Relevance of tilling in plant genomics

Muhammad Rashid1*, Guangyuan He1, Yang Guanxiao1 and Ziaf Khurram2

1China-UK HUST-RRRes Genetic Engineering and Genomics Joint Laboratory, International Science and Technology Cooperation Base (Genetic Engineering) of Chinese Ministry of Science and Technology, The key laboratory of Molecular Biophysics of Chinese Ministry of Education, College of Life Science and Technology, Huazhong University of Science and Technology(HUST), Luoyu Road 1037, Wuhan 430074, China
2College of Horticulture and Forestry Sciences, Huazhong Agricultural University, Wuhan, China

*Corresponding author: rfqmh197202@gmail.com

Abstract

With the development of rapid and inexpensive sequence technologies, the efficiency and accuracy in sequencing have interpreted the genomic information of many plant species. The emphasis on genomics has been changing from complete sequenced genomes to the study of the functional genomics. To understand the function of gene(s) many approaches like RNAi, gene knockout, site-directed mutagenesis, transposon tagging etc. have been exploited for many years. All these techniques demand the use of transgenic material which is not always possible in many commercially important crops. So it is not only impeding the functional analysis of gene(s) but also retards the improvement of existing as well as the development of improved cultivars. In this review, a nontransgenic technique called Targeting Induced Local Lesions IN Genomes (TILLING) was emphasized that determines the allelic sequence of induced point mutations in gene(s) of concern. It allocates the rapid and cost-effective detection of induced point mutations in populations of physical/chemically mutagenized individuals. Ethyl methanesulfonate (EMS) is mostly used and produce G/C to A/T transition by alkylating the G residues and the alkylated G resides to base pair with T instead of pairing with C. The crop species whose genome sequence has been finished; TILLING can be used to find the alleles in gene of interest for biotic and abiotic stresses. EcoTILLING, an extension of TILLING, is an inexpensive and rapid method of discovery and analysis of single-nucleotide polymorphism (SNP) in natural populations. In TILLING parental DNA sequence whereas in EcoTILLING genomic DNA sequence is used to identify the mutations created using endonuclease (CEL I). These techniques are independent of genome size, ploidy level and reproductive system of plants. They can be applied not only to model organisms but also to economically important crops. These provide a powerful approach for gene discovery, DNA polymorphism assessment and linkage disequilibrium. It is concluded that the induced mutations and natural polymorphism can be identified by TILLING and EcoTILLING approaches, and the next task is their implementations in crop breeding.

Keywords: dHPLC, Functional genomics, Mutagenesis, Reverse genetics.

Abbreviations: ATP = Arabidopsis TILLING Project, Az-MNU = sodium azide plus methyl-nitrosourea, bp= base pair, CEL I = Celery nuclease, CMT = Chromomethylase, CODDLE = Codons optimized to detect deleterious lesions, dHPLC = Denaturing high performance liquid chromatography, EcoTILLING = TILLING techniques to look for natural mutations in individuals, EMS = Ethyl methanesulfonate, FAD2 = fatty acid desaturase 2, Hin-a = Hordeindolina-a, HvFor1= Hordeum vulgare Floral Organ Regulator, INDELS = Insertions/Deletions, Kbp = Kilo base pair, Mbp = Mega base pair, NMU = N-methyl-N-nitrosourea, PTGS = Post transcriptional gene silencing, PARSESNP = Project aligned related sequences and evaluate SNPs, PCR= Polymerase chain reaction RNAi = RNA interference, SIFT = Sorting intolerant from tolerant, SNPs = Single nucleotide polymorphisms, SSCP = Single-strand conformation polymorphism, SSR = simple sequence repeat, T-DNA = Transfer DNA, TILLING = Targeting induced local lesions in genomes.

Introduction

Crop improvement has a long history as the key agronomic traits have been selected over hundreds of years during the domestication of crops. Recently, this progress has been accelerated as the green revolution has brought about the great rise in crop yields (Ahlouwalia et al., 2004; Khush, 2001). With the advent of genomics in the last 25 years, opportunities for crop improvement have continued to grow. This 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 of important traits and provides targets for transgenic alteration of gene expression. Completion of the genome sequence projects of Arabidopsis thaliana and rice has opened a new era of genomics for plant science research. The sequence information available in public database has highlighted the need to develop genome scale reverse genetic strategies for functional analysis (Till et al., 2003). The forward genetics can hardly meet the demand of a high throughput and large-scale survey of gene functions as most of the phenotypes are obscure. TILLING (Targeting induced local lesions in genomes) a newly developed general reverse genetic strategy helps to locate an allelic series of induced...
point mutations in genes of interest. It allows the rapid and inexpensive detection of induced point mutations in populations of physically/chemically mutagenized individuals. In addition to allowing efficient detection of mutations by TILLING approach, EcoTILLING technology is also ideal for examining natural variation. Endonuclease CEL I cut effectively the multiple mismatches in a DNA duplex. Therefore, heteroduplex DNA of unknown sequence with to that of a known sequence reveals the positions of polymorphic sites. Both nucleotide changes and small insertions/deletions are identified. It can be performed more inexpensive than full sequencing the methods currently used for the most single nucleotide polymorphism (SNP) discovery. SNP variation can provide clues to the adaptive strategies and population history that undoubtedly played roles in specie’s evolution. It is also used for screening and detection of plants with desired traits by knockdown and knockout mutations in specific genes. This makes TILLING and EcoTILLING an attractive strategy for a wide range of applications from the basic functional genomic study to practical crop breeding.

