CONCURRENT SESSION 7

High Throughput Techniques for Mutation Screening

Global TILLING Projects

B J Till[*], R Afza, S Bado, O A Huynh, J Jankowicz-Cieslak, M Matijevic & C Mba

Abstract

Induced mutation and natural nucleotide variation are powerful tools for probing gene function and improving traits in plants. Traditional mutagenesis has been widely used in forward genetic strategies and has led to the release of over 2,000 mutant plant varieties. Mutagens such as ethyl methanesulphonate (EMS) cause stable point mutations and thus produce an allelic series of truncation and missense changes that can provide a range of phenotypes. TILLING (Targeting Induced Local Lesions IN Genomes) uses traditional mutagenesis and nucleotide polymorphism discovery methods for a reverse genetic strategy that is high in throughput, low in cost, and applicable to most organisms. In less than a decade, TILLING has moved from a proof of concept to a well-accepted reverse genetic method that has been applied to over 20 different species. Large-scale TILLING services have delivered thousands of induced mutations to the international research community. Advancements in new mutation discovery techniques promise to increase further the efficiency and applicability of the TILLING method. Here, we review the progress in TILLING and describe the work of the Plant Breeding Unit of the Joint FAO/IAEA Programme to establish TILLING platforms for banana, cassava, and rice.

Background

Heritable genotypic variation is a major contributor to phenotypic diversification, and thus a fundamental driver of evolution. Naturally occurring alleles and induced mutations can be used as tools to study plant gene function and to develop crops with agronomically important traits. This is exemplified by the FAO/IAEA mutant variety database that catalogues more than 2,000 officially released crop varieties developed through the use of induced mutations (http://mvgs.iaea.org/). In recent decades, there has been a tremendous growth in the acquisition of genomic DNA sequences for plants and animals. The sequencing of entire genomes has been completed for some plant species and many more are planned (http://www.ncbi.nlm.nih.gov/sites/entrez?db=genomeprj).

With this information, it is now possible to predict genes important for crop traits based on sequence homology to well-studied model species. Reverse-genetic strategies that target lesions to specific genes, therefore, hold great promise in speeding up the process of gene-function analysis, and enhancing the efficiency of mutation breeding. While a number of reverse-genetic strategies have been described, Targeting Induced Local Lesions in Genomes (TILLING) has emerged as a robust, high-throughput, non-transgenic method that can be applied to most species [1, 2].

The TILLING method

TILLING combines traditional mutagenesis followed by high-throughput mutation discovery (**Fig. 1** [3, 4]). Mutagens such as ethylmethanesulphonate (EMS) are typically used to induce single base point mutations randomly throughout the genome. When multicellular organs, such as seed, are mutagenized, the resulting adult plant is chimeric because different cells that make up the plant harbor different mutations. Chimeras must be dissolved through self-fertilization or *in vitro* techniques before mutation screening. For sexually propagated plants, the first non-chimeric generation (the M_2 when mutagenizing seed) is selected to create the mutant population. A single-seed descent strategy is typically used, and a single M_2 plant from an M_1 parent is selected to create a mutant line. DNA is extracted from M_2 plants for mutation screening and the plants are allowed to self-fertilize to create M_3 seed for later phenotypic characterization. The DNA and seed libraries can be used for many years. Tissue culture, *in vitro* propagation and double haploidy can be employed to increase the efficiency of preparing mutant populations in species that are recalcitrant to the approach described above. Because mutagens such as EMS induce mutations randomly throughout the genome, each mutant line will harbor distinct mutations. With proper balance of mutation density and population size, multiple point alleles can be obtained in relatively small populations. For example, approximately 14 single-base mutations can be discovered when screening a 1.5kb gene fragment in a population of 3,000 mutagenized *Arabidopsis* plants (http://tilling.fhcrc.org/arab/status.html; [5]).
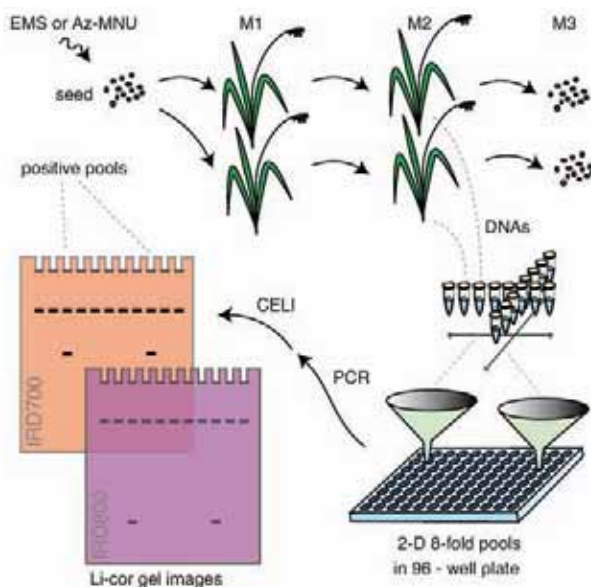
Figure 1 A typical TILLING strategy using seed mutagenesis. The first generation (M_1) is chimeric because of the multicellular composition of the seed at mutagenesis, making it unsuitable for mutation discovery. M_2 plants are non-chimeric and suitable for TILLING. DNA from each M_2 is collected for mutation screening and M_3 seed is collected and stored as the germplasm stock. DNAs are normalized to a common concentration, arrayed in 96 well plates, and pooled prior to screening to increase screening efficiency (a two-dimensional eight-fold pooling strategy is shown in this figure). DNAs are screened for mutations by PCR with gene-specific primers followed by mutation discovery via enzymatic mismatch cleavage, followed by fluorescence detection using denaturing polyacrylamide gel electrophoresis. This figure of a rice TILLING strategy is modified from [18].

Plant Breeding Unit, FAO/IAEA Agricultural & Biotechnology Laboratory, Joint FAO/IAEA Programme, International Atomic Energy Agency, Vienna, Austria

* Corresponding author. E-mail: b.till@iaea.org

Any method that can be used to identify single nucleotide mutations and small indels can be used for mutation discovery in TILLING. The first method described was denaturing HPLC [6, 7]. A variety of approaches have since been described, including direct sequencing and enzymatic mismatch cleavage with slab and capillary gel electrophoresis [3, 8-10]. A common approach is enzymatic mismatch cleavage using single-strand specific nucleases followed by denaturing polyacrylamide gel electrophoresis and fluorescence detection using the LI-COR DNA analyzer [11]. In this method, gene-specific fluorescently end-labeled PCR primers are used to amplify a ~1.5 kb genomic target. After PCR, samples are denatured and annealed to create heteroduplexed molecules from polymorphic amplicons. Mismatches in double-stranded molecules form at the location of induced mutations. Mismatches are cleaved by incubating DNA with a single-strand specific nuclease in a crude celery juice extract [12]. Samples are size fractionated and fluorescent signals detected using the LI-COR analyzer. The sum of the molecular weights of cleaved fragments equals the molecular weight of the full-length PCR product indicating the presence of a mutation. The molecular weights of fragments also indicate the position of the mutation to within ~10 base pairs [13]. Once discovered, mutations can be sequenced to determine the exact nucleotide change, and programmes such as PARSESNP can be used to estimate the effect of nucleotide changes on protein function [14]. Mutation discovery throughput can be increased and screening cost reduced by pooling samples up to eight-fold [3].

Examples of TILLING projects

TILLING projects have been reported for a large number of species including *Arabidopsis*, barley, *Drosophila*, maize, rice, soybean, and wheat (**Table 1**). The utility of TILLING for mutation breeding was shown by Slade and colleagues who used the method to develop wheat producing low amounts of amylase [15]. In addition to providing useful mutants, analysis of the spectrum and density of mutations can provide insights into the process of mutagenesis. For example, analysis of ~ 2,000 induced mutations indicates that EMS causes >99% G/C-A/T transition changes in the Arabidopsis genome and that mutations are distributed randomly across all chromosome arms [13]. Analysis of more than 2,000 induced mutations in *Drosophila* provided insight into the maintenance of mutations in populations propagated over 100 generations [16].

The high-throughput and low-cost nature of TILLING has allowed the creation of public TILLING services. The first operational TILLING service was the Arabidopsis TILLING Project (ATP; [5]). In operation since 2001, ATP has delivered more than 8,000 induced mutations to *Arabidopsis* researchers in many countries (http://tilling.fhcrc.org/arab/status.html). To streamline the service, a system for online ordering and automated generation of mutation reports was developed and incorporated into ATP and the later-developed Maize TILLING Project, and Fly-TILL screening services (http://tilling.fhcrc.org/fly/Welcome_to_Fly-TILL.html, http://genome.purdue.edu/maizetilling/). Other TILLING services include those for rice (http://tilling.ucdavis.edu/index.php/Main_Page) and barley (http://bioinf.scri.ac.uk/barley/, http://www.distagenomics.unibo.it/TILLMore/). Because methods for mutation discovery are largely the same across species, a single screening facility can provide multiple TILLING services, benefiting from an economy of scale and centralization of expertise. For example, the Seattle TILLING Project currently provides the ATP and Fly-TILL services and has completed small-scale pilot projects to evaluate the efficacy of TILLING in maize, rice, and soybean (http://tilling.fhcrc.org/, [17-19]). Likewise, RevGenUK provides TILLING services for Lotus and Medicago (http://revgenuk.jic.ac.uk/). Many other groups are involved in multi-species projects (for example, http://www.gabi-till.de/).

