CrossMark

ORIGINAL PAPER

Molecular marker associated with a deleterious recessive anomaly in *Eucalyptus grandis* seedlings

Maria C. P. Fuchs[1] · Evandro V. Tambarussi[2] · Juliana C. Lourenção[1] ·
Letícia M. Nogueira[3] · Tânia M. Bortoloto[1] · Esteban R. González[4] · Shinitiro Oda[4] ·
Celso L. Marino[1]

Abstract

· **Key message** The identification of an anomaly with deleterious effect and Mendelian segregation (3 normal: 1 abnormal) allowed the detection of a marker linked to a locus for the anomalous phenotype. This marker could be useful in marker-assisted breeding programs and for understanding this anomaly.

· **Context** Due to the mixed mating systems in *Eucalyptus* genera, outcrossing populations exhibit low inbreeding depression (abnormalities). In this study, we identified a developmental anomaly in a controlled cross of *Eucalyptus grandis*, with Mendelian segregation. The anomaly was characterized by branching, reduced height, smaller individual leaf area, and asymmetrical leaf shape.

· **Aims and methods** To investigate and evaluate the anomaly detected, we performed morphological and molecular analysis. The genitors and bulks of normal and abnormal individuals were screened with random primers. The polymorphic markers that co-segregated with the anomalous phenotype were selected, validated, converted into SCAR markers and analyzed in silico.

· **Results** Morphological analysis showed significant differences between normal and abnormal phenotypes. The co-segregating marker was present in one genitor, in all abnormal individuals, and in 31 % of normal individuals (recombinants). In silico analyses revealed a mutation of two base pairs between the contrasting phenotypes and identified the marker in an intergenic region presenting partial identity with an expressed sequence tag (EST) of Bet v1-like genes.

· **Conclusion** A marker was developed to identify the recessive allele for the anomaly in *E. grandis*. This will be important for management of crosses in eucalypt breeding programs and in anomaly studies.

Handling Editor: Ricardo ALIA

Contribution of the co-authors M.C.P. Fuchs: laboratory and field experiments, data collection, running the data analysis, analysis and discussion of the data, writing the paper.

J.C. Lourenção: laboratory experiments, data collection, analysis and discussion of the data.

E.V. Tambarussi: laboratory experiments, data collection, analysis and discussion of the data, writing the paper.

L.M. Nogueira: laboratory experiments, data collection and analysis of the data.

T.M. Bortoloto: laboratory experiments, data collection and analysis of the data.

E.R. González: experimental design and field experiments.

S. Oda: experimental design and supervising the work.

C.L. Marino: discussion of the data, writing the paper and coordinating the research project.

✉ Maria C. P. Fuchs
mceciliapf@yahoo.com.br

[1] Departamento de Genética, Instituto de Biociências, Universidade Estadual Paulista (UNESP), Distrito de Rubião Jr s/n, Botucatu, SP 18618-970, Brazil

[2] Departamento de Biologia, Campus Sorocaba, Universidade Federal de São Carlos (UFSCar), Rod. João Leme dos Santos, Km 110, Sorocaba, SP 18052-780, Brazil

[3] Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Dr Room 6E208 MSC 7248, Bethesda, MD 20892, USA

[4] Empresa Suzano Papel e Celulose SA, Av. Dr. José Lembo 1010, Itapetininga, SP 18207-780, Brazil

INRA
SCIENCE & IMPACT

Springer

1 Introduction

The genus *Eucalyptus* is the most extensively utilized forestry crop in the world due to its high adaptability, superior growth, and multipurpose applications, which are advantageous for its introduction and maintenance in a variety of environments (Grattapaglia and Kirst 2008). The genus includes a wide diversity of species, varieties and hybrids (more than 900) (Brooker and Kleinig 2004; Boland et al. 2006; Grattapaglia and Kirst 2008). In Brazil, *Eucalyptus* cultivation is particularly important, as demonstrated by the quantity of clonal and seed orchards with high eucalypt diversity in São Paulo and Espírito Santo States (Brune and Zobel 1981; Campinhos et al. 1998; Junghans et al. 1998); approximately 76.6 % of the total forest plantation area (5.10 million ha) (ABRAF 2013).

While eucalypts present mixed mating systems ($0.05 \leq t \leq 0.95$), *Eucalyptus grandis* is predominantly an outcrossing species (outcrossing rate of $t \geq 0.95$) (Chaix et al. 2003), which contributes to the significant amount of heterozygosis in the genome (Zobel and Talbert 1984; Golle et al. 2009). Outcrossing populations exhibit a low frequency of individuals with physiological and morphological anomalies (Klekowski 1988). These anomalies can be generated by allele mutations and have been observed in various plant species, e.g., *Antirrhinum majus* (Huijser et al. 1992), *Arabidopsis thaliana* (Rédei 1975), *Oryza sativa* (Koh et al. 1999) and *Cucumis sativus* (Marx 1983). According to Klekowski (1988), recessive deleterious alleles are hidden in heterozygotes of outcrossing populations. When homozygous, the recessive deleterious alleles manifest as phenotypic variations that are not restricted to the embryonic stage as they can also occur at the seedling stage where growth, morphology and physiology are altered (Zobel and Talbert 1984; Klekowski 1988). Negative effects of deleterious recessive alleles result in considerable losses to seedling production and may cause delays in improvement programs. This mortality and loss of seedling production due to recessive alleles can pose a routine problem to companies and research centers. In this context, the identification of deleterious recessive alleles by molecular markers represents a powerful tool to eliminate genotypes with unfavorable alleles in breeding programs, thus avoiding economic risks in seedling production and assisting the breeder in deciding which genitors are the best to use in crossings. Within full-sib offspring from a controlled cross of *E. grandis*, we observed an anomaly in one-quarter of the seedlings, which died within a few months. The principal characteristics of this anomaly were branching, reduced height, reduced individual leaf area, and modified leaf shape. We hypothesized that the anomaly observed was due to the presence of deleterious recessive alleles in homozygosis. Therefore, the purpose of this study was to analyze the anomalous character, as well as to obtain and characterize the molecular marker that co-segregates with the genomic region responsible for the observed phenotype in *E. grandis*.

2 Material and methods

2.1 Plant material and evaluation of characteristics

A full-sib progeny was obtained from a controlled cross between two unrelated *E. grandis* (codes G026 and G07) plants from the Suzano Paper and Cellulose Company (Suzano Papel e Celulose SA). For statistical analysis of normal and abnormal seedling frequencies, the chi-squared (χ^2) test was used to determine the segregation of normal and abnormal seedlings in a nursery trial (Table 1). These seedlings were selected for molecular marker analyses.