**TILLING antiquity**

It was first explored in the late 1990’s by the efforts of Claire McCallum and his collaborators (Fred Hutchinson Cancer Research Center and Howard Hughes Medical Institute) who was experimenting on Arabidopsis (Borevitz et al., 2003). He used T-DNA lines and antisense RNA as reverse genetic approaches to illustrate the function of two chromomethylase genes. He was impotent to successfully apply these methodologies to describe the function of CMT2 gene. This technique was developed by pooling chemically mutagenized plants together, creating heteroduplexes among the pooled DNA, intensify the region of concern and using dHPLC (denaturing high performance liquid chromatography) to identify the mutants by chromatographic variations (McCallum et al., 2000). A less expensive and faster modification of the TILLING protocol was published later, which uses a mismatch specific celery nuclease (CEL I) and LI-COR gel analyzer system (Alonso and Ecker, 2006; Olejekowski et al., 1998). In 2001, the standard procedure was developed with practical software that makes the TILLING technique a more routine method to detect mutations to get satisfactory results (Colbert, 2001). Since from its origin, it has been automated and exploited in many plant taxa. As a reverse genetic high throughput method, it purposes to detect SNPs (single nucleotide polymorphisms) and/or INDELS (insertions/deletions) in gene/genes of interest created from a mismatch in a mutagenized populace.

**Outline of TILLING technique**

To create an induced population with the use of physical/chemical mutagens is the first prerequisite for TILLING approach (Fig 1). Most of the plant species are compatible with this technique due to their self-fertilized nature and the seeds produced by these plants can be stored for long periods of time (Borevitz et al., 2003). The seeds are treated with mutagens and raised to harvest M1 plants, which are consequently, self-fertilized to raise the M2 population. DNA extracted from M2 plants is used in mutational screening (Colbert et al., 2001). To avoid mixing of the same mutation only one M2 plant from each M1 is used for DNA extraction (Till et al., 2007). The M2 seeds produce by selfing the M2 progeny can be well preserved for long term storage. Ethyl methanesulfonate (EMS) has been extensively used as a chemical mutagen in TILLING studies in plants to generate mutant populations, although other mutagens can be effective (Table 1). EMS produces transitional mutations (G/C, A/T) by alkylating G residues which pairs with T instead of the conservative base pairing with C (Nagy et al., 2003). It is a constructive approach for users to attempt a range of chemical mutagens to assess the lethality and sterility on germinal tissue before creating large mutant populations. Optimal PCR primers are designed for a functional domain target. In the next step take DNA from the population to normalize its concentration. Once samples have been standardized, they are pooled together. It is essential to confirm that all DNA samples are the same concentration so that no biasing of samples occurs. CODDELE (http://www.proweb.org) is the web based programme, and by putting the genomic, cDNA or protein sequences, it allows the researchers to evaluate the possible gene function in the induced mutant population (Gilchrist and Haughn, 2005). Generally, for diploid organisms, a pool of DNA comprising up to eight individual samples can be effective in mutation detection (Hemikoff and Comai, 2003). Thus, depending on ploidy level, heterozygosity, and the extent of certainly arising SNPs, best pooling for a species of concern should be determined practically. The pool DNA is arranged into 96 well microtiter plates. The targeting forward and reverse primers are differentially 5’ end labeled with IRD700 and IRD800 dye respectively are used for fluorescent detection at ~700 nm and ~800 nm (Fig. 1). Next, heteroduplexes and homoduplexes are produced from the PCR products of pool samples comprising of mutants and the wild form by heating (denaturing) and cooling (annealing). The endonuclease enzyme CEL I is used and a short heating is vital for the enzymatic reaction to progress. CEL I, extracted from celery, not only recognize gaps in the heteroduplex, but it also cleaves DNA on the 3’ side of the mismatch (McCallum et al., 2000). After the enzyme incubation period, digested fragments are recognized on a denaturing polyacrylamide gel attached to a LI-COR 4300 DNA analysis system (Fig 1). Pools holding an induced mutation contain a mixture of homo and heteroduplexes. Therefore, a full length product (detected in both 700 and 800 channels) and two measureable cleaved fragments (one IRD700 labeled, one IRD800 labeled) are produced. The amount of the sliced fragments should be equal the full length PCR product. The size of the cleaved fragments can be evaluated by comparison to a size standard, and the estimated position of the mutation will be recognized and further confirmed by sequencing. The PARSESNP (http://www.proweb.org/parceansnp) can be used to identify and display the positions of the polymorphisms in gene (s) in a graphical layout (Taylor and Greene, 2003).

**Boons of TILLING**

1) It provides the imprecise position within a few base pairs of the induce mutation (Borevitz et al., 2003; Colbert et al., 2001). The chemical mutagen creates a range of numerous mutations throughout the genome such as nonsense, splice site and missense. These mutations can possibly affect the protein structure and the subsequent phenotype. Therefore, through mutagenesis one can acquire partial loss or complete loss of function, which can provide valuable insight into the true role of a gene in a species of interest (Stemple, 2004).

2) The mutation detection efficiency of TILLING is credited to its high throughput screening capacity.

3) The densities of traditional chemical mutagenesis can be estimated. For example, EMS, one of the highly stable alkylations is commonly used to induce point mutation in
Chromatographic alterations were detected and confirmed to and CMT3 among 835 M2 plants in Arabidopsis. Thirteen sequence screened by dHPLC to detect mutations in CMT2 970bp in size were examined for a total of 2Mb of DNA. Seven different PCR fragments that ranged from 345 to 2006), as well as the detection of heterozygotes in a material. It is found to be highly sensitive to identify induced mutations and naturally occurring SNPs (Dahm and Geisler, 2002). The TILLING procedure is useful in the interpretation of gene function in plants without the development of transgenic suppression generates unpredictable outcomes, and the whole procedure is laborious as it requires vector construction, transformation, and transgenic analysis (Que and Jorgensen, 2007; YAN Fei and CHENG Zhuo-Min, 2005), but RNAi approaches are not equally relevant to all organisms. In rice being a model plant, there are over 200000 T-DNA insertional populations; however, only few reports have been published about rice gene knockout by T-DNA insertion (An et al., 2007; YAN Fei and CHENG Zhuo-Min, 2005). Anti-sense RNA and RNAi techniques have been commonly used to reduce expression of genes (Bagge et al., 2005). Anti-sense RNA and RNAi techniques have been commonly used to reduce expression of genes (Bagge et al., 2005). Anti-sense RNA and RNAi techniques have been commonly used to reduce expression of genes (Bagge et al., 2005). Anti-sense RNA and RNAi techniques have been commonly used to reduce expression of genes (Bagge et al., 2005). Anti-sense RNA and RNAi techniques have been commonly used to reduce expression of genes (Bagge et al., 2005).

