TILLING in the Botanical Garden: A Reverse Genetic Technique Feasible for all Plant Species

George W. Haughn* • **Erin J. Gilchrist**

Department of Botany, University of British Columbia, 6270 University Blvd, Vancouver, BC, V6T 1Z4, Canada

Corresponding author: * **haughn@interchange.ubc.ca**

Keywords: ecotilling, reverse genetics, single nucleotide polymorphisms (SNP), Targeting Induced Local Lesions IN Genomes (TILLING)

ABSTRACT

Targeting Induced Local Lesions IN Genomes (TILLING) is a powerful technology that employs heteroduplex analysis to detect which organisms in a population carry single nucleotide mutations in specific genes. Genes are amplified by PCR using pooled genomic DNA from several individuals as a template. Following denaturation and renaturation of the amplified DNA, heteroduplexes form if organisms with wild type and mutant sequence are both present in the pool. The heteroduplexes can be detected by cleavage with an endonuclease and resolution of the resulting fragments on a sequencing gel. TILLING can be an effective reverse genetic technique if it is used to screen populations mutagenized with chemical mutagens such as ethyl methane sulfonate (EMS). Since such mutagens induce a diverse array of mutant alleles at a high frequency in any organism without the need of transgenic technology, TILLING is more versatile, universal and requires a smaller mutagenized population than other reverse genetic methods. TILLING can also be used to detect naturally occurring single nucleotide polymorphisms (SNP's) in genes among accessions, varieties, ecotypes or cultivars. These SNP's can serve as genetic markers in mapping, breeding and genotyping and can provide information concerning gene structure, linkage disequilibrium, population structure or adaptation.

1. INTRODUCTION

Genetic variation is a powerful resource that humans have exploited over the millennia to advance biological knowledge and generate the crops and horticultural varieties that have become so much a part of everyday life. In recent years, the availability of genomic sequences from many plant species and the development of a wide array of molecular-genetic technologies have enhanced our ability to detect or engineer such variation at specific genetic loci (reverse genetics), greatly expanding our capacity for both probing gene function and genetic engineering. Among the variety of technologies available for detecting DNA sequence polymorphisms in specific genes, TILLING (Targeting Induced Local Lesions IN Genomes) stands out as one of the most versatile and universal. TILLING can be used to identify point mutations in specific genes to study or manipulate gene function, or to detect genetic markers in population studies. In this chapter we describe TILLING and its applications, highlighting both the strengths and weaknesses of the approach. This information can act as a guide to allow researchers to decide whether TILLING is a suitable reverse genetic approach for attaining their specific research goals. In addition, we outline the steps required to establish a TILLING program in new species. For additional information we refer the reader to several valuable reviews published recently (Henikoff and Comai 2003, Henikoff *et al*. 2004, Stemple 2004, Gilchrist and Haughn 2005, Slade and Knauf 2005).

2. THE TILLING TECHNIQUE

TILLING uses heteroduplex analysis to identify, within a population, individuals having single DNA base pair (bp) differences in a specific target sequence. The current TILLING procedure is simple, effective, universal and amenable to moderately high throughput (McCallum *et al.* 2000, Colbert *et al.* 2001, Till *et al.* 2003). The TILLING process is outlined in **Fig. 1**. As a first step, PCR forward and reverse primers are designed to amplify 1,500 bp or less of genomic DNA from a locus of interest (target sequence). The primers are synthesized such that the forward and reverse primers are tagged with different fluorescent markers. The PCR template consists of genomic DNA isolated from individual plant lines derived from a population carrying point mutations throughout the genome. Typically, the genomic DNAs from 2-8 plants are pooled prior to PCR amplification. Such pooling not only increases screening efficiency but, since only one plant of the pool is likely to carry a mutation in the target sequence, ensures that any pool with a mutated target sequence will also have the wild type target sequence. Therefore, amplification of the target sequence from a pool where at least one plant has a nucleotide polymorphism, produces a mixture of both mutant and wild type PCR products. The amplified DNA is denatured by heating and then renatured. If the pool of plants includes one with a nucleotide difference in the target sequence, mutant/wild type DNA heteroduplexes form following renaturation. Heterduplexes are detected by treatment of the amplified DNA with CEL1 endonuclease or any one of a number of single strand endonucleases (Till *et al.* 2004a) followed by electrophoretic separation

on a DNA sequencing gel. The endonuclease cleaves either strand of the heteroduplex at the site of the mismatch producing two complementary sized fluorescent DNA fragments from the original ~1,500 bp amplified product. The identification of such cleavage products can be traced back to a specific pool, one plant line of which must contain the nucleotide polymorphism in the target gene. A second round of TILLING is then used to determine which of the plant lines in the pool carries the polymorphism (DNA from individuals in the pool is mixed with wild type DNA to ensure heteroduplex formation when the mutation is present). TILLING identifies the position of the nucleotide change in the amplified fragment to within a few base pairs. Typically, DNA sequencing of the amplified fragment from the individual is used to verify the nature and exact position of any mutation detected.

