15 Bauer, D. *et al*. (2004) Constitutive photomorphogenesis 1 and multiple photoreceptors control degradation of phytochrome interacting factor 3, a transcription factor required for light signaling in *Arabidopsis*. *Plant Cell* 16, 1433–1445
16 Zhu, Y. *et al*. (2000) Phytochrome B binds with greater apparent affinity than phytochrome A to the basic helix–loop–helix factor PIF3 in a reaction requiring the PAS domain of PIF3. *Proc. Natl. Acad. Sci. U. S. A.* 97, 13419–13424
17 Matsushita, T. *et al*. (2003) Dimers of the N-terminal domain of phytochrome B are functional in the nucleus. *Nature* 424, 571–574
18 Kim, J. *et al*. (2003) Functional characterization of phytochrome interacting factor 3 in phytochrome-mediated light signal transduction. *Plant Cell* 15, 2399–2407
19 Fujimori, T. *et al*. (2004) Circadian-controlled basic/helix–loop–helix factor, PIL6, implicated in light-signal transduction in *Arabidopsis thaliana*. *Plant Cell Physiol.* 45, 1078–1086
20 Monte, E. *et al*. (2004) The phytochrome-interacting transcription factor, PIF3, acts early, selectively, and positively in light-induced chloroplast development. *Proc. Natl. Acad. Sci. U. S. A.* 101, 16091–16098
21 Salter, M.G. *et al*. (2003) Gating of the rapid shade-avoidance response by the circadian clock in plants. *Nature* 426, 680–683
22 Park, E. *et al*. (2004) Degradation of phytochrome interacting factor 3 in phytochrome-mediated light signaling. *Plant Cell Physiol.* 45, 968–975
23 Fairchild, C.D. *et al*. (2000) HFR1 encodes an atypical bHLH protein that acts in phytochrome A signal transduction. *Genes Dev.* 14, 2377–2391
24 Soh, M.S. *et al*. (2000) REP1, a basic helix–loop–helix protein, is required for a branch pathway of phytochrome A signaling in *Arabidopsis*. *Plant Cell* 12, 2061–2074
25 Duek, P.D. and Fankhauser, C. (2003) HFR1, a putative bHLH transcription factor, mediates both phytochrome A and cryptochrome signalling. *Plant J.* 34, 827–836
26 Duek, P.D. *et al.* The degradation of HFR1, a putative bHLH class transcription factor involved in light signalling, is regulated by phosphorylation and requires COP1. *Curr. Biol.* (in press)
27 Heisler, M.G. *et al*. (2001) SPATULA, a gene that controls development of carpel margin tissues in *Arabidopsis*, encodes a bHLH protein. *Development* 128, 1089–1098
28 Rajani, S. and Sundaresan, V. (2001) The *Arabidopsis* myc/bHLH gene ALCATRAZ enables cell separation in fruit dehiscence. *Curr. Biol.* 11, 1914–1922

1360-1385/$ - see front matter
doi:10.1016/j.tplants.2004.12.005

Transcriptional networks in plants

Homing into the origin of the AP2 DNA binding domain

Susan R. Wessler

Department of Plant Biology, University of Georgia, Athens, GA 30602, USA

The AP2 DNA binding domain was thought to be plant specific because of its presence in plant, but not animal, transcriptional regulators, particularly members of the AP2/ERF family. Two recent studies have identified the AP2 domain in bacteria, bacteriophage and a ciliate as part of proteins that also encode site-specific endonucleases. The association of AP2 with an enzyme known to catalyze its own movement within populations and between species explains the unusual distribution of AP2 and, as such, adds to a growing list of phenomena where mobile DNA has promoted evolutionary novelty.

Introduction

The cloning of the *Arabidopsis APETALA2* (*AP2*) gene led to the surprising finding that it did not contain a MADS domain like that in previously isolated floral regulators. Instead, the *AP2* protein harbored two copies of a 68 amino acid sequence that came to be known as the AP2 domain [1]. Soon after, this domain was recognized in four tobacco proteins where it was shown to be required to bind the so-called GCC box in the promoters of genes that encode ethylene-inducible pathogenesis-related proteins [2]. Ever since these initial discoveries, the number of AP2-containing proteins has increased dramatically, with almost 150 in the *Arabidopsis* genome alone and hundreds more in other flowering plants. However, until the recent publication of papers by Kathleen Karrer and colleagues [3] and Sarah Hake and colleagues [4], there had been no reports of an AP2 'sighting' outside of the plant kingdom. This contrasts dramatically with most other DNA binding domains (e.g. bHLH, MADS, homeo and zinc fingers), which are found in both plant and animal genomes.

