A Low-Cost Genotyping Protocol and Kit for Marker-Assisted Selection of Orange Lemma *(rob1.a*), a Feed Quality Trait in Barley *(Hordeum vulgare* L.*)*

INTRODUCTORY GUIDE AND LABORATORY PROTOCOLS

Photo

Barley field, Raasdorf Experimental Station Groß-Enzersdorf, BOKU-University of Natural Resources and Life Sciences, Vienna, Austria

Joint FAO/IAEA Division,

Plant Breeding and Genetics Laboratory,

Seibersdorf, Austria

Version__22_11_2020

Contributors to the working draft

Ingelbrecht ILW[1], Akgun E, Matijevic M, Hofinger B, Jarc L, Jankowicz-Cieslak J, Ali A, Jankuloski L, Grausgruber H, Ghanim AMA

[1]correspondence: i.ingelbrecht@iaea.org

FOREWORD

Natural and induced variation is the ultimate source of genetic variation that plant breeders use for crop improvement. Since the 1950s, induced mutagenesis using physical (gamma- or X-rays) or chemical (e.g. ethyl methanesulfonate or EMS) methods has generated numerous new crop varieties and traits that have impacted agricultural productivity worldwide.

Barley (*Hordeum vulgare* L.) is the fourth largest grain crop globally. In 2017/2018 the global production of barley was about 142 million tons with ca 70% of this used as feed for livestock. In barley, several mutant stocks with an 'orange lemma' phenotype have been developed. The orange lemma trait in barley is associated with reduced lignin content and has potential to improve forage digestibility for livestock, similar to the brown-midrib mutants in maize, sorghum and millet. However, so far this trait has not been commercially deployed in case of barley.

The Coordinated Research Project D23030 'Integrated Utilization of Cereal Mutant Varieties in Crop/Livestock Production Systems' set out to develop new tools and technologies to improve barley as a dual-purpose crop for use as feed for livestock and grain.

As a contribution to this CRP, the PBGL developed a diagnostic marker for the orange lemma trait derived from the barley mutant stock BW666, a Bowman backcross-derived line containing the *rob1.a* mutant allele, causative for the orange lemma trait. Our aim is to facilitate fast and efficient introgression of the recessive mutant trait widely in barley varieties produced for animal feed by Member States. The genotyping assay is based on allele-specific Polymerase Chain Reaction (PCR) amplification of the mutant and wild-type allele. The assay is robust, breeder-friendly and low-cost requiring only standard molecular biology equipment.

The marker assay is successfully applied for pyramiding orange lemma with the 'hooded' trait (*Kap1*) in the forage barley breeding programme of BOKU University, Vienna, Austria. The protocol has also been transferred to interested Member States with forage barley breeding programs such as Tunisia and Kuwait.

The protocols and guide presented here introduces general principles of genotyping and marker-aided breeding, as well as step-by step procedures on how to conduct the barley orange lemma marker assay from the growth of plant material to data analysis and troubleshooting.

A marker kit for introgression of the orange lemma trait in barley via Marker-Assisted Selection is available upon request. The marker kit comprises this protocol, the PCR primers for conducting the genotyping assay and the genetic stock BW666 as source of the orange lemma trait.

CONTENTS

1. INTRODUCTION

1.1. Molecular breeding

Molecular markers reveal variation at the DNA level commonly referred to as DNA polymorphisms. They can be applied at every stage of plant breeding including the characterization of genetic resources (pre-breeding), line or variety development, quality control and proof of identity (IPR) after release. A variety of markers are known in plants. They are generally classified according to their genetic or technical characteristics. For plant breeding purposes, markers that are easy to genotype, are co-dominant and are tightly linked with the trait of interest are preferred. Functional markers are developed from polymorphic sites within a gene that directly cause the variation of a trait [1]. Because functional markers show complete linkage with the trait of interest, they are the most reliable to use for e.g. in Marker-Assisted Selection (MAS). The orange lemma marker described here is derived from the causative allele underlying the orange lemma trait and is thus a functional marker.

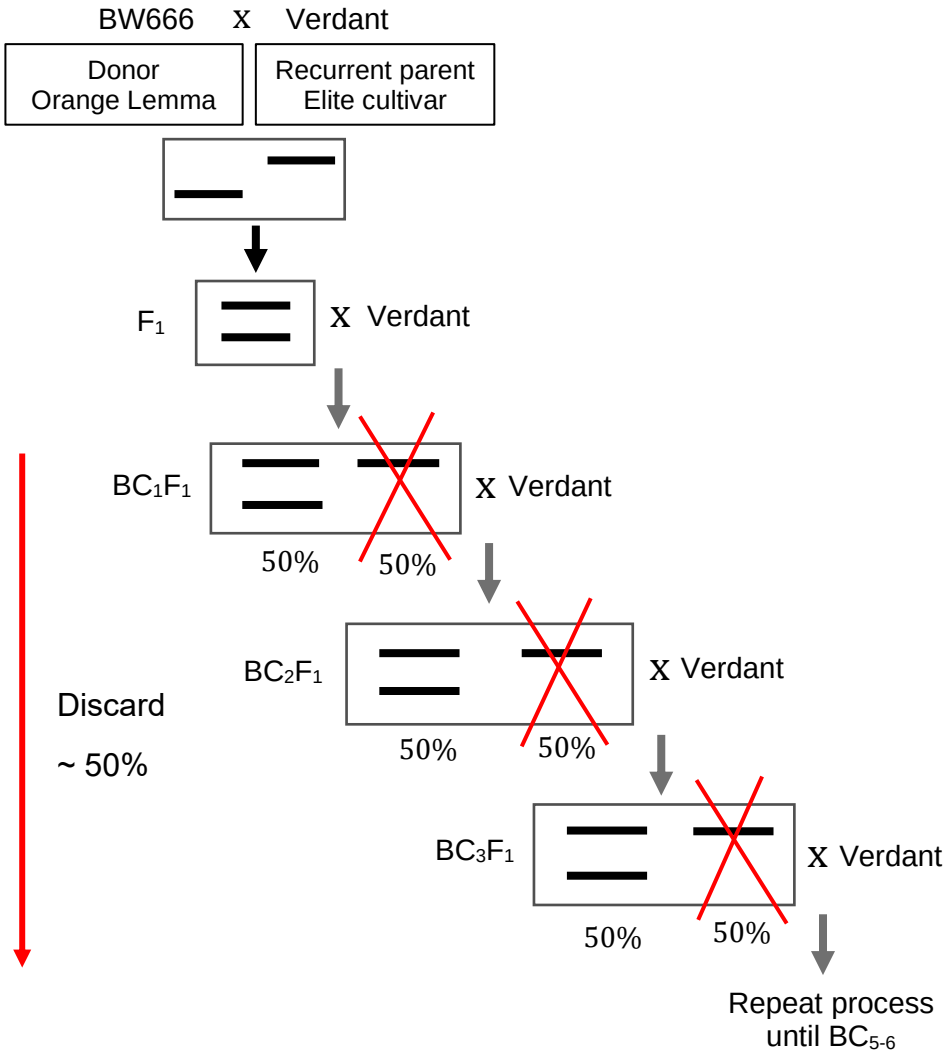
MAS in plants refers to a process whereby a molecular marker is used for indirect selection of a genetic determinant of a trait of interest. The method involves selecting plants carrying specific (mutant) alleles or genomic regions involved in the expression of traits of interest through the application of molecular markers [2]. An array of molecular markers systems has been developed for this purpose. Markers are widely used for marker-assisted backcrossing as they offer several advantages compared to conventional phenotypic selection:

(i) for traits that are difficult to evaluate phenotypically;
(ii) for recessive traits;
(iii) for stacking or pyramiding of different traits;
(iv) for stacking different alleles for the same trait.
(v) for optimal selection of recombinants by minimizing linkage drag or through the use of flanking markers;
(vi) significant time gain as the marker assay can be carried out at the seed(ling) stage;
(vii) significant reduction in the breeding cycles as there is no need for selfing cycle in the backcrossing scheme to allow for the expression of the phenotypes under homozygous states of the recessive alleles
(viii) molecular marker assays are independent of the environment

Backcrossing is a traditional breeding method commonly used to transfer alleles at one or more loci from a donor to an elite variety. It provides a precise way to improve a variety that excels in many traits but otherwise lacks an important characteristic. In a recurrent backcrossing scheme plants with the genotype (Aa) will be further backcrossed to the recurrent recipient parent (AA) to restore the elite genetic background of the recurrent parent plus the added mutant trait (aa). Recessive traits such as the orange lemma are more difficult to select for in backcross breeding, because their expression is masked by the dominant allele in each backcross. Therefore, an additional round of self-pollination is needed after each backcross to expose homozygous-recessive plants. Using a co-dominant marker, the heterozygote plants can be readily identified precluding the need for the selfing crop cycle practiced in the conventional backcrossing. Application of MAS for transfer of the recessive orange lemma trait from the donor BW666 to the barley elite cultivar Verdant through backcrossing is illustrated in Fig 1. When using MAS, the number of crop cycles is reduced, and the breeding programme can be efficiently scaled down to focus on fewer lines resulting in genetic gain while saving time, space and labor for growing plants. Note that a final selfing cycle will be applied to produce a stable homozygous line (AA) for the added trait.

Most economically important traits induced through mutagenesis are recessive with high heritability controlled by one or a few genes. Positional identification of gene(s)/QTL underlying key mutant traits will therefore lead to more efficient MAS schemes.

A.



B.

Generation	% recurrent	% donor
F_1	50	50
BC_1	75	25
BC_2	87.5	12.5
BC_3	93.7	6.3
BC_4	96.9	3.1
BC_5	98.4	1.6
BC_6	99.2	0.8

FIG 1.A. Introgression of orange lemma trait into an elite variety Verdant through Marker-Assisted Backcrossing. B. Percentage of recurrent and donor genome through backcrossing.

1.2. Genotyping procedures

Single Nucleotide Polymorphism (SNPs) and insertion-deletions (InDels) are common genetic polymorphisms in plants and other organisms. SNPs refer to a single base change (A, T, G or C) in a DNA sequence and occur when a single nucleotide in the genome of an individual or plant variety differs between members of that species or between paired chromosomes of that individual or plant variety. SNPs and InDels are common types of mutations in genes that give rise to multiple alleles. DNA sequence polymorphisms can also occur in non-genic regions. SNPs and InDels occur naturally through spontaneous mutation but can also be induced through physical or chemical mutagenesis techniques as practiced in experimental or applied mutagenesis [3].

Genetic variation between plant varieties or between alleles within the same plant can be detected using molecular tools. The process of determining differences in the genetic make-up of an individual by examining its DNA sequence and comparing it to another individual's sequence is referred to as genotyping. Since the 1980's a wide variety of molecular tools and genotyping methods have been developed to detect polymorphisms such as SNPs or InDels for use in plant breeding [4,5].

Continuous progress in high-throughput genomic technologies has resulted in numerous SNP genotyping methods [6]. These platforms combine a variety of allele discrimination techniques such as hybridization with allele-specific probes or allele-specific Polymerase Chain Reaction (PCR) amplification [7] with different detection methods such as colorimetry, fluorescence, or chemiluminescence. Overall, two major genotyping platforms can be distinguished: genome-wide SNP genotyping platforms versus genotyping for a small number of specific loci. The former platform typically uses multiplex, high-density SNP arrays and can be used for genetic mapping, estimating genetic diversity, amongst other applications. The latter is a single-plex platform that can be used for gene identification and marker-assisted selection.

To date, various single-plex genotyping platforms have been developed. One such platform is the Kompetitive Allele Specific Amplification or KASP™ technology. KASP™ is an allele-specific PCR-based technique that enables bi-allelic scoring of SNPs and InDels at specific loci using fluorescent detection. KASP™ assays are widely used in plant and animal breeding programs in the public and private sector [8]. Important factors in choosing an appropriate platform for genotyping include, *inter alia*, the intended application, ease of use, the number and quality of data points produced, flexibility, assay development requirements and cost per sample or data point.

The protocol described here is a single-plex genotyping assay based on allele-specific PCR amplification for marker-assisted breeding of orange lemma trait. Two primers are designed that differentially amplify a wild type allele and the BW666 mutant allele. The mutant and wild type allele have different lengths for easy detection on agarose gels. The protocol is technically simple and low-cost using standard PCR and agarose gel electrophoresis for easy detection and scoring. A marker kit with seed of BW666, as source of the orange lemma trait, and the two PCR primer pairs is available upon request.

2. MARKER DEVELOPMENT AND VALIDATION WORKFLOW

The key steps in the development and validation of the molecular marker assay for the barley orange lemma trait are summarized in Fig. 2. A candidate-gene approach can be followed for positional identification of causative mutations or variants. This is a relatively quick and easy way to identify causative sequence variants. A prerequisite for the candidate-gene approach is that there is *a priori* knowledge of the gene's biological functional impact on the trait or disease in question [9]. The rationale behind focusing on allelic variation in specific genes or genomic regions is that causative mutations can be uncovered which directly impact the function of the gene in question, and lead to the phenotype being studied. These conditions were met in case of barley orange lemma as brown midrib mutations with a similar phenotype in maize, sorghum and millet have been characterized at the molecular level and are known to be caused by mutations in genes of the lignin biosynthesis pathways [10]. In the absence of a priori knowledge of candidate genes, genome-wide scans can be performed to identify causative variants and establish marker-trait associations [11].

Using a candidate-gene approach, the causative mutation underlying the orange lemma trait was identified by Sanger sequencing of candidate gene(s) in a panel of barley mutant genetic stocks and their wild type parents.

