



OPEN

Genomic regions associated with herbicide tolerance in a worldwide faba bean (*Vicia faba* L.) collection

Lynn Abou-Khater^{1✉}, Fouad Maalouf^{1✉}, Abdulqader Jighly², Alsamman M. Alsamman^{3,4}, Diego Rubiales⁵, Nicolas Rispaïl⁵, Jinguo Hu⁶, Yu Ma⁷, Rind Balech¹, Aladdin Hamwiah⁸, Michael Baum⁹ & Shiv Kumar⁹

Weeds represent one of the major constraints for faba bean crop. The identification of molecular markers associated with key genes imparting tolerance to herbicides can facilitate and fasten the efficient and effective development of herbicide tolerant cultivars. We phenotyped 140 faba bean genotypes in three open field experiments at two locations in Lebanon and Morocco against three herbicide treatments (T1 metribuzin 250 g ai/ha; T2 imazethapyr 75 g ai/ha; T3 untreated) and one in greenhouse where T1 and T3 were applied. The same set was genotyped using genotyping by sequencing (GBS) which yield 10,794 high quality single nucleotide polymorphisms (SNPs). ADMIXTURE software was used to infer the population structure which revealed two ancestral subpopulations. To identify SNPs associated with phenological and yield related traits under herbicide treatments, Single-trait (ST) and Multi-trait (MT) Genome Wide Association Studies (GWAS) were fitted using GEMMA software, showing 10 and 14 highly significant associations, respectively. Genomic sequences containing herbicide tolerance associated SNPs were aligned against the NCBI database using BLASTX tool using default parameters to annotate candidate genes underlying the causal variants. SNPs from acidic endochitinase, LRR receptor-like serine/threonine-protein kinase RCH1, probable serine/threonine-protein kinase NAK, malate dehydrogenase, photosystem I core protein PsaA and MYB-related protein P-like were significantly associated with herbicide tolerance traits.

Faba bean (*Vicia faba* L.), also known as broad bean, fava bean, horse bean and field bean, was first domesticated in the Near East around 9000–10,000 BC^{1–3}. The recent estimates suggest that it is extensively grown on 2.57 M ha area distributed across 38 countries with global production of 5.4 million tonnes⁴. Faba bean is an important source of food and feed for human and animal consumption because its seeds are rich in proteins, carbohydrates, fibers and micronutrients⁵. Faba bean plays an important role in sustainable agriculture and ecosystem services because of its ability to improve soil fertility by fixing atmospheric nitrogen⁶, and its potential to enhance the grain yield of succeeding/companion crops when planted in rotation or intercropped with cereals⁷. Concerted efforts have been undertaken to improve yield, adaptation to different environments, tolerance to abiotic stresses including heat, drought, waterlogging and frost^{8–10} resistance to biotic stresses such as diseases, insect pests, viruses and parasitic weeds^{11,12}, seed quality¹³ and other agronomic traits. These efforts have more than doubled the global average yield from 0.9 tonnes/ha in 1964 to 2.1 tonnes/ha in 2019⁴. However, the current production remains insufficient to meet its global consumption. Faba bean performance is highly influenced by environments and genotype × environment (GE) interaction, making phenotypic selection for quantitative traits of breeders' interest inefficient and cumbersome.

¹International Center for Agricultural Research in the Dry Areas (ICARDA), Terbol, Lebanon. ²Agriculture Victoria, Victoria, Australia. ³Molecular Genetics and Genome Mapping, Agricultural Genetic Engineering Research Institute, Giza, Egypt. ⁴African Genome Center, Mohammed VI Polytechnic University, Ben Guerir, Morocco. ⁵Institute for Sustainable Agriculture, CSIC, 14004 Córdoba, Spain. ⁶USDA-ARS Plant Germplasm Introduction & Testing Research Unit, Pullman, USA. ⁷Department of Horticulture, Washington State University, Pullman, USA. ⁸ICARDA, Rabat, Morocco. ⁹ICARDA, Cairo, Egypt. ✉email: lynnaboukhater@hotmail.com; f.maalouf@cgiar.org

Faba bean has a relatively large genome size of 13 Gb¹⁴. Thanks to the advances in the next generation sequencing technologies (NGS) that has enabled the generation of large volumes of sequences^{15–17} and facilitated the discovery of single nucleotide polymorphisms (SNPs) that can be associated with key breeding traits either through biparental mapping or through genome wide association studies (GWAS)^{15,18}. Unlike biparental mapping, GWAS utilizes natural populations and exploits linkage disequilibrium (LD) to detect SNP-trait associations with higher resolution¹⁹. However, the power of GWAS depends on the size and structure of the population used for the analysis²⁰. While it is sometimes not feasible to phenotype large populations in a single field trial, multiple field trials, e.g. different treatments, locations or seasons, can be jointly analyzed in one model named as multi-variate or multi-trait GWAS which has shown to have higher power compared to the standard single-trait GWAS²¹. Such approach can assist conventional breeding by implementing marker assisted selections in early generations^{18,22}. Although significant progress has been made in faba bean genomics and many genetic maps are available^{23–25}, the marker density of most of them is still too low to enable accurate prediction of desired traits. SNPs correlated with traits of interest such as resistance to ascochyta and broomrape or vicine-convicine content^{26–29} have been identified, however, no study was conducted to associate SNPs with herbicide tolerance in faba bean.

Weeds are among the difficult-to-control biotic stresses that affect faba bean³⁰. When weeds are left uncontrolled, they cause severe loss on grain yield of up to 70%³¹. An integrated approach with many control measures has been recommended to provide protection against weeds^{32–34} but with limited success. Many studies have acknowledged breeding for weed resistance by selecting for morphological characteristics that promote competition and allelopathy such as early seedling emergence, seedling growth, greater plant height, greater root volume^{35–37}, but the resistance against most parasitic weeds is a difficult task because of its complex nature and low heritability^{38,39}. Thus, recent studies have focused on developing herbicide tolerant faba bean lines^{40,41}. Abou-Khater et al.⁴² evaluated faba bean germplasm for traits associated with tolerance to metribuzin and imazethapyr, two herbicides commonly available that can control the majority of weeds threatening faba bean production. They found that crop phenology, plant architecture and grain yield related traits were greatly affected by the herbicide treatments. Although useful sources for herbicide tolerance were identified by the authors, such field techniques are very laborious and require multi-environmental data. Associating the herbicide tolerance related traits⁴² with molecular markers to select for herbicide tolerance would facilitate the detection of useful markers that can be used to select herbicide tolerant lines in early generations. Keeping this in mind, the present study was undertaken to identify candidate loci significantly associated with tolerance to two post emergence herbicides, namely metribuzin and imazethapyr under different environments using GWAS and to identify associated SNP markers that can be used for introgressing such traits into desired agronomic background.

Results

Phenotyping. Multiple environmental models were fitted to obtain the best linear unbiased prediction (BLUP) values for each genotype and treatment across field trials. The genotypic effects for all studied traits and reduction indexes were significant across trials at a p -value < 0.001 except for the RI_{GCC} (Table 1) indicating a wide range of genotypic variation in faba bean. Significant differences were observed among treatments for all studied traits and reduction indexes except for RI_{PLHT} , RI_{GYPLT} and RI_{NPPLT} ; while significant Genotype \times Treatment interactions were observed across trials for DFLR, DMAT, PLHT, GYPLT and NSPLT (Table 1). The Genotype \times Treatment \times Environment interactions show that the effect of herbicide treatments on the traits and reduction indexes of the genotypes was not affected by the environment except for DFLR and NSPLT and their reduction indexes (Table 1). As for the greenhouse experiment, the DFLR, PLHT and GCC varied significantly among genotypes and treatments and significant Genotype \times Treatment interactions were observed. The reduction indexes for DFLR, PLHT and GCC varied significantly also among genotypes (Table 2). Our results showed that both herbicide treatments affected the faba bean phenology by delaying significantly the DFLR and DMAT (Tables 1, 2). In addition, the post emergence application of metribuzin and imazethapyr affected the architecture of the faba bean plants by reducing the PLHT and the GCC and increasing the NbrPLT (Tables 1, 2). Moreover, a significant reduction in the GYPLT, NPPLT and NSPLT of the genotypes treated with metribuzin or imazethapyr was observed across trials. The plant height recorded in the green house experiment at two different stages showed that the treated plants tend to recover from the herbicide effect (Tables 1, 2).

