in silico approaches to understanding the regulation of Arabidopsis clock genes

Emma Picot[1,2], Peter Krusche[3], Alexander Tiskin[3], Isabelle Carré[2] and Sascha Ott[4].
[1] Systems Biology Doctoral Training Centre, [2] Department of Biological Sciences, [3] Department of Computer Science, [4] Warwick Systems Biology Centre, University of Warwick, Coventry CV4 7AL, UK

The Arabidopsis circadian clock consists of two feedback loops centred around the partially redundant transcription factors LHY and CCA1 (Figure 1). Our current research focuses on known LHY/CCA1 targets. These genes are known to be under the control of LHY/CCA1, and contain at least one LHY/CCA1 binding site (Evening element; AAATATCT).
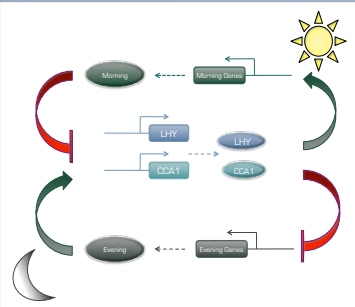
As the expression profiles of all these genes vary greatly, we hypothesise that the effects of LHY/CCA1 are modulated by other proteins. We employ a comparative genomics based approach to uncover conserved regions in orthologous plant promoters based on the assumption that conserved bases are functionally important.

We have developed a novel tool (EARS[1]), based on the Seaweed Algorithm[2] to identify evolutionarily conserved regions within orthologous promoters. In these regions we detect both known regulatory motifs and new potential transcription factor binding sites. We then test for over-representation of these novel sites within clusters of genes with similar temporal expression.

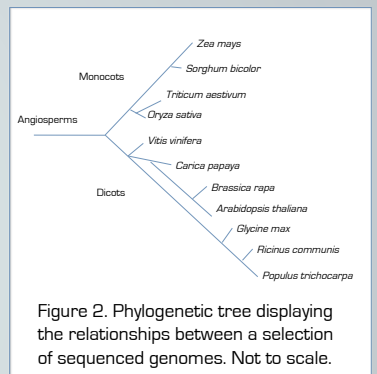Figure 1. A simplified diagram of the Arabidopsis circadian clock

Figure 2. Phylogenetic tree displaying the relationships between a selection of sequenced genomes. Not to scale.

EARS outputs a conservation profile indicating which region of an Arabidopsis promoter is conserved in orthologous promoters. Figures 3-5 show the conservation profiles for *TOC1*, *LUX* and *LHY* promoters respectively. Coloured bars indicate potential transcription factor binding sites.
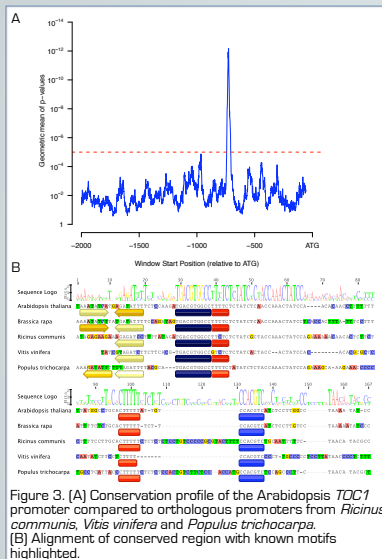

Figure 3. (A) Conservation profile of the Arabidopsis *TOC1* promoter compared to orthologous promoters from *Ricinus communis*, *Vitis vinifera* and *Populus trichocarpa*. (B) Alignment of conserved region with known motifs highlighted.
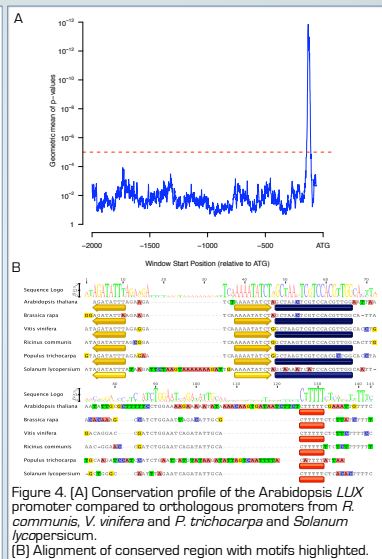