Table 1. Year-wise overview of TILLING approach among various crops

<table>
<thead>
<tr>
<th>Year</th>
<th>Organism</th>
<th>Genus species</th>
<th>Mutagen applied</th>
<th>Mutagen dose</th>
<th>Estimated genome</th>
<th>Ploidy</th>
<th>Mutation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>Arabidopsis</td>
<td>Arabidopsis thaliana</td>
<td>EMS</td>
<td>20nM to 40 mM</td>
<td>125 Mbp</td>
<td>2X</td>
<td>1/300 kb</td>
</tr>
<tr>
<td>2001</td>
<td>Arabidopsis</td>
<td>Arabidopsis thaliana</td>
<td>EMS</td>
<td>20nM to 30 mM</td>
<td>125 Mbp</td>
<td>2X</td>
<td>1000 genome</td>
</tr>
<tr>
<td>2002</td>
<td>Arabidopsis</td>
<td>Arabidopsis thaliana</td>
<td>Fast neutrons</td>
<td>60 Gy</td>
<td>125 Mbp</td>
<td>2X</td>
<td>0.7 to 12 kb</td>
</tr>
<tr>
<td>2003</td>
<td>Lotus</td>
<td>Lotus japonicus</td>
<td>EMS</td>
<td>60mL/10mL/ (v/v)</td>
<td>472Mbp</td>
<td>2X</td>
<td>1/154kb</td>
</tr>
<tr>
<td>2004</td>
<td>Barley</td>
<td>Hordeum vulgare</td>
<td>EMS</td>
<td>20nM to 30 mM</td>
<td>5300 Mbp</td>
<td>2X</td>
<td>1/Mb</td>
</tr>
<tr>
<td>2004</td>
<td>Maize</td>
<td>Zea mays</td>
<td>EMS</td>
<td>0.0625%</td>
<td>2500 Mbp</td>
<td>2X</td>
<td>0.93/ kb</td>
</tr>
<tr>
<td>2005</td>
<td>Rice</td>
<td>Oryza sativa</td>
<td>MNU</td>
<td>1mM</td>
<td>430 Mbp</td>
<td>2X</td>
<td>0.80/1kb</td>
</tr>
<tr>
<td>2006</td>
<td>Field mustard</td>
<td>Brassica rapa</td>
<td>γ radiation</td>
<td>500Gy</td>
<td>500Mbp</td>
<td>2X</td>
<td>1/6190kb</td>
</tr>
<tr>
<td>2007</td>
<td>Pea</td>
<td>Pisum sativum</td>
<td>EMS</td>
<td>4 mM</td>
<td>4300 Mbp</td>
<td>2X</td>
<td>1/669 kb</td>
</tr>
<tr>
<td>2007</td>
<td>Rice</td>
<td>Oryza sativa</td>
<td>EMS</td>
<td>1.5%</td>
<td>430 Mbp</td>
<td>2X</td>
<td>1/294 kb</td>
</tr>
<tr>
<td>2007</td>
<td>Rice</td>
<td>Oryza sativa</td>
<td>Az-MNU</td>
<td>1 to 15 mM</td>
<td>430 Mbp</td>
<td>2X</td>
<td>1/265 kb</td>
</tr>
<tr>
<td>2008</td>
<td>Ripe seed</td>
<td>Brassica napus</td>
<td>EMS</td>
<td>0.4% to 1.2% w/v</td>
<td>1150 Mbp</td>
<td>4X</td>
<td>1/41.5kb</td>
</tr>
<tr>
<td>2008</td>
<td>Sorghum</td>
<td>Sorghum bicolor</td>
<td>EMS</td>
<td>0.1 to 0.6% (v/v)</td>
<td>7.35 × 10^9 bp</td>
<td>2X</td>
<td>1/526 kb</td>
</tr>
<tr>
<td>2008</td>
<td>Soybean</td>
<td>Glycine max</td>
<td>NMU</td>
<td>2.5 mM</td>
<td>1115 Mbp</td>
<td>2X</td>
<td>~1/140 kb</td>
</tr>
<tr>
<td>2008</td>
<td>Soybean</td>
<td>Glycine max</td>
<td>EMS</td>
<td>50 mM</td>
<td>1115 Mbp</td>
<td>2X</td>
<td>1/250 kb</td>
</tr>
<tr>
<td>2009</td>
<td>Barley</td>
<td>Hordeum vulgare</td>
<td>Sodium azide</td>
<td>1.5 mM</td>
<td>5500 Mbp</td>
<td>2X</td>
<td>1/2.5 Mb</td>
</tr>
<tr>
<td>2009</td>
<td>Clover</td>
<td>Medicago truncatula</td>
<td>EMS</td>
<td>0.075% to 0.40%</td>
<td>470 Mbp</td>
<td>2X</td>
<td>1/485 kb</td>
</tr>
<tr>
<td>2009</td>
<td>Common Bean</td>
<td>Phaseolus vulgaris</td>
<td>EMS</td>
<td>20mM to 60 mM</td>
<td>625 Mbp</td>
<td>2X</td>
<td>2 to 3/Mb</td>
</tr>
<tr>
<td>2009</td>
<td>Peanut</td>
<td>Arachis hypogaea</td>
<td>EMS/γ-radiation</td>
<td>(0.5%)/ 20 &amp;30 kr</td>
<td>3.0 x 10^3 bp</td>
<td>4X</td>
<td>0.5 x10^-7</td>
</tr>
<tr>
<td>2009</td>
<td>Potato</td>
<td>Solanum tuberosum</td>
<td>γ radiation</td>
<td>0.5% to 2.0%</td>
<td>850Mbp</td>
<td>4X</td>
<td>&lt;1/1810 bp</td>
</tr>
<tr>
<td>2009</td>
<td>Wheat</td>
<td>Triticum turgidum</td>
<td>EMS</td>
<td>0.7% to-1%</td>
<td>12000 Mbp</td>
<td>4X</td>
<td>1/40 kb</td>
</tr>
<tr>
<td>2010</td>
<td>Field mustard</td>
<td>Brassica rapa</td>
<td>EMS</td>
<td>0-1%</td>
<td>500Mbp</td>
<td>2X</td>
<td>1/60kb</td>
</tr>
<tr>
<td>2010</td>
<td>Oat</td>
<td>Avina sativa</td>
<td>EMS</td>
<td>0.9% (v/v)</td>
<td>13000 Mbp</td>
<td>6X</td>
<td>1/20-40 kb</td>
</tr>
<tr>
<td>2010</td>
<td>Tomato</td>
<td>Solanum lycopersicum</td>
<td>EMS</td>
<td>0.7% to 1.0%</td>
<td>950Mbp</td>
<td>2X</td>
<td>1/322 kb</td>
</tr>
</tbody>
</table>