Our results also showed that the first (HDS1) and second (HDS2) herbicide damage scores per genotype varied from 1 to 5 across trials. Combined results of the herbicide damage scores (HDS1 and HDS2) showed that after one month of the herbicide application, 5 and 42% genotypes recovered from the damaged caused by metribuzin and imazethapyr treatments while damages in 56 and 10% genotypes exacerbated (Fig. 1). The herbicide damage on the remaining genotypes remained unchanged between the first and second recording dates.

Figure 2 shows that the herbicide treatments affected differently the plant height and grain yield of the treated genotypes. The reduction in plant height and grain reduction varied between almost negligible ($< 10\%$) and high levels ($> 40\%$).

Genotyping and population structure. The SNP calling analysis revealed 10,794 high-quality SNPs among the studied faba genotypes. The sequence variations of these SNPs were C/T (4251 SNPs), and A/G (4029 SNPs), followed by A/T (836 SNPs), G/T (761 SNPs), A/C (619 SNPs), and C/G (298 SNPs). The average CV values for the 100 replicates of the population structure started to increase directly after $K=2$ indicating the presence of two ancestral subpopulations in the germplasm set used in the present study. However, we presented the results of K up to 4 because their 100 replicate runs resulted in comparable classification of genotypes into ancestral subpopulations with top $> 20\%$ of replicates having almost the exact log-likelihood values (Fig. 3).

| | | p-value | Metribuzin 250 g ai/ha | | Imazethapyr 75 g ai/ha | | Control | | SE |
|------------|------------------|---------|------------------------|------------------|------------------------|------------------|---------|---------------|-------|
| | | | Mean | Range | Mean | Range | Mean | Range | |
| DFLR (DAS) | Geno | <0.001 | 99.51 | 90.96–121.39 | 99.44 | 90.48–121.61 | 98.10 | 90.60–117.94 | 2.57 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | <0.001 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | <0.001 | | | | | | | |
| DFLR_RI | Geno | <0.001 | – 2.37 | – 10.22–3.57 | – 2.31 | – 11.90 to 3.95 | | | 2.71 |
| | Trt | 0.749 | | | | | | | |
| | Geno × Trt | 1 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | <0.001 | | | | | | | |
| DMAT (DAS) | Geno | <0.001 | 162.92 | 159.20–167.00 | 161.88 | 158.30–167.00 | 161.05 | 158.00–166.80 | 1.17 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | <0.001 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | 0.036 | | | | | | | |
| DMAT_RI | Geno | <0.001 | – 1.63 | – 4.23 to 0.13 | – 1.07 | – 3.14 to 0.13 | | | 0.86 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | 0.22 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | 0.970 | | | | | | | |
| PLHT (cm) | Geno | <0.001 | 61.61 | 37.28–83.24 | 60.73 | 32.40–83.09 | 73.44 | 36.29–104.09 | 6.56 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | <0.001 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | 0.002 | | | | | | | |
| PLHT_RI | Geno | <0.001 | 13.93 | – 41.62–37.27 | 17.18 | – 1.16 to 43.90 | | | 11.36 |
| | Trt | 0.002 | | | | | | | |
| | Geno × Trt | 0.997 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | 0.809 | | | | | | | |
| GYPLT (g) | Geno | <0.001 | 14.68 | – 2.37 to 36.82 | 14.55 | – 1.73 to 40.57 | 20.37 | 1.25–41.94 | 5.59 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | <0.001 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | 0.190 | | | | | | | |
| GYPLT_RI | Geno | <0.001 | 15.00 | – 38.59 to 58.94 | 20.78 | – 50.26 to 65.45 | | | 25.78 |
| | Trt | 0.105 | | | | | | | |
| | Geno × Trt | 0.999 | | | | | | | |
| | Geno × Env | 0.103 | | | | | | | |
| | Geno × Trt × Env | 0.787 | | | | | | | |
| NPPLT | Geno | <0.001 | 17.41 | 6.56–36.72 | 16.99 | 8.65–29.18 | 19.43 | 7.62–44.13 | 4.76 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | 0.587 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | 0.437 | | | | | | | |
| Continued | | | | | | | | | |

| | | p-value | Metribuzin 250 g ai/ha | | Imazethapyr 75 g ai/ha | | Control | | SE |
|-----------|------------------|---------|------------------------|-------------------|------------------------|-------------------|---------|------------|-------|
| | | | Mean | Range | Mean | Range | Mean | Range | |
| NPPLT_RI | Geno | <0.001 | 4.04 | - 93.79 to 78.52 | 5.22 | - 82.43 to 49.77 | | | 32.59 |
| | Trt | 0.393 | | | | | | | |
| | Geno × Trt | 1 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | 1 | | | | | | | |
| NSPLT | Geno | <0.001 | 20.38 | 1.54–34.68 | 21.94 | 2.28–44.85 | 28.82 | 2.36–63.65 | 6.91 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | <0.001 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | <0.001 | | | | | | | |
| NSPLT_RI | Geno | 0.016 | 24.38 | - 33.02 to 75.63 | 14.1 | - 53.60 to 60.39 | | | 31.82 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | 1 | | | | | | | |
| | Geno × Env | <0.001 | | | | | | | |
| | Geno × Trt × Env | 0.007 | | | | | | | |
| NBrPLT | Geno | <0.001 | 4.16 | 0.68–8.25 | 3.31 | 0.98–6.12 | 3.14 | 0.77–7.12 | 1.46 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | 0.931 | | | | | | | |
| | Geno × Env | ND | | | | | | | |
| | Geno × Trt × Env | ND | | | | | | | |
| NBrPLT_RI | Geno | <0.001 | - 50.49 | - 308.98 to 53.28 | - 18.43 | - 307.77 to 49.14 | | | 78.70 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | 1 | | | | | | | |
| | Geno × Env | 1 | | | | | | | |
| | Geno × Trt × Env | ND | | | | | | | |
| GCC | Geno | <0.001 | 29.35 | - 0.89 to 59.23 | 25.18 | 1.71–53.46 | 34.47 | 3.12–67.22 | 11.93 |
| | Trt | <0.001 | | | | | | | |
| | Geno × Trt | 0.995 | | | | | | | |
| | Geno × Env | ND | | | | | | | |
| | Geno × Trt × Env | ND | | | | | | | |
| GCC_RI | Geno | 0.094 | 5.79 | - 180.80 to 93.46 | 16.79 | - 376.14 to 87.11 | | | 60.34 |
| | Trt | 0.027 | | | | | | | |
| | Geno × Trt | 0.954 | | | | | | | |
| | Geno × Env | ND | | | | | | | |
| | Geno × Trt × Env | ND | | | | | | | |

Table 1. Combined analysis performed for detecting differences among faba bean genotypes (Geno), herbicide treatments (Trt), Geno × Trt interaction, Genotype × Environment (Geno × Env) interaction and Geno × Trt × Env interaction expressed as p-value and means ± standard error (SE) and ranges of the genotypes under trials. *DFLR* days to flowering, *DAS* days after sowing, *DFLR_RI* DFLR reduction index, *DMAT* days to maturity, *DMAT_RI* DMAT reduction index, *PLHT* plant height, cm centimeter, *PLHT_RI* reduction index of PLHT, *GYPLT* grain yield per plant, g gram, *GYPLT_RI* GYPLT reduction index, *NPPLT* number of pods per plant, *NPPLT_RI* NPPLT reduction index, *NSPLT* number of seeds per plant, *NSPLT_RI* NSPLT reduction index, *NBrPLT* number of branches per plant, *NBrPLT_RI* NBrPLT reduction index, *GCC* green canopy cover, *GCC_RI* green canopy cover reduction index, *ND* no data.

Beyond $K=4$, the analysis started to output arbitrary results with inconsistent classifications even for the top 10 replicates with the highest log-likelihood values.

