Transgenic Research (2005) 14:109–115
DOI 10.1007/s11248-005-2770-x

Perspectives

TILLING moves beyond functional genomics into crop improvement

Ann J. Slade & Vic C. Knauf
Anawah Inc., 1102 Columbia Street, Suite 600, Seattle, WA 98104, USA

Received 17 February 2005; accepted 24 February 2005

Key words: crop improvement, SNP discovery, TILLING

Abstract

Transgenic methods have been successfully applied to trait improvement in a number of crops. However, reverse genetics studies by transgenic means are not practical in many commercially important crops, hampering investigations into gene function and the development of novel and improved cultivars. A nontransgenic method for reverse genetics called Targeting Induced Local Lesions IN Genomes (TILLING) has been developed as a method for inducing and identifying novel genetic variation, and has been demonstrated in the model plant, *Arabidopsis thaliana*. Recently, TILLING has been extended to the improvement of crop plants and shows great promise as a general method for both functional genomics and modulation of key traits in diverse crops.

Introduction

Crop improvement has a long history as key agronomic traits have been selected over thousands of years during the domestication of crops. More recently, this progress has been accelerated as the green revolution has brought about great increases in crop yields (Khush, 2001). With the advent of genomics in the last 25 years, opportunities for crop improvement have continued to grow and may help to meet future challenges of food production and land sustainability. Novel DNA sequence information allows the development of additional molecular markers for breeding as well as providing targets for transgenic alteration of gene expression and introduction of new traits. Recently, a method called Targeting Induced Local Lesions IN Genomes (TILLING) was developed to take advantage of this new DNA sequence information and to investigate the functions of specific genes. As demonstrated in a recent publication (Slade et al., 2005), TILLING also shows promise as a nontransgenic tool to improve domesticated crops by introducing and identifying novel genetic variation in genes that affect key traits.

The TILLING method

TILLING applies advances in molecular biology and genomics to the identification of genetic variation at the level of a single base pair. In the first step of the TILLING process, novel single base pair changes are induced in a population of plants by treating seeds (or pollen) with a chemical mutagen, and then advancing plants to a generation where mutations will be stably inherited (Figure 1). DNA is extracted, and seeds are stored from all members of the population to create a resource that can be accessed repeatedly over time. Establishing a good population and preparing DNA samples from what typically constitutes thousands of plants for a TILLING 'library' can easily take the better part of two years for crop plants. However, chemical mutagenesis is readily applicable to many diverse plants and animals, and so has the potential for widespread use (Kodym & Afza, 2003). In crops that propagate vegetatively or plants with long generation times (e.g., mint, coffee, and trees), creating such a population may be more difficult. However, even in these cases, natural variation in crop and