Next, a marker assay was developed for the causative variant identified in BW666, in this case a SNP. The assay should be able to differentiate homozygous from heterozygous individuals, i.e. the marker is co-dominant, to be effective as a speed breeding tool in a backcross breeding programme (foreground selection). We opted to develop an assay based on differential PCR amplification of the wild type vs mutant allele. A length polymorphism was created to enable clear separation of the two alleles on a standard 2% agarose gels which is technically simple and low-cost. The resulting marker is a functional or diagnostic marker.

Then, marker-trait linkage is established using the marker assay, for example in an F2 segregating population derived from a backcross. The phenotype of F1 plants will allow to determine whether the mutant trait is dominant or recessive. Genotyping and phenotyping the F2 lines will enable to establish the marker-trait association.

For the marker and trait to be useful for applied breeding, the marker-trait associations must be validated in different elite genetic backgrounds. The procedure involves crossing the donor BW666 and recipient elite varieties/line(s) and establishing marker-trait associations using the marker assay in, for example, an F2 population.

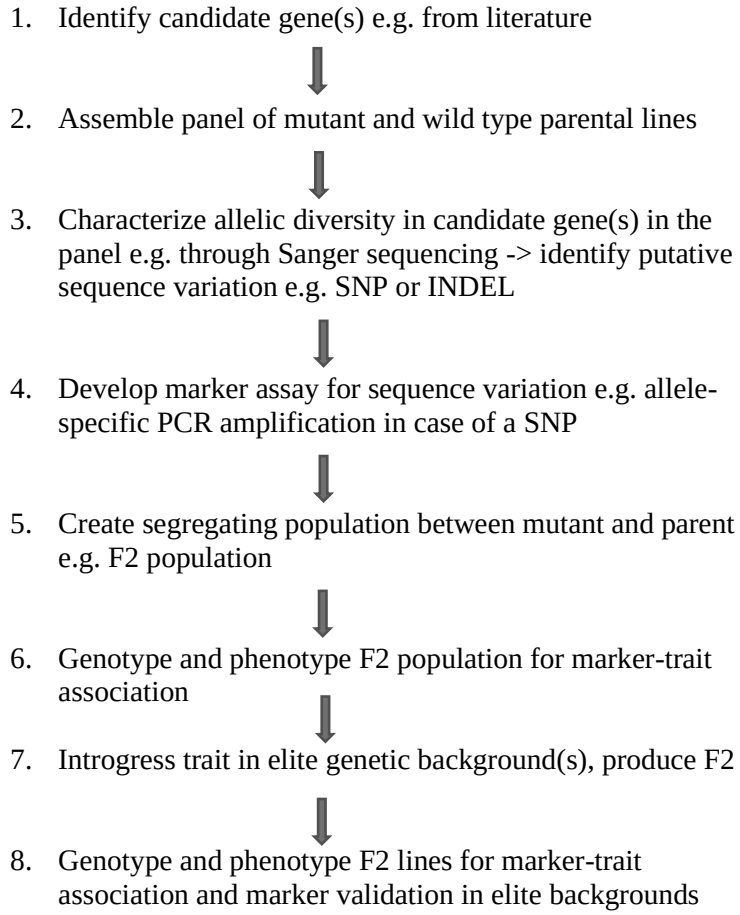


FIG. 2. Steps in the development and validation of the barley orange lemma marker assay using a candidate-gene approach

3. MATERIALS

3.1. Plant material

The diagnostic marker was developed for the orange lemma trait of the mutant barley (*Hordeum vulgare* L.) genetic stock BW666 (Fig. 3). BW666 is a Bowman backcross-derived line with the *rob1.a* gene (Bowman*8/4.1 msg24 MM). This germplasm is stored at the National Small Grains Collection (USDA) as accession GSHO 2069 and at NordGen as NGB 20752. Originally the mutant allele appeared spontaneously in a barley landrace accession (CIho 5649) collected in the Krasnodar region, Russia [12]. BW666 contains a SNP in the CAD2 gene which renders this gene non-functional (Plant Breeding and Genetics Laboratory, manuscript in preparation). Bowman is a USA, 2-rowed spring-type barley cultivar with semi-smooth awns adapted to dry land conditions [13].

FIG. 3. Barley orange lemma phenotypes: A. Flowering head from the mutant stock BW666 (left, orange lemma) and wild type barley cv Bowman (right, green); B. Immature seeds of BW666 (left, orange lemma) and wild type Bowman (right, green); C. stem of BW666 (left, orange) and Bowman (right, green).

3.2. Reagents and equipment

- Tissue Lyser II (Qiagen) for tissue homogenization
- Single and multi-channel pipettes and disposable pipette tips
- Water bath
- Protective clothing (lab coat, gloves, safety glasses)
- Liquid nitrogen
- PCR machine with heated lid (Applied Biosystems VERITI 96 Thermal Cycle)
- Agarose gel electrophoresis equipment
- Agarose
- 10x TBE buffer stock solution (dilute to 0.5x for gel electrophoresis)
- Power supply
- UV trans illuminator and box (GelDoc XR BioRad)
- 1.5 mL microfuge tubes and 0.2 mL 8-strip PCR tubes
- Bench top micro centrifuge for 1.5 mL microfuge tubes and PCR tubes
- Scissors or punches for sampling plant tissue
- Nanodrop Spectrophotometer (ND-1000 Thermo Fisher)
- DNeasy Plant Mini Kit (Qiagen; Cat No. 69104)
- DNeasy 96 Plant Kit (Qiagen; Cat No. 69181)
- Takara Ex Taq® DNA Polymerase Hot-Start Version (TaKaRaClontech: Cat No: RR006A)
- 1 kb Plus DNA ladder (Invitrogen; Cat No 10787-018)

4. EXPERIMENTAL PROCEDURES

The genotyping procedure described here consists of five steps:

(i) Growing/Planting of barley
(ii) Tissue collection
(iii) DNA extraction
(iv) Genotyping assay

(v) Data analysis

The steps below describe the low-cost genotyping of the BW666 Orange lemma mutation using standard PCR technology and agarose gel electrophoresis for detection, as routinely practiced in the Joint FAO/IAEA Division Plant Breeding and Genetics Laboratory.

4.1. Growing/planting of barley

Barley seed can be sown in individual pots or in trays in the greenhouse as required. For higher throughput such as in MAS applications, seedlings can be grown in trays in a 96-cell format to streamline planting with downstream tissue collection and DNA extraction processes which follows a 96-well microplate format (Fig. 4). Plant samples are labelled using appropriate data sheets (Annex 2). At the PBGL, seeds resulting from crosses are incubated in a petri-dish on moist filter paper for 1 to 2 days at 4°C in the dark to ensure sufficient imbibition. The petri-dish with seeds is then transferred to the greenhouse or incubator for further growth in light at ca 25°C for an additional 2-4 days until the seeds are germinated. The seeds are kept moist throughout this germination step. Once germinated, the seedlings are transplanted to individual pots or trays containing soil and peat moss (3:1:1 soil:peat:sand with 32g/100kg phosphate and 27g/100kg urea fertilizer). Note that a vernalization treatment may be required in case of winter-type barley for proper growth, flowering and seed set.