GWAS and annotation analyses. As we ended with a total of 10,794 high-quality SNPs after filtration, the Bonferroni significant threshold can be calculated as $(0.05/10,794=4.6E-6)$. Analyzing all 103 traits (including RI scores) with the ST-GWAS model resulted in only 10 highly significant associations with the

| | | p-value | Metribuzin @250 g ai/ha | | Control | | SE |
|-----------|------------|---------|-------------------------|--------------------|---------|-------------|-------|
| | | | Mean | Range | Mean | Range | |
| DFLR | Geno | <0.001 | 56.65 | 38.93–62.21 | 49.04 | 34.94–61.80 | 7.12 |
| | Trt | <0.001 | | | | | |
| | Geno × Trt | <0.001 | | | | | |
| DFLR_RI | Geno | <0.001 | 2.27 | – 65.37 to 49.79 | | | 18.26 |
| GCC | Geno | <0.001 | 5.09 | – 0.27 to 10.47 | 6.95 | 2.37–10.74 | 1.78 |
| | Trt | <0.001 | | | | | |
| | Geno × Trt | <0.001 | | | | | |
| GCC_RI | Geno | <0.001 | 18.62 | – 237.69 to 104.02 | | | 40.66 |
| PLHT_1 | Geno | <0.001 | 15.12 | 2.63–26.27 | 25.29 | 11.55–42.65 | 4.70 |
| | Trt | <0.001 | | | | | |
| | Geno × Trt | <0.001 | | | | | |
| PLHT_RI_1 | Geno | <0.001 | 28.17 | – 73.02 to 89.70 | | | 19.2 |
| PLHT_2 | Geno | <0.001 | 21.23 | 5.16–35.67 | 35.90 | 17.97–58.62 | 7.26 |
| | Trt | <0.001 | | | | | |
| | Geno × Trt | 0.004 | | | | | |
| PLHT_RI_2 | Geno | <0.001 | 53.78 | – 34.25 to 100.00 | | | 26.05 |

Table 2. Analysis of Variance performed for detecting differences among faba bean genotypes (Geno), herbicide treatments (Trt), and Geno × Trt interaction for different traits and reduction indexes, expressed as p- value and means ± standard error (SE) and ranges of the genotypes under different treatments in the pot trial. *DFLR* days to flowering, *DFLR_RI* DFLR reduction index, *GCC* green canopy cover, *GCC_RI* green canopy cover reduction index, *PLHT_1* plant height recorded at flowering (BBCH code 60), *PLHT_1_RI* reduction index of *PLHT_1*, *PLHT_2* plant height recorded at pod development (BBCH code 70), *PLHT_2_RI* *PLHT_2* reduction index.

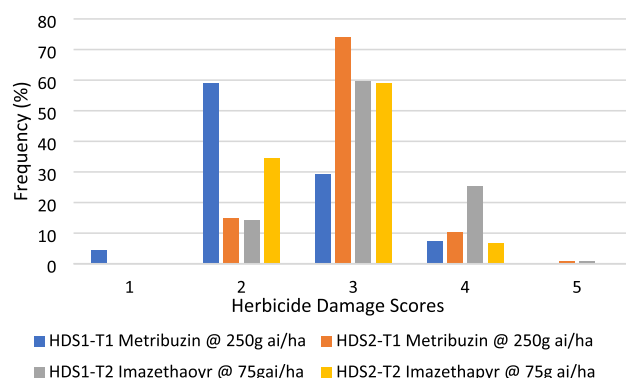


Figure 1. Distribution of faba bean genotypes for herbicide damage scores (HDS1 and HDS2) under metribuzin at 250 g ai/ha and imazethapyr at 75 g ai/ ha.

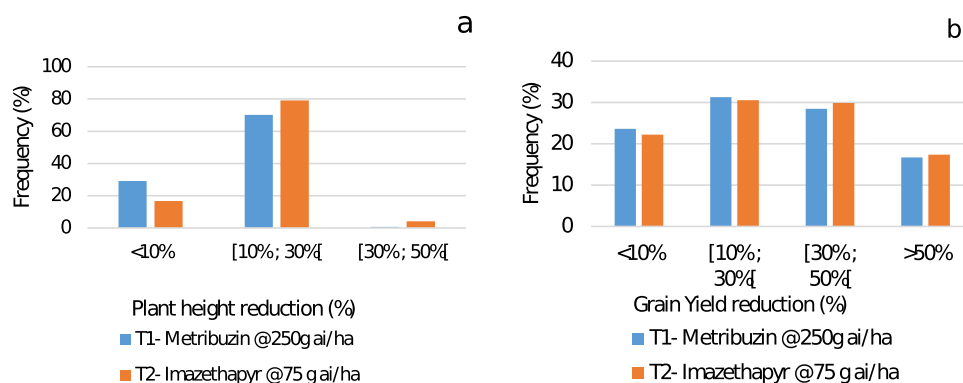


Figure 2. Distribution of faba bean genotypes for plant height reduction *PLHT_RI* (a) and grain yield per plant reduction *GYPLT_RI* (b) under metribuzin at 250 g ai/ha and imazethapyr at 75 g ai/ ha.

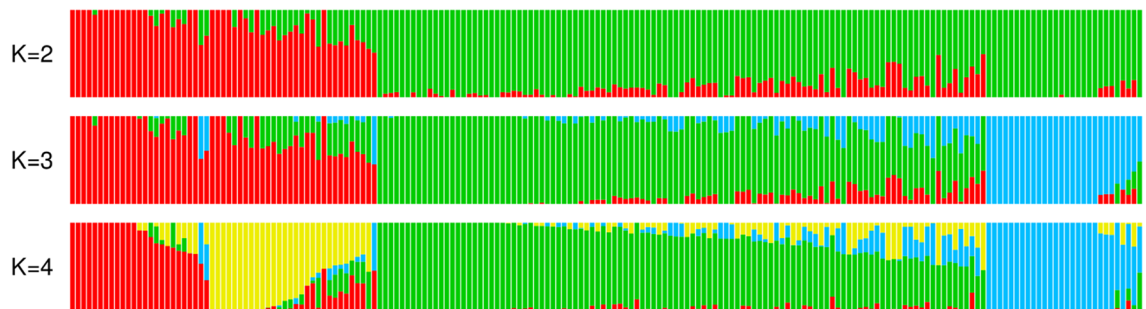


Figure 3. Population structure constructed using the SNPs data for the individual ancestry estimated using the ADMIXTURE analysis. Individuals are represented in thin vertical lines separated into segments corresponding to the assumed membership in $K = 2, 3$ and 4 genetic groups as shown by colors. Each color represents one ancestral subpopulation.

Bonferroni threshold as well as 110 suggestive associations for only 66 traits while the remaining traits had no associated SNPs (Supplementary Table S1). These associations were represented by 105 SNPs. Only one SNP (*SNODE_27970_52*) for DFLR was associated with three treatments (I, M and C) in TR16, while another five SNPs were associated with two scores for PLHT or DFLR (Supplementary Table S1). Of these, two SNPs (*SNODE_168698_34* for PLHT with treatment I, and *SNODE_23759_68* for DFLR with treatment M in TR16) were associated with a specific trait and its correspondence RI score. Another five SNPs showed association with two different traits of which *SNODE_3696_16* and *SNODE_77186_51* were associated with GYPPLT in TR16, treatments M and I respectively, while *SNODE_22383_32* was associated with RI in TR16, treatment M, for the same traits. *SNODE_239220_75* was associated with TR19_DMAT_I and TR16_NPPT_I_RI, while *SNODE_7114_58* showed associations with five DFLR and DMAT across environments/treatments (Supplementary Table S1).

The MT-GWAS model for 20 traits across environments (including RI scores) resulted in 14 highly significant associations and 64 suggestive associations for all traits represented by 72 SNPs (Table 3). The largest number of associations (12) were detected of DMAT, while GYPPLT_RI, NPPT_RI and Score1 had the lowest number with only one association each. Most of the SNPs that showed associations with multiple traits/treatments in the ST-GWAS analysis were also detected in MT-GWAS analysis. Four SNPs showed associations with a specific trait with its reduction index which were *SNODE_23759_68* for DFLR, *SNODE_14558_21* for NSPP, *SCONTIG73439_18* for PLHT, and *SNODE_103_72* for NBBR (Table 3). The SNP *SCONTIG127798_41* was associated with GCC and DFLR while the SNP *SNODE_22383_32* was associated with GYPPLT (Table 3).

Gene annotation showed that SNP *SCONTIG127798_41* associated with reduction index of GCC and DFLR is located within a gene annotated as acidic endo-chitinase annotation, *SNODE_14298_44* associated with the reduction index of PLHT is located within a gene annotated as LRR receptor-like serine/threonine-protein kinase RCH1, *SNODE_3696_16* associated with GYPPLT is located within a gene annotated as Probable serine/threonine-protein kinase NAK, *SNODE_4187_38* associated with the reduction index of DFLR is located within a gene annotated as malate dehydrogenase, *SNODE_559376_60* associated with the reduction index of NPPLT is located within a gene annotated as photosystem I core protein PsaA, while *SNODE_7114_58* associated with DFLR is located within a gene annotated as MYB-related protein P-like (Supplementary Table S2).