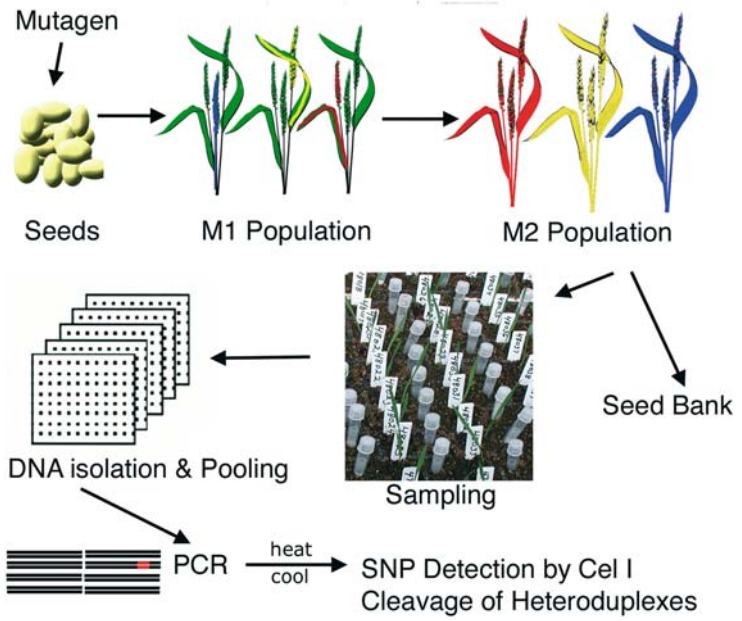

Figure 1. The TILLING Method. Seeds are treated with a chemical mutagen to induce genetic variation, and then planted. The resulting M1 population of plants is chimeric for mutations. Therefore, one seed from each M1 is planted to create the M2 population. M2 DNA is extracted from leaf tissue and M3 seeds from each plant are stored in a seed bank. DNA samples are pooled to increase throughput and PCR amplified with dye-labeled PCR primers specific to a target gene of interest. PCR products are denatured and allowed to reanneal to form heteroduplexes. Heteroduplex DNA is then cleaved by Cel I and analyzed.

non-crop species can by uncovered by 'ecoTILLING' (Comai et al., 2004).

For a TILLING assay, PCR primers are designed to specifically amplify a single gene target of interest. Specificity is especially important if a target is a member of a gene family or part of a polyploid genome. The application of bioinformatics tools can also help leverage functional genomics data across species for investigation of targets in novel crops. Next, dye-labeled primers are used to amplify PCR products from pooled DNA of multiple individuals. These PCR products are denatured and reannealed to allow the formation of mismatched base pairs (Figure 2(a)). Mismatches, or heteroduplexes, represent both naturally occurring single nucleotide polymorphisms (SNPs) (i.e., several plants from the population are likely to carry the same polymorphism) and induced SNPs (i.e., only rare individual plants are likely to display the mutation). After heteroduplex formation, the use of an endonuclease, Cel I, that recognizes and cleaves mismatched DNA is the key to discovering novel SNPs within a TILLING population (Oleykowski, et al., 1998; Colbert et al., 2001). Cleavage by Cel I occurs just 3' of the mismatch and, therefore, the sizes of the cleaved fragments indicate where the mutation resides in the fragment being analyzed (Figure 2(b)).

Using this approach, many thousands of plants (or animals) can be screened to identify any individual with a single base change as well as small insertions or deletions (1–30 bp) in any gene or specific region of the genome (Comai et al., 2004). To increase throughput, DNA is typically pooled from 4 to 8 individuals. Genomic fragments being assayed can range in size anywhere from 0.3 to 1.6 Kb. At 8-fold pooling, 1.4 Kb fragments (discounting the ends of fragments where SNP detection is problematic due to noise) and 96 lanes per assay, this combination allows up to a million base pairs of genomic DNA to be screened per single assay, making TILLING a high-throughput technique.

TILLING in wheat

TILLING was originally developed for use in the fully sequenced diploid model organism,