The Suzano Paper and Cellulose company carried out the cross between the two parents twice more and confirmed the segregation pattern. No other crossing produced the same phenotype, thus reinforcing the hypothesis that the deleterious allele is associated with a parent.

For character evaluation, we studied four traits from 19 plants of each phenotype, including: (1) seedling height (SH); (2) seedling stem-base diameter (SD); (3) leaf shape; and (4) number of lateral branches. To evaluate individual leaf width, length and area, we measured six to nine leaves from two plants of each phenotype using ImageJ software (Wayne Rasband, National Institutes of Health, Bethesda, MD). Data from SH, SD and leaf traits were analyzed using Tukey's test in the software Statistical Analysis System (SAS 2002), with statistical significance set at $P < 0.05$.

Table 1 Phenotypic analysis of seedlings originating from the controlled cross between the G26 and G07 parents from the Suzano Paper and Cellulose Company (Suzano Papel e Celulose SA)

	Normal	Value	Abnormal
Expected ratio	3	–	1
Observed plants	8,360	–	2,880
Chi-square (x^2)	–	2.32 ns[a]	–
Probability	–	3.84[b]	–

[a] Not significant

[b] χ^2 table for one degree of freedom

2.2 Genomic DNA extraction, random amplified polymorphic DNA marker identification, and sequence characterized amplified region marker development

Seeds were germinated in a root trainer (50 cm^3) using Multiplant® substrate. DNA was extracted from 30- to 40-day-old seedlings from the first leaf pair using the method described by Murray and Thompson (1980). Abnormal plants had a life span of approximately 3–4 months. The DNA of parent plants were extracted from 100 mg leaf per tree using the same method.

For random amplified polymorphic DNA (RAPD) marker identification, we used the bulk segregant analysis (BSA) technique (Michelmore et al. 1991). We created a normal bulk and an abnormal bulk using the same amount of DNA extracted from ten normal and abnormal seedlings, respectively. The two DNA bulks were screened for polymorphic markers using random primers (kits from Operon Technologies, Alameda, CA). The reaction consisted of 15 ng DNA, 1.30 μl 10× PCR Buffer (Invitrogen, Carlsbad, CA), MgCl$_2$ 1 mM, 1.04 μl bovine serum albumin (BSA) 10 mg/mL (Invitrogen), 1 mM dNTP mix, 0.3 μM primers, 1 U *Taq* DNA polymerase (Invitrogen), and water to make up a final volume of 13 μl. PCR was performed under the following conditions: initial denaturation at 94 °C for 3 min, followed by 41 cycles of 94 °C for 1 min, 35 °C for 1 min, 72 °C for 2 min, and a final extension at 72 °C for 5 min. Amplified products were visualized after electrophoresis in 1 % agarose gel in 1× TBE (Tris-borate-EDTA) stained with ethidium bromide. A total of 100 random primers was tested on normal and abnormal bulks. Of these, 13 primers (AD01, AD02, AD06, AD11, AD17, AD18, AF03, AF05, AF12, AF14, AF17, AF18 and AF20) revealed polymorphism between the bulks. However, only marker AD17 co-segregated with the AnoDel phenotype (Tambarussi et al. 2008). To validate the marker, we screened 252 individuals (both genitors, 200 normal and 50 abnormal seedlings).

We selected the polymorphic marker band that co-segregated with the abnormal characteristic phenotype and converted it into a sequence characterized amplified region (SCAR) marker. Based on the sequence information of the cloned polymorphic fragments of the marker, SCAR primers were designed (Barret et al. 1998; Hernández et al. 1999; Negi et al. 2000). The SCAR reaction consisted of 20 ng DNA, 2 μl 10× PCR Buffer (Invitrogen), MgCl$_2$ 1.5 mM, 0.5 mM dNTP mix, 0.5 μM primers, 1 U *Taq* DNA polymerase (Invitrogen), and water to make up a final volume of 20 μl. PCR was performed under the following conditions: initial denaturation at 94 °C for 5 min, followed by 34 cycles of 94 °C for 45 s, 66 °C for 30 s, 72 °C for 1 min 35 s, and a final extension at 72 °C for 10 min. Amplified products were visualized after electrophoresis in 1 % agarose gel in 1× TBE (Tris-borate-EDTA) stained with ethidium bromide.

2.3 Cloning and sequencing

The RAPD polymorphic band and PCR products of SCAR markers were extracted from the agarose gel, and DNA fragments were purified using illustra GFX PCR DNA (GE Healthcare). Purified fragments were cloned into the pGEM-T Easy Vector (Promega) and inserted into UltraMAX DH5α-FT competent *Escherichia coli* cells (Life Technologies), following the manufacturer's instructions.

We selected white bacterial colonies with the inserted RAPD products. Plasmid DNA was extracted by plasmid miniprep during which the selected bacterial colonies were grown overnight at 37 °C in 3 ml liquid medium (Circlegrow, QBiogene, Carlsbad, CA). A 1.4-ml aliquot of culture was then transferred to a new tube and centrifuged at 12,000 rpm for 1 min. The supernatant was discarded and the pellet eluted in 300 μl P1 solution (Tris–HCl 50 mM pH 8.0, EDTA 10 mM pH 8.0 and RNAse 100 μg ml^{-1}). Then, 300 μl P2 solution was added (NaOH 200 mM and SDS 1 %), followed by incubation at room temperature for 5 min. Finally, 300 μl P3 solution was added (potassium acetate 3 M, pH 5.5) and the samples were incubated on ice for 20 min. The samples were subsequently centrifuged at 14,000 rpm for 30 min and the supernatant was transferred to a new tube with 400 μl absolute isopropanol. The samples were incubated at −20 °C for 1 h and centrifuged at 12,000 rpm for 30 min. The supernatant was discarded, ethanol 70 % was added to rinse, and then the pellet was dried. Next, the DNA was eluted in deionized autoclaved water. Plasmid DNA was quantified by spectrophotometer (ND1000, NanoDrop, Wilmington, DE). White bacterial colonies with SCAR product inserts were selected and amplified with SCAR primers. The amplicons were extracted from agarose gel, and DNA fragments purified using illustra GFX PCR DNA (GE Healthcare).

Purified plasmid DNA with RAPD products, and SCAR amplicons were bi-directionally sequenced. The sequencing was performed using ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA) using BigDye Terminator Cycle Sequencing Kit v. 3.1 (Applied Biosystems) according to the manufacturer's instructions.