This story begins with the discovery by Wuitschick *et al*. [3] that three members of a family of mobile elements in the ciliate *Tetrahymena thermophila* contain insertions that encode bifunctional proteins with a site-specific (homing) endonuclease domain and an AP2 domain. Like other ciliates, *T. thermophila* has two nuclei: a silent germline micronucleus and a transcriptionally active (somatic) macronucleus. Open reading frames (ORFs) in the micronucleus are interrupted by thousands of transposable elements of different types. In a remarkable series of programmed genomic rearrangements, the development of the macronucleus from a mitotic product of the micronucleus is accompanied by the precise removal of these insertions, called internal eliminated sequences (IESs). As part of their analysis of IESs in *Tetrahymena*,

Corresponding author: Wessler, S.R. (sue@plantbio.uga.edu).
Available online 25 December 2004

Wuitschick *et al*. [3,5] noted that three members of one IES families, called Tlr (*Tetrahymena* long repeat), contain distinct but related genes that encode putative homing endonuclease and AP2 domains.

Shortly after this study was published, a paper from Hake's laboratory [4] reported the results of an extensive computer-assisted search for AP2 domain homologs that are outside the plant kingdom. In addition to identifying the three proteins from *T. thermophila*, they detected AP2 homologs in the cyanobacterium *Trichodesmium erythraeum* and in the viruses *Enterbacteria phage RB49* and *Bacteriophage Felix 01*. Enrico Magnani *et al*. [4] went on to demonstrate that the cyanobacterium AP2 domain binds preferentially to DNA substrates with poly(G)/poly(C) stretches. Binding sites of many plant AP2 domains are also rich in G and C residues, indicating that this domain has been functionally conserved. Whether in ciliates, bacteria or viruses, the AP2 domains in nonplant species are encoded by genes that also encode homing endonucleases.

What are homing endonucleases genes and why are they associated with AP2 domains?

The majority of homing endonucleases genes (HEGs) reside in group-I self-splicing introns where they are sheltered from selection because they do not disrupt the synthesis of the host protein. Homing refers to the high frequency of lateral transfer of the intervening sequence (including the HEG) from the allele with the intron (intron$^+$) to an allele without the intron (intron$^-$) (reviewed in Refs [6,7]). Homing endonucleases promote this transfer by recognizing a relatively long sequence (15 to 30 bp) that is present only in the intron$^-$ gene; they make a double-strand cut, which initiates gene conversion using the intron$^+$ allele as the repair template. In this way, the intron$^-$ allele is converted to an intron$^+$ allele; in some cases, flanking exon sequences are also converted. The presence of HEGs in members of a *T. thermophila* IES family is the first report of HEGs in a ciliate [3]. Like self-splicing introns, IESs disrupt ORFs but are removed during development of the macronucleus from the micronucleus.

Endonuclease:AP2 proteins

Four classes of HEGs have been described and are named after consensus amino acids in their catalytic core. All AP2 domains are associated with one of four classes of HEG, the HNH endonucleases, which are characterized by having separate catalytic and DNA binding domains [8]. Although it has not been demonstrated experimentally, a reasonable possibility is that the AP2 domain binds to the target allele, thus positioning the catalytic domain near its cleavage site. As discussed by Austin Burt and Vassiliki Koufopanou [6], there is selection on homing endonucleases to cleave only the intron$^-$ allele and not another (ectopic) site in the genome. However, HEGs frequently move horizontally to other species that have the target gene but might also have similar ectopic sites that need to be avoided. HEGs can increase their chance of spreading throughout the new host population by refining their DNA binding, by, for example, acquiring a new domain that restricts their activity to the target gene. Enrico Magnani *et al*. [4], demonstrated that the cyanobacterium AP2 domain, like plant AP2 domains, binds preferentially to homopolymeric stretches of poly(G)/poly(C).