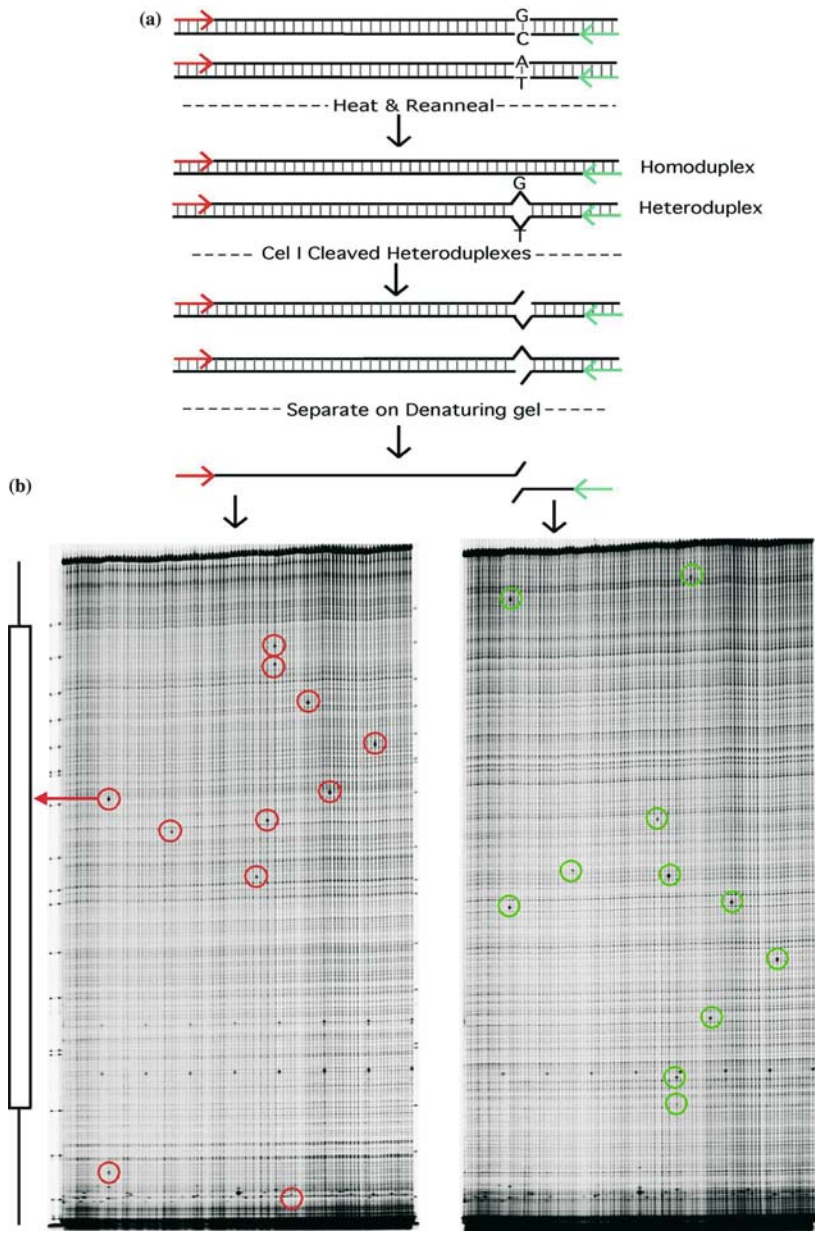

Figure 2. Cel I Cleavage and Analysis. (a) Illustration of a Cel I cleavage reaction. PCR primers that have been end-labeled with two different color dyes (red and green arrows) are used to amplify a targeted region of the genome in a pool of DNA consisting of multiple individuals. After PCR, DNA fragments are denatured and allowed to reanneal to form homoduplexes and heteroduplexes. Cel I is added to the reaction and cleaves DNA 3' of the mismatch. The cleavage reaction is concentrated, denatured and separated electrophoretically on a LI-COR DNA analyzer. (b) Example of a TILLING assay on 6-fold pooled soybean DNA. Images of the cleavage reaction are collected independently for each dye and analyzed using imaging software (Adobe Photoshop). The sizes of the cleavage products (circled) from the two dye labeled DNA strands (red or green) add up to the size of the full-length PCR product (at the top of the gel). This feature of the TILLING assay provides a built-in means to determine if a band represents a true mismatch cleavage product or a PCR artifact, thereby reducing the number of false positives. The size of the cleavage product (sizing ladder is flanking left image) indicates approximately where the SNP is in the fragment being analyzed (line drawing and arrow). Once a pool containing a SNP of interest is identified, the individual members of the pool can all be sequenced, or alternatively, they can be TILLED as individuals and only the one containing the SNP sequenced. TILLING assays typically have a background banding pattern that is apparent even without Cel I treatment. These faint bands are thought to represent aborted PCR products that are visible because even small quantities of labeled PCR product are detectable.

Arabidopsis thaliana, and thousands of mutations in hundreds of *Arabidopsis* genes have since been identified by this method (Till et al., 2003). It was uncertain whether or not TILLING could be applied successfully to species with more complex genomes and with much less sequence information available. Just such a challenging candidate for TILLING is wheat, an economically important crop with a hexaploid genome 140 times the size of *Arabidopsis*. Desirable genetic changes in wheat are difficult to identify based on phenotypic screens because these changes are often masked by redundant copies of genes within its hexaploid genome.

