Chapter 16
A Protocol for Validation of Doubled Haploid Plants by Enzymatic Mismatch Cleavage

Bradley J. Till, Bernhard J. Hofinger, Ayşe Şen, Owen A. Huynh, Joanna Jankowicz-Cieslak, Likyelesh Gugsa, and Jochen Kumlehn

Abstract Doubled haploidy is an important tool for plant breeders. It provides a rapid means of developing recombinant populations consisting of individuals that are homozygous and therefore genetically fixed. Homozygosity is also important in plant mutation breeding where many induced mutations are predicted to be recessive and mutant alleles need to be in a homozygous state before new traits are expressed. While production of doubled haploids has been described for many plant species, efficient means to validate that produced materials are indeed homozygous are needed. Polymorphism discovery methods utilizing enzymatic mismatch cleavage are ideally suited for validation of doubled haploid plants. We describe here a low-cost protocol that utilizes self-extracted single-strand-specific nucleases, standard PCR reactions and agarose gel electrophoresis that can be applied to most plant species.

Keywords Homozygosity • Genetic fixation • Single-strand-specific nuclease • Polymorphism discovery • TILLING • Ecotilling • *Hordeum vulgare* • *Eragrostis tef*

B.J. Till (✉) • B.J. Hofinger • O.A. Huynh • J. Jankowicz-Cieslak
Plant Breeding and Genetics Laboratory, Joint FAO/IAEA Division, International Atomic Energy Agency, Vienna International Centre, PO Box 100, 1400 Vienna, Austria
e-mail: b.till@iaea.org

A. Şen
Plant Breeding and Genetics Laboratory, Joint FAO/IAEA Division, International Atomic Energy Agency, Vienna International Centre, PO Box 100, 1400 Vienna, Austria

Faculty of Science, Department of Biology, Istanbul University, 34134, Vezneciler, Istanbul, Turkey

L. Gugsa
Ethiopian Institute of Agricultural Research, HARC, PO Box 2003, Holetta, Ethiopia

Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Plant Reproductive Biology, Corrensstrasse 3, 06466 Seeland, OT Gatersleben, Germany

J. Kumlehn
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Plant Reproductive Biology, Corrensstrasse 3, 06466 Seeland, OT Gatersleben, Germany

16.1 Introduction

First reported in the early 1920s, methods for the production of haploid plants have now been described for more than 250 species (Blakeslee et al. 1922; Maluszynski et al. 2003). In many cases, haploid plants either spontaneously become diploid or this state can be induced by treatment with chemicals such as colchicine. Doubled haploidy remains a popular and powerful tool in plant sciences and breeding because once plants are doubled haploid (DH), they are homozygous, genetic variants are fixed and inbreeding plants are true breeding. A wide range of methods have been described for producing DH plants, and efficiencies can vary dramatically (Maluszynski et al. 2003; Forster et al. 2007; Geiger and Gordillo 2009; Ravi and Chan 2010). Doubled haploids are extremely valuable commercially as they can accelerate the breeding process; they can be developed directly into uniform and stable cultivars of inbreeding species and provide parental lines to be used in hybrid breeding. For successful and efficient research and breeding, it is therefore necessary to validate that materials are truly DH and homozygous before plants are selected for further evaluation and use. This holds particularly true in the context of the development of new methods of haploid formation, be it the implementation of a novel principle or just the adoption of a known method in another species.

Enzymatic mismatch cleavage for discovery and genotyping of single-nucleotide polymorphisms (SNPs) and small insertion/deletions (indels) is ideally suited to detect loss of heterozygosity in putative doubled haploid plants in comparison to the progenitor lines. Widely used in Targeting Induced Local Lesions IN Genomes (TILLING) reverse-genetic studies, the procedure begins with the selection of gene-specific primers for PCR amplification of ~1–1.5 kb regions. This is followed by denaturation and annealing that creates heteroduplexed DNA in samples where heterozygous variation exists. Samples are then treated with a self-prepared mixture containing single-strand-specific nucleases that cleave DNA at mismatched regions. The resulting products are electrophoresed, and the presence of cleaved DNA fragments of lower molecular weight than the original PCR product indicates the presence and approximate location of heterozygous polymorphisms (Till et al. 2006a, b). The process is made low cost by self-extraction of nuclease and the use of standard agarose gel electrophoresis as a gel readout platform (Till et al. 2015).