Discussion

Weed menace is a serious threat to faba bean production, and the identification of herbicide-tolerant varieties is one of the most effective methods for weed control. The results obtained from the present field and greenhouse studies demonstrated how the post-emergence application of metribuzin or imazethapyr negatively affects faba bean plants. Herbicide application affected the crop phenology by delaying flowering and maturity. Although the delayed flowering helps plant escape the risk of frost in regions like Western Australia, there might be a potential yield penalty as the plants run out of moisture before it can fill its grain^{43,44}. In addition, herbicide application also affected biological and grain yields of faba bean by reducing plant height, green canopy cover, and grain yield components and by increasing the number of branches. Many studies^{30,42,45–48} reported significant reduction in plant height, grain yield and yield components while studying the effect of post-emergence herbicide application on faba bean, lentil and chickpea. On the other hand, Wall⁴⁹ and Sajja et al.⁵⁰, reported an increase in the number of branches of treated plants. The observed damage after metribuzin and imazethapyr treatments is the consequence of the growth inhibition caused by both herbicides. Metribuzin hampers photosynthesis activity by inhibiting the photosynthetic electron flow^{51,52} and imazethapyr inhibits acetolactate synthase (ALS)⁵³, the first common enzyme in the biosynthesis of the branched-chain amino acids⁵⁴ causing the death of meristematic cells. On the other hand, significant increase in the number of branches in herbicide treated plants could be caused by the plant recovery which occurs at the lateral meristem in dicots resulting in the development of new branches. The genotypic variation observed in the herbicide damage scores (HDS) highlights the difference in the reaction of each genotype toward post emergence herbicide application in faba bean. This observation was expected as the evaluated genotypes are genetically diverse²². The differences observed between the first (HDS1) and the second (HDS2) scores were due to the recovery or deterioration of the plants one month after herbicide treatment. The recovery might result from the metabolism of the herbicides into inactive compounds⁵⁵. Therefore,

| Trait | SNP | allele1 | allele0 | MAF | P |
|-----------|---|---------|---------|------|---------|
| DFLR | SNODE_7114_58 | C | T | 0.14 | 3.0E-07 |
| DFLR | SNODE_162178_22 | A | G | 0.07 | 9.8E-07 |
| DFLR | SNODE_27970_52 | C | G | 0.17 | 3.6E-05 |
| DFLR | SNODE_23759_68 | A | G | 0.11 | 7.1E-05 |
| DFLR_RI | SNODE_23759_68 | A | G | 0.10 | 8.6E-07 |
| DFLR_RI | SNODE_5725_31 | C | T | 0.41 | 3.2E-06 |
| DFLR_RI | SNODE_4187_38 | G | A | 0.05 | 3.3E-06 |
| DFLR_RI | SNODE_26501_64 | G | C | 0.13 | 9.5E-06 |
| DFLR_RI | SCONTIG127798_41 | C | T | 0.07 | 2.9E-05 |
| DFLR_RI | SNODE_1051_18 | C | T | 0.08 | 4.3E-05 |
| DMAT | SCONTIG72526_35 | A | G | 0.05 | 4.6E-06 |
| DMAT | SNODE_375879_34 | A | T | 0.08 | 5.7E-06 |
| DMAT | SCONTIG93616_28 | C | T | 0.30 | 1.3E-05 |
| DMAT | SCONTIG6418_84 | C | T | 0.22 | 1.4E-05 |
| DMAT | SNODE_61301_40 | T | C | 0.06 | 1.5E-05 |
| DMAT | SNODE_80758_20 | C | T | 0.06 | 3.3E-05 |
| DMAT | SNODE_483217_44 | G | A | 0.12 | 4.6E-05 |
| DMAT | SNODE_143506_34 | G | A | 0.05 | 6.9E-05 |
| DMAT | SCONTIG125372_89 | T | C | 0.05 | 8.3E-05 |
| DMAT | SNODE_6229_36 | C | T | 0.11 | 8.4E-05 |
| DMAT | SCONTIG79953_82 | T | C | 0.37 | 8.6E-05 |
| DMAT | SNODE_16244_35 | T | C | 0.33 | 8.8E-05 |
| DMAT_RI | SNODE_76542_45 | T | G | 0.06 | 5.1E-06 |
| DMAT_RI | SNODE_13235_37 | C | A | 0.25 | 2.5E-05 |
| GCC | SNODE_11304_24 | A | T | 0.31 | 3.5E-06 |
| GCC | SCONTIG24931_19 | G | A | 0.06 | 4.4E-05 |
| GCC | SCONTIG66488_16 | G | A | 0.07 | 7.8E-05 |
| GCC_RI | SCONTIG127798_41 | C | T | 0.07 | 5.3E-08 |
| GCC_RI | SCONTIG75553_52 | T | C | 0.34 | 8.3E-05 |
| GCC_RI | SNODE_12134_67 | T | G | 0.07 | 9.3E-05 |
| GYPLT | SNODE_77186_51 | T | A | 0.10 | 1.3E-05 |
| GYPLT | SNODE_3696_16 | G | A | 0.11 | 2.5E-05 |
| GYPLT | SNODE_54972_30 | A | C | 0.14 | 1.3E-07 |
| GYPLT | SCONTIG46666_46 | A | T | 0.06 | 2.6E-06 |
| GYPLT | SNODE_167460_49 | C | T | 0.20 | 6.7E-06 |
| GYPLT | SCONTIG90061_39 | A | C | 0.06 | 8.1E-06 |
| GYPLT | SNODE_5674_14 | G | A | 0.06 | 1.1E-05 |
| GYPLT | SNODE_4555_43 | T | A | 0.11 | 3.8E-05 |
| GYPLT | SNODE_34407_21 | A | G | 0.06 | 5.5E-05 |
| GYPLT | SNODE_4363_81 | C | T | 0.36 | 7.5E-05 |
| GYPLT | SNODE_16972_9 | A | G | 0.05 | 7.9E-05 |
| GYPLT_RI | SNODE_22383_32 | T | C | 0.14 | 5.5E-05 |
| NBrPLT | SCONTIG97891_72 | A | G | 0.40 | 7.4E-06 |
| NBrPLT | SNODE_103_72 | T | C | 0.25 | 2.6E-05 |
| NBrPLT_RI | SNODE_173108_18 | A | G | 0.06 | 2.6E-06 |
| NBrPLT_RI | SNODE_2942_50 | C | T | 0.06 | 1.0E-05 |
| NBrPLT_RI | SNODE_103_72 | T | C | 0.25 | 1.6E-05 |
| NBrPLT_RI | SNODE_144193_69 | C | A | 0.46 | 9.5E-05 |
| NPPLT | SCONTIG23347_118 | G | A | 0.07 | 2.3E-05 |
| NPPLT | SNODE_28265_65 | C | T | 0.31 | 8.0E-05 |
| NPPLT_RI | SNODE_559376_60 | A | T | 0.45 | 1.0E-11 |
| NSPLT | SNODE_14558_21 | A | G | 0.20 | 1.2E-05 |
| NSPLT | SCONTIG38056_40 | C | T | 0.07 | 2.8E-05 |
| PLHT | SNODE_43134_109 | A | G | 0.05 | 2.8E-09 |
| PLHT | SNODE_134600_32 | T | G | 0.06 | 3.8E-06 |
| PLHT | SNODE_27201_27 | T | G | 0.06 | 1.0E-05 |