We created a TILLING library in both bread and durum wheat in order to determine if TILLING could be applied to trait modification in the context of a complex genome in a commercially important crop. We targeted a well-characterized gene in wheat of economic importance encoding granule bound starch synthase I, or the *waxy* locus. Loss of all copies of this gene in hexaploid wheat results in the production of *waxy* starch (lacking amylose) that has unique physicochemical properties (Graybosch, 1998). Although of great commercial interest, production of the first *waxy* wheat by traditional breeding was delayed by the scarcity of available genetic variation at one of the *waxy* loci. Targeting the *waxy* loci by TILLING (Slade et al., 2005), we developed PCR primers specific to each locus and identified ~250 novel alleles in hexaploid and tetraploid commercial cultivars of wheat. The impact of these different genetic changes was predicted by bioinformatics analysis to range from severe to no effect, and 94 of these mutations were predicted to alter the encoded *waxy* gene products. In the paper, a triple homozygous mutant line containing severe mutations in two *waxy* loci (along with a naturally occurring deletion of the third locus) was evaluated, and a near *waxy* phenotype was detected. Neither point mutant by itself was distinguishable from wild-type by the phenotypic screen, emphasizing the difficulties of crop improvement in wheat through conventional breeding.

These experiments demonstrated that TILLING is a means to rapidly generate and identify many novel alleles, some of which were demonstrated to have a phenotypic effect. In addition, these alleles represent a rich resource of genetic diversity at the wheat *waxy* loci for potential modulation of starch quality and characteristics. Furthermore, these experiments successfully established that TILLING could be extended to polyploid genomes and to crop improvement.

Approaches for SNP discovery

Although there are many different methods to detect and genotype previously identified SNPs (Kwok, 2001), not as many methods are suited to discovering novel SNPs. A useful SNP discovery method must be high-throughput and yet sensitive enough to sift through thousands of samples for rare SNPs. In addition, the cost of analysis affects how widely a method can be used. Finally, a feature that we have found to be critical in our efforts to discover novel SNPs in various crop varieties is the ability to detect rare events even in presence of common background genetic variation.

Resequencing, or sequencing every individual in a test population over a specific genomic region, is a sensitive straightforward means to identify novel SNPs. However, resequencing is expensive, especially if many different genes are to be studied. Mass spectrometry is another means to detect changes relative to a test sample, but requires PCR fragments to be transcribed into RNA and four enzymatic reactions to be performed prior to analysis (Stanssens et al., 2004). Another method, called single-strand conformation polymorphism (SSCP) relies on differences in secondary structure within a single DNA strand to identify novel base changes. SSCP can require extensive optimization, small fragment sizes (< 250 bp), and is not conducive to pooling strategies to increase throughput (Kuhn et al., 2005).

In contrast, denaturing HPLC (DHPLC) and temperature gradient capillary electrophoresis (TGCE) are two methods for SNP discovery that can employ pooling strategies to increase throughput (Li et al., 2002; Lilleberg, 2003). Both rely on partially denaturing conditions to resolve heteroduplex DNA fragments with mismatches from homoduplex DNA. However, DHPLC requires the single running temperature to be optimized for each fragment, whereas TGCE has the advantage of analyzing multiple differently

sized fragments over a range of temperatures so that prior optimization is not necessary. One caveat to the practical application of TGCE to SNP discovery in crop plants is that common or even less frequent polymorphisms present in a population can complicate analysis rendering the identification of novel induced SNPs very difficult (Figure 3(d)).