Enzymatic mismatch cleavage has also been widely used in Ecotilling studies for discovery of natural nucleotide variation in populations (Till 2014). This approach has been shown to be highly accurate with less than 5% false discovery and false-negative error rates, even in highly heterozygous polyploids (Till et al. 2006a, 2010).

The method can be easily adapted to evaluate homozygosity in putatively haploid or DH plants. As a proof of principle, Hofinger et al. (2013) developed this method for barley and compared it to another approach for validation of DH plants: simple sequence repeat (SSR) markers (Hofinger et al. 2013). In this work, the authors showed that 11/26 primer pairs were suitable for DH validation by

enzymatic mismatch cleavage, while only 3/32 previously characterized SSR primer sets were suitable. Thus the enzymatic mismatch cleavage approach had an ~4.5× higher success rate. Furthermore, because enzymatic mismatch cleavage utilizes gene- or target-specific primers, the approach allows for accurate DH production error estimations per plant and per experiment taking into account genetic linkage. In addition, using gene-specific primers allows selection of specific alleles and for diversification of allele types in applications such as those involving F_1 hybrids. The entire process can be completed in one day (Fig. 16.1). Enzymatic mismatch cleavage has been used in over 20 plant and animal species for TILLING and Ecotilling, and thus this adaptation for haploid/DH validation is expected to be broadly applicable (Kurowska et al. 2011; Till 2014). Indeed, in addition to barley, we show data for screening putative DH *Eragrostis tef* plants (Fig. 16.3 and Hofinger et al. 2013). While the methods described in this protocol are straightforward and low cost, a key component of successful application is proper experimental design. Example data and a discussion of experimental design are provided along with a detailed protocol suitable for many crop species.

16.2   Materials

16.2.1   PCR Amplification

1. Taq DNA polymerase and Taq polymerase buffer (*see* **Note 1**)
2. dNTP mix
3. MgCl₂ (25 mM stock), if not supplied in Taq polymerase buffer
4. Forward and reverse primers (*see* **Note 2**)
5. Laboratory grade water (distilled or deionized and autoclaved)
6. Thermal cycler with heated lid and adjustable ramping

16.2.2   Enzymatic Mismatch Cleavage

1. Magnesium sulphate (MgSO₄)
2. Triton X-100
3. HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid).
4. Bovine serum albumin (BSA)
5. Potassium chloride (KCl)
6. 10× Cleavage buffer (prepare a stock solution of 5 ml 1 M MgSO₄,100 μl 10 % Triton X-100, 5 ml 1 M HEPES (pH 7.4), 5 μl 20 mg/mL BSA, 2.5 ml 2 M KCl, 37.5 ml water
7. Single-strand-specific nuclease (*see* **Note 3**)
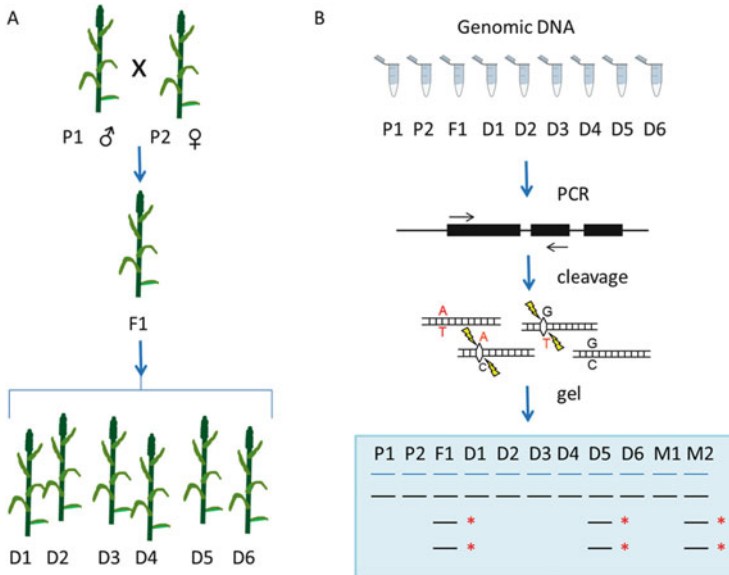8. EDTA (ethylenediaminetetraacetic acid), 0.25 M

