History of studies on transposable elements

Transposable elements (TEs) were named after their characteristics that they transpose. They change locations in a genome and increase copy numbers (Craig et al. 2002). Because not only TE transpositions alter genomic sequence, but they are probably the most abundant component in all eukaryotic genomes, they are considered as a great source for diversification of genome sizes, structures and functions (Craig et al. 2002, Feshotte et al. 2002, Kazazian Jr. 2004).

More than 60 years ago, McClintock (1950, 1951) discovered a genetic agent that is responsible for the sectors of pigmentation. Each sector of colored tissue arose where a TE, which had inserted into a gene whose expression is necessary for kernel pigmentation, was excised.

At this time, TEs were considered as a rare phenomenon of curiosity. But subsequent analyses of mutant alleles of Drosophila melanogaster, Saccharomyces cerevisiae, Caenorhabditis elegans, and other model eukaryotic organisms identified other active TEs (Craig et al. 2002). The abundance of TEs and their mutagenic potential led scientists to propose that they thrived because they had been important means of evolution of life and had been indispensable (McClintock 1984).

On the other hand, in 1980, two papers suggested that TEs were selfish and no more than junk for host genomes (Doolittle and Sapienza 1980, Orgel and Crick 1980). They suggested that the evolutionary success of TEs could be explained solely by their ability to replicate themselves faster than the host. This theory convinced many scientists to change their focus away from TEs’ impact on host evolution to characterization or mechanisms of TE transpositions.

However, a number of genome projects have now revealed that TEs are usually the largest component of the genomes of multicellular eukaryotes (International Human Genome Sequencing Consortium 2004). In addition, some human genes contain more than 100 TEs (reviewed by Wessler 2006). These findings have generated new questions that how TEs and host organisms coexist. Now investigators consider that TEs and the host genomes are in severe competition, that TEs trying to increase their copy numbers while host genome protecting its genetic information from mutations (Wessler 2006). As a result of such selective processes, TEs have survived not only because of their ability to replicate themselves but to provide the host an excellent tool to generate genetic novelties and maintain its own integrity (Labrador and Corces 2002).

Characterization of transposable elements

TEs are divided into two major groups, according to whether their transposition intermediate is RNA (class 1) or DNA (class 2) (reviewed by Feschotte et al. 2002). Class 1 elements transpose through so-called “copy-and-paste” mechanisms.
Class I elements can be divided into two subgroups on the basis of transposition mechanism and structure. One is LTR retrotransposons, which have long terminal repeats (LTRs) in direct orientation (Fig. 1A). Autonomous elements contain at least two genes, called gag and pol. The gag gene encodes a capsid-like protein and the pol gene encodes a polyprotein that is responsible for protease, reverse transcriptase, RNase H and integrase activities. Non-autonomous elements lack most or all coding sequence. Their internal region can be variable in size and unrelated to the autonomous element. The other is Non-LTR retrotransposons, which are further divided into long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). Coding regions include: ORF1, a gag-like protein; EN, endonuclease; and RT, reverse transcriptase. Both LINEs and SINEs terminate by a simple sequence repeat, usually poly(A). DNA transposons have terminal inverted-repeat (black triangles) and target-site (arrows) duplications of conserved length (and sometimes sequence) in superfamilies (for example, 8 bp for hAT; TA for Tcl1/mariner). Non-autonomous family members are usually derived from an autonomous family member by internal deletion.

Fig. 1. Structural features and classification of plant transposable elements. A. LTR retrotransposons have long terminal repeats (LTRs) in direct orientation (black triangles). Autonomous elements contain at least two genes, called gag and pol. B. Non-LTR retrotransposons are divided into long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). Coding regions include: ORF1, a gag-like protein; EN, endonuclease; and RT, reverse transcriptase. Both LINEs and SINEs terminate by a simple sequence repeat, usually poly(A). C. DNA transposons have terminal inverted-repeat (black triangles) and target-site (arrows) duplications of conserved length (and sometimes sequence) in superfamilies (for example, 8 bp for hAT; TA for Tcl1/mariner). Non-autonomous family members are usually derived from an autonomous family member by internal deletion.