TILLING is well suited as a SNP discovery method because it is high-throughput, sensitive enough to detect mutations in pooled DNA that are present in 1/16 of a pool (e.g., a heterozygous mutation in a pool of 8), cost-effective, and the size of the Cel I cleaved DNA indicates where the mutation resides in the fragment being analyzed, simplifying subsequent analysis. Importantly for SNP discovery in crops, it is a robust method for which background variability in the starting material is immaterial. Unlike in *Arabidopsis*, isogenic starting material to construct a TILLING library is not always practical for crop plants. However, common or even less frequent polymorphisms that can complicate SNP discovery by other methods are readily distinguishable from novel induced mutations by TILLING (Figure 3(a)–(c)). Even in a background of over a dozen other common SNPs, the TILLING method is sensitive enough to detect a rare SNP.

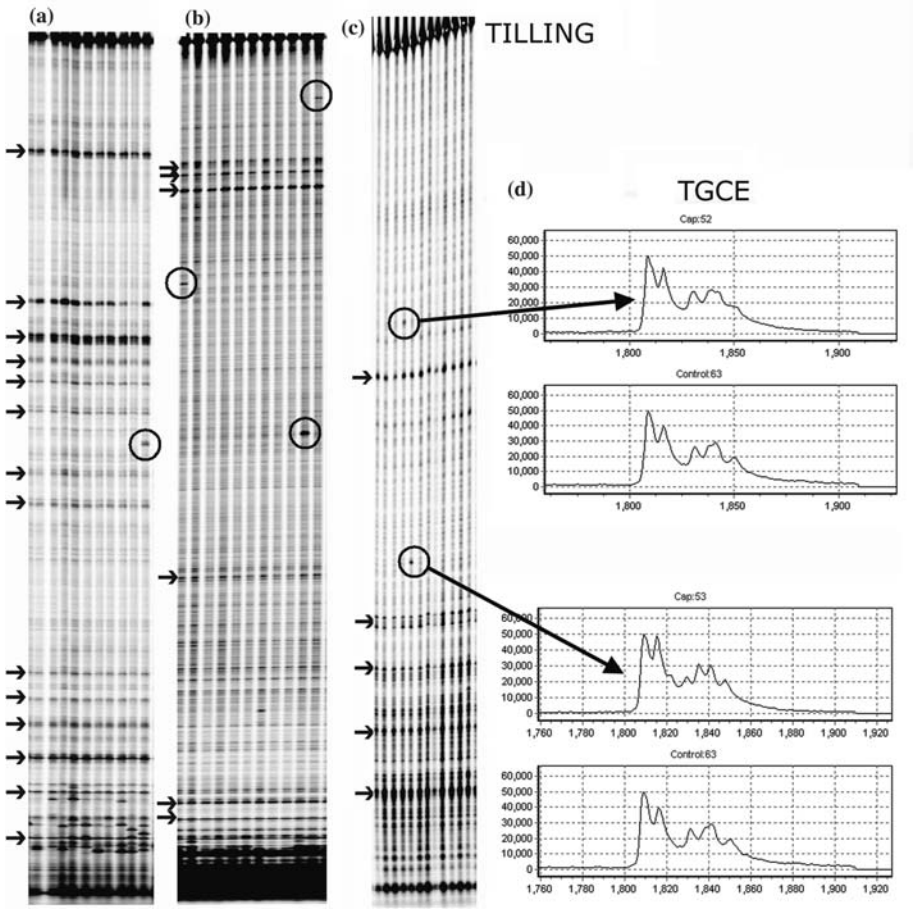

Figure 3. Novel SNP discovery by TILLING in genetic backgrounds containing frequent polymorphisms. (a) Peanut TILLING assay (4-fold) with many polymorphisms in the population (arrows). A unique SNP is still detected by TILLING (circle). (b) Soybean TILLING assay (6-fold) with several common polymorphisms in the population (arrows). Three unique SNPs are circled. (c) Zebrafish TILLING assay (4-fold) with two unique SNPs (circled) clearly evident amidst several polymorphisms. (d) Unlabeled PCR products from the same pooled DNA as in c) (arrows) were analyzed using TGCE. The resulting electropherograms were compared with electropherograms from lanes containing no additional mutations as a control. Similar results were obtained over three different temperature ranges using TGCE.