Lateral transfer of AP2

Outside the plant kingdom, only a single eukaryote (*T. thermophila*) is known to have an AP2 domain; all others are in bacteria or in their phage. There are several patterns of inheritance that could explain this unusual distribution. It is a formal possibility that AP2 was present in a common ancestor of plants and animals and was inherited vertically but lost from most animal lineages. In the past, one could argue that there were simply not enough animal sequences available to rule out this scenario. This being no longer the case, vertical inheritance can be ruled out. By contrast, several scenarios of horizontal (lateral) transfer might explain the distribution of AP2. Because AP2 is in all plants but sporadic in nonplant species, perhaps the plant AP2 domain moved laterally from plant to nonplant species. This scenario is also unlikely for at least two reasons. First, Wuitschick *et al*. [3] note that the *T. thermophila* AP2 domain has been a resident of the *T. thermophila* genome long enough to look like a ciliate gene: it has high AT content (78% versus 75% genome-wide) and an unusual codon usage (in ciliates, the stop codons TAA, TAG encode glutamine). Second, and more significantly, Magnani *et al*. [4] found that amino acid sequence similarity between one *Arabidopsis* AP2 domain (At4g39780 DREB protein) and the cyanobacterium *T. erythraeum* protein extends beyond AP2 into the HNH (endonuclease) domain. This finding also rules out plant-to-nonplant transfer. We are left with the most likely and most intriguing scenario: lateral transfer of the HNH:AP2 gene from bacteria or ciliates to plants. In support of this scenario, Magnani *et al*. [4] note that the vast majority of AP2/ERF proteins lack introns (122/145) as would be expected for a gene acquired, in part, from prokaryotes.

Much has been written about the proclivity of HEGs to jump to other species (reviewed in Refs [6,7,9]) and lateral gene transfer appears to be a necessary part of their life cycle. Like other transposons, HEGs progress through phases of activity and amplification followed by inactivity and loss. A better case can be made for a requirement for lateral transfer among HEGs than for most other transposons. HEGs are extremely efficient at spreading to a single genetic locus in all members of a population. However, once they become fixed there is no pressure to maintain HEG function because there is nowhere to go except, perhaps, into the same gene in a related species.

To account for the widespread occurrence of the AP2 domain in plant genomes, Magnani *et al*. [4] proposed that the transfer could have occurred during the evolution of chloroplasts from the endosymbiosis of an ancestral cyanobacterium. Subsequent movement of cyanobacterial genes to the plant nuclear genome followed by inactivation and loss of what would probably be a harmful endonuclease activity would account for the movement of the pioneer AP2 domain from prokaryotes to an ancestor of modern plants. Alternatively, the HNH:AP2 gene might

have undergone many lateral transfer events later in plant evolution. I would have considered this to be an extremely unlikely scenario before the discovery by Jeffrey D. Palmer and colleagues that the unusual distribution of a self-splicing group-I intron in the *cox1* gene of plant mitochondria can be explained by its horizontal transfer from a fungal donor [10]. What makes this discovery particularly relevant to the AP2 story is their evidence for not one but >1000 independent HEG-mediated transfers of this intron into the *cox1* gene during the evolution of flowering plants.

The massive transfer of DNA from fungi into plant mitochondria suggests that a similar 'portal' might have been used to shuttle new functions, such as the AP2 binding domain, into ancestral plant genomes. With the accelerated sequencing of diverse representatives from all the kingdoms of life, it is exciting to anticipate the next discovery of a gene that has jumped kingdoms and the mobile element that almost certainly made the jump possible.

References

1 Jofuku, K.D. *et al*. (1994) Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. *Plant Cell* 6, 1211–1225
2 Ohme-Takagi, M. and Shinshi, H. (1995) Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell* 7, 173–182
3 Wuitschick, J.D. *et al*. (2004) Homing endonucleases encoded by germ line-limited genes in *Tetrahymena thermophila* have APETELA2 DNA binding domains. *Eukaryotic Cell* 3, 685–694
4 Magnani, E. *et al*. (2004) From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. *Plant Cell* 16, 2265–2277
5 Wuitschick, J.D. *et al*. (2002) A novel family of mobile genetic elements is limited to the germline genome in *Tetrahymena thermophila*. *Nucleic Acids Res.* 30, 2524–2537
6 Burt, A. and Koufopanou, V. (2004) Homing endonuclease genes: the rise and fall and rise again of a selfish element. *Curr. Opin. Genet. Dev.* 14, 609–615
7 Gimble, F.S. (2000) Invasion of a multitude of genetic niches by mobile endonuclease genes. *FEMS Microbiol. Letts* 185, 99–107
8 Sitbon, E. and Pietrokovski, S. (2003) New types of conserved sequence domains in DNA-binding regions of homing endonucleases. *Trends Biochem. Sci.* 28, 473–477
9 Goddard, M.R. and Burt, A. (1999) Recurrent invasion and extinction of a selfish gene. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13880–13885
10 Cho, Y. *et al*. (1998) Explosive invasion of plant mitochondria by a group I intron. *Proc. Natl. Acad. Sci. U. S. A.* 95, 14244–14249

1360-1385/$ - see front matter
doi:10.1016/j.tplants.2004.12.007