Class I elements can be divided into two subgroups on the basis of transposition mechanism and structure. One is LTR retrotransposons, which have long terminal repeats (LTRs) in direct orientation (Fig. 1A). The other is Non-LTR retrotransposons, which are further divided into long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs) (Fig. 1B).

Class 2 elements are DNA transposons, whose transposition process is called “cut-and-paste” mechanism, that element-coded transposase excises the element itself or its deletion derivatives and inserts into a new location. They have terminal inverted-repeats (TIRs) and target-site duplications (TSDs) of conserved sequence (Fig. 1C).

In mammals, non-LTR retrotransposons dominate the genomes (Deininger and Batzer 2002, van de Lagemaat et al. 2003). Not only they comprise a large fraction of the genomes, but they are frequently associated with genes (van de Lagemaat et al. 2003). However, because most of the insertions occurred over 5 million years ago, their impact on gene expression will never be known due to the subsequent accumulation of additional mutations.

In plants, on the other hand, LTR retrotransposons comprise the largest component of the genome (Kumar and Bennetzen 1999) and largely contribute for diversity of genome size. Especially in maize, it was demonstrated that the bursts of LTR retrotransposons have doubled the genome size within the past 6 million years (SanMiguel et al. 1998). However, not like LINEs or SINEs in mammals, LTR retrotransposons in plants form clusters in intergenic regions and are not greatly associated with gene functions (Feschotte et al. 2002). On the other hand, DNA transposons in plants are preferentially found in single copy regions and therefore their association with genes is frequently found (Craig et al. 2002, Feschotte et al. 2002, Zhang et al. 2000). Among class 2 elements, miniature inverted-repeat transposable elements (MITEs) are outstanding because they predominate in the non-coding regions of grass genomes (Bureau and Wessler 1992, 1994a, 1994b, Bureau et al. 1996). The distinction of MITEs from other class 2 TEs is that the majority of characterized class 2 elements are longer than 1 kb and can amplify up to moderate copy numbers (less than 100 copies), while MITEs are short (<600 bp) and appeared to have attained over 1,000 (Wessler 2006).

**The discovery of mPing, the first active MITE**

To assess the impact of MITE insertions on gene and genome evolution, it was necessary to identify MITEs that are still transposing. Such an element was finally found in rice and named mPing (Jiang et al. 2003, Kikuchi et al. 2003, Nakazaki et al. 2003). mPing was discovered independently in three laboratories working with three different materials: long-term cell culture (Jiang et al. 2003), anther culture (Kikuchi et al. 2003) and mutant strains induced from the temperate japonica cultivar Gimbozu with gamma-irradiation (Nakazaki et al. 2003).

The identification of mPing opened the door to addressing questions concerning the impact of MITE insertions into plant genes. However, although mPing is clearly an active MITE, its copy numbers were found to be relatively low, with less than 10 copies in the subspecies indica and ~50 copies in the subspecies japonica (Jiang et al. 2003, Kikuchi et al. 2003). Thus, the available mPing-containing strains are not very useful for designing experiments to understand how MITEs attain very high copy numbers and how they impact host gene expression.
**mPing burst in ‘Gimbozu’**

Nakazaki *et al.* (2003) isolated active *mPing* from a mutant strain IM294, where *mPing* had inserted into an exon of *Rum1* gene. IM294 was generated by gamma-irradiation to a *japonica* rice cultivar Gimbozu, and showed a recessive phenotype of *slender glume* (*slg*) (Teraishi *et al.* 1999). However, the *slg* phenotype was never fixed by selfing over generations, and chimeric plants for glume shape were often observed. Those reversions were finally attributed to accurate excisions of *mPing* element from the *Rum1* gene (Nakazaki *et al.* 2003, Teraishi *et al.* 1999).

A second mutant allele isolated from another irradiated derivative of Gimbozu (HS110) was subsequently shown to contain a *mPing* insertion in an intron of *Hd1*, the rice homolog of the *Arabidopsis* flowering time regulator, CONSTANS (Yano *et al.* 2000).