More than a knockout

What does TILLING offer of unique advantage over the already available set of resources and techniques for functional genomics and suppressing gene function? With the possible exception of naturally occurring transposon systems in maize, most methods (transposon, T-DNA, antisense, and RNAi) rely on transgenic introduction of foreign DNA. For *Arabidopsis*, this is not an issue; however, the efficiency of gene transfer and subsequent plant regeneration can become a serious limitation in many crops. T-DNA insertions and/or transposon insertions may be the preferred means to obtain a specific gene knockout but are practically limited to the crops for which they are available. RNAi has the advantage of knocking down the expression of multiple related genes with one construct (Lawrence & Pikaard, 2003), whereas TILLING, like T-DNA insertions and transposons, is unlikely to affect more than one specific member of a multi-gene family in an individual plant.

The application of TILLING to crop improvement may also help with another constraint in domesticated species' genomes: limited genetic variation. During domestication and subsequent selection, much of the genetic variation available in the wild crop progenitors has been lost (Gepts & Papa, 2002). Thus, plant breeders have at times used wild relatives or landraces to introduce useful genetic variation. This practice has been successful in wheat for developing disease resistant and higher yielding varieties (Zamir, 2001), and a landrace was also used for the development of the first full *waxy* line because it carried a rare deletion allele of one of the *waxy* loci (Graybosch, 1998). As an alternative to the use of wild varieties, TILLING can be a means to introduce genetic variation in an elite germplasm without the need to acquire variation from exotic cultivars, thus avoiding introduction of agriculturally undesirable traits. In addition, the issue of biopiracy makes the use of exotic varieties to improve modern cultivars potentially fraught with complications. For example, the identification of a caffeine free Arabica coffee by Brazilian scientists in germplasm that came originally from Ethiopia has prompted disputes over ownership (Silvarolla et al., 2004).

Functional genomics by TILLING

What is unique for the TILLING approach compared to transgenic approaches is the identification of numerous mutations within a targeted region of the genome. These mutations constitute allelic series that can potentially confer a range of phenotypes from subtle to strong, and allow structure function studies. Mutations in the coding regions of genes have the potential to alter plant metabolism in ways other than changing the effective level of a target gene product. For example, a mutation may change the affinity of an enzyme for its substrate, alter regulatory domains within enzymes, or may interfere with proper subunit or other protein–protein interactions. Within a metabolic pathway, such alterations can have large effects. TILLING offers a way to investigate a target gene of interest in potentially any crop of interest without first having knowledge of the gene product, which seems to us the essence of a useful tool for functional genomics.

If a transformation system is available for a crop and there are only a few genes of interest in which one would like to have knockouts to help ascertain gene function, RNAi may be the current method of choice. However, TILLING offers many advantages in cases where transformation is difficult or if the investigation of a continuing series of unknown genes in a specific crop is desired. Once a TILLING library is set up, it becomes a renewable resource for continued analysis of many different gene targets. Thus the reiterative cost and time to analyze many different targets is much less by TILLING than by gene suppression using transgenics.

Members of an allelic series will contain a number of background mutations that will be different for each individual. If a given phenotype is evident in multiple members of the series, association of the phenotype with the mutation in the target gene is likely. To further assess phenotypes, two members of an allelic series could be crossed with each other to render all background mutations heterozygous to increase the likelihood that the phenotype results from the identified SNPs. In a polyploid such as wheat, combinations of mutant alleles at loci from all three contributing genomes may be necessary prior to detecting any phenotype. For agronomic end use of a variety acquired through TILLING, a number

of backcrosses (four or more) may also be required to purge potentially undesirable background mutations. When backcrossing, the removal of background mutations can be accelerated through marker-assisted selection using the SNP itself as a molecular marker.