Fig. 16.1 The use of enzymatic mismatch cleavage for validation of doubled haploid plants in F_1 hybrid studies. (**a**) In a typical F_1 hybrid approach, two genetically diverse parental plants (P1 and P2) are crossed to create an F_1 hybrid (F_1). Putatively doubled haploids are produced from the F_1 plant (D1 through D6). (**b**) To validate the production of DH plants, genomic DNA is prepared from all materials. PCR is performed to amplify a specific target. PCR products are denatured and annealed to form mismatches where polymorphisms exist and samples are incubated with a nuclease that cleaves DNA that is not base-paired. DNA fragments are evaluated by agarose gel electrophoresis. Suitable primer pairs show cleavage products indicating heterozygosity in the F_1 sample (marked by *asterisks*). In the example drawn, all putatively DH plants are homozygous except sample D5. Lanes marked M1 and M2 represent mixtures of samples D1 and P1 and D1 and P2 (respectively). This allows assignment of the parental allele in the DH line. In this example D1 contains the allele from P1 because there was no heterozygosity observed in the mixture of the two samples. More primer pairs should be screened to increase confidence that the remaining plants are truly DH

16.2.3 Agarose Gel Electrophoresis

1. Standard agarose
2. TBE (tris/borate/EDTA) buffer
3. Ethidium bromide (10 mg/ml) (*see* **Note 4**)
4. Horizontal gel electrophoresis system with power supply
5. Gel imaging system
6. Gel loading dye (30 % glycerol plus Orange G) (*see* **Note 5**)
7. DNA molecular weight ladder (e.g. 1 kb ladder, Invitrogen 10787-018)

16.3 Methods

16.3.1 PCR Amplification

1. Select samples for analysis based on guidelines for experimental design (*see* **Note 6**).
2. Prepare a master mix according to manufacturer's guidelines for the Taq polymerase used. For example, for 10 samples with TaKaRa Ex Taq: 109.5 µl water, 20 µl 10× Ex Taq buffer, 16 µl 2.5 mM dNTP mix, 2 µl forward primer (10 µM), 2 µl reverse primer (10 µM), 0.5 µl TaKaRa Ex Taq (5 U/µl).
3. Add 5 µL of DNA to each PCR tube (*see* **Note 7**).
4. Add 15 µl of PCR Master Mix to each PCR tube and mix by pipetting.
5. Centrifuge tubes for 1 min at $5000 \times g$.
6. Incubate samples in thermal cycler with following conditions (*see* **Note 8**):

   Step 1: Initial denaturation, 95 °C, 2 min
   Step 2: Denaturation, 94 °C, 20 s
   Step 3: Primer annealing, 73 °C (−1 °C/cycle), 30 s
   Step 4: Ramp, 0.5 °C per second to 72 °C
   Step 5: Primer extension, 72 °C, 1 min
   Step 6: Cycling, repeat steps 2–5 for 7 cycles (8 cycles in total)
   Step 7: Denaturation, 94 °C, 20 s
   Step 8: Primer annealing, 65 °C, 30 s
   Step 9: Ramp, 0.5 °C per second to 72 °C
   Step 10: Primer extension, 72 °C, 1 min
   Step 11: Cycling, repeat steps 7–10 for 44 cycles (45 cycles in total)
   Step 12: Final extension, 72 °C, 5 min
   Step 13: Denaturation, 99 °C, 10 min
   Step 14: Cooling, 72 °C, 20 s
   Step 15: Cycling and touchdown, repeat step 14 for 69 cycles (−0.3 °C/cycle)
   Step 16: Hold, 8 °C, forever

7. Samples can be stored at −20 °C for months.

16.3.2 Enzymatic Mismatch

1. Prepare an enzyme digestion master mix. For each sample add 4 µl 10× cleavage buffer, 7 µl single-strand-specific nuclease and 9 µl water (*see* **Note 3**).
2. Centrifuge PCR products for 1 min at $5000 \times g$.
3. Add 20 µl of enzyme digestion master mix to PCR reactions (*see* **Note 3**).
4. Incubate reactions in a thermal cycler at 45 °C for 15 min (no heated lid) and then hold at 8 °C.
5. Add 5 µl of 0.25 M EDTA to each tube to stop the reaction.