Global TILLING for developing countries

Success stories from TILLING suggest that this technology can be incorporated into the tool-kit of approaches for improving food security in developing nations. As a reverse-genetic approach, TILLING can improve the efficiency of using induced mutations to develop crops with improved traits. For example, knowledge of the spectrum and density of induced mutations in a population allows one to calculate the size of the population required to obtain desired alleles. In Arabidopsis, with a mutation density of ~ 1 mutation/ 200 kb, an average of 14 mutations in a 1.5 kb region can be obtained by screening ~3,000 plants (http://tilling.fhcrc.org/arab/status.html). This is sufficient to deliver one predicted deleterious allele at least 95% of the time in 75% of all gene targets (http://tilling.fhcrc.org/files/user_fees.html). Because genes in Arabidopsis are typically 3-4kb, a population size of 3,000 is sufficient to find deleterious alleles in many genes. To have the same chance of finding deleterious mutations in a maize TILLING population with a density of ~ 1/500 kb would require screening ~7,500 plants. Mutation densities can be estimated by screening a small subset of the population (~800 plants), and the decision to proceed with the population or to generate a new population can be made prior to a full-scale investment in thousands of plants. Thus, TILLING provides a means of reducing population size and increasing the chances of obtaining useful mutations when compared to traditional forward strategies where mutation densities are unknown. Furthermore, TILLING allows the discovery of mutations that provide no phenotype by themselves, but can when combined with others. This is important for duplicated genes, genes with redundant functions and in polyploid species such as wheat, where recessive traits are not likely to be obtained by traditional forward screens and mutation breeding [20, 21].

Induced mutation for plant breeding has been a powerful approach used by the FAO/IAEA Joint Programme. The Plant Breeding Unit has provided a plant mutagenesis service for decades, and maintains expertise in mutagenesis, plant growth and mutant evaluation that is shared with Member States. The Plant Breeding Unit (PBU) is therefore well-positioned to increase the efficiency of using induced mutations by establishing a TILLING facility that can serve Member States. Incorporating methods to improve efficiency is timely because a growing population and changing climate is expected to put increasing pressures on food security. We have chosen three crops: banana, cassava, and rice, to develop a TILLING platform at the PBU. Pilot-scale studies with all three crops have been successful, suggesting the efficacy of a centralized-

Table 1. Spectrum and density of induced mutations identified by TILLING.

Organism	Mutagen	Population screened	Spectrum of mutations	Mutation density (mutations/kb)	References
Arabidopsis	EMS	3000	>99% GC→AT	1/200	[5, 13]
Barley	EMS	4600	70% GC→AT	1/1000	[22]
C. elegans	EMS	1500	96% GC→AT	1/293	[23]
Drosophila	EMS (Ch2)*	5600	70% GC→AT	1/480	[16]
	EMS (Ch3)	6000	76% GC→AT	1/380	
Maize	EMS	750	>99% GC→AT	1/500	[19]
Pea	EMS	3072	100% GC→AT	1/669	[24]
Rice	EMS	768	70% GC→AT 11% AT→GC	1/290	[18]
	Az-MNU		67% GC→AT 20% AT→GC	1/270	
	MNU	767	92% GC→AT	1/135	[9]
	Gamma ray	2130	50% Transversion 17% Transition 33% small indel	1/6190 kb	[25]
Soybean	MNU	529	~90% GC→AT	1/140	[17]
	EMS	768	~90% GC→AT	1/140	
	EMS	768	~90% GC→AT	1/250	
	EMS	768	~75% GC→AT	1/550	
Wheat	EMS	1920	~99% GC→AT	1/40 (tetraploid) 1/24 (hexaploid)	[15]

* Separate populations balanced for mutations on chromosome 2 (Ch2) and 3 (Ch3) were screened.

facility approach to providing reverse-genetics for crops important to developing nations. Mutagenized banana and cassava populations have been developed in-house and initial screening is promising. We have begun to collaborate with other groups to screen the mutant populations they have developed. We hope to expand this work to include more crops.

BIBLIOGRAPHY

1. Till, B.J., Comai, L., Henikoff, S. Tilling and Ecotilling for crop improvement. In: *Genomics-Assisted Crop Improvement* (R. K. Varshney, and R. Tuberosa, Eds.), Vol. 1: Genomics Approaches and Platforms, p.333-349. Springer Netherlands (2007).

2. Tisser, A., Bourgeois, P. Reverse Genetics in Plants. *Current Genomics* **23**, 269-284 (2001).

3. Colbert, T., Till, B.J., Tompa, R., Reynolds, S., Steine, M.N., Yeung, A.T., McCallum, C.M., Comai, L., Henikoff, S. High-throughput screening for induced point mutations. *Plant Physiol.* **126**, 480-4 (2001).

4. McCallum, C.M., Comai, L., Greene, E.A., Henikoff, S. Targeted screening for induced mutations. *Nat. Biotechnol* **18**, 455-7(2000a).

5. Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., Young, K., Taylor, N.E., Henikoff, J.G., Comai, L., Henikoff, S. Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Res* **13**, 524-30 (2003).

6. Bentley, A., MacLennan, B., Calvo, J., and Dearolf, C.R. Targeted recovery of mutations in Drosophila. *Genetics* **156**, 1169-73 (2000).

7. McCallum, C.M., Comai, L., Greene, E.A., Henikoff, S. Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol.* **123**, 439-42 (2000b).

8. Perry, J.A., Wang, T.L., Welham, T.J., Gardner, S., Pike, J.M., Yoshida, S., Parniske, M. A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiol.* **131**, 866-71 (2003).

9. Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H., Moriguchi, K., Nagato, Y., Kurata, N. MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Mol. Genet. Genomics* **279**, 213-23 (2008).

10. Wienholds, E., van Eeden, F., Kosters, M., Mudde, J., Plasterk, R.H., Cuppen, E. Efficient target-selected mutagenesis in zebrafish. *Genome Res* **13**, 2700-7 (2003).

11. Till, B.J., Zerr, T., Comai, L., Henikoff, S. A protocol for TILLING and Ecotilling in plants and animals. *Nat Protoc* **1**, 2465-77 (2006).

12. Till, B.J., Burtner, C., Comai, L., Henikoff, S. Mismatch cleavage by single-strand specific nucleases. *Nucleic Acids Res* **32**, 2632-41 (2004a).

13. Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L., Henikoff, S. Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in *Arabidopsis. Genetics* **164**, 731-40 (2003).

14. Taylor, N.E., Greene, E.A. PARSESNP: A tool for the analysis of nucleotide polymorphisms. *Nucleic Acids Res* **31**, 3808-11 (2003).

15. Slade, A.J., Fuerstenberg, S. I., Loeffler, D., Steine, M. N., Facciotti, D. A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat Biotechnol* **23**, 75-81 (2005).

16. Cooper, J.L., Greene, E.A., Till, B.J., Codomo, C.A., Wakimoto, B.T., Henikoff, S. Retention of induced mutations in a *Drosophila* reverse-genetic resource. *Genetics* **180**, 661-7 (2008a).

17. Cooper, J.L., Till, B.J., Laport, R.G., Darlow, M.C., Kleffner, J.M., Jamai, A., El-Mellouki, T., Liu, S., Ritchie, R., Nielsen, N., Bilyeu, K.D., Meksem, K., Comai, L., Henikoff, S. TILLING to detect induced mutations in soybean. *BMC Plant Biol* **8**, 9 (2008b).

18. Till, B. J., Cooper, J., Tai, T. H., Colowit, P., Greene, E. A., Henikoff, S., Comai, L. Discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biol* **7**, 19 (2007).

19. Till, B.J., Reynolds, S.H., Weil, C., Springer, N., Burtner, C., Young, K., Bowers, E., Codomo, C.A., Enns, L.C., Odden, A.R., Greene, E.A., Comai, L., Henikoff, S. Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biol* **4**, 12 (2004b).

20. Enns, L.C., Kanaoka, M.M., Torii, K.U., Comai, L., Okada, K., Cleland, R.E. Two callose synthases, GSL1 and GSL5, play an essential and redundant role in plant and pollen development and in fertility. *Plant Mol Biol* **58**, 333-49 (2005).

21. Slade, A.J., Knauf, V.C. TILLING moves beyond functional genomics into crop improvement. *Transgenic Res* **14**, 109-15 (2005).

22. Caldwell, D.G., McCallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F., Waugh, R. A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare* L.). *Plant J* **40**, 143-50 (2004).

23. Gilchrist, E.J., O'Neil, N.J., Rose, A.M., Zetka, M.C., Haughn, G.W. TILLING is an effective reverse genetics technique for *Caenorhabditis elegans*. *BMC Genomics* **7**, 262 (2006).

24. Triques, K., Sturbois, B., Gallais, S., Dalmais, M., Chauvin, S., Clepet, C., Aubourg, S., Rameau, C., Caboche, M., Bendahmane, A. Characterization of *Arabidopsis thaliana* mismatch specific endonucleases: application to mutation discovery by TILLING in pea. *Plant J* **51**, 1116-25 (2007).