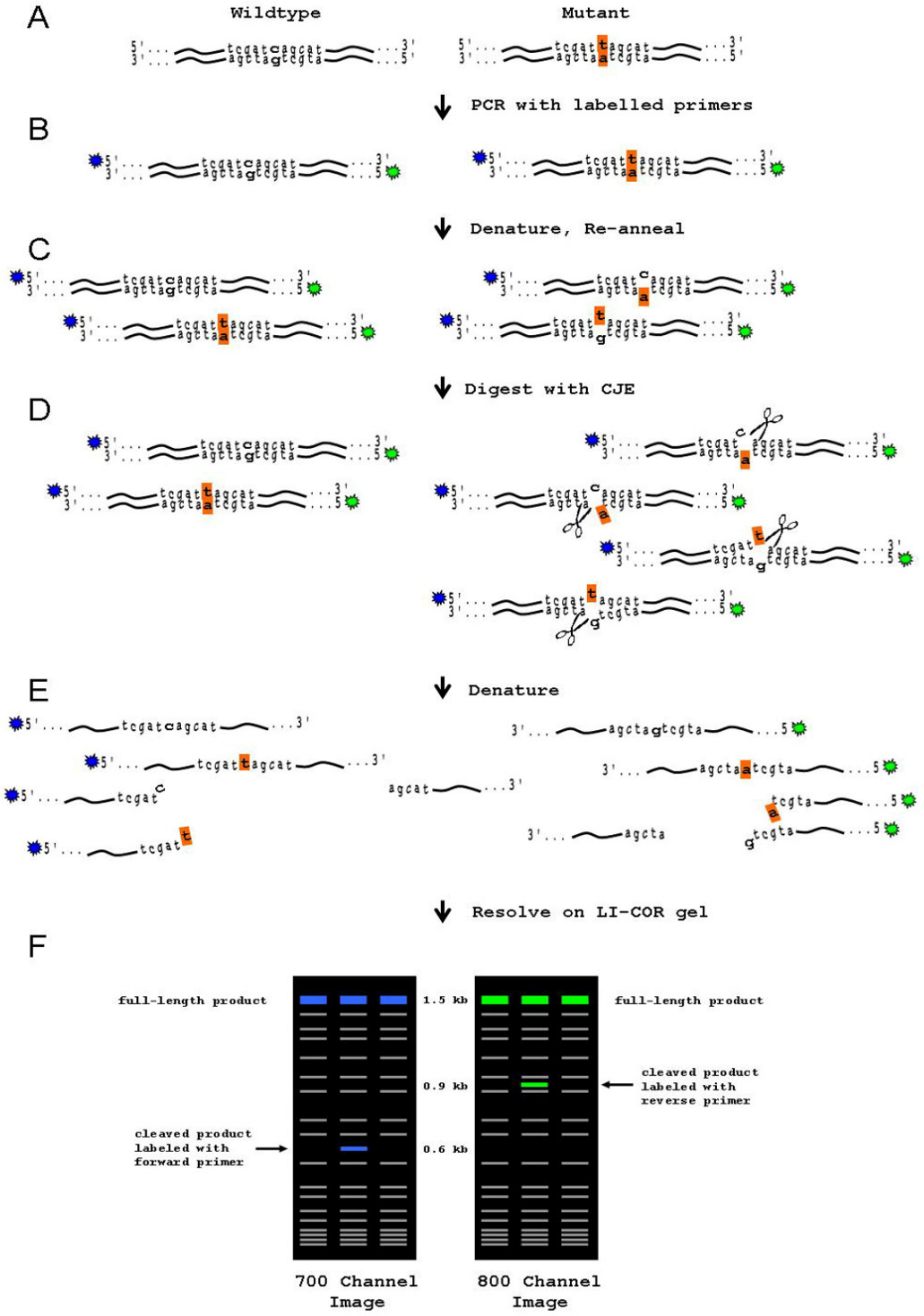

Fig. 1 Overview of the TILLING procedure. Pooled DNA (**A**) is amplified using fluorescently tagged, gene-specific primers (**B**). The forward and reverse primers are labeled with different fluorophors **(IRD700 and IRD800)** that each label a different end of the PCR fragment. The amplified products are denatured by heating and then allowed to cool slowly so that they randomly re-anneal (**C**). Heteroduplex molecules form when single DNA strands of mutant and wild-type PCR products anneal together. These heteroduplexes become targets for a single-strand-specific nuclease as found in celery juice extract. The nuclease cleaves these heteroduplex fragments at one of the two strands, 3' to the site of the mismatch in the DNA (**D**). The PCR products that retain one of the labeled primers can be detected on polyacrylamide denaturing LI-COR gels (**E, F**). Plants with a mutation in the gene of interest are identified by the fact that these individuals show a smaller, cleavage fragment as well as the wild-type product. Because the nuclease cleaves either of the two strands randomly, cleavage products can be detected in both the IRD700 and IRD800 channels of the gel image (**F**). The position of the mutation within the PCR amplicon can be calculated from the size of the fragments carrying the IRD700-labeled forward primer and the IRD800-labeled reverse primer.

3. APPLICATIONS OF TILLING

3.1. Reverse genetics

3.1.1. Reverse genetic strategies

TILLING was first developed for Arabidopsis as a novel reverse genetic technique (McCallum *et al.* 2000) to complement existing technologies. Other reverse genetic approaches for obtaining loss of gene function in plants include RNAi (formerly referred to as Post Transcriptional Gene Silencing) and detection of gene specific insertions or deletions in plants from a mutagenized population. Each of these approaches has its own strengths and weaknesses and should be considered when electing the best strategy for a given circumstance.