With these observations, many investigators considered that *mPing* was activated by irradiation stress (Lin *et al.* 2006, Moon *et al.* 2006, Naito *et al.* 2006, Nakazaki *et al.* 2003, Shan *et al.* 2005, Tsugane *et al.* 2006). However, besides the examples of Gimbozu, no other *mPing*-inserted mutations have been observed despite radiation-mutation research had been broadly performed on various modern rice cultivars and strains (Abe *et al.* 2002, Ahloowalia and Maluszynski 2001, Kang *et al.* 2003, Li *et al.* 2003, Mei *et al.* 1994, Monna *et al.* 2002, Nakamura *et al.* 1996, Ueguchi-Tanaka *et al.* 2000). This fact led us to suspect that Gimbozu would be a peculiar cultivar about *mPing* activity.

In addition, it would possibly be not only Gimbozu but its parental line(s) with unusual *mPing* activity (Naito *et al.* 2006). According to the breeding record, Gimbozu was generated from another cultivar Aikoku, by a spontaneous mutation event that had shorten plant height (Naito *et al.* 2006).

Thus, we requested the National Genebank in Japan for all the landraces of Aikoku and Gimbozu to analyze the copy numbers of *mPing* (Naito *et al.* 2006). The result was surprising that the Gimbozu in Kyoto University and the three landraces of Aikoku had accumulated more than 10 times as many copies as Nipponbare (Fig. 2) (Naito *et al.* 2006). Furthermore, we analyzed new *mPing* insertions across three generations and found *mPing* in Gimbozu was still actively transposing and was increasing its copy number by about 20 copies per plant per generation (Naito *et al.* 2006). Thus, *mPing* has already been extremely active even before irradiation, without any particular stresses.

**The behavior of mPing**

Development of pyrosequencer (Margulies *et al.* 2005) has enabled us to sequence all the insertion sites in Gimbozu. To distinguish old insertions (fixed as homozygous insertions in the population) from *de novo* insertions (appearing in a single plant as heterozygous), we sequenced 24 siblings derived from a single Gimbozu plant and identified 928 old and 736 *de novo* insertion sites (Naito *et al.* 2009). The insertion sites of *de novo* insertions should reflect the actual behavior of *mPing* element because the selection had not acted on them. When these insertion sites were mapped on the rice genome, it was shown that *mPing* was enriched in euchromatic, gene rich regions but rare in heterochromatic regions (Fig. 3A) (Naito *et al.* 2009). However, *mPing* was clearly underrepresented within ORFs, especially within CDS (Fig. 3B) (Naito *et al.* 2009). This might be due to high GC content in rice genes (Yu *et al.* 2002), because *mPing* has a target site preference for AT-rich regions (Naito *et al.* 2006, 2009).

While *mPing* avoided inserting into exons, insertions into promoter regions (within 1 kb upstream from transcription start sites) were overrepresented (Fig. 3C, 3D) (Naito *et al.* 2009), probably because of open chromatin structure around promoter regions (Naito *et al.* 2009).

**The impact of mPing**

The subsequent transcriptome analysis revealed co-relation between altered transcription and *mPing* insertions (Naito *et al.* 2009).
As expected, insertions into ORFs down-regulated the transcription, however, those within 5 kb upstream tended to up-regulate (Fig. 4A, 4B) (Naito et al. 2009). Furthermore, mPing was revealed to contain many stress-responsive cis-elements and alter expression profiles of adjacent genes by providing cold- and salt-stress inducibility (Fig. 4C, 4D) (Naito et al. 2009). Thus, mPing has provided evidence for models regarding involvement of TEs in gene regulation (Feschotte 2008). Such hypotheses were first proposed by Britten and Davidson (1969, 1971) that TE dispersal might distribute the same regulatory motif(s) at number of chromosomal locations and draw multiple genes into the same regulatory networks (Feschotte 2008).