25. Sato, Y., Shirasawa, K., Takahashi, Y., Nishimura, M., Nishio, T. Mutant Selection from Progeny of Gamma-ray-irradiated Rice by DNA Heteroduplex Cleavage using Brassica Petiole Extract. *Breeding Science* **56**, 179-183 (2006).

TILLING with TILLMore

V Talamè[1,*], R Bovina[1], S Salvi[1], M C Sanguineti[1], P Piffanelli[2], U Lundquist[3] & R Tuberosa[1]

Abstract

A TILLING (Targeting Induced Local Lesion IN Genomes) resource in barley (cv. Morex) consisting of 4,906 families was produced by sodium azide (NaN$_3$) seed treatment. This resource has been named TILLMore. TILLMore was screened for mutants at several genes based on the analysis of 8- to 12-fold DNA pools produced from DNA samples from M$_2$ or M$_3$ plants. An average of about six alleles per gene was identified, which corresponds to a rate of one mutation per every 428kb. Almost all the mutations detected were C/G to T/A transitions and several (ca. 58%) implied a change in amino acid sequence, hence possible effects on phenotype. The mutagenized barley population, although developed for reverse-genetics purposes, is also a valuable resource for forward-genetics studies. A high frequency of M$_3$ families (ca. 33%) showed morphological alterations, which have been scored regularly during the growing seasons in reference to wild-type Morex plants (www.distage-nomics.unibo.it/TILLMore/). In particular, a preliminary screening for phenotypes at the root level showed altered root morphology for ca. 7% of the families. Our results indicate the feasibility of using this collection of materials to investigate gene function in barley and closely related crops.

Introduction

One of the most widely adopted approaches in biological research is mutational analysis where chemical or physical mutagenesis is used to achieve a large-scale random gene inactivation in a population of plants. Forward genetics based on the phenotypic identification of individual mutations has been the most widely used tool in understanding the genetic control of plant development. However, this approach is not always effective in filling in the gap between gene structure and function; additionally, considerable efforts are needed to identify rare mutations [1]. These limitations can be overcome by high-throughput reverse-genetics tools that allow us to identify mutations in genes that are known only by their sequence.

Among other high-throughput reverse-genetics approaches, TILLING (Targeting Induced Local Lesion IN Genomes) is a technique combining chemical mutagenesis with a sensitive DNA screening-technique that enables the recovery of individuals carrying allelic variants at candidate genes [2]. TILLING can be applied to either chemically mutagenized populations or a collection of genotypes (e.g. cultivars, ecotypes, landraces and wild accessions), this latter approach also known as EcoTILLING [3]. TILLING can be applied essentially to any species, without significant differences in protocol. Since its first introduction, TILLING has been successfully used as a functional genomics discovery platform both in plants [2, 4, 5, 6, 7, 8, 9, 10] and animal species [11, 12, 13], and more TILLING resources are underway in model and non-model species. This study describes the utilization of TILLMore, our recently developed sodium azide-mutagenized barley population in a cv. Morex background [10], both for reverse and forward-genetics studies.

Materials and Methods

TILLING analysis

DNA preparation and pooling. Genomic DNA isolated from the whole TILLMore population, including 4,906 M$_{3/4}$ families, was quantified and diluted to a concentration of 30 ng/µl and used for serial dilution stocks of 10 and 2.5 ng/µl. Two different DNA stocks were prepared: one referred to 3,148 samples collected from M$_2$ plants, while the second DNA stock referred to the whole TILLMore population of 4,906 M$_3$ families. The first DNA stock was pooled four-fold and then combined to obtain eight-fold pools. The second DNA stock was pooled according to a two-dimensional gridding strategy combining all samples both in eight and in twelve-fold pools.

Mutation screening. The primer design, based on the Morex genomic sequence, was carried out using CODDLE (http://www.proweb.org/coddle; [14]) and Primer3 (http://frodo.wi.mit.edu/; [15]). PCR amplification was conducted as described in [16]. Amplified samples were digested with CelI nuclease according to the manufacturer's directions for the Surveyor® Mutation Detection Kit for agarose gel (Transgenomics Inc., Omaha, NE, USA) and to the protocol described in [16]. Electrophoresis was performed either on a LI-COR4200 gel analyzer (LI-COR Inc., USA) according to the protocol described in [16] or on an ABI3730 capillary DNA Sequencer (ABI, USA) according to [11] with minor modifications. Images were analyzed visually for the presence of cleavage products using Adobe Photoshop software (Adobe Systems Inc.). Presumptive mutants were subsequently confirmed by sequencing genomic DNA using the ABI BigDye Terminator V1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). Sequences of mutant alleles were analyzed with the SIFT [17] and PARSESNP [14] programmes in order to rank mutations based on probability of affecting protein function.

Forward-genetics screening

The entire TILLMore population of 4,906 M$_{3/4}$ families obtained by sodium-azide seed treatment was grown in the field and scored for visible phenotypes with reference to untreated Morex. Phenotypic information was collected as to habitus, heading date, leaf appearance, presence of necrotic spots, plant color, plant height, plant morphology, spike appearance and tillering.

To detect alterations in root morphology, a preliminary analysis of ca. 1,000 M$_3$ families was carried out using a paper-roll technique [18]. About 12 seeds/family/paper-roll were grown under controlled conditions (16/8 h photoperiod and 24/22°C day and night, respectively) in a replicated screen. Observations and measurements of seminal roots were performed on eight-day-old seedlings with reference to untreated Morex.

[1] Department of Agroenvironmental Sciences and Technology, Viale Fanin 44, 40127 Bologna, Italy
[2] Genomics Platform, CERSA, Parco Tecnologico Padano (PTP-Lodi), Via Einstein, 26900 Lodi, Italy
[3] Nordic Genetic Resource Center, SE-230 53 Alnarp, Sweden

* Corresponding author. E-mail: valentina.talame2@unibo.it

Results and Discussion

TILLMore reverse-genetics screening

We screened TILLMore with assays designed for 11 target genes involved in different aspects of barley development and metabolism. The target genes were either analyzed for internal use or for scientific collaborations or screened for other research groups (**Table 1**). The molecular screening of the first set of five genes was carried out on 3,148 samples using a *Cel*I-based heteroduplex assay coupled with gel-electrophoretic detection on a LI-COR analyzer, screening for mutations in eight-fold sample pools. For the second group of six genes, we used an ABI3730 sequencer system on the entire population of 4,906 samples, screening for mutations in eight and twelve-fold sample pools. The two methods showed comparable mutation frequencies in reference to the total number of samples screened.

For each TILLed gene, the molecular screening yielded an allelic series of mutants with an average of ca. six alleles per gene corresponding to a mutation density of ca. one mutation every 428kb. This parameter was estimated by dividing the number of base pairs screened by the total number of identified mutations. Considering that mutations placed in the terminal 80bp of both ends of the amplicon can escape identification due to PCR priming and electrophoresis artefacts, a correction for the effective screening window was applied by subtracting 160bp from the length of each amplicon [19].

A total of 69 point mutations were identified at the target genes. In 13 cases the mutations were in non-coding regions. Of the 56 point mutations identified in coding regions, some (29%) are predicted to be silent because they affect the third base of a codon which does not change the amino acid encoded by that codon, while 70% of them are missense alleles resulting in a change in one of the amino acids in the protein encoded by the gene (**Table 1**). Only one truncation mutation was identified for the targeted genes.

Table 1. TILLMore mutation screening

Gene	PCR size (bp)	Non-sense	Missense	Silent	Non coding	Total	User
BRXL	940	-	3	2	4	9	TILLMore
HvCO1	830	-	2	1	-	3	TILLMore
HvEXPB1	997	-	3	1	-	4	TILLMore
NR	990	-	7	3	-	10	TILLMore
RPD1	670	-	4	3	2	9	TILLMore
eIF4E	525	-	3	-	2	5	Collaboration
HvFZP	560	-	3	4	-	7	Collaboration
Rpg1	908	-	3	-	1	4	Collaboration
#1	677	-	5	1	1	7	Customer
#2	963	1	5	-	-	6	Customer
#3	789	-	1	1	3	5	Customer
Total	9149	1	39	16	13	69	

Of the 69 lesions, only two were C/G to A/T transversion, while 67 were G/C to A/T transitions. Although a base transition mechanism has been postulated to explain the mutagenic effect of sodium azide [20], the exact type of NaN_3-derived mutants has not been described exhaustively in terms of nucleotide changes. To our knowledge, only one study investigating the nucleotide changes induced by NaN_3 in barley has been published to date [21]. The authors proposed that NaN_3 causes mutations of transition type. Although limited to few genes, our results combined with those reported by Olsen and colleagues [21] reinforce the hypothesis that NaN_3-induced sequence alterations in barley are mainly transitions. Due to these considerations, plus the finding that all but two

of our mutations are G/C to A/T transitions, we feel confident in ruling out the possibility that they are naturally occurring polymorphisms due to seed contamination of the original Morex seed stock used for this study.

For the prediction of the impact of mutations on protein function, we utilized bioinformatic methods like PARSESNP [17, 14]. It is expected that values of SIFT or PSSM (for PARSESNP) scores above specific thresholds indicate missense mutations which are more likely to have a deleterious effect on protein function. In our case, five mutations showed significant PSSM values, while the application of the SIFT algorithm predicted possible deleterious effects for only two mutations.