Continued

| Trait | SNP | allele1 | allele0 | MAF | P |
|----------|-------------------------------|---------|---------|------|---------|
| PLHT | <i>SCONTIG73439_18</i> | G | A | 0.06 | 1.3E-05 |
| PLHT | SNODE_78412_27 | T | C | 0.11 | 2.9E-05 |
| PLHT | SNODE_124581_38 | A | C | 0.07 | 6.9E-05 |
| PLHT | SNODE_113123_17 | T | G | 0.06 | 7.5E-05 |
| PLHT | SCONTIG57859_65 | G | A | 0.07 | 7.5E-05 |
| PLHT | SCONTIG101530_33 | A | G | 0.14 | 9.6E-05 |
| PLHT_RI | <i>SCONTIG73439_18</i> | G | A | 0.06 | 6.1E-06 |
| PLHT_RI | SCONTIG157_70 | T | G | 0.16 | 9.3E-06 |
| PLHT_RI | SNODE_14298_44 | A | G | 0.06 | 1.0E-05 |
| PLHT_RI | SNODE_3358_54 | A | C | 0.18 | 1.1E-05 |
| PLHT_RI | SNODE_11304_26 | G | A | 0.12 | 2.9E-05 |
| PLHT_RI | SNODE_107804_70 | A | T | 0.05 | 3.2E-05 |
| PLHT_RI | SNODE_176979_47 | G | A | 0.08 | 4.6E-05 |
| PLHT_RI | SNODE_4904_26 | C | T | 0.17 | 9.2E-05 |
| HDS1 | SNODE_2908_40 | G | A | 0.17 | 4.4E-05 |
| HDS2 | SNODE_8269_115 | G | A | 0.07 | 5.8E-06 |
| HDS2 | SNODE_68619_39 | G | A | 0.08 | 1.3E-05 |
| HDS2 | SNODE_2018_104 | C | T | 0.07 | 9.9E-05 |
| NSPLT_RI | SNODE_2107_36 | C | T | 0.06 | 6.1E-06 |
| NSPLT_RI | SNODE_7966_59 | G | A | 0.06 | 4.2E-05 |
| NSPLT_RI | <i>SNODE_14558_21</i> | A | G | 0.20 | 8.7E-05 |

Table 3. SNP-trait associations revealed by the MT-GWAS analysis. Underscored SNPs represents the highly significant associations, while SNPs in bold italic represents the SNPs associated with multiple traits. SNP single nucleotide polymorphism, MAF minor allele frequency, DFLR days to flowering, DFLR_RI DFLR reduction index, DMAT days to maturity, DMAT_RI DMAT reduction index, GCC green canopy cover, GCC_RI green canopy cover reduction index, GYPLT grain yield per plant, GYPLT_RI GYPLT reduction index, NBrPLT number of branches per plant, NBrPLT_RI NBrPLT reduction index, NPPLT number of pods per plant, NPPLT_RI NPPLT reduction index, NSPLT number of seeds per plant, PLHT plant height, PLHT_RI reduction index of PLHT, HDS1 first herbicide damage score, HDS2 second herbicide damage score, NSPLT_RI number of seeds per plant, NSPLT number of seeds per plant, _ reduction index.

the observed differences in the genotype ability to recover might be due to differential rate of metabolic degradation for imazethapyr treatment⁴⁷ and to differential disruption of electron transfer for metribuzin treatment.

Population structure analysis revealed two major ancestral populations for the germplasm which is compatible with the original germplasm of 995 genotypes genotyped with 20 microsatellite markers, from which this population was selected^{22,22}. As expected, MT-GWAS analysis exposed higher detection power compared to ST-GWAS analysis due to the larger datapoint fitted in the model which is equivalent to increasing the population size²¹. This was revealed by the larger number of highly significant as well as suggestive association per trait detected (Table 3, Supplementary Table S1). Another advantage is the ability to detect QTL with stable effect across different environments or treatments which should have higher potential to improve the efficiency of marker assisted selection in diverse environments^{56,57}.

To the best of our knowledge, the present study is the first GWAS for herbicide tolerance in faba bean and the first for all phenotyped traits under the control treatment with no herbicide application. Thus, most of the QTL detected in the present study seem novel and have not been reported before. Very limited studies used SNP data on biparental or multi-parental faba bean populations^{15,26,58} but none aimed to dissect quantitative traits in natural diverse populations. Sallam and Martsch⁵⁸ associated 156 SNPs with frost tolerance in a population derived from 11 parental lines, while Ali et al.⁵⁹ used the same population to detect loci associated with freezing and drought tolerance using 175 SNPs and AFLP markers. The identification of QTL through GWAS in faba bean is complex due to the large undecoded genome and highly repetitive sequences. These issues have delayed the progress made towards the development of genomic resources and marker assisted selection in faba bean breeding programs⁶⁰.

Identification of key genes, mechanisms and functional markers is essential to develop herbicide tolerant faba beans. The associations between some genes identified in this study and herbicide tolerance have been reported previously in different crops. Acidic endochitinase and malate dehydrogenase which were found to be associated with the reduction indexes of DFLR and GCC were among the proteins affected by the application of sulfonylurea herbicide in soybeans⁶¹. Sulfonylurea herbicides and imazethapyr have similar mode of action; both herbicides block the biosynthesis of the branched-chain amino acids^{54,62}. The two protein kinase LRR receptor-like serine/threonine-protein kinase RCH1 and probable serine/threonine-protein kinase NAK which were found to be associated with the reduction index of PLHT and GYPLT in the present study are generally considered key regulators of plant architecture and growth behavior, and the expansion of these proteins during plant evolution has also been correlated with the specific adaptations of the species in defense and stress responses⁶³. Their direct involvement in abiotic stress resistance (drought, heat, cold, salinity) has also been demonstrated in different

studies^{64–66}. Burns et al.⁶⁷ concluded that herbicide stress is perceived similarly to other abiotic stresses and reported modification in the level of the protein kinase gene family in the multiple herbicide resistant *Avena fatua*.

The MYB-related protein P-like which was found associated with DFLR is involved in herbicide tolerance belongs to the MYB gene family that comprises one of the richest groups of transcription factors in plants. Members of this family have a well-established role in abiotic stress responses^{68,69}. Bhoite et al.⁷⁰ found also that the transcription factors MYB were significantly expressed under metribuzin stress. The photosystem I (PS I) core protein PsaA that is found in the present study to be associated with the reduction index of NPPLT is a subunit membrane protein complex involved in photosynthesis. PS I and PS II drive the light reaction of photosynthesis. The first stage of the light reaction occurs in PS II whereas the final stage of the light reaction occurs in PS I⁷¹. The metribuzin applied to faba bean plants in this study inhibits PS II by disrupting electron transfer through binding to the D1 protein of the photosystem II complex in chloroplast thylakoid membranes⁵¹. This mode of action explains the involvement of the PS I in the reaction toward herbicide application especially that PS II comes first in the path of the electron flow followed by PS I.

The described mechanism of action of the annotated genes suggests that DFLR_RI and GCC_RI are associated with tolerance to imazethapyr while DFLR and NPPLT_RI are associated with tolerance to metribuzin, and GYPLT and PLHT_RI are associated with tolerance to both herbicides.

Conclusions

Weeds represent a major problem to faba bean crop which limits its expansion in many production regions. By excluding faba bean and other legume from the cropping system, cereal monoculture will continue to deplete the soil, lowering its quality and indirectly reducing yield and quality of the produce. Herbicide tolerant faba bean lines could be a game changer in the reintegration of faba bean in modern cropping systems as it contributes to the reduction of production cost by avoiding excessive use of manual weeding. Considering the many advantages of herbicide tolerance in faba bean, it is imperative to breed elite cultivars that features this trait. However, field selection is very costly and time consuming. The integration of genomic selection and marker assisted selection will improve selection accuracy, increase the selection intensity and shorten the breeding cycle when selecting at early generations. In the present study, we identified genomic regions associated with tolerance to imazethapyr and metribuzin herbicides as highly significant associations between SNPs markers and phenological and yield traits related to herbicide tolerance were detected using multi-trait association. These markers will be useful for improving the efficiency of faba bean programs and represent important steps towards the selections for herbicide tolerance.

Materials and methods

Plant materials. A set of 134 faba bean genotypes comprising 118 landraces from 42 countries and 16 ICARDA breeding lines that were used to establish a reference set under the Generation Challenge Program (GCP) was used for phenotyping and genotyping in the present study. Previous assessment with Simple Sequence Repeat (SSR) markers showed that the set was genetically diverse and comprised 45 *major*, 17 *minor*, 63 *equina* and 9 *paucijuga* genotypes^{22,42}. In addition to the test genotypes, a total of 6 faba bean cultivars (FLIP86-98, ILB1814, Ed-damer, Hudeiba-93, Shambat-75, SML) were included in the experiments. The seeds used in the current experiments are sourced from the reserve seeds that are multiplied each year under insect-proof cages in order to ensure purity of the evaluated accessions.

Experiments. A total of four experiments were conducted: three field and one greenhouse experiments.