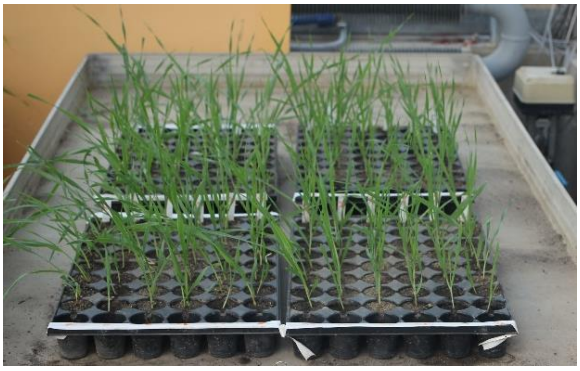


FIG. 4. Barley seedlings grown in trays in a 96-cell configuration.

4.2. Tissue collection

In this protocol, ca 10-70 mg leaf tissue is harvested and placed in a collection tube provided with the DNA extraction kit. Prior to adding the tissue, two 3-mm tungsten beads are added to the collection tube. For the genotyping protocol described here, leaf tissue is collected from individual plants using punches (leaf discs) or scissors (leaf strips) as preferred. Leaf tissue can be collected at different barley growth stages. For genotyping, we recommend collecting young leaf tissue e.g. at the 3- to 5-leaf stage for improved yield. It is important to determine the appropriate number of leaf discs or the size of the leaf strip per sample before commencing with tissue collection. Example tissue collections is shown in Fig. 5:

- a 5.0x0.5 cm leaf strip from a young leaf of a barley plant at flowering stage weighs ca 60-70 mg
- one 8 mm leaf disc from a seedling at the 3-5 leaf stage weighs approximately 10 mg

We recommend collecting maximum 70 mg per sample and to cut long strips in smaller pieces of maximum 1 cm to fit the collection tube and ensure thorough grinding of the leaf tissue in the subsequent step. The collection tubes with the tungsten beads and tissue can be placed at -80°C for subsequent DNA extraction.

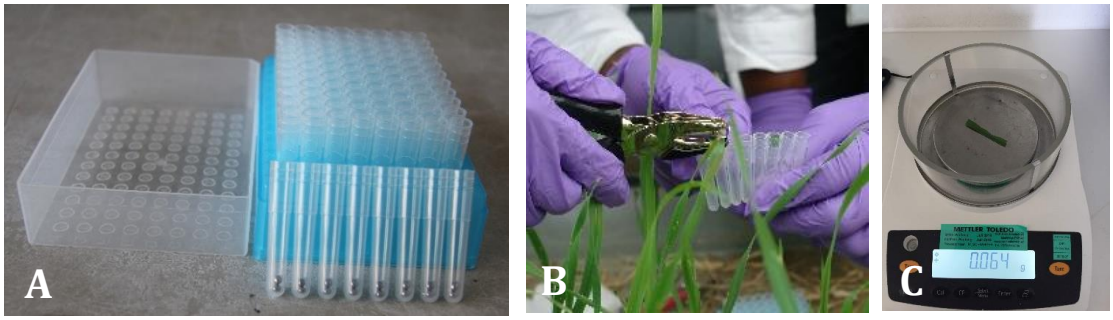


FIG. 5. A. Collection tubes with two beads for DNA extraction in 96-tube format; B. leaf tissue sampling using punches; C. weighing leaf strips.

4.3. DNA extraction and quality control

For DNA extraction, follow the steps in the protocol provided by the manufacturer (Qiagen) for single-tube or 96-tube DNA extraction for higher throughput as required. It is left to the user of the protocol to adapt the extraction method to the available DNA extraction kit. Grinding is a critical step in the DNA extraction protocol. At PBGL, the Tissue Lyser II is used for grinding the tissues (Fig. 6. A). The collection tubes with the tissues are removed from the -80°C and placed in a Styrofoam box containing enough liquid nitrogen to snap freeze the samples. The leaf tissue is ground into a fine powder using the following settings: frequency 20/s and time 30 s. Repeat as required until a fine powder is obtained. Typically, 2 to 3 rounds of grinding are sufficient (Fig. 6. B).

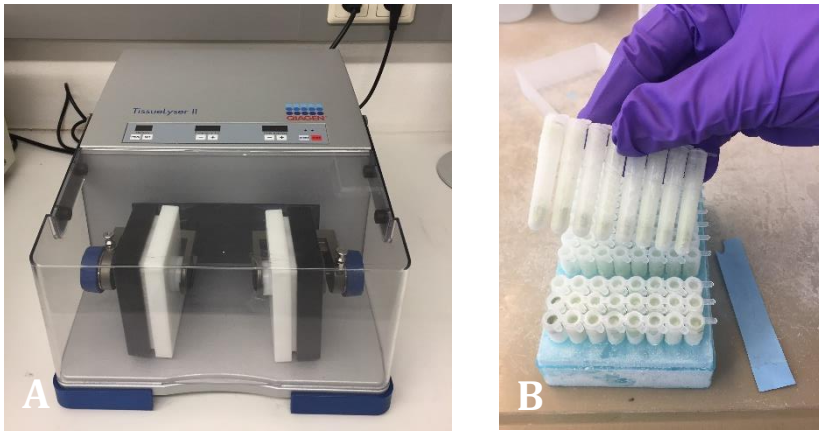


FIG. 6. A. Tissue Lyser II for disruption of leaf tissue. B. Finely ground leaf tissue in collection tubes.

The DNA is eluted in 100µL water. DNA is quantified using a Nanodrop Spectrophotometer using 1 µL. The ratio of absorbance at 260 nm and 280 nm is used to assess the purity of the DNA. A ratio of ~1.8 is generally accepted as 'pure' for DNA; if the ratio is appreciably lower, it may indicate the presence of protein, phenol or other contaminants that absorb strongly at or

near 280 nm. Typical yields are ca 0.5 to 1 µg from 5-10 mg fresh young, leaf tissue and on average 5 to 7 µg from 60-70 mg leaf tissue collected at flowering stage.

Prior to setting up the PCR reactions, the DNA concentration is adjusted to 5ng/µL using sterile, distilled water. The integrity of the DNA is checked by running 100 ng DNA on a standard 0.8% agarose gel next to a DNA size marker. A distinct band of high molecular weight (size ca 15 kb) should be obtained (Fig. 7). In our experience, slight degradation of the DNA, as visualized by a smear on the gel, does not interfere with the genotyping assay.