DNA and produces primarily C to T changes resulting in C/G to T/A transition mutation in Arabidopsis. Ninety-nine percent of mutations from alkylation of guanine induced by EMS are reported as C/T-to-A/T transitions (Greene et al., 2003). From these consequences, the most appropriate fragment is selected in a specific gene of interest. The ability of chemical mutagenesis is to induce high density of mutations in multiple loci. So, genome wide saturated mutagenesis can be achieved using a relatively small mutant population. According to the general estimation made by the Arabidopsis TILLING Project (ATP), approximately 7 mutations per 1 Mb are identified after screening the mutant Arabidopsis plant lines. Based on the above estimation, a total of 100000 mutant plants can achieve satisfied mutant densities (Henikoff et al., 2004). The chemical mutagenesis, coupled with LI-COR analyzer and TILLING procedure is useful in the interpretation of gene function in plants without the development of transgenic material. It is found to be highly sensitive to identify induced mutations and naturally occurring SNPs (Dahn and Geisler, 2006), as well as the detection of heterozygotes in a population. This has been proven in the original work by McCallum and her colleagues (McCallum et al., 2000). Seven different PCR fragments that ranged from 345 to 970bp in size were examined for a total of 2Mb of DNA sequence screened by dHPLC to detect mutations in CMT2 and CMT3 among 835 M2 plants in Arabidopsis. Thirteen chromatographic alterations were detected and confirmed to be mutations by amplification and sequencing; no PCR errors were found, indicating an error rate of <10^-6. Combination of CEL I, double- end fluorescent dyes labeling and LI-COR system as an alteration to dHPLC maintained and secured the high sensitivity of the modified high-throughput TILLING. The noteworthy benefit of TILLING is the saving of time and money as it does not demand resequencing of all the individuals in a population to comb frequent or rare SNPs. For a diploid organism, it is processed by pooling eight individuals of a population in one time and finding variations due to mismatches in a heteroduplex. The purpose of reverse genetics is to determine the function of a gene with known sequence by phenotypic analysis of cells or organisms in which the function of this gene is diminished. In plants, the most commonly used reverse genetic approaches are insertional mutagenesis, anti-sense RNA, and double-stranded RNA interference (RNAi). These reverse genetic approaches are not equally relevant to all organisms. In rice being a model plant, there are over 200000 T-DNA insertional populations; however, only few reports have been published about rice gene knockout by T-DNA insertion (An et al., 2005). Anti-sense RNA and RNAi techniques have been commonly used to reduce expression of genes (Bagge et al., 2007; YAN Fei and CHENG Zhuo-Min, 2005), but RNAi suppression generates unpredictable outcomes, and the whole procedure is laborious as it requires vector construction, transformation, and transgenic analysis (Que and Jorgensen, 2007; YAN Fei and CHENG Zhuo-Min, 2005).
The larger-scale mutant detection using these reverse genetic technologies is hampered.

6) The TILLING technique is a permutation of the traditional chemical mutagenesis, and the double dye far-red fluorescent detecting technique. This technique requires no complicated manipulations and expensive apparatus. It enables to screen the mutant pools easily for investigating the functions of specific genes, avoiding both the confounding gene separation steps and tedious tissue culture procedures as involved in anti-sense RNA and RNAi.