Forward-genetics screening

The main purpose of our mutagenized barley population was the implementation of a TILLING resource facility to be used for reverse-genetics screenings. However, the same population can also be used for forward-genetics studies.

For this purpose, the 4,906 M_3 families were grown in the field and visible phenotypes were scored during the growing season with reference to untreated Morex plants. A visible variant phenotype was recorded for 32.7% of the M_3 families (1,605/4,906) either fixed or segregating within the family (**Figure 1**). Changes in plant color, including families showing segregation for albino seedlings, was the phenotype most frequently observed (27% of mutated families; 12% of total families).
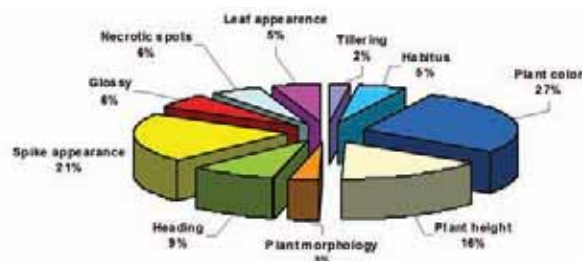
Figure 1 Proportions of different types of altered phenotypes in reference to cv. Morex recorded among M_3 families grown in field conditions.

Furthermore, a preliminary study of root morphology in 1,000 M_4 families by means of a paper-roll technique allowed the detection of putative mutants showing, for instance, a reduction of root length and severe modifications in root appearance and growth. About 7% of the families tested showed a clearly altered root phenotype. In order to confirm the observed phenotype the paper-roll screening will be repeated for the putative mutants identified during this preliminary screening.

An online database with further information on phenotypes observed in the forward-genetics screening is publicly accessible at www.distagenomics.unibo.it/TILLMore/ and seed is available upon request.

Conclusions

Our report confirms that TILLMore, a TILLING resource recently developed in Morex barely, is suitable for both forward- and reverse-genetics screening. The use of NaN_3 provided an efficient alternative to more commonly used mutagenic agents to obtain a high mutation density suitable for TILLING. TILLMore is available to the research community, both as forward and reverse-genetics resources, on a cost-recovery basis and/or through collaborations (for details, see www.distagenomics.unibo.it/TILLMore/).

BIBLIOGRAPHY

1. Stemple, D.L. TILLING - a high-throughput harvest for functional genomics. *Nat. Rev. Genet.* **5**, 145-150 (2004).

2. McCALLUM, C.M. *et al*. Targeted screening for induced mutations. *Nat. Biotechnol.* **18**, 455-457 (2000).

3. Comai, L. *et al*. Efficient discovery of DNA polymorphisms in natural populations by EcoTILLING. *Plant J.* **37**, 778-786 (2004).

4. Perry, J.A. *et al*. A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiol.* **131**, 866-871 (2003).

5. Till, B.J. *et al*. Discovery of induced point mutations in maize genes by TILLING. *BMC Plant Biol.* **28**, 4-12 (2004).

6. Caldwell, D.G. *et al*. A structured mutant population for forward and reverse genetics in barley (*Hordeum vulgare* L.). *Plant J.* **40**, 143-150 (2004).

7. Till, B.J. *et al*. TILLING and EcoTILLING for crop improvement. In: *Genomics-Assisted Crop Improvement. Vol. 1: Genomics Approaches and Platforms* (Varshney. R.K., Tuberosa. R., Eds.). Springer. Dordrecht, 333-349 (2007).

8. Slade, A.J. *et al*. A reverse genetic. nontransgenic approach to wheat crop improvement by TILLING. *Nature Biotechnol.* **23**, 75-81 (2005).

9. Suzuki, T. *et al*. MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Mol. Genet. Genomics* **279**, 213-223 (2008).

10. Talamè, V. *et al*. TILLMore. a resource for the discovery of chemically induced mutants in barley. *Plant Biotechnol. J.* **6**, 477-485 (2008).

11. Wienholds, E. *et al*. Efficient target-selected mutagenesis in zebrafish. *Genome Res.* **13**, 2700-2707 (2003).

12. Winkler, S. *et al*. Target-selected mutant screen by TILLING in *Drosophila. Genome Res.* **15**, 718-723 (2005).

13. Gilchrist, E.J. *et al*. TILLING is an effective reverse genetics technique for *Caenorhabditis elegans. BMC Genomics* **7**, 262 (2006).

14. Taylor, N.E.. Greene. E.A. PARSESNP: A tool for the analysis of nucleotide polymorphisms. *Nucleic Acids Res.* **31**, 3808-3811 (2003).

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

16. Colbert, T. *et al*. High-throughput screening for induced point mutations. *Plant Physiol.* **126**, 480-484 (2001).

17. Ng, P.C., Henikoff. S. Predicting deleterious amino acid substitutions. *Genome Res.* **11**, 863-874 (2001).

18. Woll, K. *et al*. Isolation, characterization and pericycle-specific transcriptome analyses of the novel maize lateral and seminal root initiation mutant rum1. *Plant Physiol.* **139**, 1255-1267 (2005).

19. Greene, E.A. *et al*. Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. Genetics 164 (2003) 731-740.

20. Kleinhofs, A. *et al*. Azide. Mutat. Res. **55**, 165-195 (1978).

21. Olsen, O. *et al*. Sodium azide mutagenesis: preferential generation of A·T·G·C transitions in the barley *Ant18* gene. *Proc. Natl. Acad. Sci. USA* **90**, 8043-8047 (1993).

EMAIL - A Highly Sensitive Tool for Specific Mutation Detection in Plant Improvement Programmes

L S Lee[1,*], M J Cross[2,3] & R J Henry[2,3]

Abstract

TILLING (Targeting Induced Local Lesions IN Genomes) is a useful tool for discovery of specific point mutations in genes of interest to plant breeders. It employs mismatch cleavage detection using endonucleases, particularly CEL I and CEL II. However, conventional protocols are limited in their ability to detect mismatch cleavage due to non-specific removal, by the nuclease, of 5' end-labelled termini. Mutation detection is further limited by the high background characteristic of PCR-based end-labelling mismatch scanning techniques. Here we show that, as nuclease activity increased, internal signal was maintained while 5' signal decayed and that internal labelling reduced the background signal. A new mismatch scanning assay called 'Endonucleolytic Mutation Analysis by Internal Labelling' (EMAIL), was developed using capillary electrophoresis, involving internal amplicon labelling by PCR incorporation of fluorescently-labelled deoxynucleotides. Multiple mutations amongst allelic pools were detected when EMAIL was applied with the mismatch nucleases CEL I and CEL II. This technique offers greatly increased sensitivity in the detection of mutations in specific genes in pooled samples, enabling enlarged pool sizes and improving throughput and efficiency. We are investigating the limits of pool sizes to deliver a highly efficient mutation detection and analysis strategy for plant breeders and geneticists.

Introduction

Over the past decade, the TILLING technique (Targeting Induced Local Lesions IN Genomes) has proven to be a useful tool for the discovery of specific point mutations in genes of interest to plant breeders and other researchers [1, 2]. The approach employs mismatch cleavage detection using endonucleases, particularly the enzymes CEL I and CEL II. During PCR annealing, heteroduplexes in dsDNA arise in pooled genomic DNA samples which contain one or more Single Nucleotide Polymorphisms (SNP), resulting, for instance, from induced mutation. The cleaved fragments can be distinguished from the larger perfectly-matched homoduplex DNA of the unmutated wild types in the sample of pooled individuals. The ability to efficiently detect individuals with specific mutations within pools of samples, provides plant breeders with a powerful early-screening tool with which to greatly reduce the numbers of plants requiring phenotypic assessment. Further, it enables geneticists to analyze gene function and associate genotype with phenotype.

TILLING was developed for scanning populations harboring point mutations derived primarily by chemical mutation, typically using ethyl methanesulfonate (EMS) [3], and extended to detection of natural mutants [4, 5]. Recently, however, TILLING has also proven useful in scanning gamma-irradiated mutant populations [6].

Mismatch detection protocols suffer from a limited ability to detect the mismatch cleavage signal due to non-specific removal, by the nuclease, of 5' end-labelled termini used in the conventional approach. Mutation detection is further limited by the high background that is characteristic of PCR-based end-labelling mismatch scanning techniques. By studying the activity of CEL nucleases using amplicon substrates labelled both internally and at each 5' terminus, we showed that, as nuclease activity increased, internal signal was maintained while 5' signal decayed. Furthermore, the background using internal labelling was reduced relative to conventional end-labelling techniques. The loss of end-signal constitutes a fundamental problem with the conventional approach to mismatch scanning with CEL nucleases.

An improved technique

A new mismatch scanning assay has been developed using capillary electrophoresis, in which amplicon labelling is achieved by PCR incorporation of fluorescently-labelled deoxynucleotides. We have named this strategy 'Endonucleolytic Mutation Analysis by Internal Labelling' (EMAIL) [7]. Multiple mutations amongst allelic pools have been detected when the EMAIL assay was applied with the mismatch nucleases CEL I and CEL II.

