# STARR-seq for high-throughput identification of plant enhancers

Ling Zhang,[1] Wai-Shing Yung,[2,*] and Mingkun Huang [1,2,*]

[1]Lushan Botanical Garden Jiangxi Province and Chinese Academy of Sciences, No. 9 Zhiqing Road, Jiujiang, Jiangxi, P.R. China
[2]School of Life Sciences and Center for Soybean Research of the State Key Laboratory of Agrobiotechnology, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, P.R. China

Trends in Plant Science
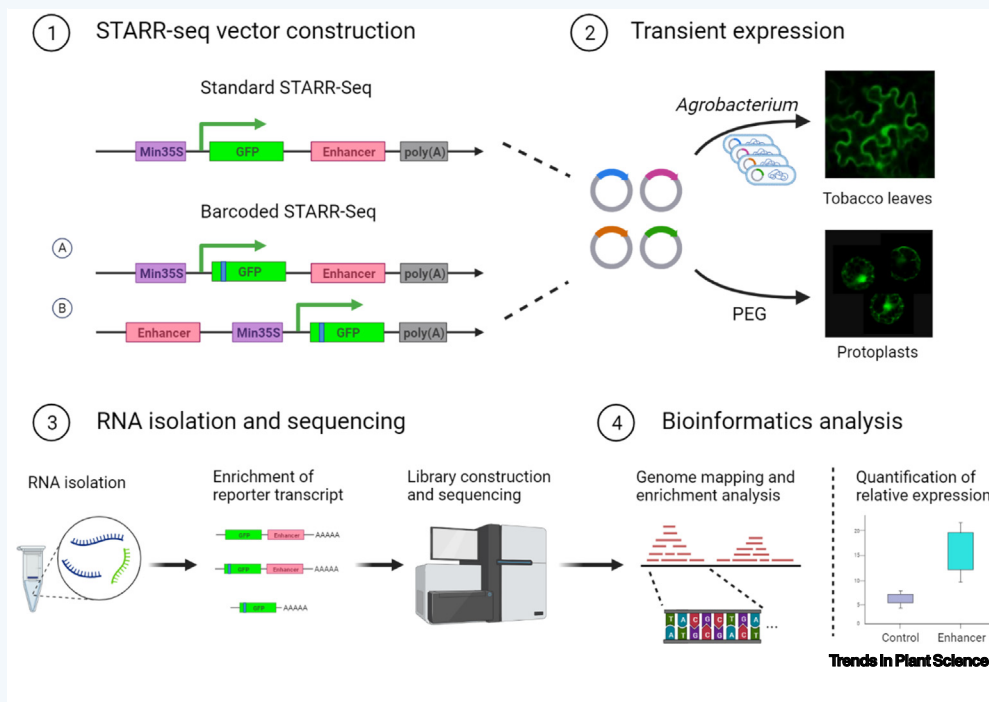
ADVANTAGES:
Based on the well-designed vector coupled with the transient transformation system, STARR-seq is an efficient method which allows parallel identification and analysis of enhancers in different plant species.

Enhancer activity in different environmental conditions (e.g., light regime) can be assessed through the STARR-seq method.

Integrative analyses using STARR-seq and other technologies such as assay for transposase-accessible chromatin using sequencing (ATAC-seq) and ChIP-seq enable verification of functional enhancers. This approach can also elucidate the relationships between chromatin modification and enhancer elements.
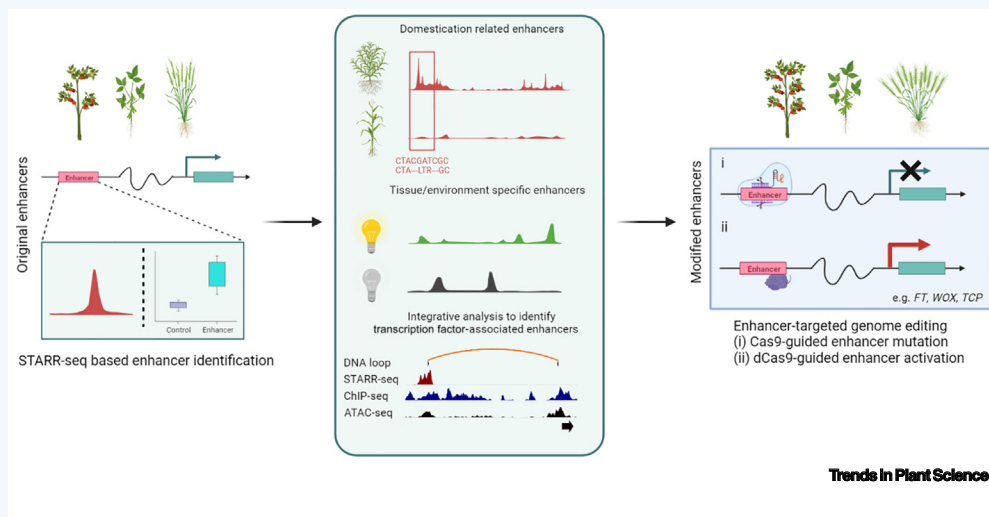
CHALLENGES:
The method for enrichment of the GFP mRNA with poly(A) tail through the PCR-based strategy needs to be improved in order to distinguish other types of mRNA during the sequencing library construction.