16.3.3   Agarose Gel Electrophoresis and Data Analysis

1. Prepare a 1.5 % agarose gel in 0.5x TBE buffer. Heat mixture in microwave until all agarose is dissolved. Take care to avoid boiling agarose.
2. Cool agarose to approximately 55 °C (at this temperature you can hold the flask in your hand comfortably for 5 s).
3. Add ethidium bromide to the warm gel solution to obtain a final concentration of 0.5 µg/mL in the gel. Mix by swirling.
4. Pour gel solution into gel apparatus and insert comb according to the instructions of the manufacturer.
5. Add 2 µl 6× loading dye to 10 µl PCR product.
6. Run gel at 100 V for 1 h or until suitable separation of bands is achieved (*see* **Note 9**).
7. Photograph the gel.
8. Evaluate data (*see* Figs. 16.2 and 16.3 and **Note 10**).

16.4   Notes

1. Most Taq polymerases should be suitable for this method. Compare your favourite Taq with the least expensive you can purchase. If the two produce
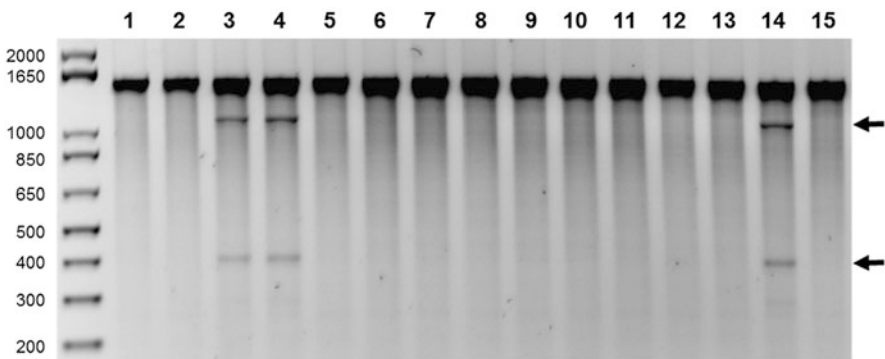

Fig. 16.2 Example data of validation of doubled haploid plant production in barley. A 1476-bp fragment of the barley *Mlo9* gene was PCR amplified and digested with a crude celery juice extract (CJE) containing single-strand-specific nuclease activity followed by agarose gel analysis. The top band in lanes 1–15 represents undigested PCR product. The cleavage products present in heterozygous samples are marked with *arrows*. Parental lines cv. Golden Promise (GP) and HOR1606 are homozygous for this gene region (*lanes 1 and 2*, respectively). A synthetic mixture of parental DNA and also the F_1 sample from crossing of the two parents show cleavage fragments resulting from a heterozygous SNP (*lanes 3 and 4*). Doubled haploid plants (*lanes 5–13*) are homozygous. Mixtures of genomic DNA from a DH plant and GP show cleavage products, while mixture of the same material with HOR1606 does not, indicating the DH harbours the HOR1606 allele (*lanes 14 and 15*). This figure and legend are reproduced from (Hofinger et al. 2013)
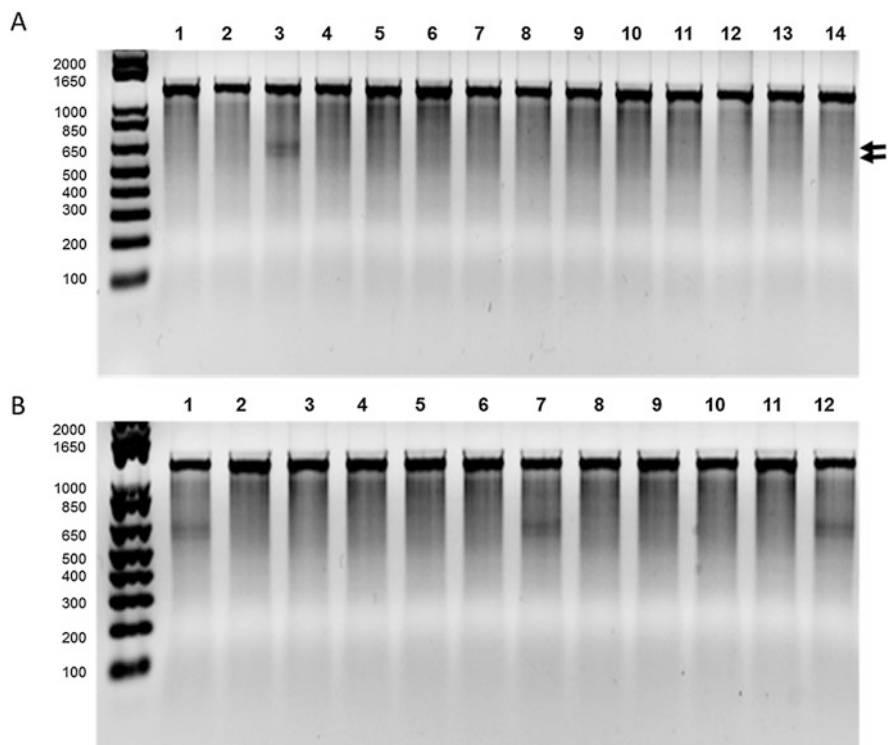
Fig. 16.3 Example data of validation of DH production in Tef primers were designed to amplify a 1400-bp region of the Tef genome. Screening for loss of heterozygosity was first performed with putative doubled haploid plants (panel **a**). Lanes 1 and 2 are parents. Lane 3 represents a synthetic F_1 hybrid created by mixing equal amounts of genomic DNA from each parent. Two cleavage products are produced indicating heterozygosity in the F_1 (marked by *arrows*). Lanes 4–14 represent putative DH plants. Experiments were then performed to determine the parental origin of the allele in putative DH plants. Panel **b** shows mixtures of samples with parent 1 (lane 1 of panel **a**). The first lane of panel **b** is the positive control. Lanes 2–12 are putative DH plants mixed with parent 1. This shows that two plants (lanes 7 and 12) inherited the allele from parent 2, while all others inherited the allele from parent 1. The reciprocal experiment was done by mixing samples with parent 2 (not shown)

