Gene editing of *Arabidopsis thaliana* cytosolic/nuclear subclass of HSP70 family using CRISPR/Cas

Andreja Škiljaica

Division of Molecular Biology, Department of Biology, Faculty of Science, University of Zagreb, Croatia

- In the last decade, the **CRISPR/Cas technology** has evolved into a highly specific tool for generating plants with site-specific mutations, free of unwanted mutations scattered across the genome.
- The main objective of this STSM was a **theoretical introduction** to CRISPR/Cas gene editing in *Arabidopsis thaliana.*
- Virtual reagents and protocols were prepared for introduction of CRISPR/Cas9-induced **knockout mutations** of single (***HSP70-4***) and multiple ***HSP70*** genes.

CRISPR/Cas GENE EDITING PROTOCOL

Phylogenetic analysis of the HSP70 family

HSP70s have two major functional domains, an **ATPase domain** in the N-terminal part of the protein (~ 40 kDa) and a **peptide-binding domain** in the C-terminal part of the protein (~ 25 kDa). Both domains are vital for protein function, therefore, the sgRNA sequences were designed to bind within either of these domains.

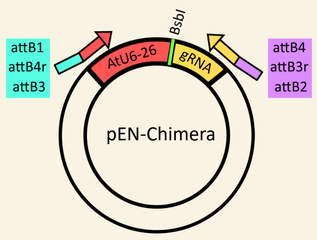
To obtain full-length sequences of *Arabidopsis thaliana HSP70* family, I performed a BLAST(n) search of the TAIR database using the cDNA sequence of ***HSP70-4*** (AT3G12580.1) as query. Fifteen protein-coding sequences with highest similarity were pooled and their CDS were retrieved from the TAIR database. Multiple sequence alignments were obtained and maximum likelihood phylogeny was inferred. The alignments and the phylogenetic tree were the basis for selection of regions suitable for design of sgRNAs targeting multiple genes of the *HSP70* subclass.

single guide RNA (sgRNA) design

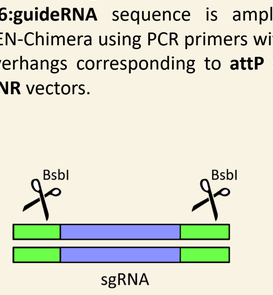
The ***HSP70-4*** CDS is shown below with two exons highlighted in dark and light blue. Three different **sgRNAs** were designed for introduction of a knockout mutation within the *HSP70-4* gene (white). All sgRNAs are compatible with **Cas9** protein from *Streptococcus pyogenes,* which recognizes the PAM sequence **NGG**, where N is any nucleotide base (pink). sgRNAs with highest specificity and efficiency score were selected. The same protocol was used for generation of sgRNAs targeting multiple members of the *HSP70* family (not shown).

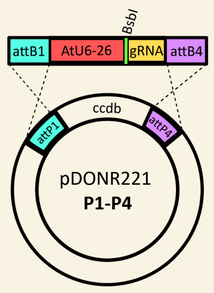
```
ATGGCGGGTAAAGGTGAAGGTCCAGCTATCGGTATCGATCTGGGTACAACCTACTCTTGCGTCGGTGTTTC
GCAACATGACCGGCGTCGAAATCATCGCCAACGATCAAGGCAACCGCACCACTCCTTCCTACGTCGCTTTCA
CTGACGAGCGACGCTCTCATCGGGGATGCTGCCAAGAATCAAGTCGCCATGAACCCTACCAACACCGTCTTC
GATGCTAAGCGTCTAATCGGAAGAAGATACAGTGATCCCTCTGTTCAAGCGGATAAGAGTCACTGGCCTTT
TAAGGTTGTTTCCGGTCCAGGTGAGAAGCCTATGATTGTGGTTAACCACAAGGGAGAGGAGAAACAGTTCT
CTGCTGAGGAAATCTCGTCGATGGTTCTTATTAAGATGCGGGAGATTGCAGAAGCTTTCCTTGGTTCTCCT
GTTAAGAACGCTGTCGTTACAGTTCCTGCTTATTTCAACGACTCTCAGGTCAAGCAGACTAAGGAGCGCTGG
AGTTATCTCTGGTCTCAACGTGATGCGTATCATCAATGAGCCAACTGCTGCTGCTATTGCTTACGGTCTTG
ACAAGAAGGCGTCGAGTGTTGGCGAGAAGAATGTTTTGATCTTTGATTTGGGAGGTGGTACTTTTGATGTG
TCTTTGCTTACGATTGAGGAAGGTATCTTTGAAGTCAAGGCAACTGCTGGTGACACGCATCTTGGTGGTGA
GGACTTCGACAACAGGATGGTTAATCATTTCGTTCAGGAGTTTAAGAGGAAGAACAAGGATATTACTG
GGAACCCGAGAGCTTTGAGGAGGCTTAGGACAGCTTGTGAGCGGGCGAAGAGAACTCTTTCTTCGACTGCT
CAGACGACTATAGAGATTGACTCTCTTTTTGAGGGTATTGATTTCTACACTACCATCACTCGTGCTAGGTT
CGAGGAGCTCAACATGGATTTGTTTAGGAAGTGTATGGAGCCAGTGGAGAAGTGTTTGAGGGATGCTAAGA
TGGACAAGAGCAGTGTCATGATGTTGTTCTTGTTGGTGGCTCTACAAGGATTCCCAAAGTGCAGCAGCTT
TTGCAGACTTCTTCAATGGGAAAGAGCTCTGTAAAGCATTAACCCGACCAGCTGTTGCTTACGGAGC
AGCTGTGCAAGCTGCAATCTTGAGCGGTGAAGGGAATGAGAAGGTCCAGGACTTGTTACTTCTTGATGTCA
CTCCTCTGTCCTTGGGTTTGGAAACTGCCGGTGGTGTTATGACTGTTTTGATTCCGAGGAACACCACAATT
CCGACCAAGAAAGAGCAGATATTCTCTACCTATTCAGACAACCAGCCCGGTGTACTGATCCAGGTCTACGA
AGGAGAGAGGGCACGAACAAAGACACAACACTTTTGGGAAAGTTCGAGCTCAGTGGATAACACCTCGCTC
CACGAGGTGTACCGCAGATTACTGTCTGTTTCGACATCGACGCCAATGGTATCCTGAATGTGTCGGCTGAG
GACAAGACGACTGGTCAGAAGAACAAGATCACAATCACAAACGACAAGGGAAGGTTATCAAAGGAAGAGAT
CGAGAAGATGGTACAAGAGGCAGAGAAGTACAAGGCTGAGGATGAAGAACACAAGAAGAAGGTGGATGCAA
AGAACGCTCTCGAGAACTATGCATACAACATGAGGAACACGATCAAGGACGAGAAGATCGCATCTAAGCTT
GACGCAGCTGACAAGAAGAAGATTGAGGATGCAATCGACCAAGCTATTGAATGGTTAGATGGTAATCAACT
GGCTGAGGCAGATGAGTTCGAGGATAAGATGAAGGAGCTCGAGTCTCTTTGCAACCCTATTATTGCAAGAA
TGTACCAAGGAGCTGGGCCTGATATGGGTGGTGCAGGAGGAATGGATGACGACACACCTGCTGGTGGTAGC
GGCGGTGCTGGCCCAAAGATTGAAGAAGTTGATTAA
```