2.4 Southern blot

Total DNA (20 μg) from the two genitors was digested with restriction enzymes *Nde*I and *Xba*I, electrophoresed on a 0.7 % agarose gel, and transferred to Hybond-N+ nylon membrane (Amersham). For probe construction, the SCAR marker was amplified then electrophoresed on a 0.7 % agarose gel and the fragment was purified using PureLink Quick Gel Extraction Kit (Invitrogen). The fragments were labeled using AlkPhos Direct Labeling and Detection Systems Kit (Amersham Pharmacia, Little Chalfont, UK). Hybridization

and detection were performed as recommended by the instruction manual (Amersham Pharmacia).

2.5 In silico analysis

We analyzed the sequencing SCAR data using the software ChromasPro 1.5 (Technelysium; http://technelysium.com.au) and ClustalX 2.0.11 (Thompson et al. 1997) to form a consensus sequence. The sequence information has been submitted to the DNA Data Bank of Japan (DDBJ, http://www.ddbj.nig.ac.jp/) with the accession number AB889687 (Fuchs et al. 2013a). The sequences were compared with *Eucalyptus* EST ForEST (*Eucalyptus* Genome Sequencing Project Consortium) and GenBank (http://www.ncbi.nlm.nih.gov/genbank/) databases using the BLASTN tool (Altschul et al. 1997). The identified ESTs with high identity were translated by ExPASy translate tool (http://expasy.org) (Gasteiger et al. 2003) and analyzed using FGENESH (Softberry, http://www.softberry.com) to predict the protein. The results were compared with the PFAM 27.0 database (http://pfam.sanger.ac.uk) (Finn et al. 2008) for domain identification.

The SCAR sequence was compared with the eucalypt genome in the Phytozome v10.2 database (*E. grandis* V2.0 - http://www.phytozome.net) using BLAST (Altschul et al. 1997). Genomic regions with high identity were analyzed by FGENESH to find predicted genes in the region. We translated the identified predicted genes using the ExPASy translate tool (Gasteiger et al. 2003) and we compared the protein to the PFAM 27.0 database (Finn et al. 2008) for domain identification.

3 Results

Two phenotypic classes (abnormal and normal) were identified in full-sib progeny of a total of 11,240 seedlings. The 2,880 abnormal seedlings were identified based on the following traits: reduced root volume (Fig. 1a), shoot branching, reduced height (Fig. 1b), reduced individual leaf area, changes in leaf shape (Fig. 1c), and mortality at 3–4 months. Because of these features, we named the abnormal seedlings AnoDel (Anomaly and Deleterious). Measurement of seedling height, seedling stem-base diameter and individual leaf width, length and area showed significant differences ($P \leq 0.05$) (Fig. 1d, Table 2) between normal and AnoDel phenotypes. The significant differences detected indicate that these traits can be used to detect AnoDel plants. Based on our results (Table 2), abnormal plants are three times smaller and with a stem-base diameter that is two times smaller than normal plants. The width and length of abnormal plants were three and five times

smaller than normal plants, respectively. Furthermore, the individual leaf area also showed significant differences, with normal plants presenting an area 18 times greater than AnoDel plants.

In *E. grandis*, the leaves of seedlings (juveniles) are oval with a flat border; they are broader below the middle, pointed at the end, and the ratio of length to breadth can be up to 3:1 (Brooker and Kleinig 2004). These leaf characteristics were observed in all normal seedlings, but not in AnoDel seedlings (Fig. 1c). The AnoDel seedlings showed lanceolate leaves that were lance-shaped, pointed at the tip, with a ratio of length to breadth of up to 8:1 (Brooker and Kleinig 2004), and an undulate border (Fig. 1c).

In general, eucalypt trees are single-stemmed (Brooker and Kleinig 2004) and our observations in the nursery showed that the majority of eucalypt seedlings possess a single stem after germination. In natural habitats, branching can occur just above ground level (Brooker and Kleinig 2004). We observed that all normal seedlings were single-stemmed, whereas AnoDel seedlings had lateral branches (Fig. 1b) ranging in number from 1 to 13.

The Chi-square (χ^2) test showed that the progeny segregated in a ratio of normal to abnormal of 3:1, confirming a Mendelian segregation (Table 1). To identify markers linked to abnormal traits, a total of 100 random primers was tested. Of these, 13 (AD01, AD02, AD06, AD11, AD17, AD18, AF03, AF05, AF12, AF14, AF17, AF18 and AF20) revealed polymorphism between abnormal and normal bulks. However, only AD17 co-segregated with the AnoDel phenotype (Fig. 2a) (Tambarussi et al. 2008). Of the 252 individuals screened (both genitors, 200 normal and 50 abnormal seedlings), the marker (band at 500 bp) was present in one genitor, in all abnormal individuals, and in 31 % of normal individuals (normal and recombinants). The marker was absent from the other genitor and 69 % of normal individuals (normal and non-recombinants). The marker that was linked to the abnormal traits was converted into a SCAR marker and named AnoDel (Anomaly and Deleterious), as noted above. AnoDel marker (SCAR) genotyping of normal and abnormal individuals showed the same size (500 bp) and pattern of the polymorphic marker detected (Fig. 2c).

Sequence SCAR marker analysis detected a polymorphism of two base pairs between the reference genome of *E. grandis* and the mutant (abnormal) individuals. The polymorphism corresponds to the last two bases (CC) from the 3′ end of the AnoDel sequence of abnormal individuals (513–514 bp) (Fig. 2d). SCAR sequence (514 bp) comparisons in ForEST and GenBank (est_others) databases showed high levels of similarity (identity of 84 %) with a 3′-UTR sequence of a eucalypt putative protein (E-value 3e-40) and an EST of *Eucalyptus gunnii* (E-value 9e-62), respectively (Fig. 2f). EST of the putative protein and *E. gunnii* EST, when translated by the ExPASy tool (Gasteiger et al. 2003), resulted in a