similar results, use the cheaper version. Some Taq polymerases, like TaKaRa Ex Taq, come supplied with dNTPs and buffer containing $MgCl_2$.

2. Primer pairs should be selected to different genomic regions that are not genetically linked. The total number of primer pairs needed depends on the percentage of primer pairs where heterozygosity is observed. Primers are typically designed with the program Primer3 (Rozen and Skaletsky 2000) with a melting temperature of 70 °C (*see* **Note 8**).

3. The optimal amount of nuclease varies depending on the activity of the batch being used. Analysis of activity can be performed by screening PCR amplicons with heterozygous SNPs and titrating the amount of enzyme to produce clear

bands on the gel. See Chap. 15 for a protocol on enzyme preparation and activity optimization. When making master mixes, it is best to prepare more volume than is needed (e.g. for one additional sample) to compensate for minor pipetting errors. Volumes of reactions can be reduced from 40 to 20 μl in some cases.

4. *Caution*: Ethidium bromide is hazardous. Wear gloves and avoid contamination. Consult material safety data sheet (MSDS) for proper handling and disposal procedures. Alternative DNA stains can be used, e.g. SYBR Safe DNA gel stain, but detection may vary with different dyes.

5. Avoid loading dyes containing bromophenol blue or other dyes that migrate in a molecular weight range where you expect to observe DNA fragments. The presence of loading dyes can reduce the intensity of bands.

6. It is important to note that methods described here test only for homozygosity versus heterozygosity. Enzymatic mismatch cleavage does not distinguish between haploid and DH plants. An alternative method such as flow cytometry is required for this. The optimum experimental design includes DNA from the progenitor plant along with DNA from each putative DH plant produced from the progenitor. The only useful primer pairs will be those where heterozygosity is discovered in the progenitor plant. For example, if making doubled haploids from F_1 hybrid material, DNA from the F_1 hybrid is ideally evaluated along with DNA from the parents that were used to make the F_1. If the F_1 harbours heterozygous SNPs at a particular locus, and the putative DH plants do not, this is evidence that the plant is DH (*see* **Note 10** for more on data analysis and interpretation). In this example, if F_1 material is not available, a synthetic F_1 sample can be prepared by mixing an equal concentration of DNA from each parent in a 1:1 ratio prior to PCR amplification. Screening the parental material alone is informative to learn if parents are heterozygous in any interrogated regions. It is not ideal to screen only putative DH material as it is difficult to estimate the probability that plants are truly DH rather than being homozygous because progenitor material was homozygous at that locus. In some cases, however, only putative DH material is available. In this circumstance it is advised to mix two to four putative DH samples together prior to PCR. If heterozygosity is observed in the mixture but not in individual samples, this suggests that the progenitor material was heterozygous. The set of primers used in this experiment would then be considered suitable (*see* **Note 10**).

