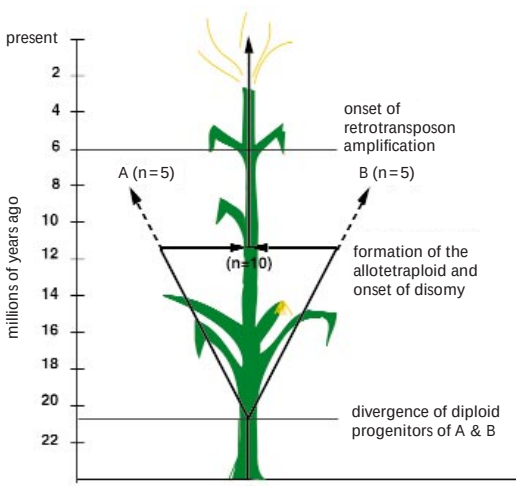
Rapid flux in plant genomes

Daniel F. Voytas & Gavin J.P. Naylor

Department of Zoology & Genetics and the Iowa Computational Biology Laboratory at Iowa State University, Ames, Iowa 50011, USA.
e-mail: voytas@iastate.edu & gnaylor@iastate.edu

The sizes of plant genomes vary remarkably—the well-studied *Arabidopsis* genome is 10^8 base pairs while those of some lilies are over one thousand times this size[1,2]. Genome sizes vary even among closely related plants. The plum genome, for example, is three times larger than that of peaches (both are members of the genus *Prunus*), suggesting that fluctuations in genome size occur over relatively short periods of evolutionary time. While transposon insertion is recognized as a force underlying genome fluidity, the pace at which it contributes to genome evolution has remained obscure. On page 43, Phillip SanMiguel and colleagues document the rapidity by which transposons can restructure genomes; by analysing retrotransposon end-sequences, they have determined that an eruption of transposon activity over the past six million years has led to the plethora of transposons that currently litter the maize genome[3].

The most abundant transposable elements in maize are retrotransposons—mobile genetic elements that replicate by reverse transcription. They proliferate by making copies of themselves; parental elements are transcribed, reverse transcribed, and the resulting progeny are 'seeded' to new genomic sites. Maize retrotransposons have amplified to such an extent that they are densely packed in intergenic regions and, in total, occupy over 50% of the nuclear DNA (ref. 4). To explore the 'population' dynamics of retrotransposon activity, SanMiguel *et al.* took advantage of the fact that the retrotransposon can be used to gauge evolutionary time; it can be thought of as a molecular clock that is set to zero when it integrates into a genome. Like retroviral proviruses, the maize retrotransposons are flanked by two long terminal repeats (LTRs). These repeats are reverse transcribed from an mRNA template that has only one copy of the LTR sequence, so both LTRs are typically identical after DNA synthesis. Time since integration is measured by the extent of nucleotide sequence divergence that has occurred between the two LTRs of a single element. This 'clock' can be calibrated relative to sequence divergence among orthologous genes between species, whose rates of change, in turn, are set by fossil data. By this method, SanMiguel *et al.* have calculated that most maize retrotransposons have integrated within the past 3 million years.



Plotting and ploiding—genome dynamics in the evolutionary history of maize. n=chromosome number

A genome riddled with retrotransposons would seem to pose serious challenges to an organism's well-being. From a mechanistic standpoint, the replication machinery would have to process more DNA. Chromosome pairing may be challenged by the asymmetric accumulation of elements. Finally, and perhaps most importantly, the cell must have an efficient means of recognizing vital coding sequences against a changing background of inserted elements.

How do so many elements get incorporated so quickly, and why are they tolerated? These two questions may be interrelated: tolerance may promote amplification and amplification may, in turn, foster tolerance. Assume, for example, that there is just one target site on a chromosome into which an element can insert. When an element occupies this site, its sequence may, in turn, provide multiple new targets for further insertion. As the percentage of elements increases in the genome, there is an exponential increase in the number of potential sites for integration. This may explain both the rapid amplification and the nested organization of the maize retrotransposons.