Conclusions

TILLING is a powerful reverse genetic approach that has the unique advantage of allowing the generation of an allelic series for potentially any target gene, including essential genes. TILLING represents a high-throughput, sensitive, cost-effective, and rapid means to find genetic variation to enable functional genomics studies and crop improvement. This technique works well even if a population contains many pre-existing polymorphisms that would compromise SNP discovery by other methodologies. TILLING is generally applicable—its use has been demonstrated in *Arabidopsi*s, *Lotus*, maize, zebrafish, rat, and *Drosophila* (reviewed in Henikoff et al., 2004). In addition, we have successfully applied TILLING to soybean, maize, romaine and iceberg lettuce, tomato, rice, peanut, bread and durum wheat, and castor. These successes indicate that TILLING is generally applicable to genomes whether small or large, diploid or even allohexaploid, and thus has great potential to identify both induced and naturally occurring variation in many species. When such variation has commercial implications for crop development, TILLING offers the added advantage that resulting crop varieties are not subject to the same regulatory approval requirements as transgenic crops.

References

Colbert T, Till BJ, Tompa R, Reynolds S, Steine MN, Yeung AT, McCallum CM, Comai L and Henikoff S (2001) High-throughput screening for induced point mutations. *Plant Physiol* **126**: 480–484.

Comai L, Young K, Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR and Henikoff S (2004) Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J* **37**: 778–786.

Gepts P and Papa R (2002) Evolution during domestication. *Encycl Life Sci*: 1–7.

Graybosch RA (1998) Waxy wheats: Origin, properties, and prospects. *Trends in Food Science and Technology* **9**: 135–142.

Henikoff S, Till BJ and Comai L (2004) TILLING. Traditional Mutagenesis Meets Functional Genomics. *Plant Physiol* **135**: 630–636.

Khush GS (2001) Green Revolution: The way forward. *Nature Rev Genet* **2**: 815–822.

Kodym A and Afza R (2003) Physical and chemical mutagenesis. *Methods Mol Biol* **236**: 189–204.

Kuhn D, Borrone J, Meerow A, Motamayor J, Brown J and Schnell R (2005) Single-strand conformation polymorphism analysis of candidate genes for reliable identification of alleles by capillary array electrophoresis. *Electrophoresis* **26**: 112–125.

Kwok P-Y (2001) Methods for genotyping single nucleotide polymorphisms. *Annu Rev Genomics Hum Genet* **2**: 235–258.

Lawrence RJ and Pikaard CS (2003) Transgene-induced RNA interference: a strategy for overcoming gene redundancy in polyploids to generate loss-of-function mutations. *Plant J* **36**: 114–121.

Li Q, Liu Z, Monroe H and Culiat CT (2002) Integrated platform for detection of DNA sequence variants using capillary array electrophoresis. *Electrophoresis* **23**: 1499–4511.

Lilleberg SL (2003) In-depth mutation and SNP discovery using DHPLC gene scanning. *Curr Opin Drug Discov Devel* **6**: 237–252.

Oleykowski CA, Bronson Mullins CR, Godwin AK and Yeung AT (1998) Mutation detection using a novel plant endonuclease. *Nucleic Acids Res* **26**: 4597–4602.

Silvarolla MB, Mazzafera P and Fazuoli LC (2004) A naturally decaffeinated arabica coffee. *Nature* **429**: 826.

Slade AJ, Fuerstenberg SI, Loeffler D, Steine MN and Facciotti D (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat Biotechnol* **23**: 75–81.

Stanssens P, Zabeau M, Meersseman G, Rernes G, Gansemans Y, Storm N, Hartmer R, Honisch C, Rodi CP, Bocker S and Boom Dvd (2004) High-throughput MALDI-TOF discovery of genomic sequence polymorphisms. *Genome Res* **14**: 126–133.

Till BJ, Reynolds SH, Greene EA, Codomo CA, Enns LC, Johnson JE, Burtner C, Odden AR, Young K, Taylor NE, Henikoff JG, Comai L and Henikoff S (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. *Gen Res* **13**: 524–530.

Zamir D (2001) Improving plant breeding with exotic genetic libraries. *Nature Rev Genet* **2**: 983–989.