Fig. 1 **a** Root of normal (*lower panel*) and abnormal (*upper panel*) seedlings. **b** Shoot of normal (*left*) and abnormal (*right*) seedlings. **c** Leaf of normal (*left*) and abnormal (*right*) seedlings. **d** Graphic representation of height, stem-base diameter and individual leaf area of abnormal and normal seedlings
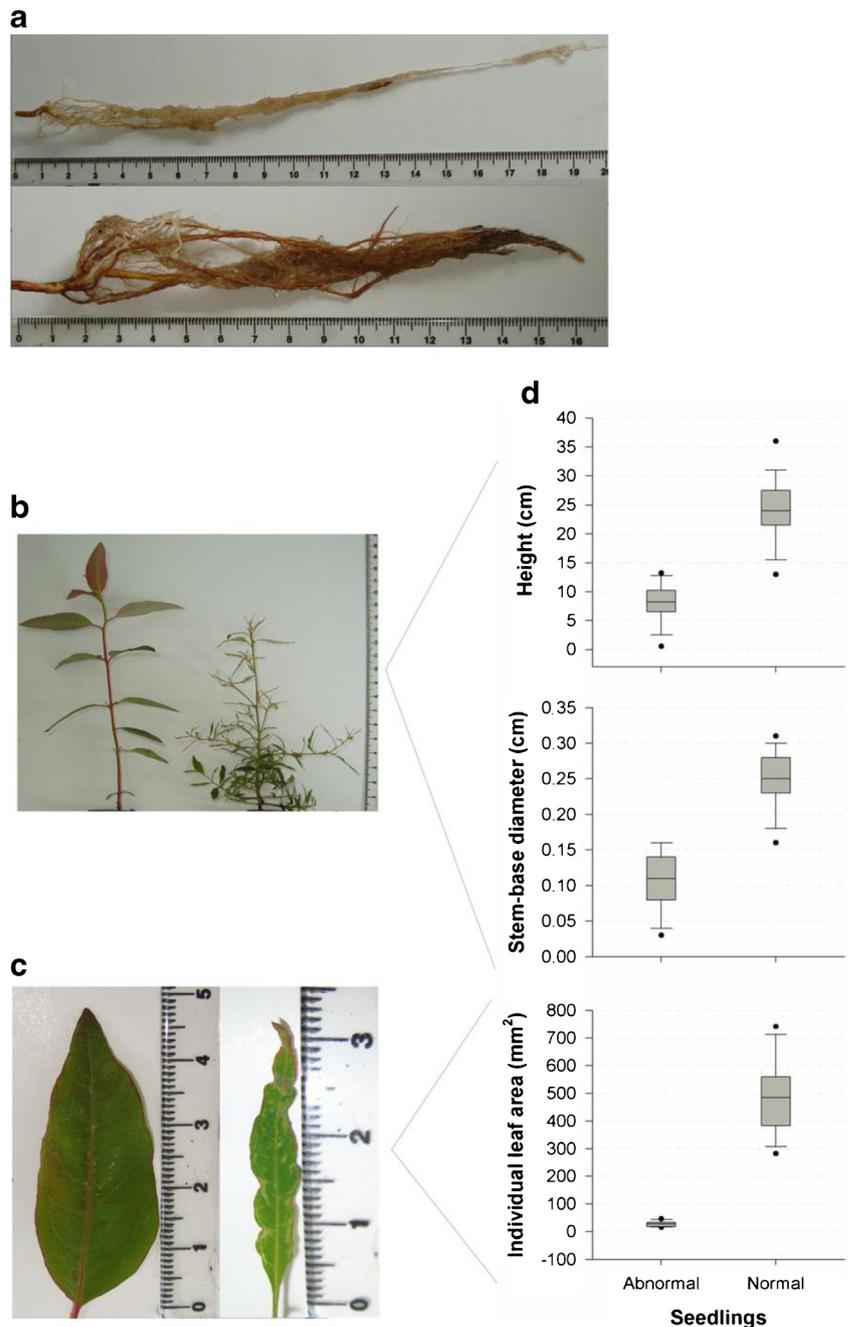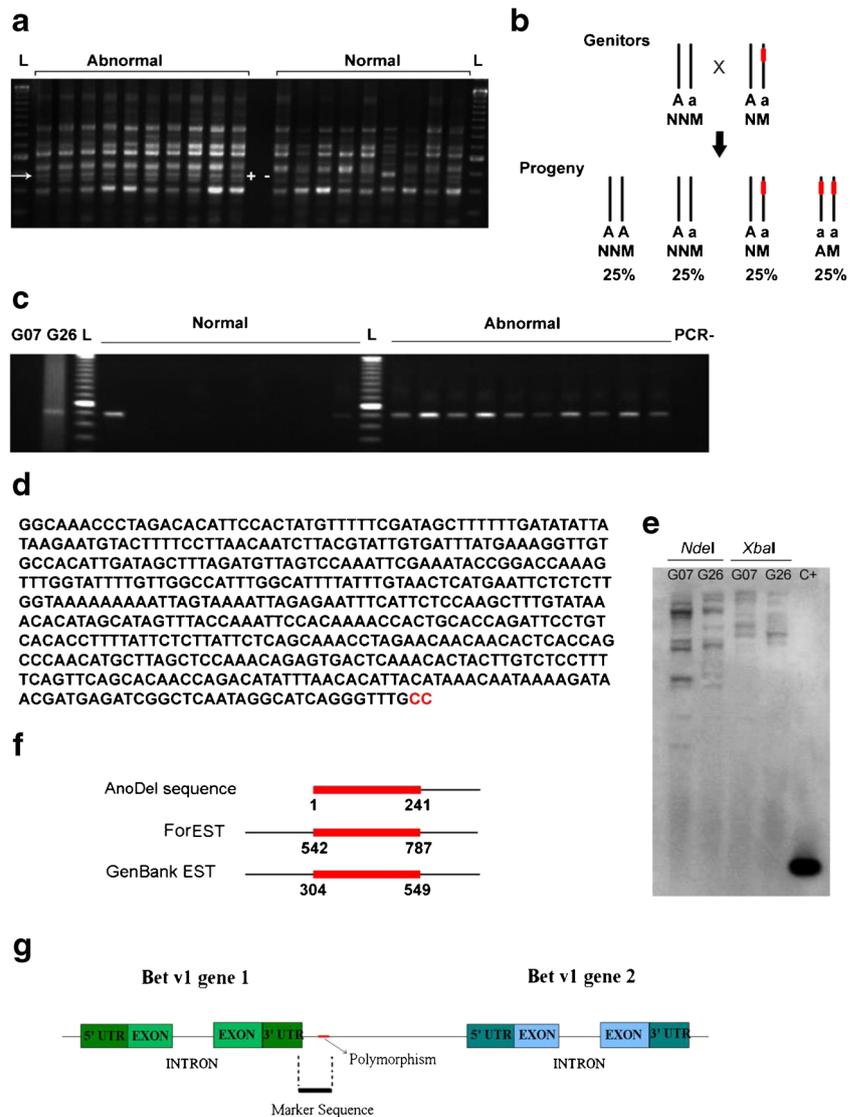