7. The optimal amount of genomic DNA to be used should be determined empirically. Typically, a PCR amplicon yield of 10 ng/μl is sufficient. The amount of genomic DNA needed to produce this amount can be roughly estimated by size of the genome (Till et al. 2006b). The yield of PCR product should be sufficiently high to produce cleavage fragments visible by agarose gel electrophoresis (see Figs. 16.2 and 16.3).

8. PCR conditions may need to be optimized. For example, primers with a melting temperature (Tm) of 70 °C were used to develop this protocol. Higher Tm primers increase specificity of amplification, but may not be necessary for all species. If lower Tm primers are used, the annealing temperature must be adjusted accordingly.

9. Optimal run time may vary. The Orange G DNA dye migrates at approximately 50 base pairs. A migration distance of 6–8 cm of Orange G from the comb typically provides suitable separation of fragments.

10. When analysing data, fragments observed of lower molecular weight than the full-length PCR product are typically the result of cleavage of heteroduplexed DNA at the site of a mismatch due to a nucleotide polymorphism. Truly doubled haploid plants are homozygous and therefore should not show cleavage products. However, cleavage fragments can also be observed due to a homopolymeric stretch of adenosine residues (Till et al. 2004). Evaluation of the parental material is therefore advised. Mixtures of parental material with putative DH plants allow assignment of alleles to a specific parent (see Figs. 16.2 and 16.3b). An estimation of the number of suitable primers can be made by calculating the probability that the data results by chance from self-fertilization. For example, the probability that the offspring of a self-fertilized heterozygous F_1 is homozygous for a specific locus is 0.5 (assume a Mendelian 1:2:1 ratio). The probability that two genetically unlinked loci are homozygous is $0.5^2 = 0.25$. By screening seven primer pairs from unlinked loci that show heterozygosity in the F_1, one can achieve 99 % confidence.

References

Blakeslee AF, Belling J, Farnham ME, Bergner AD (1922) A Haploid Mutant in the Jimson Weed, "*Datura stramonium*". Science 55(1433):646–647

Forster BP, Heberle-Bors E, Kasha KJ, Touraev A (2007) The resurgence of haploids in higher plants. Trends Plant Sci 12(8):368–375

Geiger HH, Gordillo GA (2009) Doubled haploids in hybrid maize breeding. Maydica 54:485–499

Hofinger BJ, Huynh OA, Jankowicz-Cieslak J, Muller A, Otto I, Kumlehn J, Till BJ (2013) Validation of doubled haploid plants by enzymatic mismatch cleavage. Plant Methods 9(1):43

Kurowska M, Daszkowska-Golec A, Gruszka D, Marzec M, Szurman M, Szarejko I, Maluszynski M (2011) TILLING—a shortcut in functional genomics. J Appl Genet 52(4):371–390

Maluszynski M, Kasha KJ, Forster BP, Szarejko I (2003) Doubled haploid production in crop plants. A manual. Dordrecht/Kluwer Academic, Boston/London

Ravi M, Chan SW (2010) Haploid plants produced by centromere-mediated genome elimination. Nature 464(7288):615–618

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132:365–386

Till BJ (2014) Mining genetic resources via Ecotilling. In: Tuberosa R, Graner A, Frison E (eds) Genomics of plant genetic resources. Springer, Dordrecht, pp. 349–365

Till BJ, Burtner C, Comai L, Henikoff S (2004) Mismatch cleavage by single-strand specific nucleases. Nucleic Acids Res 32(8):2632–2641

Till BJ, Zerr T, Bowers E, Greene EA, Comai L, Henikoff S (2006a) High-throughput discovery of rare human nucleotide polymorphisms by Ecotilling. Nucleic Acids Res 34(13):e99

Till BJ, Zerr T, Comai L, Henikoff S (2006b) A protocol for TILLING and Ecotilling in plants and animals. Nature Protocols 1(5):2465–2477

Till BJ, Jankowicz-Cieslak J, Sagi L, Huynh OA, Utsushi H, Swennen R, Terauchi R, Mba C (2010) Discovery of nucleotide polymorphisms in the Musa gene pool by Ecotilling. Theor Appl Genet 121(7):1381–1389

Till BJ, Jankowicz-Cieslak J, Huynh OA, Beshir M, Laport RG, Hofinger BJ (2015) Low-cost methods for molecular characterization of mutant plants. Springer, Cham