The RNAi approach requires the construction of a chimeric gene consisting of both sense and antisense sequence from the target gene, the transcript of which forms a double stranded mRNA. When this chimeric gene is transformed into the species of interest this double stranded mRNA induces degradation of the endogenous (target) gene mRNA, resulting in post-transcriptional gene silencing (Fagard and Vaucheret 2000, Matthew 2004, Helliwell and Waterhouse 2005). Unlike other approaches, RNAi silencing can be achieved without the need for a large mutagenized population (see below). The dominant nature of RNAi allows for detection of phenotypes even in heterozygous individuals but reduces the usefulness of the approach for studying genes where loss of function leads to lethality or sterility. The level of RNAi silencing and, therefore, the severity of phenotype typically varies between independent lines transformed with the same construct. This variation in silencing ranges from none to partial loss-of function but rarely is a functional null observed. Unfortunately, the labour- intensive nature of transformation and molecular analysis of transgenic lines makes high throughput challenging for RNAi.

Both transposon and T-DNA mutagenesis have also been commonly used for reverse genetics in plants. Transposon mutagenesis for reverse genetics requires an active transposon of known sequence, present in the species of interest, while T-DNA mutagenesis requires an efficient *Agrobacterium* transformation system to generate an insertional mutagenized population big enough for reverse genetics. Thus the usefulness of these approaches is limited to those species that meet these criteria. In plants that are amenable to this technique, insertions in a specified target can be detected from a mutagenized population by PCR amplification using one gene-specific and one insertion-specific primer (Walden 2002, Alonso *et al.* 2003, May and Martienssen 2003). Insertions typically occur relatively randomly in the genome and can result in strong loss-of-function mutations (knockouts). However the frequency of mutation tends to be low (**Table 1**), necessitating the screening of a large mutagenized population of plants (Arabidopsis requires a population of 100,000-200,000; species with larger genomes need proportionately more.) for a high probability of finding insertions in any gene.

Another commonly used reverse genetic technique involves detecting deletions in target genes, using PCR amplification. Primers are used to amplify the target sequence from pooled genomic DNA isolated from plants mutagenized with ionizing radiation (which induces deletions). Deletions are then detected by the presence of a PCR product shorter than the one amplified from wild type sequence (Li *et al*. 2001 2002). Since virtually any organism can be mutagenized with radiation the application of this approach is more universal than insertional mutagenesis and, like insertions, deletions typically result in knockout mutations. However, deletions vary in size and can eliminate function of more than a single gene at a given locus. Elimination of more than one gene can be an advantage in situations of tandem duplications, but can make it more difficult to ascribe individual gene function on the basis of mutant phenotype. Since the induced frequency of deletion is relatively low (**Table 1**), large populations (>100,000) of mutagenized plants also need to be screened to ensure the isolation of a mutation in the target gene.

Table 1 Mutation frequencies induced by different mutagens for reverse genetics in Arabidopsis and by EMS for TILLING in a variety of different organisms.

Organism	Mutagen	Mutations/1kb/ind.	Reference
Arabidopsis	Fast neutron	0.27 x 10^{-5} (deletion) (25 genes, 52,000 ind.)	Li *et al*. 2001
Arabidopsis	T-DNA	0.55 x 10^{-5} (insertion) (whole genome, 127,706 ind.)	Alonso *et al*. 2003
Arabidopsis	EMS	328 x 10^{-5} (point mutation) (125 genes; 3000 ind.)	Greene *et al*. 2003
Barley	EMS	175 x 10^{-5} (point mutation) (2 genes; 4608 ind.)	Caldwell *et al*. 2004
maize	EMS	255 x 10^{-5} (point mutation) (11 genes; 750 ind.)	Till *et al*. 2004
tetraploid wheat	EMS	2466 x 10^{-5} (point mutation) (2 genes; 768 ind.)	Slade *et al*. 2005
hexaploid wheat	EMS	5084 x 10^{-5} (point mutation) (2 genes; 1152 ind.)	Slade *et al*. 2005
Drosophila	EMS	730 x 10^{-5} (point mutation) (3 genes; 2086 ind.)	Draper *et al*. 2005