Field experiments. A total of three field experiments were conducted at two ICARDA research stations: Marchouch (33.558°N 6.693°W, altitude 255 m) in Morocco and Terbol (35.98°N, 33.88°E, altitude 890 m) in the Bekaa Valley of Lebanon. Marchouch station is characterized by the semi-arid environment with a Vertisol soil, mostly silty clay, while Terbol station is characterized by cool and high rainfall winter and moderate wet spring with a deep and rich clay loam soil. Each experiment comprised three treatments applied at the pre-flowering stage: T1-Metribuzin @250 g ai/ha, T2-Imazethapyr @ 75 g ai/ha and T3- No herbicide application. Faba bean genotypes were sown in rotation with cereals in mid-December at Marchouch 2014/2015, late November at Terbol 2014/2015 and 2018/2019 main seasons. Each genotype was planted in 2 m long two-row plot with 0.5 m spacing between rows. At Marchouch, the crop received 291 mm of precipitation during the cropping season in addition to 30 mm irrigation during early vegetative phase; the crop was exposed to intermittent drought and heat. 120 genotypes along with the three following cultivars FLIP86-98, ILB1814 and Hudeiba-93 were evaluated at Marchouch using Augmented design⁴². At Terbol, a total precipitation of 343 mm and 810 mm was recorded respectively during 2015/2016 and 2018/2019 cropping seasons. Supplemental irrigation (30 mm) was provided at Terbol station in 2015/2016 season during dry-spell periods, while no irrigation was provided in case of highly and well distributed rains in 2018/2019. A total of 134 genotypes were evaluated at Terbol using Alpha lattice design. In 2015/2016 season, the field experiment was conducted with two replicates and 15 blocks and with the cultivar FLIP86-98⁴² and in 2018/2019 the field experiment was conducted with 3 replicates and 14 blocks and the following cultivars FLIP86-98, ILB1814, Ed-damer, Hudeiba-93, Shambat-75 and SML.

Greenhouse experiment. The germplasm genotypes along with six checks (FLIP86-98, ILB1814, Ed-damer, Hudeiba-93, Shambat-75, SML) were evaluated in an alpha design with two replicates and two treatments: 250 g ai/ha of metribuzin and untreated treatment during 2017/2018 cropping season. Three seeds per pot for each genotype were sown in this experiment. Irrigation was provided regularly to maintain 100% soil water capacity

in pots. Temperature inside the greenhouse was fixed at 24 to 28 °C the optimal day time temperature of faba bean.

The herbicide treatments applied in all experiments are metribuzin (M, T1), imazethapyr (I, T2) and the control treatment (C, T3) in which no herbicide was applied. The doses of herbicides applied are the recommended doses as per the labels of metribuzin (Sencor: Bayer) and imazethapyr (Pursuit: BASF). Both herbicides were uniformly sprayed at the rate of 250 g ai ha⁻¹ and 75 g ai ha⁻¹ respectively at the inflorescence stage BBCH code 5072⁷² for the field experiments and at the stem elongation stage BBCH code 30⁷² for the greenhouse experiment using an electric sprayer with automated flow (375 L/ha). In the field, the herbicide was sprayed early in the morning to ensure a low wind speed. Details of traits scored in each trial can be found in Supplementary Table S3. Traits were coded as the environment, followed by the trait, the treatment and “RI” if the score describe a reduction index. For the multi-trait GWAS analysis, the trait name does not have the name of the environment or the treatment.

Phenotyping for herbicide tolerance. Observations (Supplementary Table S3) were recorded on days to 50% flowering (DFLR) and maturity (DMAT) on plot basis for the untreated treatment, and plant height (PLHT) and grain yield per plant (GYPLT) on three plants selected randomly for all the three treatments at Marchouch 2014/2015. At Terbol station, the following additional traits were also recorded on three plants selected randomly for the three treatments: number of pods per plant (NPPLT), number of seeds per plant (NSPLT), number of branches per plant (NBrPLT) and green canopy cover (GCC). Green canopy cover expressed as the average percentage of green coverage of three plants was quantified using the Canopeo application developed by Oklahoma State University using Matlab. Under the greenhouse conditions where temperature was controlled at optimal conditions, PLHT was recorded at flowering PLHT_1 (BBCH code 60) and pod development PLHT_2 (BBCH code 70) stages⁷². The herbicide damage score (HDS) was recorded in all the four experiments using a 1–5 scale⁴² (Supplementary Table S4) at flowering (HDS1) and pod development (HDS2) stages. The ratio of each quantitative trait was calculated for each plot using the following formula described by Abou-Khater et al.⁴²:

$$RI\% = 100 - \left(\frac{T}{C} \times 100 \right)$$

where RI%, the reduction index, represents the reduction or penalty in traits of herbicides treated plots compared to the control untreated plots, *T* is the average of plots treated with herbicide (metribuzin or imazethapyr); *C* is the mean of genotypes under untreated conditions.

DNA extraction and genome by sequencing analysis. Genomic DNA was extracted from young leaf tissues for each tested genotype using the DNeasy 96 Plant Kit (QIAGEN, Valencia, CA, USA) Qiagen Plant DNA Preparation Kit. For the preparation of the GBS library, the two restriction enzymes, *PstI* and *MspI*, were used to generate fewer variation in the distribution of read depth and higher number of scorable SNPs. GBS libraries were prepared with 48 barcode adapters with 4–9 bp sequence⁷³. The single read (100 base pairs) sequencing on an Illumina HiSeq 2500 produced approximately 4 million reads per genotype. Raw read sequences were processed using TASSEL-GBS 5.0 with the default parameters⁷⁴. A faba bean sequence database was constructed using 223,801 genomic and transcriptomic faba bean sequences downloaded from NCBI and pulsedb databases (www.ncbi.nlm.nih.gov and <https://www.pulsedb.org/analysis/136>) and additional faba bean sequences constructed using the Trinity assembler from one run of the GBS files. These sequences were used as a reference to align GBS sequence tags and indexed using Bowtie2 version 2.2.4⁷⁵. Bowtie2 was used to align GBS tags to faba sequences using the—very-sensitive-local option. Resulting SNPs were filtered with 20x coverage, where SNPs with more than 15% missing data or less than 5% minor allele frequency (MAF) were removed. SNPs were named by contig base pair position.

Statistical analysis of phenotyping data. The spatial statistical model was applied for variance analysis for all quantitative data using the Automatic Spatial Analysis of Row-Column modules of Genstat 19 edition⁷⁶. Significance of variation among genotypes and treatments was assessed in terms of *P*-values. The analysis of variance (ANOVA), means of genotypes, means of treatments and interactions between genotypes and treatments were estimated with standard errors using best linear unbiased prediction (BLUP) values using GenStat software. BLUPs were used to conduct all downstream analyses. Multi environment trials analysis (META) were conducted to evaluate variation among genotypes, treatments and the genotype × treatment interaction across trials for the traits recorded in more than one trial. Genotype and treatment were fitted as fixed parameters while environment (year-location) were fitted as random parameter.

Genome-wide association analysis. ADMIXTURE software⁷⁷ was used to infer population structure with the number of underlying subpopulations (*K*) ranges between 2 and 20. The analysis was run with 100 random replicates and 20 cross validations. The most probable *K* was determined at the point when the average cross validation (*CV*) values across the 100 replicates started to increase. Single-trait (ST) and Multi-trait (MT) GWAS was fitted using GEMMA software²¹ by fitting each trait independently (for the ST analysis) or fitting all field or greenhouse records together (for the MT analysis) with the default parameters and by fitting the genomic relatedness matrix as a covariate to control for population stratification⁷⁸. Bonferroni correction was used to determine the significant threshold at *p* < 0.05 but all SNPs with *p* < 1E–4 were presented as suggestive associations. Pairwise linkage disequilibrium (LD) between associated SNPs within each trait was estimated with the *r*² statistics following Weir⁷⁹ to determine the SNPs that are associated with the same quantitative trait

locus (QTL). Genomic sequences containing herbicide tolerance associated SNPs were aligned against the NCBI database using BLASTX tool using default parameters to annotate potential candidate genes underlying the causal variants.

Ethical approval. The authors confirm that the study complies with local and national regulations. The seeds were collected from the genebank of the International Center for Agricultural Research in the Dry Areas (ICARDA) for research purposes according to the International Treaty of Plant Genetic Resources for Food and Agriculture (ITPGRFA). For the collection of seeds, all relevant permits or permissions have been obtained. The seeds flow from ICARDA GenBank at Terbol to Morocco was made following the phytosanitary regulations of both countries and using the Standard Material Transfer Agreement (SMTA) governed by ITPGRFA. The experiments were conducted at ICARDA sites at Terbol and Marchouch in accordance to National and International regulations.