Table 2 Average measurements of morphological traits from abnormal and normal phenotypes. Significant differences are indicated with lowercase letters ($P<0.05$)

Morphological trait	Abnormal	Normal
Height (cm)	8.21 b	24.07 a
Stem-base diameter (cm)	0.11 b	0.25 a
Leaf area (mm^2)	27.49 b	484.78 a
Length (mm)	15.87 b	48.4 a
Width (mm)	3.68 b	16.48 a

predicted protein of 159 and 92 amino acids, respectively. Analysis using the PFAM database (Finn et al. 2008) demonstrated that these putative proteins possess a Bet v1-type domain. Similar results were obtained when protein sequences were predicted with a FGENESH tool. Analysis of ESTs using BLASTX (Altschul et al. 1997) showed significant homology (67 % identity, E-value 8e-33) with a sequence in the *Vitis vinifera* protein database in GenBank (nr). The identified *Vitis vinifera* hypothetical protein had 159 amino acids and showed a Bet v1-type domain using PFAM database (Finn

Fig. 2 **a** Polymorphic marker (*arrow*) obtained with primer AD17. The marker occurs in all abnormal individuals (*left*) and in some normal individuals (*right*). **b** Schematic of the marker presence in the genitors and full-sib progeny that presents three classes NNM (normal without marker), NM (normal with marker), and AM (abnormal with marker). **c** Sequence characterized amplified region (SCAR) marker validation. The marker is shown in all abnormal individuals (*right*) and in two normal individuals (recombinants) (*left*). G07 and G26 are the genitors. *L* (**a**, **c**): 1 kb ladder (Invitrogen, Carlsbad, CA). **d** SCAR [AnoDel, (Anomaly and Deleterious)] sequence shows the two-base polymorphism (*red*). **e** Southern blot analysis showing the presence of more than one copy of SCAR marker in the genome of genitors. **f** Diagram showing alignment between AnoDel, Phytozome EST, and GenBank EST sequence analyses. Similarity between the three sequences (*red*). **g** Diagram of marker sequence alignment with the eucalypt genome

d

```
GGCAAACCCTAGACACATTCCACTATGTTTTTCGATAGCTTTTTTGATATATTA
TAAGAATGTACTTTTCCTTAACAATCTTACGTATTGTGATTTATGAAAGGTTGT
GCCACATTGATAGCTTTAGATGTTAGTCCAAATTCGAAATACCGGACCAAAG
TTTGGTATTTTGTTGGCCATTTGGCATTTTATTTGTAACTCATGAATTCTCTCTT
GGTAAAAAAAAATTAGTAAAATTAGAGAATTTCATTCTCCAAGCTTTGTATAA
ACACATAGCATAGTTTACCAAATTCCACAAAACCACTGCACCAGATTCCTGT
CACACCTTTTATTCTCTTATTCTCAGCAAACCTAGAACAACAACACTCACCAG
CCCAACATGCTTAGCTCCAAACAGAGTGACTCAAACACTACTTGTCTCCTTT
TCAGTTCAGCACAACCAGACATATTTAACACATTACATAAACAATAAAAGATA
ACGATGAGATCGGCTCAATAGGCATCAGGGTTTGCC
```

et al. 2008) analysis, which is identical to the ForESTs and *Eucalyptus gunnii* ESTs analysis.

Conducting a BLAST (Altschul et al. 1997) analysis with the eucalypt genome database enabled us to locate the genome region of the AnoDel marker. As with Southern blot (Fig. 2e), the AnoDel sequence showed similarity with more than one region in the eucalypt genome and these regions were located on chromosome 1. The region with most similarity (94.2 %, E-value 0) to the AnoDel sequence plus the surrounding 2 kb genomic sequence on either side, was analyzed using the FGENESH tool; the results showed that the AnoDel sequence is located in an intergenic region between two protein-coding regions. In other words, it occurs between the 3′-UTR of the predicted protein similar to the AnoDel sequence and another protein-coding region (Fig. 2g). According to the PFAM database (Finn et al. 2008) analysis, both predicted proteins have a Bet v1-type domain.

4 Discussion

The *Eucalyptus grandis* controlled cross showed a progeny segregated in the ratio of 3 (normal): 1 (abnormal) seedlings. Based on this Mendelian segregation ratio, we hypothesized that the anomaly (AnoDel phenotype), with undesirable features and premature death, is controlled by a gene with major effects. The major gene would have a monogenic inheritance pattern with recessive deleterious allele effects (Klekowski 1988; Charlesworth et al. 1990; Young et al. 1996; Byers and Waller 1999). Given that *Eucalyptus* has an outcrossing mating system, open-pollinated progenies that confer a high level of heterozygous loci (Klekowski 1988; Zobel and Talbert 1984; Golle et al. 2009), as well as the unrelated genitors used in this study, these recessive deleterious alleles are allozygous, with the identical state of the alleles having independent origins through convergent evolution (Keller et al. 2011).

The anomaly is characterized by a set of morphological modifications, including a reduction in root volume (Fig. 1a), branching, reduced height (Fig. 1b), and a reduction in leaf size (Fig. 1c). These morphological changes suggest a pleiotropic effect of the major gene, resulting in a pleiotropic syndrome in which a set of characteristics appearing together constitute a syndrome as the anomaly (Ramalho et al. 2004).

To identify markers linked to the abnormal phenotype, a combination of BSA and molecular marker techniques were used. This combination has been shown to be effective in identifying markers linked to specific genes or traits (Michelmore et al. 1991; Barret et al. 1998; Hyten et al. 2009; Lan et al. 2010; Liu et al. 2011). Since some species, such as *Eucalyptus*, have had their genome completely sequenced, these combined techniques are effective in locating genomic regions of interest or molecular characterization (Hyten et al. 2009; Schneeberger and Weigel 2011).

Since the search for a marker correlated with the AnoDel phenotype involved an unknown region in the genome, it was necessary to use a technique, such as RAPD, with which the whole genome could be screened rapidly to locate several polymorphic loci simultaneously. RAPD is a PCR-based method that uses short (10-nt) arbitrary primers to access random region of the genome, providing polymorphism information between the phenotypes, caused by differences in nucleotide sequence at the priming sites or by structural rearrangements within the amplified sequence (Williams et al. 1990; Paran and Michelmore 1993). This technique, associated with BSA, offers other advantages to provide the first information on the molecular basis of the polymorphisms in this study, i.e., it is technically easy and has low cost.