3.1.2. TILLING as a reverse genetic technique

EMS mutagenesis

Like several other reverse genetic approaches, TILLING is used to detect induced mutations following mutagnesis, but it is unique in its detection of point mutations. Point mutations can be induced at high frequency by a number of different mutagens but, to date, the chemical alkylating agent ethylmethane sulfonate (EMS) has been the mutagen of choice for the creation of TILLING populations in plants. EMS can be used to induce primarily recessive loss-of-function mutations at frequencies several orders of magnitude higher than insertions or deletions (Greene *et al.* 2003). Such high mutation frequencies allow the use of much smaller populations which in turn reduces the time and cost associated with producing and TILLING them. For example, TILLING just 3,000 EMS-mutagenized Arabidopsis plants identifies approximately ten different mutations in a 1 kb region of a target sequence, and one to two of these mutations are expected to completely eliminate the function of the gene (**Fig. 2**). A T-DNA-mutagenized population 30 times this size would be needed to identify one to two alleles. The upper limit of mutation frequency within any mutagenized population is determined largely by the cumulative detrimental effects of deleterious mutations which limit viability and fertility. The EMS induced mutation frequencies currently achieved in self pollinating diploid species such as Arabidopsis (Greene *et al.* 2003) or barley (Caldwell *et al.* 2004) may be approaching this limit (**Table 1**). However, since either gene redundancy or heterozygosity generated by out-crossing can reduce the impact of deleterious mutations, it should be possible to achieve much higher mutation frequencies in species that are tetraploid and hexaploid, or are out-crossed following mutagenesis. Indeed, EMS mutation frequencies in tetraploid and

hexaploid wheat were found to be an order of magnitude higher than that for Arabidopsis or barley (**Table 1**; Slade *et al.* 2005). High mutation frequencies do come at some cost. Since any individual from such a population may have up to 500 distinct mutations in its genome, backcrossing and segregation analysis may be required to eliminate unwanted background mutation and determine whether or not an observed phenotype is due to mutation of the target gene. With TILLING, however, the generation of multiple alleles of the same target gene simplifies the correlation between mutation and phenotype since more than one allele should give rise to the same phenotype if it is caused by a disruption of the target gene's function.

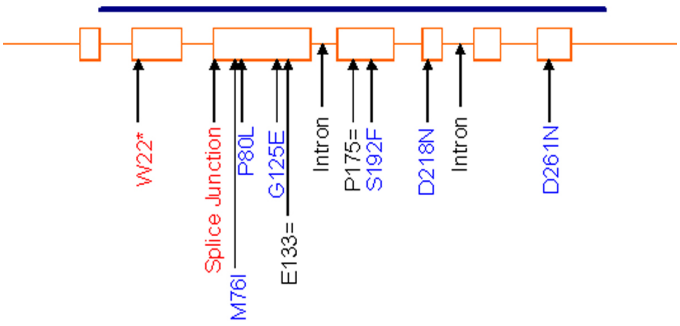

Fig. 2 Gene model, designed from PARSESNP output figure, showing results of TILLING on a candidate gene in approximately 3,000 mutagenized Arabidopsis plants. Orange open boxes denote exons, and orange lines, intron and intergenic sequences. Solid blue bar indicates the extent of the target sequence amplified. Mutations that introduce a premature stop codon or splice-site error into the coding region of the gene are indicated in red, missense mutations are indicated in blue, and mutations that have no effect on the protein product are indicated in black. The first letter in the text for each mutation indicates the wild-type amino acid at the site, followed by the position of this amino acid within the protein, and then the amino acid change induced by the mutation. Mutations in introns are indicated as such, "=" means no change in the amino acid encoded by that codon, and "*" indicates the introduction of a stop codon.

EMS induced mutations are primarily G to A transitions and occur randomly throughout the genome (Greene *et al.* 2003). Consequently, an advantage of TILLING is that it has the potential to detect hundreds of distinct alleles in every gene of an organism (**Fig. 2**). Approximately 5% of the mutations in the coding region will cause truncations of the protein, either through introduction of stop codons (nonsense mutations) or splice junction defects. The vast majority of such alleles will be functional gene knockouts. Roughly 50% of EMS-induced mutations within coding regions result in an amino acid substitution (missense). Such mutations can have one of a wide variety of effects on the function of a gene product, ranging from knockout to gain-of function to no effect, depending on the specific amino acid substitution. However missense mutations can inactivate specific domains of a protein without affecting the other domains making them extremely valuable in dissecting the function of a protein. **TILLING is the only reverse genetic technique that can deliver such allelic diversity.** Many point mutations will not have deleterious affects on gene function including most mutations detected in noncoding regions, some missense mutations (described above) and mutations in coding regions altering the third position of a codon. For this reason multiple alleles of each gene must be identified by TILLING in order to increase the probability of obtaining an allele with deleterious effects.

Universality of TILLING for reverse genetics

TILLING can reliably detect point mutations that can easily be induced at a high frequency by chemical mutagens in virtually any organism. Thus as a reverse genetic approach, TILLING is one of the most universal. Genomes can be mutagenized at high frequency regardless of size and, as discussed above, large redundant genomes can tolerate higher mutation frequencies. Consequently the size of a mutagenized population required for TILLING is relatively modest (<10,000) compared to other methods (>100,000). In addition, outbreeding species do not pose a significant problem because heterozygous mutations can be detected. Since transformation is not required, TILLING is available for species for which transformation is impossible or inefficient, and for commercial applications where transgenic organisms are undesirable. The wide applicability of TILLING is evident in its popularity and the diversity of organisms to which it has been applied. In just the last five years, mutagenized populations have been developed and used for TILLING in a remarkably diverse array of plants (*Arabidopsis*, lotus, barley, wheat and maize; McCallum *et al.* 2000, Colbert *et al.* 2001, Perry *et al.* 2003, Caldwell *et al.* 2004, Till *et al.* 2004b Slade *et al.* 2005) and animals (zebrafish, rat, *Drosophila*; Wienholds *et al.* 2003, Draper *et al.* 2004, Smits *et al.* 2004, Winkler *et al.* 2005), and new species are being added to this array constantly.