The electropherogram in **Fig. 1** shows the effect of internal labelling compared to 5' end-labelling. The upper panel shows a sample pool detected with end-labelling; the lower panel shows the result of internal labelling. The sample comprised a six-fold genomic DNA pool consisting of four homozygous wild-type rice (*Oryza sativa*) individuals, but with inclusion of two additional homozygous mutant individuals. The rice samples consisted of cultivars highly characterized for their SNP content in exon 8 of starch synthase IIa. The mutant cultivars were known to contain SNP of A>G (SNP3) and GC>TT (SNP4) at positions 2412 and 2543-2544, respectively [8] (NCBI cDNA accession AF419099).

The electropherogram panels represent different dye traces from a single injection of purified heteroduplex digestion products following 18 minutes of CEL I activity. The single PCR product was amplified from the genomic DNA pool, consisting of homozygous wild-type individuals plus the homozygous mutant individuals, SNP3 and SNP4. The PCR product is in one instance end-labelled at both 5'-termini with HEX (hexachloro-6-carboxyfluorescein), and in the other case, internally-labelled with fluorescent dUTP [R110] (6-carboxyrhodamine), represented by the green and blue traces respectively. [R110]-labelled nucleotides were present in the PCR at 4μM. Peaks are noted for each site of mismatch cleavage. In each case, the ~5nt shorter internally labelled peak (lower panel) is of a considerably higher signal strength than its end-labelled counterpart (upper panel). Both cleavage fragments from SNP3, and the larger fragment from SNP4 are detectable, however, detection of the ~100bp cleavage fragment from SNP4 in the end-labelled sample is limited due to significant background in the small size range. In the internally labelled trace, all cleavage fragments from SNP3 and SNP4 are clearly detectable, both due to increased signal strength and reduced background, especially in the small size range. Although

[1] Center for Tropical Crops and Biocommodities, Queensland University of Technology, Brisbane, Australia
[2] Center for Plant Conservation Genetics, Southern Cross University, Lismore, Australia
[3] Grain Foods CRC Ltd, 1 Rivett Road, North Ryde, Australia

* Corresponding author. E-mail: slade.lee@qut.edu.au

incomplete removal of [R110]dUTP artefacts (see lower panel) has resulted in carry-over, this does not pose significant problems for the purposes of mutation scoring, since the resultant interference is at two single-points, the 'size' of which is expected at ~76 and ~141 'nt'. For this individual assay, CEL I digestion time has been increased to optimize the internally-labelled approach, however, detection of the ~100nt cleavage fragment may not necessarily be improved simply by reducing digestion time, since increased background in the small size range will also result.

In summary, non-specific digestion of the end-labels results in significantly reduced signal from the cleaved amplicons (upper panel) compared to the result with internally labelled amplicons.

This new technique offers an increased degree of sensitivity in detection of mutations in specific genes in pooled samples, thereby enabling enlarged pool sizes to improve throughput and efficiency of the mutation scanning process. We are now investigating the limits of pool sizes in order to deliver a highly efficient mutation detection and analysis strategy for plant breeders and geneticists.
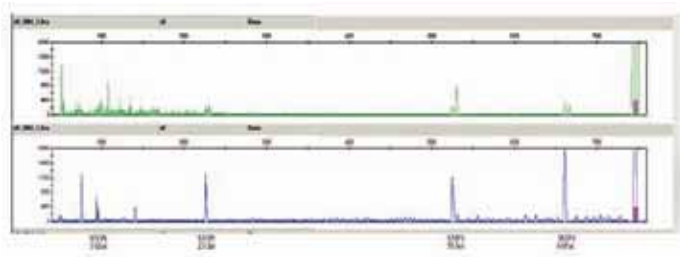
Figure 1 Comparisons of starch synthase IIa electropherograms of a pooled sample of rice genomic DNA digested with CEL I. The sample included two individuals with SNP mutations, detectable as cleavage fragments. The upper panel shows 5' end-labelled PCR products, whilst the lower panel shows the effect of internal labelling in increasing relative signal strength.

Practical implications

Previous work in our laboratories [4] demonstrated the improved detection of point mutations in polyploid sugarcane specimens by taking advantage of the sensitivity of capillary electrophoresis systems. This equated to fourteen-fold pooling due to the high ploidy level. Our research to date on the EMAIL technology has distinguished discrete SNP individuals in pools of 16 and we anticipate being able to perform routine detection in pools of at least 20. Developments in DNA sequencing have made sequencing the most cost-effective method for detection of most common mutations. However, the TILLING approach is an attractive option for discovery of rare alleles or mutations if large pools of individuals can be surveyed. EMAIL has now provided a protocol that has the capacity to meet this requirement and may prove to be the method of choice for the detection of rare mutations.

The advent of the EMAIL technique introduces a significant improvement in the efficiency of scanning pools of samples potentially containing point mutations in specific genes of interest. Furthermore, because of the degree of resolution of capillary electrophoresis, information is simultaneously obtained on the location of the mutation in the DNA sequence. Accordingly, the technique offers the plant breeder a new tool for the efficient screening of induced mutant populations at an early stage for variants in genes of interest before taking plants to field trial. It has the added advantage of providing information to assist the molecular characterization of mutations in genes of interest.

BIBLIOGRAPHY

1. Comai, L., Henikoff, S. TILLING: practical single-nucleotide mutation discovery, *Plant J.* **45**, 684-694 (2006).

2. Caldwell, D.G., Mccallum, N., Shaw, P., Muehlbauer, G.J., Marshall, D.F. Waugh, R. A structured mutant population for forward and reverse genetics in Barley (*Hordeum vulgare* L.). *Plant J.* **40**, 143-150 (2004).

3. Henikoff, S., Comai, L. Single-nucleotide mutations for plant functional genomics. *Ann. Rev. Plant Biol.* **54**, 375-401 (2003).

4. Cordeiro, G., Eliott, F.G., Henry R.J. An optimized ecotilling protocol for polyploids or pooled samples using a capillary electrophoresis system. *Anal. Biochem.* **355**, 145-147 (2006).

5. Comai, L., Young, K., Till, B.J., Reynolds, S.H, Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., Henikoff, S. Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant J.* **37**, 778-786 (2004).

6. Sato, Y., Shirasawa, K., Takahashi, Y., Nishimura, M., Nishio, T. Mutant selection from progeny of gamma-ray-irradiated rice by DNA heteroduplex cleavage using Brassica petiole extract. *Breeding Science* **56**, 179-183 (2006).

7. Cross, M.J., Waters, D.L.E., Lee, L.S., Henry R.J. Endonucleolytic mutation analysis by internal labelling (EMAIL). *Electrophoresis* **29**, 1291-1301 (2008).

8. Waters, D.L.E., Henry, R.J., Reinke, R.F., Fitzgerald, M.A. Gelatinization temperature of rice explained by polymorphisms in starch synthase. *Plant Biotech. J.* **4**, 115-122 (2006).

DNA Sequence Analysis of Induced Mutants in Soybean

Kyujung Van[1] & Suk-Ha Lee[1,2,*]

Abstract

Chemical mutagens, such as ethylmethanesulfonate (EMS), that cause point mutations were commonly used to induce mutations in both plants and animals due to the high frequency of heritable mutations. Saturated mutagenized populations could be generated with relatively few individuals. High-throughput genome sequencing is now available using the GS-20 or GS-FLX machines from Roche/454 Life Sciences, which are based on the advantages of emulsion PCR and pyrosequencing. Because of the wide availability of several high throughput genotyping methods, single nucleotide polymorphisms (SNPs) have drawn the attention of researchers. Degenerate oligonucleotide primed PCR (DOP-PCR) has been used for SNP genotyping in various organisms. In this study, we report the massive DNA sequence analysis of soybean mutants induced by EMS mutagenesis. Three soybean genotypes were used: 'Sinpaldalkong 2,' 'SS2-2' and '25-1-1.' Sinpaldalkong 2 is a recommended soybean variety in Korea and SS2-2 was generated by EMS mutagenesis of Sinpaldalkong 2. Genotype 25-1-1 is a regenerated M_3 plant from EMS-treated immature embryo culture of Sinpaldalkong 2. After genomic DNA from the three soybean genotypes was amplified with modified DOP-PCR, nucleotide sequences were determined using the GS-FLX. Assuming only orthologues show 100% identity in BLASTN, an average of 1,100 contigs and 7,000 singlets were formed in each soybean genotype. A total of 1,187 SNPs were detected, with a frequency as 1 SNP per 272bp, using POLYBAYES to survey sequence polymorphisms.

Introduction

Chemical mutagens, such as EMS, have been commonly used to induce mutations for forward genetics and enable saturated mutagenized populations to be assembled with relatively few individuals [1]. These chemical mutagens can be applied to many organisms [2-4]. The EMS treatment causes point mutations and any genome size is suitable for chemical mutagenesis by EMS [1, 5-7].

Several new 'next-generation' sequencing instruments, including GS-20 or GS-FLX from Roche/454 Life Sciences [8], Illumina's Solexa 1G sequencer [9], and the SOLiD system from Applied Biosystems (http://solid.appliedbiosystem.com), are available for high-throughput sequencing [10, 11]. The 454/Roche Genome Sequencers are based on emulsion PCR and pyrosequencing. They are used for various purposes. Initially, resequencing microbial genomes, studying phylogenic relationships among microbial species and the identification of mutation sites in bacteria were performed with the 454/Roche platform [11-13]. Now, complicated large genomes are being studied for repetitive DNA and genomic copy number by high-throughput 454 sequencing [14, 15]. Many studies have been undertaken with maize using this Genome Sequencer, such as the identification of polymorphic sequences with aligned cDNA transcriptome [16] and AFLP [17] and the detection of additional maize ESTs expressed in the shoot apical meristems collected via laser capture microdissection [18].