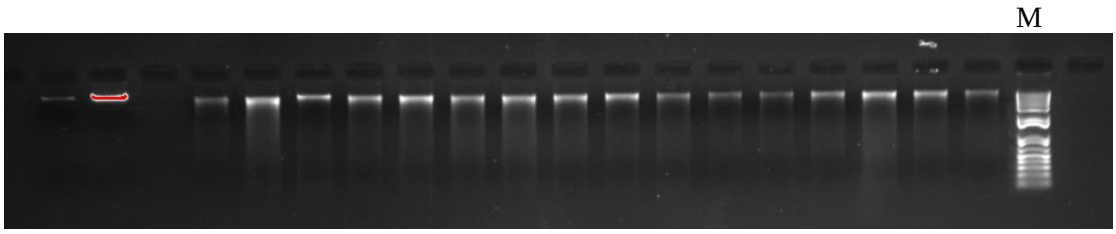


FIG. 7. Agarose gel electrophoresis of barley genomic DNA next to DNA size marker M.

4.4. Low-cost allele-specific PCR protocol

Recommendations before you start:

- Familiarize yourself with the PCR and gel electrophoresis workflow summarized in Annex 1 to ensure that you understand all the steps.
- The below protocol and reactions are described for PCR analysis of eight plant samples.
- Four PCR primers are used in the protocol: primers 1 and 2 amplify the 164-bp mutant allele while primers 3 and 4 amplify the 212-bp wild type allele.
- Two separate PCR reactions are set up for each DNA sample, therefore two 8-strip PCR tubes with duplicate plant DNA samples are prepared.
- An initial PCR master mix excluding primers is prepared for 20 reactions which is divided in two.
- The PCR master mix is then split in two and the primers are added to their respective mix, enough for 10 PCR reactions.
- After PCR, 5 µL of each PCR reaction for a given plant DNA sample is combined in one tube, loading buffer is added and the sample is loaded on a 2% agarose gel.
- Always include known homozygous wild type and mutant genotypes as controls for the PCR and as reference for scoring.

4.4.1. Setting up the PCR reactions

A standard PCR machine with heated lid is used (Fig. 8). We use the Takara Ex Taq® DNA Polymerase Hot-Start Version for amplification. PCR reactions are performed in 8-tube strips in a final volume of 15 µL. The PCR conditions are: (i) initial denaturation: 95°C (30 s); (ii) 35 cycles of: 95°C (20 s); 60°C (30 s); 72°C (60 s); (iii) 72°C (300 s); (iv) 8°C (∞)

FIG. 8. PCR machine with heated lid to set up the PCR reactions.

STEPS

1. Thaw the template DNAs (5ng/µL) and the following PCR components on ice:
    - 10x Ex Taq buffer,
    - 2.5 mM dNTP mix,
    - the four primers (each primer is 10 µM)

    Note: keep the TaKaRa Ex Taq HS (5 U/µL) at -20°C until use

2. Prepare a PCR master mix with all the components except for the template DNA and the primers in a 1.5 mL microfuge tube. For example, for 20 PCR reactions:

	For 1 reaction	**For 20 reactions**
H_2O	10.46	209.2
Buffer (10x)	1.5	30
dNTPs (2.5 mM)	1.2	24
Taq HS (5 u/µL)	0.0375	0.75
Forward primer (10 µM)	0.15	1.5
Reverse primer (10 µM)	0.15	1.5
DNA (5 ng/µL)	1.5	
Final volume (µL)	15	

3. Pipet all the components to the master mix, gently mix

4. Split up the PCR master mix in two equal parts 1 and 2 in a 1.5 ml microfuge tube; label 1 and 2

5. Thaw all four primers, keep on ice.

6. Add 1.5 µL of primers Fw_1 and Mu_re to PCR master mix 1 and 1.5 µL of primers Fw_3 and WT_re to PCR master mix 2; mix, store on ice

7. Label 0.2 ml 8-tube PCR strips with the name of the DNA sample; prepare strips in duplicate with two tubes for each DNA sample; label each strip 1 and 2 according to the PCR master mix 1 and 2 that will be added

8. Pipet 13.5 µL of each master mix 1 and 2 to the PCR reaction tubes

9. Add 1.5 µL template DNA (5 ng/µL) to the tubes

10. Close the PCR reaction tubes, place into the PCR machine and start the programme

4.4.2. Agarose gel electrophoresis

Standard agarose gel electrophoresis equipment is used for detection of the PCR fragments including the gel casting tray, the gel running tray, 10x stock TBE electrophoresis buffer and a power supply. The PCR amplicons (2 times 5 µL) are separated on a 2 % agarose gel (Fig. 9). A loading dye (1.0 µL of a 10x stock) is added to the sample to track migration of the DNA in the gel. A DNA size marker is loaded next to the samples. The agarose gel is stained using ethidium bromide (EtBr) for visualizing the amplified DNA fragments under a UV illuminator box. Alternatively, GelRed can be used for staining or any other available staining dye.

FIG. 9. Agarose gel electrophoresis.

STEPS

1. Prepare 2% agarose gel by adding 6 g agarose to 300 mL 0.5x TBE buffer in a 500 mL bottle (adjust the volume based on the size of the tray)
2. Heat in a microwave oven until the agarose is completely dissolved
3. Cool down to approximately 60°C in a water bath
4. Add 6 µL of EtBr stock solution (10 mg/mL) and mix carefully
5. Pour gel into a casting tray with the appropriate comb
6. Let the gel solidify at room temperature (for at least 30 min)
7. Label tubes for gel loading
8. Add 5 µL of PCR reaction 1 and 5 µL of PCR reaction 2 to the same tube; add 1 µL of 10x loading dye. *Note: use a new clean tip for each PCR sample*
9. Quick spin the tubes to bring liquid to bottom
10. Place gel tray into gel tank
11. Add 0.5x TBE buffer to the gel tank until the gel is completely covered by buffer
12. Carefully remove the combs
13. Load the samples
14. Run gel for about 1.5 – 2 hours at 130V
15. Place the gel on a UV transilluminator and take picture using Gel Doc system

5. DATA ANALYSIS

5.1. Scoring and data recording

The marker for the BW666 orange lemma mutant allele is 164 bp while the wild type CAD2 allele is 212 bp.

Genotypes are scored based on the banding pattern as homozygous mutant (one band of 164 bp), homozygous wild type (one band of 212 bp), or heterozygous (one band each of 212 and 164 bp) and entered in a data sheet (see Annex 2) as following:

- AA: homozygous wild type
- Aa: heterozygous
- aa: homozygous mutant

The presence of a clearly visible band of the predicted size is considered a positive reaction and indicates the presence of the wild type or mutant allele. Bands that are absent or faint are considered a negative reaction and indicate the absence of the allele. Samples with unclear banding pattern are scored as missing data (should be less than 5%).

A typical result of F2 progeny analysis from a BW666 x Verdant cross showing donor and recipient genotypes and F2 progeny genotypes is illustrated in Fig. 10.