3.2. EcoTILLING

The genomes of individuals within a single species contain significant genetic variation that has arisen from spontaneous mutation. The vast majority of this diversity is in the form of single nucleotide changes commonly referred to as simple nucleotide polymorphisms (SNPs). Such naturally occurring SNPs are of great interest to scientists because they are useful as genetic markers in mapping, breeding and genotyping and can provide information concerning gene structure, linkage disequilibrium, population structure or adaptation.

A number of different techniques for identifying SNPs have been developed. Some of these detect differences in denaturation or single stand structure that result from changes in nucleotide sequence but such techniques fail to identify the number or position of mutations within the DNA fragment examined (DeFrancesco and Perkel 2001) so detection must be followed by sequencing to distinguish between different polymorphisms. The more direct methods of array hybridization or sequencing are currently expensive when applied to multiple loci in large numbers of individuals. TILLING provides an alternate approach to identification of naturally occurring SNPs in large populations that is both robust and relatively inexpensive. This application of TILLING has been termed Ecotilling (Comai *et al.* 2004).

In Ecotilling, genomic DNA from individuals of a single or several populations is isolated. As in TILLING, the pooling of genomic DNA from more than one individual is necessary to allow for detection of SNPs that are homozygous since CEL1 only cleaves heteroduplexes. If SNPs in a population occur relatively rarely (fewer than one polymorphic individual per pool), the DNA of up to eight such individuals can be pooled, as is done in TILLING. However, when most individuals within a population differ at one or more base pairs in any given specific target sequence, 8-fold pooling will complicate genotyping. For this reason, in highly heterozygous species, the genomic DNA from each individual is usually pooled only with DNA from a reference individual for which the target has been sequenced. In addition, to detect those loci that were heterozygous prior to pooling, unpooled genomic DNA from an individual is Ecotilled separately (**Fig. 3**).

Ecotilling detects the number and relative position of all SNP's, including point mutations, and small insertions and deletions, within a target

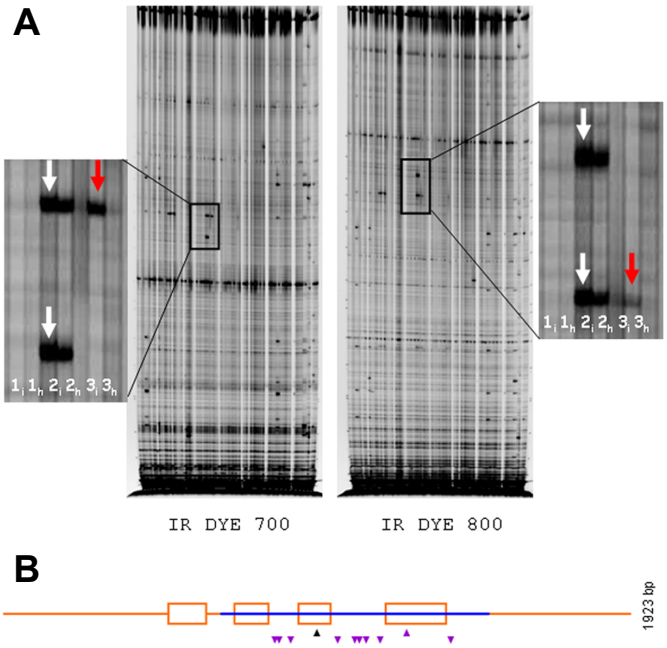

Fig. 3 (**A**) Images of Ecotilling gels obtained from each of the two fluorescent channels of the LI-COR sequencing machine. The IR DYE 700 channel shows the PCR products that carry the labeled forward primer and the IR DYE 800 channel shows the PCR products that carry the labeled reverse primer. Insets depict fragments that have been cleaved by celery juice extract endonuclease indicating a SNP in the amplicon. Multiple cleavage bands may be seen in each lane representing several polymorphic sites in one individual. DNA from each sample was run individually (lanes 1i, 2i, 3i...) and heteroduplexed with the reference DNA (1h, 2h, 3h...) in adjacent lanes on the gel. White arrows show heterozygous alleles identified by the presence of a cleavage product when self annealed. Red arrows show a homozygous SNP identified by a cleavage product only when heteroduplexed with DNA from the reference plant (Nisqually-1). (**B**) A schematic diagram from the PARSESNP output file shows the positions of the SNPs within the gene. Blue lines indicate the extent of the amplified region which is approximately 1 kb. Orange open boxes denote exons, and orange lines denote introns or intergenic DNA. Purple arrows indicate changes in the DNA sequence that do not affect the amino acid sequence of the protein product. The black arrow indicates a change that results in an amino acid polymorphism in the individuals carrying this SNP. Most of the polymorphisms in natural populations are found in non-coding DNA or are silent at the amino acid level.

sequence in each individual tested (**Fig. 3**). Thus both the spectrum of natural variation within the target sequence and the distribution of that variation throughout the population can be established. If knowledge of the specific nucleotide changes is required then DNA sequencing must be done following Ecotilling. However, since the number of different genotypes will normally be much smaller than the number of individuals examined, the target DNA from only a few representative individuals will need to be sequenced to establish the exact array of genotypes thus reducing the cost of SNP detection relative to direct sequencing.