Point mutations, in the form of SNPs, can be used as molecular markers. These mutations of a single nucleotide have been studied extensively in both plants and animals following the recent availability of several high throughput genotyping methods [19-24]. Locus-specific primers were generally used for detecting these mutations, after comparing directly sequenced amplicons of individuals [25]. However, a new strategy, DOP-PCR, was used for SNP genotyping in human, mouse and *Arabidopsis thaliana* by massive DNA sequencing because of the requirement for sequence information and the high cost of designing primers for a limited number of SNPs [25-26].

Here we report the massive DNA sequence analysis of soybean mutants induced by EMS mutagenesis using the GS-FLX and DOP-PCR.

Materials and Methods

Three soybean genotypes were used: 'Sinpaldalkong 2,' 'SS2-2' and '25-1-1.' Sinpaldalkong 2 is a recommended soybean variety in Korea [27] and SS2-2 was generated by EMS mutagenesis of Sinpaldalkong 2 [28]. Genotype 25-1-1 is a regenerated M_3 plant from EMS-treated immature embryo culture of Sinpaldalkong 2. After genomic DNA from these three soybean genotypes was extracted [29], DOP-PCR was performed as described by Janiak et al. [25]. Amplified nucleotide sequences were determined using the GS-FLX.

Results and Discussion

Custom PERL scripts and the TGICL tool were used for trimming primers and alignment and assembly, respectively [30]. First, sequences were aligned and assembled, only if the aligned sequences showed 98% or 100% identity. Subsequently, we used only orthologues showing 100% identity to avoid possible inclusion of paralogue sequences. On average 1,100 contigs and 7,000 singlets were formed in each soybean genotype (**Fig. 1**).
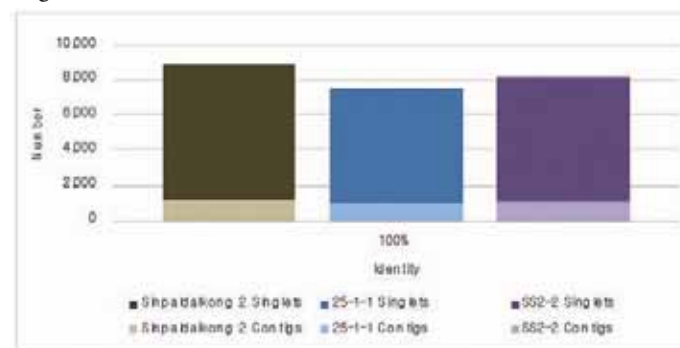

Figure 1 Number of singlets and contigs for each soybean genotype after alignments with sequences showing 100% identity.

[1] Department of Plant Science and Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul, 151-921, Korea
[2] Plant Genomic and Breeding Research Institute, , Seoul National University, Seoul, 151-921, Korea

* Corresponding author. E-mail: sukhalee@snu.ac.kr
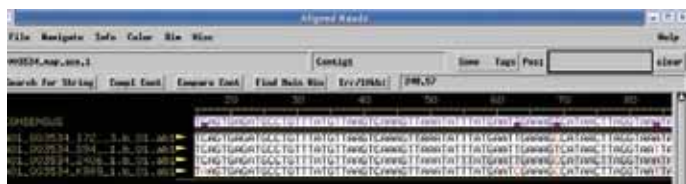
Figure 2 A portion of aligned GS-FLX sequences of soybean genotypes. SNPs are indicated by arrows.

The best BLASTN matches were used for comparing nucleotide sequences among soybean genotypes. Mutations of single nucleotides were surveyed using POLYBAYES (**Fig. 2**). A total of 1,187 SNPs were detected, with a frequency of 1 SNP per 272bp.

Further study will include the reconfirmation of sequence polymorphisms by direct sequencing after homology searches against GenBank databases [29].

BIBLIOGRAPHY

1. Henikoff, S., Comai, L. Single-nucleotide mutations for plant functional genomics. *Annu. Rev. Plant Biol.* **54**, 375-401 (2003).

2. Mccallum, C.M., Comai, L., Greene, E.A., Henikoff, S. Targeted screening for induced mutations. *Nat. Biotechnol.* **18**, 455-457 (2000).

3. Bentley, A., Maclennan, B., Calvo, J., Dearolf, C.R. Targeted recovery of mutations in Drosophilla. *Genetics* **156**, 1169-1173 (2000).

4. Wienholds, E., Schulte-Merker, S., Walderich, B., Plasterk, R.H. Target-selected inactivation of the zebrafish *rag1* gene. *Science* **297**, 99-102 (2002).

5. Koornneef, M., Dellaert, L.W., Van Der Veen, J.H. EMS-and radiation-induced mutation frequencies at individual loci in *Arabidopsis thaliana*. (L.) *Heynh. Mutat. Res.* **93**, 109-123 (1982).

6. Goll, M.G., Bestor, T.H. Histone modification and replacement in chromatic activation. *Genes Dev.* **16**, 1739-1742 (2002).

7. Zhu, B.G., Gu, A.Q., Deng, X.D., Geng, Y.Z., Lu, Z.X. Effects of caffeine or EDTA post-treatment on EMS mutagenesis in soybean. Mutat. Res. 334, 157-159 (1995).

8. Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.-J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L.I., Jarvie, T.P., Jirage, K.B., Kim, J.-B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., Mcdade, K.E., Mckenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376-380 (2005).

9. Bennett, S.T., Barnes, C., Cox, A., Davies, L., Brown, C. Toward the 1,000 dollars human genome. *Phamacogenomics* **6**, 373-382 (2005).

10. Parameswaran, P., Jalili, R., Tao, L., Shokralla, S., Gharizadeh, B., Ronaghi, M., Fire, A.Z. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Res.* **35**, e130 (2007).

11. Schuster, S.C. Next-generation sequencing transforms today's biology. *Nature Methods* **5**, 16-18 (2008).

12. Liu, Z., Lozupone, C., Hamady, M., Bushman, F.D., Knight, R. Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res.* **35**, e120 (2007).

13. Velicer, G.J., Raddatz, G., Keller, H., Deiss, S., Lanz, C., Dinkelacker, I., Schuster, S.C. Comprehensive mutation identification in an evolved bacterial cooperator and its cheating ancestor. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8107-8112 (2006).

14. Macas, J., Neumann, P., Navrátilová, A. Repetitive DNA in the pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and Medicago truncatula. *BMC Genomics* **8**, 427 (2007).

15. Swaminathan, K., Varala, K., Hudson, M.E. Global repeat discovery and estimation of genomic copy number in a large, complex genome using a high-throughput 454 sequence survey. *BMC Genomics* **8**, 132 (2007).

16. Barbazuk, W.B., Emrich, S.J., Chen, H.D., Li, L., Schnable, P.S. SNP discovery via 454 transcriptome sequencing. *Plant J.* **51**, 910-918 (2007).

17. Van Orsouw, N.J., Hogers, R.C.J., Janssen, A., Yalcin, F., Snoeijers, S., Verstege, E., Schneiders, H., Van Der Poel, H., Van Oeveren, J., Verstegen, H., Van Eijk, M.J.T. Complex reduction of polymorphic sequences (CRoPSTM): a novel approach for large-scale polymorphism discovery in complex genome. *PLoS ONE* **11**, e1172 (2007).

18. Ohtsu, K., Smith, M.B., Emrich, S.J., Borsuk, L.A., Zhou, R., Chen, T., Zhang, X., Timmermans, M.C.P., Beck, J., Buckner, B., Janick-Buckner, D., Nettleton, D., Scanlon, M.J., Schnable, P.S. Global gene expression analysis of the shoot apical meristem of maize (*Zea mays* L.). *Plant J.* **52**, 391-404 (2007).

19. Kuppuswamy, M.N., Hoffman, J.W., Kasper, C.K., Spitzer, S.G., Groce, S.L., Bajaj, S.P. Single nucleotide primer extension to detect genetic disease: Experimental application to hemophilia B (factor IX) and cystic fibrosis genes. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1143-1147 (1991).

20. Collins, F.S., Brookes, L.D., Charkravati, A. A DNA polymorphism discovery resource for research on human genetic variation. *Genome Res.* **8**, 1229-1231 (1998).

21. Brookes, A.J. The essence of SNPs. *Gene* **234**, 177-186 (1999).

22. Cho, R.J., Mindrinos, M., Richards, D.R., Sapolsky, R.J., Anderson, M., Drenkard, E., Dewdney, J., Reuber, T.L., Stammers, M., Federspiel, N., Theologis, A., Yang, W.H., Hubbell, E., Au, M., Chung, E.Y., Lashkari, D., Lemieux, B., Dean, C., Lipshutz, R.J., Ausubel, F.M., Davis, R.W., Oefner, P.J. Genome-wide mapping with biallelic markers in *Arabidopsis thaliana*. *Nature Genet.* **23**, 203-207 (1999).

23. Kwok, P.-Y., GU, Z. Single nucleotide polymorphism libraries: why and how are we building them? *Mol. Med. Today* **5**, 538-543 (1999).