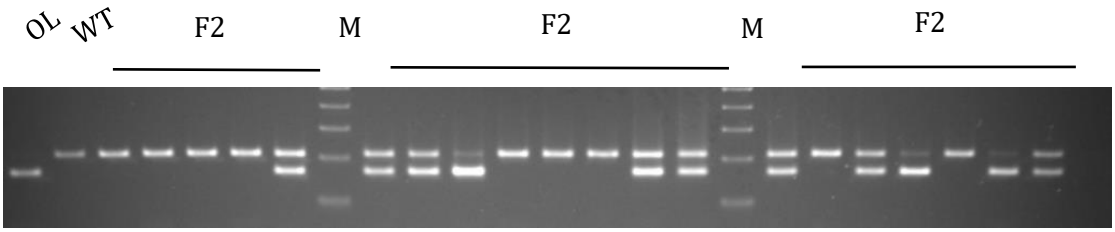


FIG. 10. Genotyping of F2 progeny differentiating homozygous WT, homozygous OL and heterozygous plants; WT = wild type allele; OL = orange lemma mutant allele; F2 = F2 plants; M = size marker (100-bp ladder).

5.2. Statistical analysis of segregation ratios: the Chi-square test

The orange lemma trait in BW666 (allele *rob1*.a) is a mono-factorial recessive trait. The genotype frequency for a single recessive trait in an F2 population is 1:2:1 corresponding to 25% homozygous wild type (AA), 50% heterozygous (Aa) and 25% homozygous mutant allele (aa). Therefore, in an F2 population derived from a cross between BW666 (aa) and a wild type barley (AA) the expected genotypic frequencies as determined through the marker analysis are 1:2:1.

To check for conformance of the observed segregation ratios with the expected segregation ratios or genotypic frequencies, a chi-square goodness-of-fit test can be performed. Briefly, the chi-square goodness-of-fit test is a statistical test used to see if observed values for a set of data are similar or significantly different from expected values. The test can be applied when studying the inheritance of certain traits, for example to determine whether a trait follows simple Mendelian ratios or whether something more complex is happening. In the case of barley orange lemma, the genotypes are expected to follow ratios for a single recessive trait (1:2:1). In a chi-square test, the null hypothesis (H_0) states that any differences between the observed and expected genotypic frequencies are due to chance only. The following formula is used:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

where O = the observed value and E = the expected value.

Here, the null hypothesis states that the data follow a 1:2:1 segregation. Example data from three F2 populations derived from crossing BW666 with its parent Bowman and two elite barley cultivars is summarized in Tables 1 and 2. All three F2 populations were genotyped using the marker described in this protocol and χ^2 values calculated for each F2 progeny using above formula. The observed values *O* are the genotypes that occur in each of the three categories: AA, Aa or aa. The expected value *E* is derived by multiplying the total number of observations with their expected frequencies for each category as shown in Table 1. Next, the χ^2 value is used to find the p-value in a χ^2 distribution table (Annex 3). For this, the degrees of freedom (DF) must be determined. The DF is the number of categories minus 1. In this example, there are three categories, i.e. the three different marker genotypes. The DF is therefore 3-1= 2. The level of significance is set at 0.05 which is a commonly adopted standard for biological applications. Using the χ^2 values, the probability (*p*)-values can now be determined from the chi-square distribution table in Annex 3. The *p*-value will indicate whether there is a significant difference between the expected and observed data sets. If $p>0.05$, there is no significant

difference between the observed and expected genotypic frequencies and the null hypothesis can be accepted. If $p<0.05$, then the null hypothesis is rejected.

In the below example analysis of the BW666xBowman F2 progeny, the *p*-value is (much) greater than 0.05, therefore the null hypothesis can be accepted. One can conclude that the marker data fit the expected segregation ratio for a mono-factorial, recessive trait (1:2:1). A similar analysis was conducted for BW666xOptic F2 and for BW666xVerdant F2 with results summarized in Table 2.

Table 1. Chi-square calculation for BW666 x Bowman F2

BW666 x Bowman						
marker genotype	observation	expected frequency	EXPECTED	deviation	$(O-E)^2$	$(O-E)^2/E$
AA	14	0.25	13	1	1.00	0.08
aa	13	0.25	13	0	0	0.00
Aa	25	0.5	26	-1	1.00	0.04
	52					$\chi^2=0.12$

Table 2. Summary chi-square and *p*-value calculations for different barley BW666 F2 crosses

F2 progeny	**chi-square value**	
BW666xBowman BW666xOptic BW666xVerdant	0.12 1.19 0.63	
Segregation ratio DF=2	1:2:1	
P<0.05 P<0.01	5% 1%	5.991 9.21

6. VERIFYING MARKER-TRAIT ASSOCIATIONS IN ELITE BACKGROUND

For a marker to be useful for applied breeding, the marker-trait associations must be verified in elite genetic backgrounds. In the case of orange lemma, the barley mutant stock BW666 was crossed to several elite barley varieties. The F2 lines are genotyped and phenotyped for presence of the orange lemma trait. The genotyping can be performed at the seedling stage as described above. The plants can be phenotyped visually at the booting or grain filling stage for the presence of an orange lemma. Alternatively, the stem can be peeled and visually inspected; an orange/brownish colored stem is indicative of the orange lemma trait compared to the wild type stem which is green. The phenotypes predicted by the marker assays are then compared to the phenotypes that have been scored visually. A typical result is presented in Annex 4 for F3 lines derived from a BW666 x Optic cross which shows that the Y28 marker assay accurately predicted the orange lemma phenotype, i.e. there was a 100% match between the phenotype predicted by the marker assay and the observed phenotype.

7. DATA INTERPRETATION AND TROUBLESHOOTING

The marker assay relies on selective amplification of a wild type versus the specific mutant allele present in BW666. The selective amplification is based on a 2-bp sequence difference between the wild type and mutant allele at the 3'end of the primer. Given that the primer pairs for the wild type allele are designed on sequence information from one specific barley genotype, it is possible that the PCR conditions or primers presented in this protocol are suboptimal when amplifying the wild type allele from unrelated barley genotypes, due to sequence polymorphisms at the primer binding site. This may affect amplification of the wild type allele in the recipient line.

So far, the protocol has been validated and proved highly reliable in two different, unrelated barley genetic backgrounds: cv Optic and cv Verdant. Nevertheless, some caution is warranted when applying the assay to an unknown barley genotype for the first time. In case amplification is suboptimal using the conditions and primers described here, the conditions for PCR and/or primers sequence can be further optimized. To rule out any artefacts due to experimental error or variation, it is important to always include a homozygous mutant and homozygous wild type as controls in the experiment.