The efficacy of Ecotilling has been demonstrated by two recent studies involving representative plants from different ecotypes of *Arabidopsis thaliana* (Comai *et al.* 2004) and different sub-populations of the poplar *Populus trichocarpa* (Gilchrist *et al.* 2006). Both studies were effective in rapidly identifying numerous target sequence SNP's within the populations examined. The poplar study provided information on population heterozygosity and linkage disequilibrium, identified a conserved potential regulatory domain in an intron and generated ecotype and species specific markers for genotyping. The fact that *A. thaliana* is a small inbreeding annual while *P. trichocarpa* is a large out-breeding perennial underscore the universality of Ecotilling.

4. CONSIDERATIONS FOR ESTABLISHING TILLING OR ECO-TILLING IN A NEW SPECIES

If a researcher is fortunate enough to work on an organism where a public TILLING or Ecotilling facility is already available (e.g. Arabidopsis http://tilling.fhcrc.org:9366/ or maize http://genome.purdue.edu/maizetilling/), then the only decision for TILLING will revolve around choosing the best target sequence within the gene of interest (see below). However, some scientists will be faced with the task of establishing a new TILLING initiative and must consider the requirements for doing so.

4.1. The TILLING population

The first consideration in producing a TILLING population is mutagenesis. Mutagenizing seed plants with EMS can be done effectively by soaking seeds in an EMS solution followed by planting. Typical conditions are 14-18 h in a 30-100 mM (0.3%-1%) EMS solution although optimal conditions are dependent on many variables and need to be established empirically for each species. The resulting mutagenized population (M1 generation) is grown to maturity and allowed to self-fertilize to produce M2 seeds. Plants that do not self-fertilize efficiently or suffer from severe inbreeding depression may need to be out-crossed to an unmutagenized parent or, alternatively, pollen could be mutagenized as has been done for maize (Till *et al.* 2004b). M2 seeds can be maintained as lines (seeds harvested from individual M1 plants) or bulked. If the M2 seed is bulked, then lines need to be established using M3 seed harvested from individual M2 plants. In either case, when sampling M2 plants to establish the TILLING population (see below) care must be taken to minimize the number of lines derived from the same M1 parent. Leaf tissue, for isolating genomic DNA, and M3 seeds to establish plant lines for distribution are harvested from individual M2 plants. Based on typical EMS mutation frequencies (**Table 1**) an M2 population of fewer than 10,000 M2 plants is required. To maximize the genetic diversity, the size of the M1 population should be relatively high (up to 10,000). If the seed set per plant is low there may not be enough M3 seed for distribution. In this situation it may be necessary to increase the seed from each line by growing multiple M3 plants for each line and bulking the resulting M4 seed.

4.2. Arraying the population for TILLING

The TILLING procedure has been adapted to a 96-well format to facilitate high throughput (Till *et al*. 2003). Genomic DNA is isolated from individuals and arrayed in a 96-well microtiter plate. The quality of genomic DNA greatly affects TILLING results and longevity of the samples so DNA isolation protocols that produce clean high quality DNA samples are worth the extra time and money. With current TILLING protocols, a single mutated DNA sequence can be detected among 15 normal ones allowing up to 8-fold pooling of diploid plants to increase TILLING throughput. Typically, individual genomic DNA samples from eight 96-well plates are pooled to produce a plate containing DNA from 768 plants that is used in the first round of TILLING. When pooling, it is important to normalize genomic DNA from individuals to ensure equal representation within the pool. In polyploid species with highly redundant sequences, amplification of mulitiple copies of a target sequence may

decrease the ratio of polymorphic sequences below the 1:16 threshold making their detection more difficult. This problem can be addressed either by reducing the pool size or redesigning the PCR primers to make amplification specific to a pair of homologous genes.

4.2.1. Choosing a target sequence

The optimal length of target sequence that can be TILLED in a single reaction is 1,500 bp. However most eukaryotic genes are over 2,000 bp and sometimes much longer requiring choices to be made concerning what portion of a gene will be targeted. Since the objective in TILLING is to identify plants with deleterious mutations in the target gene, and most mutations in non-coding sequences such as introns, untranslated regions, and promoters will have no effect on gene function, the target sequence should be chosen to minimize such sequences and maximize coding regions. As discussed above over 50% of mutations within a coding region will have no effect on gene function. The probability of finding deleterious mutations can be increased by selecting coding regions within the target that encode conserved domains. Such domains can be identified by comparison to homologous genes from other organisms. A program, named CODDLE (Till *et al.* 2003) has been written to aid in selecting target sequences for TILLING (http://www.proweb.org/coddle). The program constructs exon-intron gene models from submitted genomic sequence and coding information, and then identifies conserved domains within the target gene where mutations would be most likely to have an effect on gene function. CODDLE uses the Primer3 programme (Rozen and Skaletsky 2000) to design primers that will amplify the desired target sequence.

The selection of targets for Ecotilling, where identification of deleterious alleles is unlikely to be the major objective, may be based on other criteria. For example, a researcher may choose to maximize intron sequence and poorly conserved coding domains in order to maximize the likelihood of finding natural polymorphisms.