**EcoTILLING**

An extension of TILLING is EcoTILLING, which uncover natural alleles at a locus contrary to induced mutations. Many crop species cannot be exposed to induced mutation, and EcoTILLING can be used to find the natural variants and their putative gene functions in these crops. This can be done at low cost of SNP/haplotyping methods, and one can screen many samples with a gene of interest (Comai et al., 2004). It does not require all the population to be screened to find the polymorphism which can be arduous and time consuming in TILLING. In this strategy, each ecotype is pooled in 1:1 ratio with the standard (Fig 1). Moreover, in open pollinated crops, EcoTILLING can be used to find the heterozygosity level within a gene fragment. As the CEL I can digest small proportion of the hetroduplexes, it can be used to find the multiple polymorphisms in a single gene of interest. Furthermore, phylogenetic diversity estimates, selection and linkage disequilibrium can be estimated. It can be used to detect the DNA polymorphism in satellite repeat numbers. It can also be effectively used as an efficient, rapid technique to identify DNA polymorphisms in populations with high genetic similarity and to mine for SNPs in collections of plant germplasm. The EcoTILLING approach first time used on Arabidopsis thaliana in 2004 to mine for natural polymorphisms (Comai et al., 2004). From five different genes that were almost 1 Kb in length, 55 haplotypes were identified in the introns. This study showed that CEL I could cut hetroduplexes to detect SNPs, INDELS and microsatellite repeat polymorphism. Another important application is to find alleles of resistant genes that could provide immunity to various diseases. EcoTILLING in Mla resistance genes of Hordeum vulgare (barley) was used in 2006 to examine the allelic variation (Mejlhede et al., 2006). It demonstrates how effectively it can be used for the evaluation of diversity in natural populations (Gilchrist et al., 2006). It was employed to identify polymorphisms in mung bean (Vigna radiata (L.) R. Wilczek). The majority of the polymorphisms were detected between sublobata and radiata in putative introns (Fery, 2002). In general, EcoTILLING shows great promise in the process of identifying natural disease resistance alleles, which can be used in crop improvement. Besides, this technique is applicable to any organism even those that are heterozygous and polyploidy in nature.

**TILLING Applications**

**For gene discovery**

This technique was first utilized in Arabidopsis TILLING Project (ATP) during 2001. The ATP project has detected, sequenced, and delivered over 1000 mutations in more than 100 genes (Till et al., 2003). Through the workshop, mutant materials, DNA samples and mutant information were fully shared by all researchers working on Arabidopsis. Well-developed and tested protocols of TILLING are available for Arabidopsis (McCallum et al., 2000; Till et al., 2003), Lotus japonicas (Perry et al., 2003), maize (Zea mays L.) (Till et al., 2004), wheat (Triticum aestivum L.) (Slade et al., 2005) and rice (McCallum, 2000). In maize 17 independent induced mutations from 11 genes were obtained from a population of 750 pollen-mutagenized plants (Till et al., 2004). The result obtained from maize was consistent with that from Arabidopsis, indicating that TILLING is a broadly applicable and efficient reverse genetic strategy for large genome. The TILLING strategy also succeeded to create and identify genetic variation in wheat, thereby showing a great potential as a tool for genomic research in polyploidy plants.

**For DNA polymorphism assessment**

DNA polymorphism widely exists in various plant species and plays an important role in biological evolution. The methods currently available for revealing DNA polymorphism encompass DNA sequencing, single-strand conformation polymorphism (SSCP), hybridization, and microarray. These techniques possess some advantages and limitations. Although DNA sequencing is simple and straightforward, but it is rather costly and time consuming. SSCP provides a high throughput strategy for polymorphism detection. However, it has low efficiency in detecting novel mutations with a limit of 200 to 300bp length of target DNA sequence. Microarray holds two disadvantages. One is high cost of operation, and the other is the low detecting frequency(< 50%) (Caldwell et al., 2004; Triques et al., 2007). The TILLING can detect DNA variations from single nucleotide polymorphism (SNP), small fragment insertion and deletions to simple sequence repeat (SSR). It can be performed as a high throughput, low cost, and high accuracy approach compared with the other methods.

**Approach for functional genomics**

Two principal approaches forward and reverse genetics has been extensively used to determine the function of gene(s) and how genotypes are linked to phenotypes. Conventionally in forward genetics (phenotype to genotype) the gene sequence is finally inferred through selecting large numbers of mutagenized individuals for phenotypic variants. In forward genetic methodologies for genome wide analysis primarily for gene coding to a particular phenotype needs a lot of time and work (Alonso and Ecker, 2006). In reverse genetics (from genotype to phenotype), the gene sequence is known, and mutants are identified and screened with structural alterations in the gene of interest (Nagy et al., 2003). In this approach generally less time is needed than forward genetics, and its strategies have been effectively used for functional genomics in many plant species. The ubiquitous availability of sequence data from different databases permits researchers to design swiftly their reverse genetic schemes to decide gene function. Some of the reverse genetic approaches employed in plants comprise homologous recombination, Agrobacterium mediated insertional mutagenesis, transposon tagging, RNAi (RNA interference) or PTGS (post transcriptional gene silencing), and chemical mutagenesis. Among all these TILLING is a more efficient mutation detection method, grosses the advantage of chemical mutagenesis to generate induced mutations in a population. What is unique for the TILLING approach compared to transgenic approaches is the identification of numerous mutations within a targeted region of the genome.
Fig 1. Schematic diagram of the TILLING and EcoTILLING techniques. In TILLING seeds are mutagenized with chemical/physical mutagens to produce M₁ plants. M₁ plants are self to produce the M₂ from which DNA is extracted for analysis. The M₂ is allowed to produce seed, which can be easily stored for future analysis. After the DNA is extracted from the mutant population, the DNA concentration of all samples is standardized and pooled together. The number of individuals in a pool depends on the ploidy level of the plant and the amount of naturally occurring SNPs, which may require the number of individuals in the pool to be reduced. The desired gene is amplified using a forward primer with 700 nm dye label and a reverse primer with an 800 nm dye label attached to the 5’ ends. The PCR products are heated and cooled to form a mixture of homoduplexes and heteroduplexes among the genotypes in the pool. Any mismatches (SNPs or small INDELS) will be detected by a mismatch endonuclease (CEL I) and cleaved into two separate products, which will be detected in the 700 and 800 dye channel of a LICOR DNA Analyzer. The total size of the cleaved fragments should equal the total length of the entire product. Once the cleaved fragments and their respective polymorphic site are identified, these individuals are sequenced to verify the induced mutation. EcoTILLING is performed in the same manner except that the seed is not mutagenized; therefore, the process begins by extracting DNA from a reference plant and members of the population and continuing with the remaining steps to determine natural polymorphisms.

