hi

## Computational analysis:

testCite (*1*)

RNA-Seq: Mapping and quantification

Raw reads were mapped either with the Tophat pipeline or hisat2+stringtie pipeline due to logistic reasons, and is recorded in the supplementary TPM table. We used "Bdistachyon\_314\_v3.1" from Phytozome (!CITE!PHYTOZOME) as reference genome throughout the study.

**Hisat2+Stringtie** pipeline: Adapters were trimmed off from raw reads with Trimmomatic (!CITE! TRIMMOMATIC) with argument "ILLUMINACLIP:${{FA\_ADAPTER}}:6:30:10 LEADING:3 TRAILING:3 MINLEN:36 SLIDINGWINDOW:4:15". Raw reads were mapped to the transcriptome using HISAT2 (!CITE! HISAT2) with argument:"--no-mixed --rna-strandness RF --dta --fr" . Duplicate reads were removed with Picard (!CITE! PICARD) using default setting. Transcripts were quantified according to this alignment with StringTie (!CITE! STRINGTIE) in TPM values (Transcripts per Million mapped transcripts) with argument "--rf".

**Tophat+Cufflinks** pipeline: Adapters were trimmed off from raw reads with Trimmomatic (!CITE! TRIMMOMATIC) with argument "ILLUMINACLIP:${{FA\_ADAPTER}}:2:10:5:1". Raw reads were mapped to the transcriptome using TOPHAT (!CITE! TOPHAT) with argument:"--no-mixed --max-multihits 1". Duplicate reads were removed with Picard (!CITE! PICARD) using default setting. Transcripts were quantified according to this alignment with cufflinks (!CITE! CUFFLINKS) in TPM values (Transcripts per Million mapped transcripts) with argument "--frag-bias-correct --multi-read-correct".

The TPM values were transformed into log abundances through

( indexes genes while indexes conditions).

Any gene with a maximum log abundance smaller than 3.0 was discarded from downstream analysis to avoid introducing noisy variation. Detailed TPM table can be found in the supplementary. Out of 34310 reference genes, 20586 were kept.

RNA-Seq: Clustering and target calling

To investigate transcriptomic response towards a particular treatment, time-course perturbation matrices were constructed as the difference of log abundances between paired conditions. For example,

Clustering: The selected perturbation matrices were concatenated into a data matrix, against which a von Mises-Fisher mixture of increasing concentration is fitted (similar to !CITE!clusterVMF). Optimal concentration is manually selected to be ~3.080 according to the diagnostic plot and clusters with an average uncertainty higher than 2.5 were considered non-significant. Out of 20586 genes included, 7301 were assigned a significant cluster.

The following perturbation matrices were selected:

* [LD/SD, WT, Wk2]
* [LD/SD, WT, Wk3]
* [SD, elf3-1/WT,Wk2]
* [SD, elf3-1/WT,Wk3]
* [LD, phyC-4/WT,Wk2]
* [LD, ppd1-1/WT,Wk3]

Selection of genes responsive to *elf3-1* and *phyC-4*: A continous timecourse function was interpolated linearly from each of the perturbation matrices [SD, elf3-1/WT,Wk2] and [LD, phyC-4/WT,Wk2], which is then integrated over 24hr to give a temporal average, and . Genes that satisfy 1.0 were considered transcriptionally perturbed.

ChIP-Seq: Mapping and quantification

Adapters were trimmed off from raw reads with Trimmomatic (!CITE!TRIMMOMATIC) (!CITE!trimmomatic) with argument "ILLUMINACLIP:${{FA\_ADAPTER}}:6:30:10 LEADING:3 TRAILING:3 MINLEN:36 SLIDINGWINDOW:4:15". Raw reads were mapped to the genome "Bdistachyon\_314\_v3.1" with Bowtie2 ((*2*)) under argument:"--no-mixed --no-discordant --no-unal -k2". Any read that mapped to more than one genomic location was discarded. Duplicate reads were removed with Picard using default setting (!CITE!PICARD).

Genomic binding profile was quantified in RPKM (Reads Per Kilobase per Million mapped reads) using a bin-size of 10bp. For each bin,

For each treated ChIP-Seq library, peaks were called against a control using MACS2 (!CITE!MACS2) with argument "--keep-dup 1 -p 0.1". Peaks from all ChIP-Seqs were fitlered for FC>2.0.

ChIP-Seq: Defining regions differentially bound by ELF3

Differential occupancy: Peaks from all ELF3 ChIP-Seqs were filtered and merged into contiguous regions. Abundances were computed for 100bp windows within each merged region, and log-transformed with

( denotes the window, denotes the 10bp bin, indexes condition) Differential occupancy was quantified for each window by taking its dot-product with a signature vector to get , where the signature vector is computed from a set of marker windows so that

Cross-condition variance is also defined for each window . Any window with > 1.5 and > 0.5 is called differentially bound.

Any peak with a differentially bound window is considered a differentially-bound peak. Any gene with a differential bound window within +/- 6000bp of its start codon is considered a differentially bound gene. Pile-up of ELF3 occupancy: was performed on the set of peaks that contains at least 1 differentially bound window.

ChIP-Seq: Peak-level overlap

Both peak lists of ELF3 and of phyC were filtered for "FC>3.0". Each ELF3 peak was considered overlapped if it’s within 1000bp of a phyC peak, and vice versa. The Venn diagram is constructed by specifying the number of non-overlapping ELF3 and phyC peaks, and the number of intersection peaks as . See Dataset S5 for details.

Availability and external packages

Code is available from this Github repo: https://github.com/shouldsee/BrachyPhoto, where dependencies were detailed.

RNA-Seq data are available from: {{GEO\_RNA}}

ChIP-Seq data are available from: {{GEO\_CHIP}}

GNU-parallel (!CITE!GNUPARA) was used in paralleling the computational analysis.

Binding profiles near marker genes were visualised with *fluff* (!CITE!FLUFF), and visualised with Integrated Genomic Browser during investigation (!CITE!IGV).

Bedtools (!CITE!BEDTOOLS) was used for intersection of peaks, making genomic windows.

1. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. **30**, 2114–2120 (2014).

2. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nature methods*. **9**, 357–359 (2012).