Data availability

The datasets generated and analyzed during the current study are available from the corresponding authors on request.

Received: 23 June 2021; Accepted: 9 December 2021

Published online: 07 January 2022

References

- Cubero, J. I. On the evolution of *Vicia faba* L. *Theor. Appl. Genet.* **45**, 47–51 (1974).
- Tanno, K. I. & Willcox, G. The origins of cultivation of *Cicer arietinum* L. and *Vicia faba* L.: early finds from Tell el-Kerkh, north-west Syria, late 10th millennium BP. *Veg. Hist. Archaeobot.* **15**, 197–204 (2006).
- Caracuta, V. *et al.* 14,000-year-old seeds indicate the Levantine origin of the lost progenitor of faba bean. *Sci. Rep.* **6**, 1–6 (2016).
- FAOSTAT. Food and Agriculture Organization of the United Nations. www.fao.org. Accessed 25 April 2021 (2020).
- Crépon, K. *et al.* Nutritional value of faba bean (*Vicia faba* L.) seeds for feed and food. *Field Crops Res.* **115**, 329–339 (2010).
- Haugaard-Nielsen, H., Mundus, S. & Jensen, E. S. Nitrogen dynamics following grain legumes and subsequent catch crops and the effects on succeeding cereal crops. *Nutr. Cycl. Agroecosyst.* **84**, 281–291 (2009).
- Stagnari, F., Maggio, A., Galieni, A. & Pisante, M. Multiple benefits of legumes for agriculture sustainability: An overview. *Chem. Biol. Technol. Agric.* **4**, 1–13 (2017).
- Zhou, R. *et al.* Phenotyping of faba beans (*Vicia faba* L.) under cold and heat stresses using chlorophyll fluorescence. *Euphytica* **214**, 1–13 (2017).
- Anjum, N. A. Book review: Legumes under environmental stress: Yield, improvement and adaptations. *Front. Plant Sci.* **7**, 798 (2016).
- Stoddard, F. L. *et al.* Screening techniques and sources of resistance to abiotic stresses in cool-season food legumes. *Euphytica* **147**, 167–186 (2006).
- Sillero, J. C. *et al.* Faba bean breeding for disease resistance. *Field Crops Res.* **115**, 297–307 (2010).
- Stoddard, F. L., Nicholas, A. H., Rubiales, D., Thomas, J. & Villegas-Fernández, A. M. Integrated pest management in faba bean. *Field Crops Res.* **115**, 308–318 (2010).
- Khamassi, K. *et al.* A baseline study of vicine–convicine levels in faba bean (*Vicia faba* L.) germplasm. *Plant Genet. Resour.* **11**, 250–257 (2013).
- Cooper, J. W. *et al.* Enhancing faba bean (*Vicia faba* L.) genome resources. *J. Exp. Bot.* **68**, 1941–1953 (2017).
- Carrillo-Perdomo, E. *et al.* Development of new genetic resources for faba bean (*Vicia faba* L.) breeding through the discovery of gene-based SNP markers and the construction of a high-density consensus map. *Sci. Rep.* **10**, 1–14 (2020).
- Mardis, E. R. The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24**, 133–141 (2008).
- Morozova, O. & Marra, M. A. Applications of next-generation sequencing technologies in functional genomics. *Genomics* **92**, 255–264 (2008).
- Gnanasambandam, A. *et al.* Impact of molecular technologies on faba bean (*Vicia faba* L.) breeding strategies. *Agron.* **2**, 132–166 (2012).
- Flint-Garcia, S. A., Thornsberry, J. M. & Buckler, E. S. IV. Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* **54**, 357–374 (2003).
- Korte, A. & Farlow, A. The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* **9**, 1–9 (2013).
- Zhou, X. & Stephens, M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* **11**, 407–409 (2014).
- Maalouf, F. *et al.* Breeding and genomics status in faba bean (*Vicia faba*). *Plant Breed.* **138**, 465–473 (2019).
- Webb, A. *et al.* A SNP-based consensus genetic map for synteny-based trait targeting in faba bean (*Vicia faba* L.). *Plant Biotechnol. J.* **14**, 177–185 (2016).
- Satovic, Z. *et al.* A reference consensus genetic map for molecular markers and economically important traits in faba bean (*Vicia faba* L.). *BMC Genom.* **14**, 1–15 (2019).
- Ellwood, S. R. *et al.* Construction of a comparative genetic map in faba bean (*Vicia faba* L.); conservation of genome structure with *Lens culinaris*. *BMC Genom.* **9**, 1–1 (2008).
- Kaur, S. *et al.* SNP discovery and high-density genetic mapping in faba bean (*Vicia faba* L.) permits identification of QTLs for ascochyta blight resistance. *Plant Sci.* **217**, 47–55 (2014).
- Ocaña-Moral, S., Gutiérrez, N., Torres, A. M. & Madrid, E. Saturation mapping of regions determining resistance to Ascochyta blight and broomrape in faba bean using transcriptome-based SNP genotyping. *Theor. Appl. Genet.* **130**, 2271–2282 (2017).
- Khazaei, H. *et al.* Flanking SNP markers for vicine–convicine concentration in faba bean (*Vicia faba* L.). *Mol. Breed.* **35**, 1–6 (2015).
- Song, M. Preventing favism by selecting faba bean mutants using molecular markers. *SFJ.* **3**, 2–6 (2017).
- Maalouf, F. *et al.* Breeding for post-emergence herbicide tolerance in cool-season food legumes. In *Proceedings of the International Conference on Pulses*, 65 (Marrakesh, 18–20 April 2016).
- Babiker, A. G. T. & Khalid, M. E. Chemical weed control in faba bean at Wad Hamid. In *Annual Report, Hudieba Research Station* 183–185 (Agric. Res. Corp (ARC), 1990).
- Rubiales, D. & Fernández-Aparicio, M. Innovations in parasitic weeds management in legume crops. A review. *Agron. Sustain. Dev.* **32**, 433–449 (2012).
- Pérez-de-Luque, A. *et al.* Broomrape management in faba bean. *Field Crops Res.* **115**, 319–328 (2010).