These mutations constitute allelic series that can potentially confer a range of phenotypes from subtle to strong, and allow structure and functional 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 as 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 determine gene function, RNAi may be the current method of choice. However, TILLING offers many advantages in cases where the 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 different targets is much less by TILLING than by gene suppression using transgenics.

More than a knockout

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. TDNA insertions and/or transposon insertions may be the preferred means to obtain a specific gene knockout. These 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 and 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 having limited genetic variation. During domestication and subsequent selection, much of the genetic variation available in the wild crop progenitors has been lost (Gepts and Papa, 2002). Thus, plant breeders have at times used wild relatives or land races to introduce useful genetic variation. This practice has been successful in wheat for developing disease resistant and higher yielding varieties (Zamir, 2001) and a land race 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 (Gilchrist and Haughn, 2005). 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, consequently, avoiding the introduction of agriculturally undesirable traits. In addition, the issue of bio-piracy makes the use of exotic varieties to improve modern cultivars potentially filled with complications. The identification of caffeine free Arabica coffee by Brazilian scientists in germplasm that came originally from Ethiopia has prompted a dispute over ownership (Silvarolla, 2004).

For crop breeding

Conventional mutation breeding, either by radiation or by chemical treatment, has a proven influence on production of many varieties, including high-yielding rice, barley and wheat, etc. (Gilchrist and Haughn, 2005). Unlike conventional mutation breeding in which the mutation frequency is unknown or estimated from mutations conveying a visible phenotype, TILLING provides a direct measure of induced mutations. Besides, it allows not the prompt, parallel selection of numerous genes but also a forecasting of the number of alleles that will be recognized based on the mutation frequency and library size. The efforts done on plants are deliberated as follows.

Arabidopsis

The TILLING Project (ATP) had mentioned 1,890 mutations in 192 target gene, and it was detected that heterozygote mutations were twice folded more than homozygote mutations (Tilli and Mirzabekov, 2001). The several mutations in Arabidopsis italiana have been recognized through this technique that clearly explains the function of gene (s) and protein for Arabidopsis researchers.

Barley

Barley is also used as an important cereal crop having a fairly large genome size of ~5,300 Mb. It was used for TILLING experiments to find the induced mutations in two genes (Caldwell et al., 2004). Hin-an and HvFor1 genes were studied, and 10 mutants were identified. Among these ten mutants six have missense mutations.

Clover

Medicago truncatula has been extensively adopted as a model plant for crop legume species of the Vicieae. Regardless the convenience of transformation and regeneration protocols, there are presently inadequate tools accessible in this species for the systematic investigation of gene function. M. truncatula was treated with chemical mutagens to create mutant population that provides a TILLING (targeting induced local lesions in genomes) platform and a phenotypic database for both reverse and forward genetics screens. Fifty-six targets were identified and screened for polymorphism. About 546 point mutations were recognized with a mutation frequency of 1/485 kb in clover (Signor et al., 2009).

Common Bean

Phaseolus vulgaris is the main food legume used worldwide, making it an important target for innovative methodologies of genetic analysis. Genotype BAT 93 was used for TILLING approach. It was found that 40 mM EMS was an appropriate concentration for the generation of a mutant population. Higher the concentrations of EMS, lower the survival rate (~10%) of mutants. The low concentrations of EMS result in the generation of fewer mutants (Porch et al., 2009) with a mutation rate of 2-3 mutations/Mbp.

Field Mustard

B. rapa was used as the first EMS TILLING source in the diploid Brassica species. It has a genome size of 625Mbp. The mutation frequency in this population is ~1 per 60 kb, which makes it the most densely mutated diploid organism (Stephenson et al., 2010).

Lotus

Lotus japonicus, a model plant has also been emphasized for explaining gene function through TILLING. It is a perennial temperate legume and is used as a model plant for genomic studies due to its short life cycle, a diploid nature (2n = 2x = 12), with a small genome size (472Mbp), and is self-fertilized nature (Sato and Tabata, 2006). To discover induced mutations in the protein kinase domain of the SYMRK gene, this approach was used and six missense mutations were discovered in the splice acceptor site. The overall mutation rate of 1/154kb in Lotus japonicus was recorded (Horst et al., 2007).

Maize

Maize having a large genome size was found to be promising for this study. During 2005 a TILLING project was launched at Purdue University on maize crop. In this project 319 mutation in 62 genes were identified (Weil and Monde, 2007) having a mutation rate of 0.93/kb. In another study a population of 750 mutagenized plants was used to illustrate
the function of 11 genes and six genes having the visible mutation were screened. In this investigation among six genes, a chromomethylase gene DMT102 was studied. This gene played a vital role in Arabidopsis for non-CpG DNA methylation and gene silencing (Waterhouse et al., 1998).

Oat

The oat (Avena sativa) having a genome size of 13000Mbp was treated with chemical mutagen EMS to establish a TILLING population. On an average, it exhibited hundreds of mutations in every individual gene having a mutation rate of 1/20-40kb in the genome (Bagge et al., 2007).

Peanut

A TILLING population was generated in peanut having a tetraploid genome and screened for mutations in genes for allergenic proteins (Ara h 1, Ara h 2) and for the oil biosynthesis enzyme FAD2. The silencing of Ara h 2 by RNA interference has delivered evidence that this protein and its related family member Ara h 6 may be dispensable for peanut seed growth, development, and viability. Therefore, recovery of knockout mutations in the two genes of Ara h 2 should allow elimination of this most severe allergen from peanut seed. Up to now one copy of potential knockout mutation in each of Ara h 1, Ara h 2, and FAD2 has been identified in peanut (Knoll et al., 2009).