8. SAFETY INFORMATION

- Wear lab coat, gloves and goggles throughout the DNA extraction and gel electrophoresis procedures
- Liquid nitrogen (boiling point: -195.8 °C) and other cryogenic materials can cause severe frostbite upon contact with skin surfaces. Wear proper personal protective equipment to prevent contact or inhalation
- When using or decanting liquid nitrogen, a face shield or safety goggles must be used; thermal protective gloves; full length trousers/pants, and footwear that cover the entire foot

- EtBr is mutagenic and carcinogenic; gloves must always be worn during the gel electrophoresis procedure, when taking pictures and disposal

9. MARKER KIT

A molecular marker kit with the following components is available upon request:

1. Seed of the barley mutant stock BW666 with the orange lemma mutant allele

2. Four primers used for selective amplification of the wild type and mutant alleles:

    (1) Fw1
    (2) Mu_re
    (3) Fw3
    (4) Wt_re

    - primers (1) and (2) amplify the 164-bp mutant allele = primer set 1
    - primers (3) and (4) amplify the 212-bp wild type allele = primer set 2

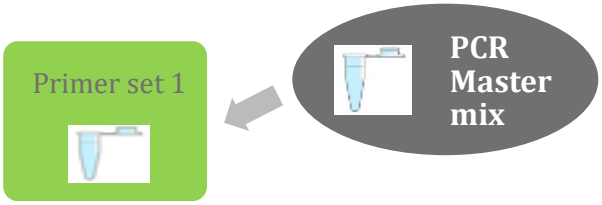
3. This protocol

1. PCR and gel electrophoresis workflow

1. Prepare the PCR master mix & split in two equal parts

Reagents	1x	20x
H2O	10.46	209.20
Buffer (10X)	1.50	30.00
dNTPs (2 mM)	1.20	24.00
Taq (5 units/µL)	0.0375	0.75
Total		264
split in two		132

2. Add primer Set 1 and 2 to each half of the PCR master mix



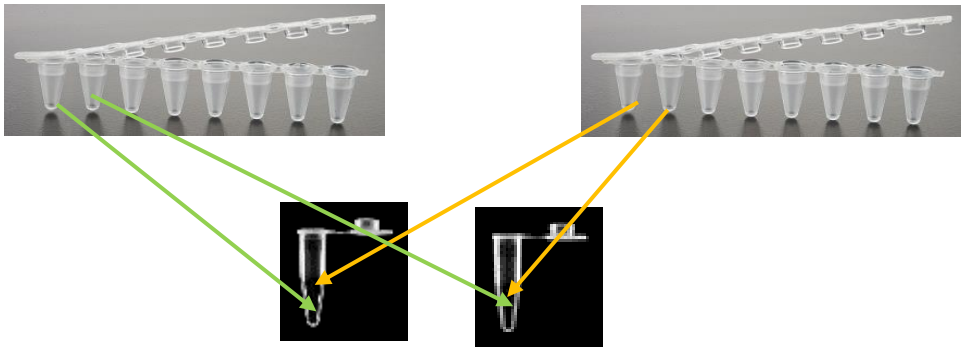
	1x	10x
Fw_1 (10 µM)	0.15	1.50
Mu_re (10 µM)	0.15	1.50
		132.0

	1x	10x
Fw_3 (10 µM)	0.15	1.50
WT_re (10 µM)	0.15	1.50
		132.0

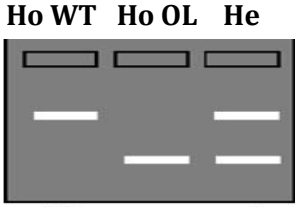
3. Add 1.5 µL DNA (5 ng/µL)

4. Run PCR

5. Combine the two PCR reactions into one tube



6. Run the combined PCR reactions on a 2% agarose gel



Co-dominant marker

2. Data sheets

The following data sheets are useful in the PBG Laboratory and we recommend them for keeping a record of all information during the entire genotyping procedure from the growth of plants in trays in a 96-cell format through tissue collection and DNA extraction in 96-collection tubes or microwell plates, agarose gel electrophoresis and data entry into a genotyping and phenotyping score sheet.

96-well/cell map for barley growth, tissue collection & DNA extraction

Name:

Plate nr:

Project:

Date planting, tissue collection, or DNA extraction:

Date and location storage:

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B												
C												
D												
E												
F												
G												
H												

Genotyping scoresheet

52xO	Sample ID	DNA	phenotyp	genotype		52xO	Sample ID	DNA	phenotyp	genotype
1	190	Y	WT	WT-he		29	218	Y	WT	WT-ho
2	191	Y	WT	WT-ho		30	219	Y	WT	WT-ho
3	192	Y	OL	OL-ho		31	220	Y	WT	WT-he
4	193	Y	WT	WT-he		32	220	Y	WT	WT-ho
5	194	Y	WT	WT-he		33	222	Y	OL	OL-ho
6	195	Y	OL	OL-ho		34	223	Y	WT	WT-he
7	196	Y	WT	WT-he		35	224	Y	WT	WT-he
8	197	Y	OL	OL-ho		36	225	Y	WT	WT-he
9	198	Y	WT	WT-he		37	226	Y	WT	WT-ho
10	199	Y	OL	OL-ho		38	227	Y	OL	OL-ho
11	200	Y	WT	WT-ho		39	228	Y	WT	WT-ho
12	201	Y	DIED	WT-ho		40	229	Y	OL	OL-ho
13	202	Y	OL	OL-ho		41	230	Y	WT	WT-he
14	203	Y	WT	WT-he		42	231	Y	OL	OL-ho
15	204	Y	WT	WT-ho		43	232	Y	OL	OL-ho
16	205	Y	WT	WT-ho		44	233	Y	OL	OL-ho
17	206	Y	WT	WT-he		45	234	Y	WT	WT-he
18	207	Y	WT	WT-he		46	235	Y	WT	WT-he
19	208	Y	WT	WT-he		47	236	Y	WT	WT-ho
20	209	Y	WT	WT-he		48	237	Y	WT	WT-he
21	210	Y	WT	WT-ho		49	238	Y	WT	WT-he
22	211	Y	OL	OL-ho		50	239	Y	WT	WT-he
23	212		WT			51	240	Y	WT	WT-he
24	213	Y	WT	WT-ho		52	241	Y	WT	WT-ho
25	214	Y	OL	OL-ho		53	242	Y	WT	WT-ho
26	215	Y	WT	WT-he		54	243	Y	OL	OL-ho
27	216	Y	OL	OL-ho		55	244	Y		WT-ho
28	217	Y	WT	WT-he						

Agarose gel sheet

Agarose Gel Format																										
Name:													% agarose:													
Date:													# of lyers/ # of wells:													
Project:													# of samples:													
Materials:													Running conditions:													
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	LINE #
																										Sample ID
27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	LINE #
																										Sample ID
53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	LINE #
																										Sample ID

3. Chi-square distribution table

The *p*-values can be estimated from a chi-square distribution table once the chi-square values have been calculated. If the *p*-value is higher than 0.05, then the null hypothesis can be accepted. If the *p*-value is lower, then the null hypothesis can be rejected.