24. Cai, C.M., Van, K., Kim, M.Y., Lee, S.-H. Optimization of SNP genotyping assay with fluorescence polarization detection. *Kor. J. Crop Sci.* **50**, 361-367 (2005).

25. Janiak, A., Kim, M.Y., Van, K., Lee, S.-H. Application of degenerate oligonucleotide primed PCR (DOP-PCR) for SNP discovery in soybean. *Euphytica* **162**, 249-256 (2008).

26. Jordan, B., Charest, A., Dowd, J.F., Blumenstiel, J.P., Yeh, R.-F., Osman, A., Housman, D.E., Landers, J.E. Genome complexity reduction for SNP genotyping analysis. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2942-2947 (2002).

27. Kim, S.D., Hong, E.H., Song, Y.G., Kim, Y.H., Lee, Y.H., Hwang, Y.H., Kim, H.S., Lee, S.H., Kim, W.H., Ryu, Y.H., Park, R.K. New soybean variety resistant to disease and lodging, with adapted high yielding "Sinpaldalkong 2". *RDA J. Agric. Sci.* **36**, 153-157 (1994).

28. Lee, H.S., Chae, Y.A., Park, E.H., Kim, Y.W., Yun, K.I., Lee, S.H. Introduction, development, and characterization of supernodulating mutant. I. Mutagenesis of soybean and selection of supernodulating mutant. *Kor. J. Crop Sci.* **42**, 247-253 (1997).

29. Van, K., Hwang, E.-Y., Kim, M.Y., Kim, Y.H., Cho, Y.-I., Cregan, P.B., Lee, S.-H. Discovery of single nucleotide polymorphisms in soybean using primers designed from ESTs. *Euphytica* **139**, 147-157 (2004).

30. Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., Lee, Y., White, J., Cheung, F., Parvizi, B., Tsai, J., Quackenbush, J. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* **19**, 651-652 (2003).

Driving Forward in Reverse

T L Wang[1,*], J Perry[1], T Welham[1], J Pike[3], M Parniske[2], C Rogers[1], G E D Oldroyd[1], P Stephenson[1], L Ostergaard[1], B Mccullagh[1], D Baker[1], S Walsh[1] & J Clarke[1]

Abstract

We describe the use of TILLING in *Lotus japonicus* and the development of deletion (De)-TILLING in *Medicago truncatula*. The evolution of *RevGen*UK has been driven by the development of reverse genetics technologies in these two model legumes and *Brassica rapa*, which functions as a translational species for brassica crops. TILLING and De-TILLING are underpinned by populations of plants mutagenized with either EMS (that causes point mutations) or fast neutrons (that cause deletions), respectively. They permit the isolation of either allelic series of mutants or knockouts. Mutation detection will be developed from a number of independent gel-based systems to be carried out on a single platform – capillary electrophoresis. We are currently TILLING in both model legumes, but these developments will be applied to all three species. The resource will develop an open source database-driven system to support laboratory information management, analysis and the cataloguing of mutants in a genome context across all the species.

Introduction

We are currently experiencing a revolution in plant genomics since improvements to technologies and adoption of novel approaches have greatly reduced the cost of sequencing the genome of an organism. Assigning function, however, to the thousands of plant genes that have and will be discovered presents a major challenge to the research community. Reverse genetics allows the identification of plants carrying mutations in known genes and thereby provides the means to ascertain gene function by exploiting available genome sequences [1], [2].

Lotus japonicus TILLING

Several years ago, we set up the first reverse genetics TILLING platform in Europe [3] based on an EMS-mutagenised population of ca. 5000 plants of the model legume, *Lotus japonicus*, with help from the Seattle TILLING group. This platform consisted of pooled DNA samples from single M_2 plants representing each family in the whole population plus potential mutants isolated via forward screens for morphology, nodulation ability and starch content, plus a population for eco-TILLING [3], [4], [5]. The forward populations have been used as an enriched source of deleterious mutations. By the end of 2007, we had examined more than 150 gene targets for 20 collaborators representing 11 countries worldwide. On average, we have identified ca. 2 mutations per TILLed kb per 1,000 plants, equivalent to a mutation load of ca. one mutation per 0.5Mb (**Table 1**). This platform has allowed us to examine legume-specific processes in detail, especially the symbiosis with rhizobia. In several instances, however, we have been unable to recover homozygous plants for mutations that have been found; frequently they died at the seedling stage. This indicates that some genes may have a very important role in plant growth and not one that is specific to symbiosis; loss of function in the nodule should not be lethal. In such circumstances, we can use the TILLING platform to isolate alleles that are not lethal and thus investigate gene function further. This is currently underway and demonstrates a major advantage of the TILLING platform over other types of reverse genetics that generate knockout mutants. Relying on TILLING alone for reverse genetics is limiting, since TILLING generates null alleles at a low frequency. As predicted for the original mutagen, 98% of all mutations caused G/C to A/T changes using our platform (**Table 2**), but this only mines a small proportion of the genome since of the 380 (20 amino acids x 19 alternatives) theoretical amino acid interconversions only 26 are possible through EMS mutagenesis [4]. Across all the fragments investigated, only ca. 5% of mutations identified lead to knockout mutants by causing premature stop codons or splicing abnormalities (**Table 3**) [4]. Hence, as a complement to TILLING, we have generated a novel reverse genetics platform initially in another model legume, *M. truncatula*. This platform utilizes deletions generated by fast-neutron (FN) mutagenesis and allows the identification of plants carrying deletions in target genes [6], [7]. Hence, TILLING and De-TILLING provide complementary reverse genetic platforms allowing reliable identification of allelic series and null mutants.

Table 1. Statistics of the *L. japonicus* TILLING population

Population	Total
Number of lines	4904
Number of genes TILLed	61
Number of fragments TILLed	84
Total fragment length TILLed [kb]	100
Non-overlapping fragment length TILLed [kb]	95.1
Number of mutations obtained	576
Mb screened[1]	294.6
Number of mutations per Mb[2]	2
Mutation load [kb]	511

[1] Length of non-overlapping fragments multiplied with number of plants TILLed. [2] Number of mutations obtained divided by length of non-overlapping fragments screened.

De-TILLING in *Medicago truncatula*

A major challenge to the identification of deletion mutants by PCR is overcoming the amplification of wild-type sequences to allow rare deletion-containing alleles to be preferentially amplified. Central to this platform, therefore, is our development of a detection strategy that allows a mutant amplicon, possessing an internal deletion, to be amplified in pools where the genomic target sequence is present at a 20,000-fold excess. This allows great efficiencies in time and cost in comparison to the standard eight-fold pooling of TILLING. This detection sensitivity has been achieved by combining two approaches that suppress the amplification of the undeleted wild-type fragment - restriction suppression and a

poison primer approach (**Fig. 1**). Restriction enzyme suppression relies upon the pre-digestion of highly complex DNA pools with a restriction enzyme which cuts once within the target sequence. This prevents a vast majority of the wild-type sequence from acting as a PCR template. This step relies on the fact that the deletion in the target gene will remove the restriction enzyme site and thus the deletion allele is protected from the restriction enzyme suppression. Wild-type target sequences escaping restriction enzyme suppression are subject to 'poison primer' suppression. In this strategy a third functional 'poison' primer is included in the first round of PCR. A shorter fragment, known as the suppressor fragment, is produced more efficiently and acts to suppress amplification of the longer fragment. Amplification from a mutant template present within the DNA pool, in which the poison primer-binding site has been deleted, produces a single amplicon from the external primers. Using a combination of poison primer and restriction enzyme PCR suppression we have been able to obtain detection sensitivities that allows us to screen very large pools of M_2 plants from which we can recover mutants carrying deletions in target genes [6]. Where conventional TILLING is carried out on pools of eight DNAs from individual plants, the De-TILLING platform has been optimized for the detection of a single heterozygous mutant in tower pools of 6,000 plants. Such a screening strategy is essential for exploiting fast neutron mutagenesis by reverse genetics because much larger populations of plants (50,000-100,000) are required compared to the EMS-populations required for TILLING (3,000-5,000). This is because the large deletions introduced by FN mutagenesis are tolerated at a much lower frequency than point mutations. Because there are fewer mutations per genome, however, much less clean up is required before the phenotype can be fully analyzed.

Table 2. Mutation types by nucleotide change detected in the TILLING population.				
Change	**G to A**	**C to T**	**other**	**Sum**
Absolute number	298	264	14	576
Percentage of total	51.74	45.83	2.43	100

Table 3. Distribution of mutation types in the TILLING population				
Mutation type[1]	**Total mutants**	**HOM**[1]	**HET**[1]	**Ratio HOM:HET**[1]
All	576	51	55	1:10
Missense	275	20	255	1:12
Stop	16	2	14	1:7
Splice	4	0	4	0
Deletion	0	0	0	0
Silent (intron)	163	16	147	
Silent (no aa change)	118	13	105	1:9

[1] HOM, homozygous mutants; HET, heterozygotes

Pooling lines of mutants to very high depths has the obvious advantage of reducing the amount of screening required. De-convolution of these pools, once a mutant has been detected can be time-consuming. To avoid this problem the De-TILLING platform employs a three-dimensional pooling strategy. The population is segregated into tower structures consisting of five 96-well plates of DNA extractions. Screening of the *M. truncatula* De-TILLING population is carried out on half tower pools, each representing 6,000 M_2 plants. Each tower is also pooled to create three-dimensional pools of rows, columns and plates. When a mutation is identified within a tower, these 25 pools can be screened simultaneously to identify a single seed lot within the tower in a single step. A population of 13 towers (156,000 M_2 plants) has now been established in *M. truncatula* at the JIC. A further population of around 20 towers has

also been established at the Noble Foundation, where there are plans to expand this population to 40 towers. We have focused on detecting deletions within 2-3kb PCR fragments. The screening that we have carried out on the JIC population alone indicates that we can recover mutants for 20% of targeted genes. Extrapolating this to the Noble population should allow a >50% recovery of knockout mutants for targeted genes.