Potato

Three potato (Solanum tuberosum) cultivars were treated with different doses of gamma radiations to create a platform for TILLING and EcoTILLING studies. Three genes specific primer pairs were used to amplify a sequence of ~1 to 1.5 kb of targeted gene and 15 putative nucleotide polymorphisms per kb were found. Among three tetraploid cultivars, nine polymorphisms out of fifteen were found distinctive (Elias et al., 2009).

Rapeseed

Two EMS mutant populations of the semi-winter rapeseed were created for functional genomics in Brassica napus, and for introduction of novel allelic variation in rapeseed breeding. Forward genetic selection of mutants from the M1 populations helped in identification of a large number of unique phenotypes. The existing SNPs were used as positive control to find the distinguishing novel mutations. The method was used to 1344 M2 plants, and 19 mutations were identified (Wang et al., 2008). Among them three were functionally conceded with reduced seed erucic acid content.

Rice

Rice is an important economic and staple food crop providing about 80% of the caloric intakes of three billion people of the world (Storozhenko et al., 2007). Its genome has ~50, 000 genes but the function of all genes is empirically not yet determined. To detect the mutations and identify the function of genes, TILLING studies were done on indica rice (Wu et al., 2005) using chemical mutagens EMS and Az-MNU respectively. In this study among 10 target genes 57 polymorphism were identified (Till et al., 2007). The use of agarose gel and LI-COR DNA analyzer was also used in rice to find the induced mutations (Raghavan et al., 2007).

Sorghum

Sorghum bicolor (L.) Moench is used as a most important grain crop and fodder resource for most of the arid and semi-arid regions of the world. A sorghum inbred line BTx623 was treated with chemical mutagen EMS to create a mutant population. Out of 1600 lines, 768 mutant lines were analyzed by TILLING using four target genes and only five mutations were identified resulting in a calculated mutation rate of 1/526 kb (Zhanguo et al., 2008).

Soybean

Soybean (Glycine max) an important economic crop and a rich source of protein (35-50%) is beneficial for human health (Krishnan, 2005). It also improves soil quality by fixing nitrogen. Two cultivars viz; Forrest and Williams 82 were used to create four mutagenized populations by treatment with EMS or NMU to identify induced mutations. For seven targeted genes, about 116 mutations were identified through TILLING approach. Most of the mutations discovered were found to be the estimated as G/C to A/T transition (Cooper et al., 2008).

Tomato

Using TILLING approach, a set of new mutants was created in tomato (Solanum lycopersicum) cultivar, i.e., Red Setter at 0.7%-1.0% EMS dose. To confirm the Red Setter TILLING platform, induced point mutations were investigated in 7 tomato genes with the mismatch specific ENDO1 nuclease. About 66 nucleotide substitutions were identified from 9.5 kb of tomato genome. The overall mutation rate was estimated to be 1/322 kb and 1/574 kb for the 1% EMS and 0.7%-1.0% EMS treatments respectively (Minoia et al., 2010).

Wheat

Wheat being hexaploid have a large genome size of 12. 000Mbp is used as an important staple crop, yielding annually 600 million tons. To make a good in quality, partial waxy wheat cultivars are desirable (Graybosch, 1998) which is good for noodles and superior flour. Using TILLING approaches in wheat 246 allelic series were identified in the waxy gene homologues. Among this allelic series 84 missense, 3 non-sense and 5 spile site mutations were identified (Stemple, 2004) with mutation rate of 1/40kb. Albeit TILLING was primarily designed in Arabidopsis, but it has been recognized as an exceptionally flexible approach as compared to many other reverse genetic techniques. It has proven to be very successful in interpreting the gene function in hexaploid and diploid organisms having such large genome size. Besides the use of physical/chemical mutagens in diploid and polyploid plants, it yields a series of various allelic mutations throughout the genome (Table 1).
Experimental encounters in application of TILLING

There are some scientific challenges in TILLING experiments;

1. To develop high quality mutant population about two to three years may be predictable to be required (Caldwell et al., 2004; Slade and Knauf, 2005).

2. The first step to generating a population is to use different concentration of the chemical mutagen to assess lethality to find an optimal concentration for conducting the experiments (An et al., 2005). It is challenging because the lethality of species and varieties respond differently to physical/chemical mutagenesis (Till et al., 2007). An ideal population would exploit mutational load i.e, more than 50% survival of mutant population is prerequisite (Weil and Monde, 2007).

3. Creating mutant populations in vegetative propagated plants (Slade and Knauf, 2005) also slow down the progress of generating a mutant population.

4. The species that are highly heterozygous may confuse mutation detection for researchers due to natural polymorphisms in the genome, which may deter in finding of rare induced mutations (Till et al., 2006).

5. The production and maintenance of clones of vegetative propagated plants for future analysis are somewhat problematic (Stemple, 2004).

6. After the creation of mutant population, it is essential that all DNA extracts be equivalent in concentration so that they are all corresponding characterized in the pools being investigated. Otherwise, unique induced mutations may not be recognized as the amount of mutant DNA decreases in contrast to others in a pool of DNA.

7. Another task, particularly in plants is the selection of target genes that sometimes exist as a single copy throughout the genome. This leads to a problem when experimenting on polyploid plants who have complex genomes such as wheat or peanut. To overcome this contest primer need to be designed that is precise to single gene of interest, which may entail some extra effort (Ramos et al., 2006).

8. Another strategy is to sequence the multiple alignments of the homologous target genes to find restriction site differences between the target genes. Proceeding to TILLING, the DNA can be digested, which may cleave the annoying target leaving the desired gene unbroken for analysis (Cooper et al., 2008).

