Phylogenetic footprinting: Using the natural variation among related species to identify transcription factor binding sites upstream of developmentally regulated genes

Kenneth Morales and Brian G. Ayre
1 Texas Academy of Math and Science, University of North Texas, Denton TX
2 Department of Biological Sciences, University of North Texas, Denton, TX

Background: A Prelude to Footprinting

Phloemeomics (flō-mē-om'iks): the science of understanding phloem structure and function by cataloguing and characterizing all genes expressed in the phloem.

Identifying phloem specific genes

To identify genes that are specifically expressed in leaf phloem, an enhancer trap genetic screen was carried out.

➢ Tissue-specific gene expression is mediated by transcription factors binding to enhancers in DNA to promote transcription (Figure 1).

➢ In an enhancer trap screen, a promoterless reporter gene is randomly integrated into the genomes of a large population of individuals. In the model plant Arabidopsis thaliana, this random insertion is achieved by Agrobacterium tumefaciens mediated transformation, and the reporter gene encodes the β-glucuronidase, or GUS, enzyme (Figure 2).

➢ The integrated GUS gene is expressed in response to nearby enhancer elements, and may assume the expression pattern of a tissue-specific, flanking gene (Figure 3).

➢ Of 2000 independent enhancer-trap lines, 1% demonstrated GUS expression in leaf veins. A sampling: 64-1F, minor vein expression following the sink-to-source transition; 64-8A, expression in large veins of mature and immature leaves; 59-B4, expression in large veins of mature leaves; 64-12C, expression in all veins of mature and immature leaves.

The enhancer-trap DNA of line 64-1F was inserted between two oppositely orientated genes on chromosome 4: At4g28630, encoding an ATP-binding cassette (ABC) transporter, and At4g28640, an auxin inducible transcription factor (IAA 11). The intergenic sequence was fused to GUS and resulted in phloem-specific gene expression in both orientations (Figure 4). See poster by McGarry, Turgeon, and Ayre for more information.

Phylogenetic Footprinting: The science of identifying potential transcription factor binding sites by comparing promoter sequences from related species, based on the principle that these sites are conserved over evolutionary time to maintain the proper expression of orthologous genes (Figure 5; Ayre et al. 2003).

Future Work: Sequencing orthologous regions from other Brassicaceae family members continues (Figure 7). These sequences will be aligned using several different algorithms. We anticipate that addition of more species to the alignment will increase the resolution of conserved sequences. If sequence conservation in the intergenic region is too high (i.e., insufficient genetic drift at non-essential sequences), DNA from species outside the Brassicaceae will be obtained [other Brassicales (capers), Malvales (cotton), or Sapindales (citrus)].