We identified a molecular marker co-segregating with the AnoDel phenotype (Fig. 2a). This marker (band at 500 bp) was present in only one genitor, in all 50 abnormal individuals, and in 31 % of normal individuals; it was absent from the other genitor used in the study and 69 % of normal individuals. The observed presence of the co-segregation marker was consistent with our expectations (marker presence in all abnormal individuals and in 33 % of normal individuals) (Fig. 2b), indicating that it is linked in repulsion to the gene related with the AnoDel trait of one genitor. Notably, no recombinant phenotype (anomalous plants without the marker) was observed. Thus, the molecular marker present in 31 % of normal plants is due to the recessive allele, from the parent with the marker, segregating in the progeny and is not due to a recombination process (Fig. 2b). This result indicates that the marker is so near the responsible gene that there is no recombination between them. Therefore, the marker demonstrated high efficiency to detect anomalous plants, and detected normal plants with the deleterious recessive allele hidden in the heterozygous state.

Sequence analysis detected a polymorphism of two base pairs between the reference genome of *E. grandis* and mutant (abnormal) individuals. The polymorphism corresponds to the last two bases (CC) from the 3′ end of the AnoDel sequence of abnormal individuals (513–514 bp) (Fig. 2d). This coincides with the first two bases from the 5′ end of the RAPD primer, but has no similarity to the corresponding bases (GA) in the wild-type genome. This indicates that non-amplification of the RAPD primer in normal individuals is due to non-pairing of these two bases of the primer in the wild-type genome. The knowledge generated in genomic studies will enable the development of a marker derived from the polymorphic site that can be used in marker-assisted selection (MAS) techniques (Andersen and Lübberstedt 2003; Kettener et al. 2014; Bhowmicka et al. 2014). MAS has been shown to increase selection efficiency in genetic improvement programs; however, in relation to quantitative traits, MAS is more effective for selecting genes with major effects (Lande and Thompson 1990; Bernardo 2001; Andersen and Lübberstedt 2003). Utilizing markers that have been identified as major effect genes related to phenotypic variation will reduce the risk of linkage drags (with undesirable effects) and the time and costs associated with breeding programs (Andersen and Lübberstedt 2003).

However, the RAPD technique is sensitive to changes in reaction conditions, having disadvantages of poor reproducibility; comigration of fragments with equal or very close size and the presence of a band at the same position on two different individuals (amplified fragments of the same size at different loci) can lead to errors in interpretation data (Paran and Michelmore 1993). In this context, to overcome these problems, the detected marker was converted into a SCAR marker based on the nucleotide polymorphism observed. SCAR markers are obtained easily from RAPD fragments with 500 to 1500 bp, in which the RAPD sequence is used to design more stringent primers (about 20–24 nt in length) and consequently more specific and reproducible amplifications (Paran and Michelmore 1993; Hernández et al. 1999; Kasai et al. 2000; Negi et al. 2000). SCAR markers are advantageous over RAPD markers because they are more informative, distinguishing homozygous and heterozygous genotypes (in a codominant marker) (Paran and Michelmore 1993; Barret et al. 1998; Negi et al. 2000; Agarwal et al. 2008); SCAR amplification reactions are faster and less sensitive than RAPD reactions (Paran and Michelmore 1993; Barret et al. 1998); and finally SCAR markers detect only a single locus, generating more straightforward results (Paran and Michelmore 1993; Kasai et al. 2000).

The genotyping of normal and abnormal individuals using the AnoDel marker (SCAR) showed the same band pattern and marker size (500 bp) in the same individuals analyzed with RAPD polymorphic marker (Fig. 2d), showing the same efficiency to detect anomalous plants and normal plants that carry the deleterious recessive allele in heterozygous state. This result confirms the link between the AnoDel marker

and the abnormal traits. For a more precise location between the marker and the gene responsible for the trait, construction of a genetic linkage map using additional molecular markers is necessary. Nevertheless, the AnoDel marker developed can be used to screen plant material in eucalypt genetic improvement programs. Since this marker is specific to *E. grandis*, validation of the marker for other populations will enable the detection of these recessive deleterious alleles across many eucalypt species to avoid problems with breeding programs in eucalypt production. On the other hand, the information obtained from this work may also help us to understand the anomaly with more detailed study of the region, searching for candidate genes upstream and downstream of the AnoDel marker and investigating inherent variability in gene expression.

Analysis of sequences using BLASTN (Altschul et al. 1997) from the ForEST and GenBank (est_others) databases, showed high levels of similarity with the 3′-UTR sequence of an EST encoding a putative protein with Bet v1-type domain (Fig. 2f). Furthermore, BLAST (Altschul et al. 1997) analysis using the eucalypt reference genome allowed us to locate the AnoDel marker in the genome. The AnoDel sequence showed similarity with 19 regions in chromosome 1, and the region with the most similarity to the AnoDel sequence (94.2 %, E-value 0) is located in an intergenic region between Bet v1-like genes (Fig. 2g). Of the 19 regions with AnoDel sequence similarity, 14 showed Bet v1 proteins downstream. All regions are proximal, suggesting that the Bet v1 multigene family (Breiteneder et al. 1989; Hoffmann-Sommergruber et al. 1997; Wen et al. 1997) is organized in tandem arrangements, like other multigenic families (Capesius 1997; Passardia et al. 2004; Valério et al. 2004).

Since the discovery of the Bet v1 sequence by Breiteneder et al. (1989), this protein has been compared with other proteins to identify its homologues and, consequently, its function (Breiteneder et al. 1989; Swoboda et al. 1995; Walter et al. 1996; Wen et al. 1997). Bet v1-homologous proteins form the PR10 family (Pathogenesis-related protein 10). In general, PR10 family proteins are associated with plant defense functions in response to biotic and abiotic stresses (Hoffmann-Sommergruber et al. 1997; Wen et al. 1997; Hoffmann-Sommergruber 2002; Liu and Ekramoddoullah 2006). However, the exact function of Bet v1-like protein is still unclear. PR10 proteins are reported to have important roles in biological processes such as plant development because they have high homology to ribonucleases (Walter et al. 1996; Liu and Ekramoddoullah 2006) and the identification of PR protein isoforms in specific organs at certain stages of development has been observed (Hoffmann-Sommergruber et al. 1997). In addition, Poupard et al. (2001), studying birch trees (*Betula pendula*), reported the presence of a Bet v1-like protein at sites of emerging secondary primordial root after induction by indole-3-acetic acid, suggesting that Bet v1 could be involved in processes of lateral root initiation.

