015. <https://github.com/FelixKrueger/TrimGalore>.
86. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009;25:2078–9.
87. Broad Institute. Picard tools. Broad institute, GitHub repository Accessed: 2018/02/21; version 2.17.8. <http://broadinstitute.github.io/picard/>.
88. Jarvis DE, Ho YS, Lightfoot DJ, Schmöckel SM, Li B, Borm TJA, et al. The genome of *Chenopodium quinoa*. *Nature.* 2017;542:307.
89. Maughan PJ, Chaney L, Lightfoot DJ, Cox BJ, Tester M, Jellen EN, et al. Mitochondrial and chloroplast genomes provide insights into the evolutionary origins of quinoa (*Chenopodium quinoa* Willd.). *Sci Rep.* 2019;9:1–1.
90. Li H, Durbin R. Fast and accurate long-read alignment with burrows–wheeler transform. *Bioinformatics.* 2010;26:589–95. <https://doi.org/10.1093/bioinformatics/btp698>.
91. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20:1297–303.
92. Money D, Migicovsky Z, Gardner K, Myles S. LinkImputeR: user-guided genotype calling and imputation for non-model organisms. *BMC Genomics.* 2017;18:523.
93. Maruki T, Lynch M. Genotype-frequency estimation from high-throughput sequencing data. *Genetics.* 2015;201:473–86.
94. Zhang C, Dong S-S, Xu J-Y, He W-M, Yang T-L. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics.* 2018;35:1786–8.
95. Liu X, Huang M, Fan B, Buckler ES, Zhang Z. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* 2016;12:1–24.
96. Malik PL, Janss L, Nielsen LK, Borum F, Jørgensen H, Eriksen B, et al. Breeding for dual-purpose wheat varieties using marker–trait associations for biomass yield and quality traits. *Theor Appl Genet.* 2019;132:3375–98.
97. Voicheck Y, Weigel D. Identifying genetic variants underlying phenotypic variation in plants without complete genomes. *Nat Genet.* 2020;52:534–40.
98. Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, et al. Efficient control of population structure in model organism association mapping. *Genetics.* 2008;178:1709–23.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