However, it should be noted that more than 80% of mPing insertions had no detectable effect on adjacent gene transcriptions (Naito et al. 2009). Thus, MITEs such as mPing had benign effect to the host and this might be one of the reasons why MITEs could attain high-copy numbers in the host genome (Naito et al. 2009).

**Application of mPing**

The great copy number difference of mPing between Gimbozu and other rice cultivars indicated a great potential of this MITE as DNA markers, especially in japonica × japonica population where even SSR polymorphisms are very rare. In addition, a mPing insertion produces 433 bp difference in length, which can be easily resolved by a simple PCR and electrophoresis in agarose gels. Thus we designed primers flanking mPing insertions and developed a marker system with which we constructed a linkage map and successfully detected QTLs for heading date and culm length (Monden et al. 2009).
mPing: The bursting transposon

In addition, the very high activity and the unique insertion preference of mPing can be a valuable source of mutations. Because mPing is already highly active, this system does not need cell culture, irradiation or chemical mutagens. We cultivated 11,520 Gimbozu plants in the paddy field (Yasuda et al. 2013) and identification of approximately 50,000 mPing insertion sites using next-generation sequencer is now underway. mPing was also introduced into soybean and a tagging population was successfully developed (Hancock et al. 2011).

Conclusion

We have caught a transposon in the act of rapid amplification. The case of mPing has demonstrated that populations of rice can survive rapid and massive increases in TE copy numbers within genic regions, because (successful) TEs have evolved target preferences that are largely neutral (Naito et al. 2009). Furthermore, our case indicated that a large subset of the new alleles might actually benefit the host by creating potentially useful allelic variants and novel regulatory networks (Naito et al. 2009). Taken together, mPing amplification can potentially create populations of rice with hundreds of thousands of new alleles (Naito et al. 2009). For rice and other selfing plants, TE bursts may be one of the critical solutions to rapidly generate genetic diversity in the face of an ever-changing environment. Furthermore, as evidence for the rapid and massive amplification of MITEs has been found in virtually all sequenced eukaryotic genomes (Feschotte et al. 2002), features of TE amplification documented for mPing are likely to be widespread in nature.

Acknowledgement

We appreciate Japanese Society of Breeding for awarding Young Scientist Award to Ken Naito. This study was supported by NSF Plant Genome grant, the University of Georgia Research Foundation and JSPS Grant-in-Aid for Scientific Research (B).

Literature Cited

Abe, T., T. Matsuyama, S. Sekido, I. Yamaguchi, S. Yoshida and T. Kameya (2002) Chlorophyll-deficient mutants of rice demonstrated the deletion of a DNA fragment by heavy-ion irradiation. J. Radiat. Res. 43: S157–S161.

Ahloowalia, B.S. and M. Maluszynski (2001) Induced mutations—a new paradigm in plant breeding. Euphytica 118: 167–173.

Britten, R.J. and E.H. Davidson (1969) Gene regulation for higher cells: a theory. Science 165: 349–357.

Britten, R.J. and E.H. Davidson (1971) Repetitive and non-repetitive DNA sequences and a speculation on the origins of evolutionary novelty. Q. Rev. Biol. 46: 111–138.

Bureau, T.E. and S.R. Wessler (1992) Tourist: a large family of small inverted repeat elements frequently associated with maize genes. Plant Cell 10: 1283–1294.

Bureau, T.E. and S.R. Wessler (1994a) Mobile inverted-repeat elements of the Tourist family are associated with the genes of both monocotyledonous and dicotyledonous plants. Proc. Natl. Acad. Sci. USA 91: 907–916.

Craig, N.L., R. Craigie, M. Gellert and A.M. Lamboqitz (2002) Mobile DNA II, ASM Press, Washington DC.

Deininger, P.L. and M.A. Batzer (2002) Mammalian retroelements. Genome Res. 12: 1455–1465.

Doolittle, W.F. and C. Sapienza (1980) Selfish genes, the phenotype paradigm and genome evolution. Nature 284: 601–603.