34. Singh, N. P. & Yadav, I. S. Herbicide tolerant food legume crops: possibilities and prospects. In *Herbicides-properties, synthesis and control of weeds* (ed. Hasaneen M. N.) 435–452. (InTech, 2012)
35. Minotti, P. L. & Sweet, R. D. Role of crop competition in limiting losses from weeds. In *Handbook of Pest Management in Agriculture* (ed. Pimental, D.) 351–367 (CRC Handbook of Pest Management in Agriculture, 1981).
36. Berkowitz, A. R. Competition for resources in weed-crop mixtures. In *Weed Management in Agrosystems: Ecological Approaches* (eds Altiera, M. A. & Liebman, M.) 89–119 (CRC Press, 1981).
37. Gaudet, C. L. & Keddy, P. A. A comparative approach to predicting competitive ability from plant traits. *Nature* **334**, 242–243 (1988).
38. Rubiales, D. Can we breed for durable resistance to broomrapes?. *Phytopath. Mediterranea* **57**, 170–185 (2018).
39. Rispaill, N. *et al.* Plant resistance to parasitic plants: Molecular approaches to an old foe. *New Phytol.* **173**, 703–712 (2007).
40. Beckie, H. J. *et al.* A decade of herbicide-resistant crops in Canada. *Can. J. Plant Sci.* **86**, 1243–1264 (2006).
41. Mao, D. *et al.* *Developing Improved Herbicide Tolerance in Pulse Crops*. (Grain Research and Development Corporation, 2016). <https://grdc.com.au/resources-and-publications/grdc-update-papers/tab-content/grdc-update-papers/2016/02/developing-improved-herbicide-tolerance-in-pulse-crops> (accessed on 24 April 2021).
42. Abou-Khater, L. *et al.* Identification of tolerance to metribuzin and imazethapyr herbicides in faba bean (*Vicia faba* L.). *Crop Sci.* **1–19** (2021).
43. GRDC. *Grownotes Faba Bean Section 17 Environmental Issues*. <https://grdc.com.au/> (accessed on 5 April 2021). (2017).
44. Pulse Australia. *Managing Pulses to Minimize Frost Damage*. <http://www.pulseaus.com.au/> (accessed on 5 April 2021).
45. García-Torres, L. & López-Granados, F. Control of broomrape (*Orobanche crenata* Forsk.) in broad bean (*Vicia faba* L.) with imidazolinones and other herbicides. *Weed Res.* **31**, 227–235 (1991).
46. Sharma, S. R. *et al.* Effect of post-emergence herbicide metribuzin application on morpho-physiological traits, yield and yield components in Lentil (*Lens culinaris* Medik.). In *Proceedings of the International Conference on Pulses for Health, Nutrition, and Sustainable Agriculture in Drylands*, 100 (Marrakesh, Morocco, 18–20 April 2016).
47. Sharma, S. R. *et al.* Genetic variation for tolerance to post-emergence herbicide, imazethapyr in lentil (*Lens culinaris* Medik.). *Arch. Agron. Soil Sci.* **64**, 1818–1830 (2018).
48. Taran, B., Warkentin, T. D., Vandenberg, A. & Holm, F. A. Variation in chickpea germplasm for tolerance to imazethapyr and imazamox herbicides. *Can. J. Plant Sci.* **90**, 139–142 (2010).
49. Wall, D. A. Lentil (*Lens culinaris*) and faba bean (*Vicia faba*) tolerance to post-emergence applications of imazethapyr. *Can. J. Plant Sci.* **76**, 525–529 (1996).
50. Sajja, S. *et al.* Effect of post-emergence herbicide imazethapyr on phenological and agronomic traits in chickpea breeding lines. In *Proceedings of the International Plant Breeding Congress (IPBC) and Eucarpia—Oil and Protein Crops Section Conference* (Antalya, Turkey, 1–5 November 2015).
51. Senseman, S. A. *Herbicide Handbook*. 9th ed.; Weed Science Society of America. (Lawrence, US, 2007) ISBN 1891276565, 9781891276569.
52. Trebst, A. & Wietoska, H. Mode of action and structure-activity-relationships of the aminotriazinone herbicide Metribuzin. Inhibition of photosynthetic electron transport in chloroplasts by Metribuzin (author's transl). *Zeitschrift für Naturforschung, Section C, Biosciences* **30**, 499–504 (1975).
53. Shaner, D. L., Anderson, P. C. & Stidham, M. A. Imidazolinones: potent inhibitors of acetohydroxyacid synthase. *Plant Physiol.* **76**, 545–546 (1984).
54. Saari, L. L. & Mauvais, C. J. Sulfonylurea herbicide-resistant crops. In *Herbicide-Resistant crops* (ed. Duke, S. O.) 127–142 (CRC Press, 1996).
55. Teclé, B., Da Cunha, A. & Shaner, D. L. Differential routes of metabolism of imidazolinones, basis for soybean (*Glycine max*) selectivity. *Pestic. Biochem. Phys.* **46**, 120–130 (1993).
56. Manickavelu, A. *et al.* Genome wide association mapping of stripe rust resistance in Afghan wheat landraces. *Plant Sci.* **252**, 222–229 (2016).
57. Oladzad, A. *et al.* Single and multi-trait GWAS identify genetic factors associated with production traits in common bean under abiotic stress environments. *G3* **9**, 1881–1892 (2019).
58. Sallam, A. & Martsch, R. Association mapping for frost tolerance using multi-parent advanced generation inter-cross (MAGIC) population in faba bean (*Vicia faba* L.). *Genetica* **143**, 501–514 (2015).
59. Ali, M. B. *et al.* Association analyses to genetically improve drought and freezing tolerance of faba bean (*Vicia faba* L.). *Crop Sci.* **56**, 1036–1048 (2016).
60. Mulugeta, B., Tesfaye, K., Keneni, G. & Ahmed, S. Genetic diversity in spring faba bean (*Vicia faba* L.) genotypes as revealed by high-throughput KASP SNP markers. *Genet. Resour. Crop Evol.* **68**, 1971–1986 (2021).
61. Yu, X., Yang, A. & James, A. T. Selecting soybeans for sulfonylurea herbicide tolerance: A comparative proteomic study of seed germinations. *Crop Pasture Sci.* **68**, 27–32 (2017).
62. Brown, H. M. Mode of action, crop selectivity, and soil relations of the sulfonylurea herbicides. *Pestic. Sci.* **29**, 263–281 (1990).
63. Lehti-Shiu, M. D., Zou, C., Hanada, K. & Shiu, S. H. Evolutionary history and stress regulation of plant receptor-like kinase/pelle genes. *Plant Physiol.* **150**, 12–26 (2009).
64. Boller, T. & Felix, G. A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annu. Rev. Plant Biol.* **60**, 379–406 (2009).
65. Chae, L., Pandey, G. K., Luan, S., Cheong, Y. H. & Kim, K. N. Protein kinases and phosphatases for stress signal transduction in plants. In *Abiotic Stress Adaptation in Plants* (eds Pareek, A. *et al.*) 123–163 (Springer, 2009).
66. Shehzad, M. *et al.* Genome-wide mining and identification of protein Kinase Gene Family impacts salinity stress tolerance in highly dense genetic map developed from interspecific cross between *G. hirsutum* L. and *G. darwinii* G. Watt. *Agronomy* **9**, 560 (2019).
67. Burns, E. E., Keith, B. K., Refai, M. Y., Bothner, B. & Dyer, W. E. Constitutive redox and phosphoproteome changes in multiple herbicide resistant *Avena fatua* L. are similar to those of systemic acquired resistance and systemic acquired acclimation. *J. Plant Physiol.* **220**, 105–114 (2018).
68. Roy, S. Function of MYB domain transcription factors in abiotic stress and epigenetic control of stress response in plant genome. *Plant Signal. Behav.* **11**, e1117723 (2016).
69. Katiyar, A. *et al.* Genome-wide classification and expression analysis of MYB transcription factor families in rice and Arabidopsis. *BMC Genom.* **13**, 1–19 (2012).
70. Bhoite, R., Si, P., Siddique, K. H. & Yan, G. Comparative transcriptome analyses for metribuzin tolerance provide insights into key genes and mechanisms restoring photosynthetic efficiency in bread wheat (*Triticum aestivum* L.). *Genomics* **113**, 910–918 (2021).
71. Rachna, C. *Difference Between Photosystem I and Photosystem II*. *Bio Differences*. <https://bioidifferences.com/difference-between-photosystem-i-and-photosystem-ii.html> (accessed on 10 May 2021) (2019).
72. Lancashire, P. D. *et al.* A uniform decimal code for growth stages of crops and weeds. *Ann. App. Biol.* **119**, 561–601 (1991).
73. Poland, J. A., Brown, P. J., Sorrells, M. E. & Jannink, J. L. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* **7**, e32253 (2012).
74. Glaubitz, J. C. *et al.* TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS ONE* **9**, e90346 (2014).
75. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357 (2012).
76. Goedhart, P. W. & Thissen, J. T. N. M. *Biometris Genstat procedure library manual* 19th edn. (Biometris, Wageningen, 2018).

77. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
78. VanRaden, P. M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**, 4414–4423 (2008).
79. Weir, B. S. *Genetic Data Analysis. Methods for Discrete Population Genetic Data* (Sinauer Associates Inc Publishers, Sunderland, 1990).

Acknowledgements

This research was funded by CRP-grain legume program 2012–2017, Grant number 100231 and Arab Fund for Economic and Social Development (AFESD) for the operationalization of decentralized Center in five Arab countries. This publication falls within the framework of the doctoral study of Lynn Abou-Khater at the University of Córdoba in Spain. We acknowledge the efforts made by M. Wehbe, G. Al-Khatib, S. B. Patil and K. Hejjaoui in the implementation of the field trials and in the data collection.

Author contributions

F.M. developed the core idea, designed the experiments, L.A.K. conducted the experiments. R.B. performed the herbicide tolerance evaluation during 2018–2019. F.M., L.A.K. analyzed the phenotypic data. A.J. and A.A., conducted the genotypic analysis. A.H., F.M. and M.B. assembled the GCP reference set. J.H. and Y.M. conducted the DNA extraction and the genotyping. L.A.K., A.J., and A.A. wrote the original draft. F.M., D.R., N.R., J.H., Y.M., M.B. and S.K. edited and reviewed the article. M.B., S.K. and F.M. contributed to fund raising.

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/s41598-021-03861-0>.

Correspondence and requests for materials should be addressed to L.A.-K. or F.M.

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) 2022