4.2.2. Technical Considerations for TILLING

The list of essential equipment for TILLING is not extensive. A central item is the sequencing gel apparatus. The LI-COR slab-gel analyzer system (LI-COR Biosciences, Lincoln, NE) with a 96-well loading format has proved to be well suited for TILLING because of the sensitivity of the LI-COR laser detectors (Colbert *et al.* 2001, Till *et al.* 2003). However the potential savings in time and gel costs that would result from the use of a capillary detection system make the development of this approach a desirable goal for the future. CEL1 endonuclease can be bought commercially or substituted with a crude celery juice extract prepared from 0.5 kg of celery in a kitchen-quality juicer followed by ammonium sulfate fractionation and dialysis (Till *et al.* 2004a). Other key items of equipment include a PCR thermocycler and centrifuge capable of holding 96 well microtitreplates, multichannel pipets (or robot liquid handler) and some means of purifying the samples in 96-well format before running on the TILLING gel (either sephadex purification or ethanol precipitation). A good database management system is also essential for tracking the samples and this has recently been simplified by the release of GelBuddy, a programme designed to simplify analysis of the TILLING LI-COR gels (Zerr and Henikoff 2005).

TILLING throughput can depend on many variables. A rough estimate for a modest operation involving one to two LI-COR slab-gel analyzer systems and at least one trained person working full-time would be about 20-50 target genes, or 200-500 mutations per year. At least six months should be allowed to set up and test the facility and more if a population needs to be generated.

TILLING is not cheap relative to other reverse genetic techniques. Materials (fluorescent primers, microtiter plates, LI-COR gel supplies and consumables for PCR and sequencing reactions) needed to TILL a single target using a population of 3,000 plants with 8-fold pooling and verify mutations by sequencing costs approximately $US 1000 (including the cost of sequencing representative individuals). If the costs of labour, equipment and incidentals are factored in the cost per target would rise to approximately $US 2,500.

4.2.3. TILLING output

TILLING identifies individual plants with mutations of known position in the target gene. If the M2 plant is heterozygous for the mutation, homozygous M3 plants must be first identified before attempting to find an associated phenotype. Typically a researcher will save time by pursuing only those mutations that are likely to result in a phenotype. Plants with nonsense or splice junction mutations will almost certainly be deleterious but many missense mutations will not. The probability that any given missense mutation will be deleterious can be predicted on the basis of the specific amino acid exchange that has occurred. The PARSESNP program (Till *et al.* 2003), made available for use on the web (http://www.proweb.org/parsesnp/) by the Seattle TILLING Project, can be used to predict the potential effect of missense mutations on gene function.

Because each individual mutagenized plant can have more than 200 mutations scattered throughout the genome and even strong deleterious mutations may not result in a phenotype due to redundancy, it is important to verify that any phenotype observed is due to the mutation detected by TILLING. The simplest method of doing so is to demonstrate that more than one deleterious independent allele in the target gene result in the same phenotype. Alternatively the mutant can be backcrossed to wild type to determine whether the mutation and the phenotype co-segregate among the F2 progeny. The mutation can be detected among the F2 progeny using TILLING or by developing a CAPS or dCAPS marker for this purpose (Neff *et al.* 1998, Neff *et al.* 2002, Komori and Nitta 2005), or by polymorphic restriction endonuclease sites. Co-segregation does not eliminate the possibility that the phenotype is due to mutation in a separate gene closely linked to the target and steps should be taken to make certain that this is not the case. Complementation of the mutant phenotype by transformation with the target gene represents a third approach to linking the phenotype with the target gene.

5. CONCLUSION

Clearly TILLING is a powerful new approach to SNP detection that can be used for reverse genetics or SNP analysis in mutagenized or natural plant populations respectively. The virtually universal applicability of TILLING, the high degree of certainty of identifying mutations in targets and the broad array of alleles that can be generated has attracted considerable attention to TILLING and resulted in an impressive array of TILLING initiatives in many different organisms including plants and animals, model systems, crops and trees. Drawbacks include the cost and labour-

intensity of the technique, as well as uncertainty surrounding the deleterious nature of missense mutations. However for many researchers working on organisms with few alternatives to reverse genetics or SNP detection in large natural populations, TILLING may represent the only practical solution to their research needs. Therefore, until the cost of sequencing is substantially reduced, interest in TILLING is likely to remain high.

ACKNOWLEDGEMENTS

We would like to thank Drs. Steven Henikoff and Luca Comai for generous support and advice during the establishment of our TILLING facility CAN-TILL, Drs. Gillian Dean, Xin Li and Clifford Weil for comments on the manuscript and UBC Michael Smith Laboratories, Genome BC, Genome Prairie and Genome Canada for financial support.