The De-TILLING method can be used therefore, to isolate mutants at a fraction of the time of conventional TILLING. Fast neutron mutagenesis generates complete knockout mutants which posses a lower number of background mutations than those from TILLING platforms.
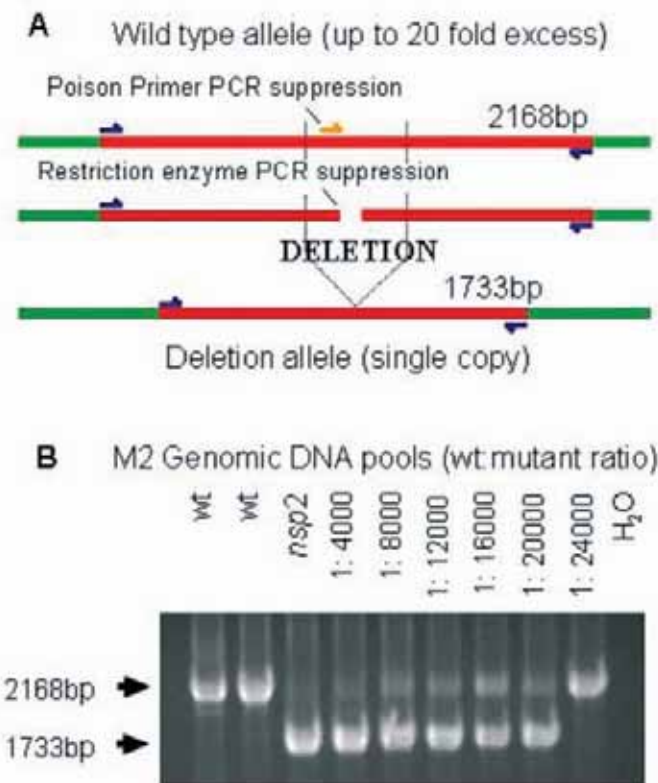
Figure 1 The De-TILLING deletion detection strategy. A. Amplification from a vast excess of wild-type sequences in DNA pools containing a deletion allele is suppressed by template restriction and production of the suppressor fragment from the poison primer. Amplification from an allele in which the restriction site and poison primer binding site is unaffected. B. A reconstruction experiment showing the amplification of the 435 bp deletion allele of nsp2-1 in dilutions of wild-type genomic DNA. In the primary amplification only the relatively short poison primer product is visible. In a nested PCR the deletion allele is amplified from a pool of 20,000 wild-type genomes [6].

Fast neutron mutagenesis can also address the problems of targeting small genes as well as recovering mutations in tandemly duplicated genes, problems that are intrinsic to all methods based on insertion and point mutation. As the cost of sequencing continues to fall, the scalability, potential low cost and technical simplicity of fast neutron-based reverse genetics is likely to be exploited for a wide variety of plant species.

Investigating the rhizobium-legume symbiosis using TILLING
As mentioned above, we have made extensive use of reverse genetics to investigate symbioses. In recent years, our understanding of the root nodule symbioses has advanced considerably. Much detail has been added by map-based cloning of forward screened mutants and by reverse genetics using both RNAi and mutants isolated using our TILLING platform. Receptor-kinases (NFR1, NFR5 and SYMRK), ion channels (CASTOR and POLLUX), a nucleoporin (NUP133) and nuclear proteins (CCaMK, CYCLOPS, NIN, NSP1 and NSP2) are all involved (**Fig. 2**; see [8, [9][8] for further details), although the interplay between these

genes has not been established. Some of the genes for which we are currently undertaking TILLING with collaborators will help to elucidate the mechanisms involved. In *L. japonicus* serine/threonine receptor kinases recognize properly decorated rhizobial lipochitin-oligosaccharide signal molecules at their extracellular LysM domains and initiate the signal transduction cascade resulting in dedifferentiation of cortical cells and initiation of cytokinin-stimulated cell divisions (that establish the nodule meristem), organ formation and invasion of the organ tissue by the microsymbiont. To date, we have isolated mutants in a number of key components of the signaling pathway that lead to the formation of a nodule as shown in **Fig. 2** (see also [4]).

As the nodule is formed, bacteria re-differentiate to form bacteroids that rely on the plant for all their nutrients, in return supplying fixed atmospheric nitrogen to the plant. Carbon is obtained via the breakdown of sucrose either by sucrose synthase or invertase as these are the only enzymes that carry out this process in plants. The carbon is believed to enter bacteroids in the form of dicarboxylic acids. Since the TILLING platform permits targeting of specific isoforms, it was also used to isolate mutants including null alleles for two of the family of six sucrose synthases in *L. japonicus* that are the main ones present in the nodule. Analyses of these lines demonstrated that both isoforms were required for nitrogen fixation and assimilation in *L. japonicus*; only double mutants that knocked out both genes were unable to assimilate nitrogen via their nodules [11].
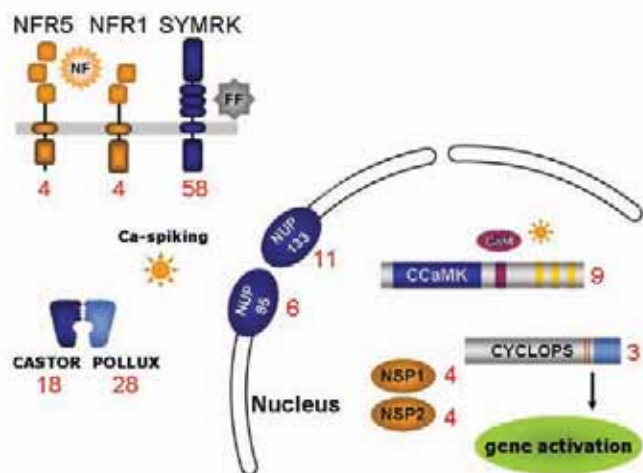
Figure 2 Components of the legume-rhizobium symbiosis signaling pathway in *Lotus japonicus*. Numbers in red after each component are the number of mutant alleles obtained by TILLING to date (see [8], [9], [10] for further details of genes.)

*RevGen*UK

There is an increasing demand for mutant alleles in conventional breeding programmes. TILLING and De-TILLING strategies provide an alternative to transgenic approaches. Using these technologies as the main pillars, in May 2008 we launched a 'single-stop' shop for use in functional genomics research based around the EMS and FN populations of *L. japonicus*, *M. truncatula* and *B. rapa* and a single capillary electrophoresis platform. It is hoped that the platform (*RevGen*UK), supported by the UK's Biotechnology and Biological Sciences Research Council, will eventually provide TILLING and De-TILLING services across three model species: our two legumes and *B. rapa*; initially only TILLING in the model legumes is being offered to the community. The platform will also have a web-accessible informatics capability and be of sufficient flexibility to allow for future expansion as the research community demands.

BIBLIOGRAPHY

1. Østergaard, L.,Yanofsky, M. Establishing gene function by mutagenesis in *Arabidopsis thaliana*. *Plant Journal* **39**, 682 (2004).

2. Alonso, J., Ecker, J. Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis. Nature Reviews Genetics* **7**, 524 (2006).

3. Perry, J. *et al*. A TILLING reverse genetics tool and a web-accessible collection of mutants of the legume *Lotus japonicus*. *Plant Physiology* **131**, 866 (2003).

4. Perry, J. *et al*. TILLING in *Lotus japonicus* identified large allelic series for symbiosis genes and revealed a bias in non-functional alleles towards hits in glycine codons. *Plant Physiology* (submitted).

5. Comai, L. *et al*. Efficient discovery of DNA polymorphisms in natural populations by Ecotilling. *Plant Journal* **37**, 778-86 (2004).

6. Rogers, C., Oldroyd, G.E.D. Fast neutron mutagenesis for functional genomics. In: *The Handbook of Plant Functional Genomics* (Kahl, G., Meksem, K., Eds.), Wiley-VCH, Weinheim, 291-306 (2008).

7. Li, X. *et al*. A fast neutron deletion mutagenesis-based reverse genetics system for plants. *Plant Journal* **27**, 235 (2001).

8. Oldroyd, G.E.D., Downie, J.A. Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annual Reviews of Plant Biology* **59**, 519 (2008).

9. Yano, K. *et al*. CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proceedings of the National Academy of Sciences* **105**, 20450-20545 (2008).

10. Parniske, M. Arbuscular mycorrhiza: the mother of all endosymbioses. *Nature Reviews Microbiology* **6**, 763-775 (2008).

11. Horst, I. *et al*. TILLING mutants of *Lotus japonicus* reveal that nitrogen assimilation and fixation can occur in the absence of nodule-enhanced sucrose synthase. *Plant Physiology* **144**, 806-820 (2007).