Studies assessing phenotypes similar to AnoDel, such as dwarfism and branching, are associated with genes related to a defense response (Knight 1947; Bomblies and Weigel 2007; Bomblies et al. 2007). However, despite the fact that the molecular marker linked with the anomaly is located in a region of Bet v1-like genes and part of its sequence shares identity with these genes, the direct or indirect involvement of Bet v1-like genes in the development of anomalous characteristics remains unclear. On the other hand, the polymorphism of two base pairs differentiates normal and abnormal individuals, enabling a molecular marker related to the anomalous characteristic to be developed.

Due to their recent domestication, *E. grandis* genomes have a high degree of heterozygosity (Myburg et al. 2014). Thereby, recessive deleterious alleles can be maintained in populations of this species. These unfavorable alleles can appear in advanced stages of breeding programs as the result of crosses between related individuals or reduction of genetic variability. Selection based on molecular markers to eliminate unfavorable alleles in tree species have been encouraged. However, few alleles are associated with a molecular marker closely linked to phenotype, and used in practice in breeding programs. In this context, our study provides a molecular marker that can detect the recessive deleterious allele related to a deleterious anomaly in *E. grandis*. The screening procedure with the AnoDel marker can help breeders in the appropriate management of individuals or populations under improvement, eliminating carriers of the recessive allele or avoiding crosses that could cause abnormalities. Besides, this work can provide relevant information for future studies to better understand the anomaly.

References

ABRAF (2013) Yearbook statistical ABRAF 2013 - base year 2012. ABRAF - Brazilian Association of Forest Plantation Producers, Brasília

Agarwal M, Shrivastava N, Padh H (2008) Advances in molecular marker techniques and their applications in plant science. Plant Cell Rep 27:617–631. doi:10.1007/s00299-008-0507-z

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402

Andersen JR, Lübberstedt T (2003) Functional markers in plants. Trends Plant Sci 8:554–560. doi:10.1016/j.tplants.2003.09.010

Barret P, Delourme R, Foisset N, Renard M (1998) Development of a SCAR (sequence characterized amplified region) marker for

molecular tagging of the dwarf BREIZH (Bzh) gene in *Brassica napus* L. Theor Appl Genet 97:828–833

Bernardo R (2001) What if we knew all the genes for a quantitative trait in hybrid crops? Crop Sci 41:1–4

Bhowmicka BK, Nandab S, Nayakb S, Jhaa S, Joshib RK (2014) An APETALA3 MADS-box linked SCAR marker associated with male specific sex expression in *Coccinia grandis* (L). Voigt. Sci Hortic 176:85–90

Boland DJ, Broker MIH, Chippendale GM, Hall N, Hyland BPM, Johnson RD, Kleining DA, McDonald MW, Turner JD (2006) Forest trees of Australia, 5th edn. CSIRO, Collingwood

Bomblies K, Lempe J, Epple P, Warthmann N, Lanz C, Dangl JL, Weigel D (2007) Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. PLoS Biol 5:1962–1972. doi:10.1371/journal.pbio.0050236

Bomblies K, Weigel D (2007) Hybrid necrosis: autoimmunity as a common barrier to gene flow in plants. Nat Rev Genet 8:382–393. doi:10.1038/nrg2082

Breiteneder H, Pettenburger K, Bito A, Valenta R, Kraft D, Rumpold H, Scheiner O, Breitenbach M (1989) The gene coding for the major birch pollen allergen BetvI, is highly homologous to a pea disease resistance response gene. EMBO J 8:1935–1938

Brooker MIH, Kleinig DA (2004) Field guide to *Eucalyptus*, vol 3, 2nd edn. Bloomings Books, Melbourne

Brune A, Zobel B (1981) Genetic base populations, gene pools and breeding populations for *Eucalyptus* in Brazil. Silvae Genet 30:146–149

Byers DL, Waller DM (1999) Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. Annu Rev Ecol Syst 30:479–513

Campinhos EN, Peters-Robinson I, Bertolucci FL, Alfenas AC (1998) Interspecific hybridization and inbreeding effect in seed from a *Eucalyptus grandis* x *E. urophylla* clonal orchard in Brazil. Genet Mol Biol 21:369–374

Capesius I (1997) Analysis of the ribosomal RNA gene repeat from the moss *Funaria hygrometrica*. Plant Mol Biol 33:559–564

Chaix G, Gerber S, Razafimaharo V, Vigneron P, Verhaegen D, Hamon S (2003) Gene flow estimation with microsatellites in a Malagasy seed orchard of *Eucalyptus grandis*. Theor Appl Genet 107:705–712. doi:10.1007/s00122-003-1294-0

Charlesworth D, Morgan MT, Charlesworth B (1990) Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. Evolution 44:1469–1489

Finn RD, Tate J, Mistry J, Coggill PC, Sammut JS, Hotz HR, Ceric G, Forslund K, Eddy SR, Sonnhammer EL, Bateman A (2008) The Pfam protein families database. Nucleic Acids Res 36:281–288. doi:10.1093/nar/gkm960

Fuchs MCP, Lourenço JC, Tambarussi EV, Nogueira LM, Bortoloto TM, Gonzalez ER, Oda S, Marino CL(2013a) Molecular marker associated with a recessive anomaly in Eucalyptus Grandis Hill Ex Maiden. TreeGenes https://dendrome.ucdavis.edu/tgdr

Fuchs MCP, Lourenço JC, Tambarussi EV, Nogueira LM, Bortoloto TM, Gonzalez ER, Oda S, Marino CL (2013b) Molecular characterization of a genomic region associated with a natural mutant of Eucalyptus grandis. DDBJ TGDR020 http://getentry.ddbj.nig.ac.jp/getentry/na/AB889687. AB889687

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784–3788. doi:10.1093/nar/gkg563

Golle DP, Reiniger LRS, Curti AR, Bevilacqua CB (2009) Forestry improvement: emphasis on biotechnology application. Ciênc Rural 39:1606–1613

Grattapaglia D, Kirst M (2008) *Eucalyptus* applied genomics: from gene sequences to breeding tools. New Phytol 179:911–929. doi:10.1111/j.1469-8137.2008.02503.x

Hernández P, Martín A, Dorado G (1999) Development of SCARs by direct sequencing of RAPD products: a practical tool for the introgression and marker-assisted selection of wheat. Mol Breeding 5:245–253

Hoffmann-Sommergruber K (2002) Pathogenesis-related (PR)-proteins identified as allergens. Biochem Soc Trans 30:930–935