Figure 4. (A) Conservation profile of the Arabidopsis *LUX* promoter compared to orthologous promoters from *R. communis*, *V. vinifera* and *P. trichocarpa* and *Solanum lycopersicum*. (B) Alignment of conserved region with motifs highlighted.
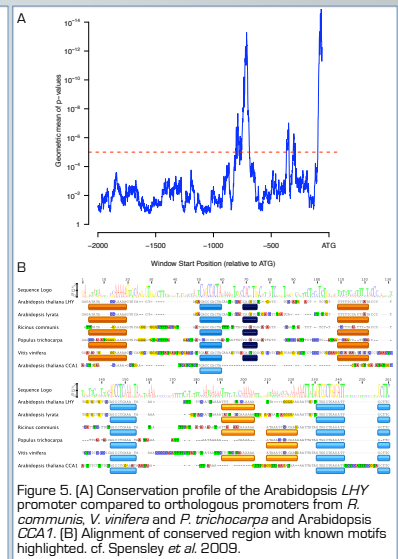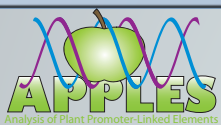

Figure 5. (A) Conservation profile of the Arabidopsis *LHY* promoter compared to orthologous promoters from *R. communis*, *V. vinifera* and *P. trichocarpa* and Arabidopsis *CCA1*. (B) Alignment of conserved region with known motifs highlighted. cf. Spensley *et al.* 2009.

Phase specific gene lists were extracted from Diurnal (http://diurnal.cgrb.oregonstate.edu). Each phase bin was then grouped with the two adjacent bins. Using the APPLES software package currently in development at Warwick Systems Biology Centre, 2000bp promoters were retrieved and then tested for over-representation of time-of-day specific motifs. Figures 6-9 show results for motifs generated in the analysis of EARS results and known matrices.
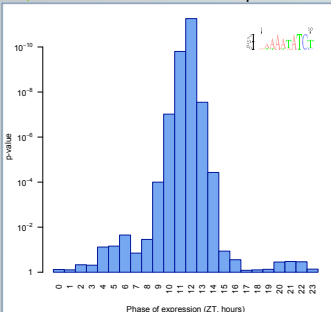

Figure 6. Temporal enrichment of the evening element motif (inset) in promoters of Arabidopsis genes. The evening element is significantly enriched in promoters of genes which peak between ZT8 and ZT14 under constant light.
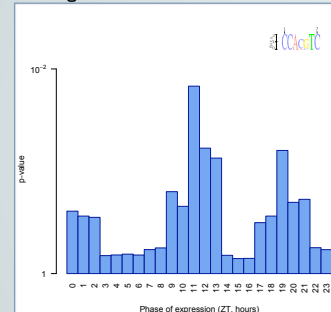

Figure 7. Temporal enrichment of a G-box-like motif (inset) in promoters of Arabidopsis genes. This putative motif is significantly enriched in the promoters of genes which peak at ZT12 under constant light.
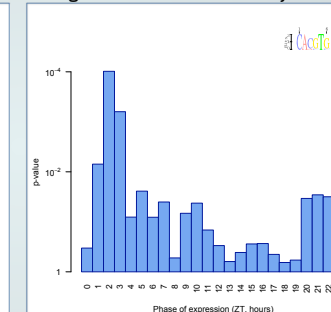

Figure 8. Temporal enrichment of a G-box motif (inset) in promoters of Arabidopsis genes. This motif, a binding site for bHLH proteins, is significantly enriched in the promoters of genes which peak at ZT1 – ZT3 under constant light.
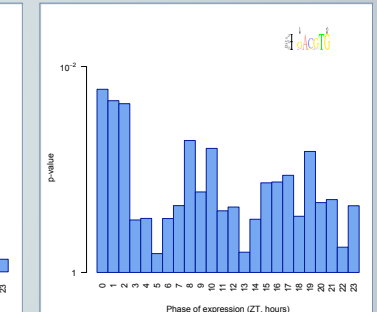

Figure 9. Temporal enrichment of a G-box-like motif (inset) in promoters of Arabidopsis genes. This putative motif is significantly enriched in the promoters of genes which peak at ZT0 – ZT2 under constant light.

References
1. Picot et al. Evolutionary Analysis of Regulatory Sequences (EARS) in Plants [in review].
2. Tiskin (2008). Semi-local String Comparison: Algorithmic Techniques and Applications. Math. Comput. Sci **1**: 571-603.
3. Spensley (2009). Evolutionarily conserved regulatory motifs in the promoter of the *Arabidopsis* clock gene late elongated hypocotyl. Plant Cell **21**: 2606-2623.