Feschotte, C. (2008) Transposable elements and the evolution of regulatory networks. Nat. Rev. Genet. 9: 397–405.
Feschotte, C., N. Jiang and S.R. Wessler (2002) Plant transposable elements: where genetics meets genomics. Nat. Rev. Genet. 3: 329–341.

Hancock, C.N., F. Zhang, K. Floyd, A.O. Richardson, P. LaFayette, D. Tucker, S.R. Wessler and W.A. Parrott (2011) The rice miniature inverted repeat transposable element mPing is an effective inser- tional mutagen in soybean. Plant Physiol. 157: 552–562.

International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. Nature 431: 931–945.

Jiang, N., Z. Bao, X. Zhang, H. Hirochika, S.R. Eddy, S.R. McCouch and S.R. Wessler (2003) An active DNA transposon family in rice. Nature 421: 163–167.

Kang, H.J., I.K. Hwang, K.S. Kim and H.C. Choi (2003) Comparative structure and physicochemical properties of ilpumbyeo, a high-quality japonica rice, and its mutant, Suweon 464. J. Agric. Food Chem. 51: 6598–6603.

Kazazian Jr., H.H. (2004) Mobile elements: drivers of genome evolution. Science 303: 1626–1632.

Kikuchi, K., K. Terauchi, M. Wada and H.Y. Hirano (2003) The plant retrotransposon mPing in rice. Plant Cell Physiol. 47: 1147–1153.

Kumar, A. and J.L. Bennetzen (1999) Plant retrotransposons. Annu. Rev. Genet. 33: 479–532.

Labrador, M. and V. Corces (2002) Interactions between transposable elements and the host genome. In: Craig, N.L., R. Craigie, M. Gellert and A.M. Lambowitz (eds.) Mobile DNA II, ASM Press, Washington DC, pp. 1008–1023.

Li, X., Q. Qian, Z. Fu, Y. Wang, G. Xiong, D. Zeng, X. Wang, X. Liu, S. Teng, H. Fujimoto et al. (2003) Control of tillering in rice. Nature 422: 618–621.

Lin, X., L. Long, X. Shan, S. Zhang, S. Shen and B. Liu (2006) In planta mobilization of mPing and its putative autonomous element Pong in rice by hydrostatic pressurization. J. Exp. Bot. 57: 2313–2323.

Margulies, M., E. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bemben, J. Berka, M.S. Braverman, Y.J. Chen, Z. Chen et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437: 376–380.

Melintosh, B. (1950) The origin and behavior of mutable loci in maize. Proc. Natl. Acad. Sci. USA 36: 344–355.

Melintosh, B. (1951) Chromosome organization and genic expression. Cold Spring Harb Symp. Quant. Biol. 16: 13–47.

Melintosh, B. (1984) The significance of responses of the genome to challenge. Science 226: 792–801.

Mei, M., H. Deng, Y. Lu, C. Zhuang, Z. Liu, Q. Qiu, Y. Qiu and T.C. Yang (1994) Mutagenic effects of heavy ion radiation in plants. Adv. Space Res. 14: 363–372.

Monden, Y., K. Naito, Y. Okumoto, H. Saito, N. Oki, T. Tsukiyama, O. Ieda, T. Nakazaki, S.R. Wessler and T. Tanisaka (2009) High potential of a transposon mPing as a marker system in japonica × japonica cross in rice. DNA Res. 16: 131–140.

Monna, L., N. Kitazawa, R. Yoshino, J. Suzuki, H. Masuda, Y. Maehara, M. Tanji, M. Sato, S. Nasu and Y. Minobe (2002) Positional cloning of rice semidwarfing gene, sd-1: rice “green revolution gene” encodes a mutant enzyme involved in gibberellin synthesis. DNA Res. 9: 11–17.

Moon, S., K.H. Jung, D.E. Lee, W.Z. Jiang, H.J. Koh, M.H. Heu, D.S. Lee, H.S. Suh and G. An (2006) Identification of active transposon dTol, a member of the hAT family, in rice. Plant Cell Physiol. 47: 1473–1483.