Hoffmann-Sommergruber K, Vanek-Krebitz M, Radauer C, Wen J, Ferreira F, Scheiner O, Breiteneder H (1997) Genomic characterization of members of the Bet v 1 family: genes coding for allergens and pathogenesis-related proteins share intron positions. Gene 197:91–100

Huijser P, Klein J, Lönnig WE, Meijer H, Saedler H, Sommer H (1992) Bracteomania, an inflorescence anomaly, is caused by the loss of function of the MADS-box gene squamosa in *Antirrhinum majus*. EMBO J 11:1239–1249

Hyten DL, Smith JR, Frederick RD, Tucker ML, Song Q, Cregan PB (2009) Bulked segregant analysis using the GoldenGate assay to locate the *Rpp3* locus that confers resistance to soybean rust in soybean. Crop Sci 49:265–271

Junghans TG, Peters-Robinson I, Bertolucci FL, Alfenas AC (1998) The use of self-incompatibility in the production of hybrid eucalyptus seed by 'Aracruz Celulose' in Brazil. Genet Mol Biol 21:375–379

Kasai K, Morikawa Y, Sorri VA, Valkonen JPT, Gebhardt C, Watanabe KN (2000) Development of SCAR markers to the PVY resistance gene Ry_{adg} based on a common feature of plant disease resistance genes. Genome 43:1–8

Keller MC, Visscher PM, Goddard ME (2011) Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. Genetics 189:237–249. doi:10.1534/genetics.111.130922

Kettener K, Fuchs MCP, Madacki ACA, Gonzales E, Souza ICG, Oda S, Marino CL (2014) Development of SCAR Molecular Markers in *Eucalyptus saligna* and *Eucalyptus tereticornis*. Int J Agric For 4:435–439

Klekowski EJ Jr (1988) Genetic load and its causes in long-lived plants. Trees 2:195–203

Knight RL (1947) The genetics of blackarm resistance v. dwarf-bunched and its relationship to B1. J Genet 48:43–51

Koh HJ, Son YH, Heu MH, Lee HS, McCouch SR (1999) Molecular mapping of a new genic male-sterility gene causing chalky endosperm in rice (*Oryza sativa* L.). Euphytica 106:57–62

Lan C, Liang S, Zhou X, Zhou G, Lu Q, Xia X, He Z (2010) Identification of genomic regions controlling adult-plant stripe rust resistance in Chinese landrace Pingyuan 50 through bulked segregant analysis. Phytopathology 100:313–318. doi:10.1094/PHYTO-100-4-0313

Lande R, Thompson R (1990) Efficiency of marker-assisted selection in the improvement of quantitative traits. Genetics 124:743–756

Liu F, Yao J, Wang X, Hu Z, Duan D (2011) Identification of SCAR marker linking to longer frond length of *Saccharina japonica* (Laminariales, Phaeophyta) using bulked-segregant analysis. J Appl Phycol 23:709–713. doi:10.1007/s10811-010-9567-x

Liu J-J, Ekramoddoullah AKM (2006) The family 10 of plant pathogenesis-related proteins: their structure, regulation, and function in response to biotic and abiotic stresses. Physiol Mol Plant P 68:3–13. doi:10.1016/j.pmpp.2006.06.004

Marx GA (1983) Developmental mutants in some annual seed plants. Annu Rev Plant Physiol 34:389–417

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4325

Myburg AA, Grattapaglia D, Tuskan GA, Hellsten U, Hayes RD, Grimwood J et al (2014) The genome of *Eucalyptus grandis*. Nature 510:356–362. doi:10.1038/nature13308

INRA SCIENCE & IMPACT   Springer

Negi MS, Devic M, Delseny M, Lakshmikumaran M (2000) Identification of AFLP fragments linked to seed coat colour in *Brassica juncea* and conversion to a SCAR marker for rapid selection. Theor Appl Genet 101:146–152

Passardia F, Longetb D, Penela C, Dunand C (2004) The class III peroxidase multigenic family in rice and its evolution in land plants. Phytochemistry 65:1879–1893. doi:10.1016/j.phytochem.2004.06.023

Paran I, Michelmore RW (1993) Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor Appl Genet 85:985–993

Poupard P, Brunel N, Leduc N, Viémont J-D, Strullu D-G, Simoneau P (2001) Expression of a Bet v 1 homologue gene encoding a PR 10 protein in birch root: induction by auxin and localization of the transcripts by *in situ* hybridization. Aust J Plant Physiol 28:57–63. doi:10.1071/PP00040

Ramalho MAP, Santos JB, Pinto CABP (2004) Genetics in agriculture, 3rd edn (in Portuguese). Universidade Federal de Lavras, Lavras

Rédei GP (1975) *Arabidopsis* as a genetic tool. Annu Rev Genet 9:111–127

SAS (2002) SAS user's guide: statistics. SAS Institute, Cary

Schneeberger K, Weigel D (2011) Fast-forward genetics enabled by new sequencing technologies. Trends Plant Sci 16:282–288. doi:10.1016/j.tplants.2011.02.006

Swoboda I, Dang TCH, Heberle-Bors E, Vicente O (1995) Expression of Bet v 1, the major birch pollen allergen, during anther development. An in situ hybridization study. Protoplasma 187:103–110

Tambarussi EV, Lourenção JC, Bortoloto TM, Sassaki FT, Oda S, Marino CL (2008) Desenvolvimento de marcador SCAR (Sequence Characterized Amplified Region) relacionado a uma anomalia de viveiro em *Eucalyptus grandis*. Sociedade Brasileira de Genética. http://web2.sbg.org.br/congress/sbg2008/pdfs2008/24722.pdf. Accessed 1 December 2013

Tompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Valério L, De Meyer M, Penel C, Dunand C (2004) Expression analysis of the *Arabidopsis* peroxidase multigenic family. Phytochemistry 65:1331–1342. doi:10.1016/j.phytochem.2004.04.017

Walter MH, Liu J-W, Wünn J, Hess D (1996) Bean ribonuclease-like pathogenesis-related protein genes (Ypr10) display complex patterns of developmental, dark-induced and exogenous-stimulus-dependent expression. Eur J Biochem 239:281–293

Wen J, Vanek-Krebitz M, Hoffmann-Sommergruber K, Scheiner O, Breiteneder H (1997) The potential of Bet v 1 homologues, a nuclear multigene family, as phylogenetic markers in flowering plants. Mol Phylogenetic Evol 8:317–333

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. Tree 11:413–418

Zobel B, Talbert J (1984) Applied forest tree improvement. Wiley, New York