REFERENCES

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen HM, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Guilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) Genome-wide Insertional mutagenesis of Arabidopsis thaliana. *Science* **301,** 653-657

Caldwell DG, McCallum N, Shaw P, Muehlbauer GJ, Marshall DF, Waugh R (2004) A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare* L.). *Plant Journal* **40,** 143-150

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

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

DeFrancesco L, Perkel JM (2001) In search of genomic variation. *Scientist* **15,** 24-26

Draper BW, McCallum CM, Stout JL, Slade AJ, Moens CB (2004) A high-throughput method for identifying N-ethyl-N-nitrosourea (ENU)-induced point mutations in zebrafish. *Zebrafish* (2nd ed) *Genetics Genomics and Informatics* **77,** 91-112

Fagard M, Vaucheret H (2000) (Trans)gene silencing in plants: How many mechanisms? *Annual Review of Plant Physiology and Plant Molecular Biology* **51,** 167-194

Gilchrist EJ, Haughn GW (2005) TILLING without a plough: a new method with applications for reverse genetics. *Current Opinion in Plant Biology* **8,** 211-215

Gilchrist EJ, Haughn GW, Ying CC, Otto SP, Zhuang J, Cheung D, Hamberger B, Aboutorabi F, Kalynyak T, Johnson L, **Bohlmann J, Ellis BE, Douglas CJ, Cronk QCB** (2006) Use of Ecotilling as an efficient SNP discovery tool to survey genetic variation in wild populations of *Populus trichocarpa*. *Molecular Ecology* **15**, 1365-1376

Greene EA, Codomo CA, Taylor NE, Henikoff JG, Till BJ, Reynolds SH, Enns LC, Burtner C, Johnson JE, Odden AR, Comai L, Henikoff S (2003) Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. *Genetics* **164,** 731-740

Helliwell CA, Waterhouse PM (2005) Constructs and methods for hairpin RNA-mediated gene silencing in plants. *RNA Interference* **392,** 24-35

Henikoff S, Comai L (2003) Single-nucleotide mutations for plant functional genomics. *Annual Review of Plant Biology* **54,** 375-401

Henikoff S, Till BJ, Comai L (2004) TILLING. Traditional mutagenesis meets functional genomics. *Plant Physiology* **135,** 630-636

Komori T, Nitta N (2005) Utilization of the CAPS/dCAPS method to convert rice SNPs into PCR-based markers. *Breeding Science* **55,** 93-98

Li X, Lassner M, Zhang YL (2002) Deleteagene: a fast neutron deletion mutagenesis-based gene knockout system for plants. *Comparative and Functional Genomics* **3,** 158-160

Li X, Song YJ, Century K, Straight S, Ronald P, Dong XN, Lassner M, Zhang YL (2001) A fast neutron deletion mutagenesis-based reverse genetics system for plants. *Plant Journal* **27,** 235-242

Matthew L (2004) RNAi for plant functional genomics. *Comparative and Functional Genomics* **5,** 240-244

May BP, Martienssen RA (2003) Transposon mutagenesis in the study of plant development. *Critical Reviews in Plant Sciences* **22,** 1-35

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. *Plant Physiology* **123,** 439-442

Neff MM, Neff JD, Chory J, Pepper AE (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in Arabidopsis thaliana genetics. *Plant Journal* **14,** 387-392

Neff MM, Turk E, Kalishman M (2002) Web-based primer design for single nucleotide polymorphism analysis. *Trends in Genetics* **18,** 613-615

Perry JA, Wang TL, Welham TJ, Gardner S, Pike JM, Yoshida S, Parniske M (2003) A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume Lotus japonicus. *Plant Physiology* **131,** 866-871

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz, Misener S (eds) *Bioinformatics Methods and Protocols*, Humana Press, Totowa, NJ, pp 365-386

Slade AJ, Fuerstenberg SI, Loeffler D, Steine MN, Facciotti D (2005) A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nature Biotechnology* **23,** 75-81

Slade AJ, Knauf VC (2005) TILLING moves beyond functional genomics into crop improvement. *Transgenic Research* **14,** 109-115

Smits BMG, Mudde J, Plasterk RHA, Cuppen E (2004) Target-selected mutagenesis of the rat. *Genomics* **83,** 332-334

Stemple DL (2004) TILLING - a high-throughput harvest for functional genomics. *Nature Reviews Genetics* **5,** 145-U19

Till BJ, Burtner C, Comai L, Henikoff S (2004a) Mismatch cleavage by single-strand specific nucleases. *Nucleic Acids Research* **32,** 2632-2641

Till BJ, Renolds SH, Weil C, Springer N, Burtner C, Young K, Bowers E, Codomo CA, Enns LC, Odden AR, Greene EA, Comai L, Henikoff S (2004b) Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biology* **4,** 12

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

Walden R (2002) T-DNA tagging in a genomics era. *Critical Reviews in Plant Sciences* **21,** 143-165

Wienholds E, van Eeden F, Kosters M, Mudde J, Plasterk RHA, Cuppen E (2003) Efficient target-selected mutagenesis in zebrafish. *Genome Research* **13,** 2700-2707

Winkler S, Schwabedissen A, Backasch D, Bokel C, Seidel C, Bonisch S, Furthauer M, Kuhrs A, Cobreros L, Brand M, Gonzalez-Gaitan M (2005) Target-selected mutant screen by TILLING in Drosophila. *Genome Research* **15,** 718-723

Zerr T, Henikoff S (2005) Automated band mapping in electrophoretic gel images using background information. *Nucleic Acids Research* **33,** 2806-2812