9. Additionally, potential difficulty in TILLING may be the increased number of SNPs per fragment. The identification, scoring and tracking of cleaved fragments hence becomes more challenging. Single SNPs discovered in a heteroduplex needs a high focused as compared to multiple mismatch sites (Comai et al., 2004; Raghavan et al., 2007; Till et al., 2004). Moreover, care should be taken during scoring fragments as large numbers of SNPs are existing in a gene portion.

10. Designing an Eco-TILLING or TILLING experiment, the selection of the nuclease is very important to digest the mismatches in the heteroduplexes. CEL I identify and cleave mismatched fragments in a heteroduplex also contain 5’ to 3’ exonuclease/lytic activity (An et al., 2005; Yeung et al., 2005) can as well digest the full length PCR product starting with the 5’ fluorescent label. Therefore, care should be taken not to over digest DNA samples to avoid the loss of the fluorescent signal of the PCR products.

11. The last challenge is allocating a particular phenotype to a genotype and suppose the putative function of a gene. Chemical mutagenesis sometimes creates background mutations, which can make phenotype analysis more difficult (McCullum et al., 2000). This may take several generations of out crossing or backcrossing (An et al., 2005; Nagy et al., 2003). Obviously, to assign a function to a gene will be more challenging if there are any epistasis or pleiotropic effects created from the background mutations (Weil and Monde, 2007).

Viewpoint

As a novel reverse genetics technique, TILLING has been put into practices since its origin. It has been convincingly proven that this technique has a considerable potential for crop improvement. It represents the use of spontaneous and induced mutants in plant breeding. It allows the direct identification of beneficial nucleotide and amino-acid changes in genes with known functions and their use as the genetic markers for selection. The range of alleles can be developed in a short time is matchless and unlikely to be found elsewhere in the pool of germplasm accessible to plant breeders, including land races and undomesticated relatives. As the TILLING population is a stable source, the results of basic scientific research can be well interpreted into crop improvement as new information about the functions of potential gene targets becomes available. There are at least two instantaneous applications in plant breeding using TILLING as a haplotype tool for detection of genetic loci that are putatively associated with agronomically important traits. The first application is the identification of allelic variation in genes exhibiting expression correlating or cosegregating with phenotypes. This will link gene expression with DNA variation. It helps to overcome the main difficulty of finding DNA variation based on restriction site polymorphism or linkage to hypervariable markers such as SSR. The second application is the establishment of an allelic series at genetic loci for the traits of interest in germplasm or induced mutants. Allelic series at such loci will provide confirmatory evidence of the relationship between the phenotypes and candidate gene sequences. A large collection of alleles at a locus will provide patterns of association to deduce the functional significance of certain SNPs. It has been suggested that the recent progress in the area of plant molecular biology and plant genomics have the potential to initiate a new Green Revolution. These findings need to be executed in new cultivars to realize that potential. Now, TILLING service centers are accessible for Arabidopsis thaliana, Lotus japonicus, barley (Hordeum vulgare), common bean (Phaseolus vulgaris), field mustard (Brassica rapa ), maize (Zea mays), oat (Avena sativa), pea (Pisum sativum), peanut (Arachis hypogaea L), potato (Solanum tuberosum L), rice (Oryza sativa), rape seed (Brassica napus), Sorghum (Sorghum bicolor L), soybean (Glycine max), Medicago truncatula, , tomato (Solanum lycopersicum), and wheat (Triticum turgidum). Many of these aforesaid species already have widespread genomic information publicly and now the emphasis for these species has shifted from genome to genomics (to empirically find the function of gene (s). With the passage of time more information on genomic will be available for other plant species, and thus focuses on the use of reverse genetics approaches to allocate the putative gene’s functions. This ambition of geneticists to find and explain the function of coded DNA may ultimately lead to the development of public TILLING services in numerous plant species.
species, which will enable to streamline the procedure of functional genomics for all researchers. As a unique approach for genome wide functional genomics, coupled with other recently developed genomic resources, including detailed genetic maps, large numbers of ESTs and complete genome sequencing and annotation, will allow for the efficient identification of gene(s) controlling phenotypes in both model systems and economically important plants. It can be predicted that more and more direct or indirect benefits will be revealed through continuous applications of TILLING in the near future.

Conclusion

Nucleotide sequences contain hidden information about the forces for conservation and variation that shaped their evolutionary history. It is known that detection of a gene responsible for a mutation is a challenging process. TILLING and EcoTILLING are inexpensive and swift natural polymorphism detection and genotyping methods. They have advantages for determining the range of variation for genetic mapping based on linkage analysis. In these techniques, if a mutation is detected in a pool, the individual DNA samples that went into the pool can be individually analyzed to identify the individual that carries the mutation. Once this individual has been identified, its phenotype can be determined. These techniques work with good results, even if a population contains preexisting mutations that would compromise SNP discovery by other methodologies. This review illustrates how mutagens, particularly chemical mutagenesis is becoming a powerful tool, especially for reverse genetics in plant species by using the TILLING approach. The use of these approaches can facilitate the handling of the large-scale discovery of induced point mutations through populations that are required. These new screening methods can be applied to several plant species, whether small or large, diploid or allohexaploid in nature. They may serve as rapid approaches to identify the induced and naturally occurring variation in many species. Now the successes have been reported in a variety of important plant species, the next challenge is to use these technologies to develop improved crop varieties. The utility of induced mutations and natural polymorphism has already been established for crop breeding, and so the task is mostly one of the implementations.

Acknowledgements

This work was supported by the National Major Specific Project of the People’s Republic of China (2008 ZX 08002-004, 2009 ZX 08016-001A, 2009 ZX 08002-013B) and the International Science and Technology Cooperation of China MoST (2009 DFB 20290).

References:


McCallum CM, Comai L, Greene EA, Henikoff S (2000 a or

McCallum CM, Comai L, Greene EA, Henikoff S (2000 a or