The influences of native chromatin structure on enhancer activities cannot be reflected in STARR-seq.

Different host systems may favor different vector systems or species origins of candidate fragments.

High levels of repetitive elements in plant genomes may lead to a low signal-to-noise ratio during enrichment analysis. Bioinformatics analysis methods and pipelines need to be optimized for plant studies. What is currently lacking is a software designed specifically for processing barcoded STARR-seq data.

Self-transcribing active regulatory region sequencing (STARR-seq) is a convenient method for genome-wide identification of plant enhancers. The basic principle is that an enhancer is able to activate the reporter gene expression (e.g., *GFP*) in a well-designed transient expression vector. The constructed vectors, which harbor the enhancer candidates, are transiently transformed into tobacco leaves or various protoplasts. The resulting reporter transcripts are enriched for sequencing. In the original design, the locations and activities of the enhancers are predicted by mapping to the reference genome and enrichment analysis. Recent research in plants showed the robustness of adding barcode sequences to link transcripts to the candidates placed upstream. The enhancer activities can, thus, be evaluated as the relative reporter expression level driven by the candidate versus the negative control.

Trends in Plant Science

https://doi.org/10.1016/j.tplants.2022.08.008

CellPress

Check for updates

Integrative analysis of STARR-seq and other sequencing data is an effective strategy for identification of essential plant enhancers, which might serve as candidates for genome-editing systems to modify the downstream gene expression for crop improvement.

Declaration of interests

No interests are declared.

Literature

1. Arnold, C.D. *et al.* (2013) Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* 339, 1074–1077
2. Muerdter, F. *et al.* (2018) Resolving systematic errors in widely used enhancer activity assays in human cells. *Nat. Methods* 15, 141–149
3. Liu, Y. *et al.* (2017) Functional assessment of human enhancer activities using whole-genome STARR-sequencing. *Genome Biol.* 18, 219
4. Liu, S. *et al.* (2017) Systematic identification of regulatory variants associated with cancer risk. *Genome Biol.* 18, 194
5. Dao, L.T.M. *et al.* (2017) Genome-wide characterization of mammalian promoters with distal enhancer functions. *Nat. Genet.* 49, 1073–1081
6. Vanhille, L. *et al.* (2015) High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq. *Nat. Commun.* 6, 6905
7. Inoue, F. and Ahituv, N. (2015) Decoding enhancers using massively parallel reporter assays. *Genomics* 106, 159–164
8. Ricci, W.A. *et al.* (2019) Widespread long-range cis-regulatory elements in the maize genome. *Nat. Plants* 5, 1237–1249
9. Sun, J. *et al.* (2019) Global quantitative mapping of enhancers in rice by STARR-seq. *Genomics Proteomics Bioinforma.* 17, 140–153
10. Jores, T. *et al.* (2020) Identification of plant enhancers and their constituent elements by STARR-seq in tobacco leaves. *Plant Cell* 32, 2120–2131
11. Jores, T. *et al.* (2021) Synthetic promoter designs enabled by a comprehensive analysis of plant core promoters. *Nat. Plants* 7, 842–855
12. Niu, L. *et al.* (2020) Resolving a systematic error in STARR-seq for quantitative enhancer activity mapping. *bioRxiv* Published online October 21, 2020. https://doi.org/10.1101/2020.10.20.346908