Ploidy and proliferation

Plants are known to adapt to other large-scale genome perturbations. Changes in ploidy, for example, are commonplace, and over 50% of all plant species are polyploid or have undergone periods of polyploidy in their evolutionary history[5]. Although modern maize behaves as a strict diploid, there are relics of a past polyploidization event, including large duplicated segments of the genome[6] and a chromosome number that is twice that of many of its relatives[7]. The increased availability of maize genome sequence has

allowed the timing and history of the tetraploid event to be reconstructed—analyses of sequence divergence among multiple loci suggest that the two diploid progenitors of maize diverged approximately 20 million years ago and then united to form the allotetraploid 11 million years ago[8] (see figure).

An obvious outcome of polyploidy is the duplication of gene loci, which may or may not confer a selective advantage to the polyploid. As one might expect, polyploids are not simply the sum of their parts; artificial polyploids of *Brassica* (cabbage and its cousins) and wheat, for example, undergo extensive changes in their nuclear genome within a few generations after formation[9,10]. Although the mechanisms underlying these changes are obscure, transposition is clearly one possibility: quiescent transposable elements in a diploid may be activated in the new genetic environment of the polyploid. Alternatively, genetic redundancy in the polyploid may buffer the potential deleterious effects of transposition. A non-plant example of transposon activation through hybridization has recently been described in wallabies; retrotransposons were unleashed in a hybrid of two wallaby species, resulting in their rapid proliferation[11].

Transposable element spread through hybridization is also seen in cotton[12]. Cultivated cotton is tetraploid and has arisen from a cross between two diploid species—one from Africa/Asia (Old World cotton) and one from the Americas (New World cotton). Chromosomes originating from either donor are distinguishable in the tetraploid, and repeats specific to Old World cottons, including transposable elements, have colonized New World cotton chromosomes in the million or so years since polyploid formation[13]. Interestingly, one diploid New World species carries repeats from the Old World, suggesting that it is either a direct descendent of tetraploid cotton or was spawned from the same polyploidization event that led to the formation of the tetraploid. Co-opting genetic information through wide outcrosses and episodes of polyploidy is a recurrent theme in plant evolution, and only recently have we begun to appreciate how such processes influence transposable element population dynamics.

In maize, there is presently no evidence linking polyploidy and the observed burst in retrotransposition; element amplification occurred at least five million years after formation of the tetraploid. Nonetheless, retrotransposons may have been seeded by the tetraploid event, which have slowly amplified over time—like a modest investment in a mutual fund that has now grown to an appreciable fortune. Bursts of transposition and changes in ploidy attest to the fluidity of plant genomes. Ongoing plant genome projects and studies such as the one described in this issue will shed further insight into the nature of these events and their consequences for plant genome evolution. □

1. Bennett, M. & Smith, J. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **274**, 227–274 (1976).
2. Arumuganathan, K. & Earle, E. *Plant Mol. Biol. Reporter* **9**, 211–215 (1991).
3. SanMiguel, P. *et al. Nature Genet.* **20**, 43–45 (1998).
4. SanMiguel, P. *et al. Science* **274**, 765–768 (1996).
5. Masterson, J. *Science* **264**, 421–424 (1994).
6. Helentjaris, T., Weber, D. & Wright, S. *Genetics* **118**, 353–363 (1988).
7. Anderson, E. *Chron. Bot.* **9**, 88–92 (1945).
8. Gaut, B. & Doebley, J. *Proc. Natl Acad. Sci. USA* **94**, 6809–6814 (1997).
9. Song, K., Lu, P., Tang, K. & Osborn, T. *Proc. Natl Acad Sci. USA* **92**, 7719–7723 (1995).
10. Feldman, M. *et al. Genetics* **147**, 1381–1387 (1997).
11. Waugh O'Neill, R., O'Neill, M. & Marshall Graves, J. *Nature* **393**, 68–72 (1998).
12. Zhao, X. *et al. Genome Res.* **8**, 479–492 (1998).
13. Wendel, J. F. *Proc. Natl Acad. Sci. USA* **86**, 4132–4136 (1989).