Naito, K., E.Y. Cho, G. Yang, M.A. Campbell, K. Yano, Y. Okumoto, T. Tanisaka and S.R. Wessler (2006) Dramatic amplification of a rice transposable element during recent domestication. Proc. Natl. Acad. Sci. USA 103: 17620–17625.

Naito, K., F. Zhang, T. Tsukiyama, H. Saito, C.N. Hancock, A.O. Richardson, Y. Okumoto, T. Tanisaka and S.R. Wessler (2009) Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature 461: 1130–1134.

Nakamura, Y., T. Uemoto, Y. Takahara, K. Komae, E. Amano and S. Satoh (1996) Changes in structure of starch and enzyme activities affected by sugary mutations in developing rice endosperm. Possible role of starch debranching enzyme (R-enzyme) in amylopectin biosynthesis. Physiol. Plant. 97: 491–498.

Nakazaki, T., Y. Okumoto, A. Horibata, S. Yamahira, M. Teraishi, H. Nishida, H. Inoue and T. Tanisaka (2003) Mobilization of a transposon in the rice genome. Nature 421: 170–172.

Ogul, E.E. and F.H. Crick (1980) Selfish DNA: the ultimate parasite. Nature 284: 604–607.

SanMiguel, P., B.S. Gaut, A. Tikhonov, Y. Nakajima and J.L. Bennetzen (1998) The paleontology of intergene retrotransposons of maize. Nat. Genet. 20: 43–45.

Shan, X., Z. Liu, Z. Dong, Y. Wang, Y. Chen, X. Lin, L. Long, F. Han, Y. Dong and B. Liu (2005) Mobilization of the active MITE transposons mPing and Pong in rice by introgression from wild rice (Zizania latifolia Griseb.). Mol. Biol. Evol. 22: 976–990.

Teraishi, M., Y. Okumoto, H. Hirochika, A. Horibata, H. Yamagata and T. Tanisaka (1999) Identification of mutable slender glume gene in rice (Oryza sativa L.). Mol. Gen. Genet. 261: 487–494.

Tsugane, K., M. Maekawa, K. Takagi, H. Takahara, Q. Qian, C.H. Eun and S. Iida (2006) An active DNA transposon nDart causing leaf variegation and mutable dwarfishism and its related elements in rice. Plant J. 45: 46–57.

Ueguchi-Tanaka, M., Y. Fujisawa, M. Kobayashi, M. Ashikari, Y. Iwasaki, H. Kitano and M. Matsuoka (2000) Rice dwarf mutant d1, which is defective in the α subunit of the heterotrimeric G protein, affects gibberellin signal transduction (2000) Proc. Natl. Acad. Sci. USA 97: 11638–11643.

van de Lagaemait, L.N., J.R. Landry, D.L. Mager and P. Medstrand (2003) Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. Trends Genet. 19: 530–536.

Wessler, S.R. (2006) Eukaryotic transposable elements: teaching old genomes new tricks. In: Caporale, L. (ed.) The Implicit Genome, Oxford University Press. USA.

Yano, M., Y. Katayose, M. Ashikari, U. Yamanouchi, L. Monna, T. Fuse, T. Baba, K. Yamamoto, Y. Umehara, Y. Nagamura et al. (2000) Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the Arabidopsis flowering time gene CONSTANS. Plant Cell 12: 2473–2484.

Yasuda, K., M. Ito, T. Sugita, T. Tsukiyama, H. Saito, K. Naito, M. Teraishi, T. Tanisaka and Y. Okumoto (2013) Utilization of transposable element mPing as a novel genetic tool for modification of the stress response in rice. Mol. Breed. 32: 505–516.

Yu, J., S. Hu, J. Wang, G.K. Wong, S. Li, B. Liu, Y. Deng, L. Dai, Y. Zhou, X. Zhang et al. (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. indica). Science 296: 79–92.

Zhang, Q., J. Arbuckle and S.R. Wessler (2000) Recent, extensive, and preferential insertion of members of the miniature inverted-repeat transposable element family Heartbreaker into genomic regions of maize. Proc. Natl. Acad. Sci. USA 97: 1160–1165.