Chi-Square Table

	PROBABILITY (p)								
DF	**0.95**	**0.9**	**0.5**	**0.2**	**0.1**	**0.05**	**0.025**	**0.02**	**0.01**
1	0.00393	0.0158	0.455	1.642	2.706	3.841	5.024	5.412	6.635
2	0.103	0.211	1.386	3.219	4.605	5.991	7.378	7.824	9.21
3	0.352	0.584	2.366	4.642	6.251	7.815	9.348	9.837	11.345
4	0.711	1.064	3.357	5.989	7.779	9.488	11.143	11.668	13.277
5	1.145	1.61	4.351	7.289	9.236	11.07	12.833	13.388	15.086
6	1.635	2.204	5.348	8.558	10.645	12.592	14.449	15.033	16.812
7	2.167	2.833	6.346	9.803	12.017	14.067	16.013	16.622	18.475
8	2.733	3.49	7.344	11.03	13.362	15.507	17.535	18.168	20.09
9	3.325	4.168	8.343	12.242	14.684	16.919	19.023	19.679	21.666
10	3.94	4.865	9.342	13.442	15.987	18.307	20.483	21.161	23.209
11	4.575	5.578	10.341	14.631	17.275	19.675	21.92	22.618	24.725
12	5.226	6.304	11.34	15.812	18.549	21.026	23.337	24.054	26.217
13	5.892	7.042	12.34	16.985	19.812	22.362	24.736	25.472	27.688
14	6.571	7.79	13.339	18.151	21.064	23.685	26.119	26.873	29.141
15	7.261	8.547	14.339	19.311	22.307	24.996	27.488	28.259	30.578
16	7.962	9.312	15.338	20.465	23.542	26.296	28.845	29.633	32
17	8.672	10.085	16.338	21.615	24.769	27.587	30.191	30.995	33.409
18	9.39	10.865	17.338	22.76	25.989	28.869	31.526	32.346	34.805
19	10.117	11.651	18.338	23.9	27.204	30.144	32.852	33.687	36.191
20	10.851	12.443	19.337	25.038	28.412	31.41	34.17	35.02	37.566

Accept null hypothesis (variation due to chance only)

Reject null hypothesis (not chance only)

4. Verification of marker-trait association in an elite barley cultivar

Example data of F3 progeny (horizontal, nr 1 through 16) derived from different BW666xOptic F2 lines (vertical, nr 5, 7, 8, 23, and 24); x indicate missing data

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
5	Genotype	Aa	Aa	Aa	aa	aa	aa	aa	aa	Aa	AA	AA	Aa	AA	Aa	aa	aa
	Predicted phenotype	WT	WT	WT	OL	OL	OL	OL	OL	WT	WT	WT	WT	WT	WT	OL	OL
	Observed phenotype	WT	WT	WT	OL	OL	x	OL	OL	WT	WT	WT	WT	WT	x	x	OL
7	Genotype	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa
	Predicted phenotype	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL
	Observed phenotype	OL	OL	OL	OL	OL	OL	OL	x	OL	OL	OL	x	OL	OL	OL	OL
8	Genotype	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
	Predicted phenotype	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT
	Observed phenotype	WT	WT	WT	WT	WT	WT	WT	x	WT	WT	WT	x	WT	WT	WT	WT
23	Genotype	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA
	Predicted phenotype	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT	WT
	Observed phenotype	x	WT	WT	WT	WT	WT	WT	WT	WT	x	WT	WT	WT	x	WT	WT
24	Genotype	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa
	Predicted phenotype	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL
	Observed phenotype	OL	OL	x	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	OL	x

ACKNOWLEDGEMENTS

This work was supported by funding from the International Atomic Energy Agency under the CRP 'Integrated Utilization of Cereal Mutant Varieties in Crop/Livestock Production Systems' (D23030). We gratefully acknowledge the support of PBGL staff and fellows, and the team involved in the implementation of the CRP in developing this protocol.

REFERENCES

[1] Anderson JR, and Lübberstedt T. 2003. Functional markers in plants. Trend Plant Sci 8:554-560. https://doi.org/10.1016/j.tplants.2003.09.010

[2] Lema M. 2018 Marker assisted selection in comparison to conventional plant breeding: Review article. Agric Res & Technol Open Access J 14:555914. https://doi.org/10.19080/ARTOAJ.2018.14.555914

[3] FAO/IAEA. 2018. Manual on Mutation Breeding – Third Edition. Spencer-Lopes, MM, Forster, BP and Jankuloski, L. (eds), Food and Agriculture Organization of the United Nations. Rome, Italy. 301 pp. http://www.fao.org/3/I9285EN/i9285en.pdf

[4] Collard BCY, and Mackill DJ. 2008. Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. Philos Trans R Soc Lond B Biol Sci 363:557-572. https://doi.org/10.1098/rstb.2007.2170

[5] Ingelbrecht I, Jankowicz-Cieslak J, Mukhtar Ali Ghanim A, Anwar Y, Forster BP and Till BJ. 2018. DNA marker and genotyping applications for mutation breeding. In: Spencer-Lopes MM, Forster BP, and Jankuloski L (eds.), Manual Mutation Breeding, 3rd Ed., pp 249-264. Food and Agriculture Organization of the United Nations. Rome, Italy.

[6] Broccanello C, Chiodi C, Funk A, McGrath JM, Panella L, and Stevanato P. 2018. Comparison of three PCR-based assays for SNP genotyping in plants. Plant Methods 14:28. https://doi.org/10.1186/s13007-018-0295-6

[7] Bui M, and Liu Z. 2009. Simple allele-discriminating PCR for cost-effective and rapid genotyping and mapping. Plant Methods 5:1. https://doi.org/10.1186/1746-4811-5-1

[8] Semagn K, Babu R, Hearne S, and Olsen M. 2014. Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): Overview of the technology and its application in crop improvement. Mol Breed 33:1-14. https://doi.org/10.1007/s11032-013-9917-x

[9] Patnala R, Clements J, and Batra J. 2013. Candidate gene association studies: A comprehensive guide to useful *in silico* tools. BMC Genetics 14:39. https://doi.org/10.1186/1471-2156-14-39

[10] Gorthy S, Mayandi K, Faldu D, and Dalal M, et.al. 2014. Molecular characterization of allelic variation in spontaneous brown midrib mutants of sorghum (*Sorghum bicolor* (L.) Moench). Mol Breed 31:795-803. https://doi.org/10.1007/s11032-012-9834-4

[11] Mascher M, Jost M, Kuon JE, Himmelbach A, Aßfalg A, Beier S, Scholz U, Graner A, and Stein N. 2014. Mapping-by-sequencing accelerates forward genetics in barley. Genome Biol 15: R78. https://doi.org/10.1186/gb-2014-15-6-r78

[12] Dahleen L, Franckowiak JD, and Lundqvist U. 2007. BGS 254, Orange lemma. Barley Genet Newsl 37:255-256. https://wheat.pw.usda.gov/ggpages/bgn/37/Volume37.pdf

[13] Franckowiak JD, Foster AE, Pederson VD, and Pyler RE. 1985. Registration of 'Bowman' barley. Crop Sci 25:883. https://doi.org/10.2135/cropsci1985.0011183X002500050037x