Cloning into vectors - Gateway cloning

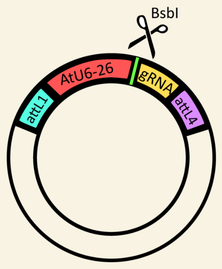


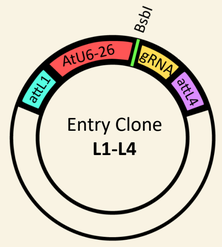
The three **sgRNAs** first need to be cloned into a suitable transcription context. Entry vector **pEN-Chimera** contains the **AtU6-26** promoter and a **guideRNA** sequence. The **AtU6-26:guideRNA** sequence is amplified from pEN-Chimera using PCR primers with 5' **attB** overhangs corresponding to **attP** sites in **pDONR** vectors.



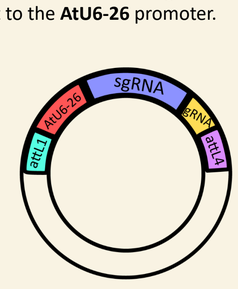


Three **BP Cloning** reactions occur between a **pDONR** vector with **attP** sites and an **attB-flanked PCR product** containing **AtU6-26:guideRNA** insert.

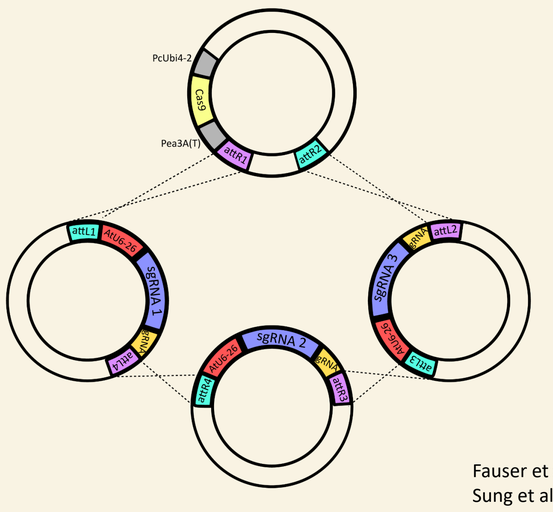




The three BP Cloning reactions will yield three **entry clones** with the **AtU6-26:guideRNA** sequence flanked by **attL** and/or **attR** sites. Between the **AtU6-26** sequence and the **guideRNA** sequence is a **BbsI** restriction site, which enables cloning of **sgRNAs** immediately adjacent to the **AtU6-26** promoter.



Each of the three **sgRNA sequences** were used to design a pair of forward and reverse primers. When annealed, the primers form double stranded **sgRNA oligos** with **BbsI** 5' overhangs**,** compatible to overhangs of previously described **entry clones** cut by **BbsI.**



Restriction and ligation reactions of **entry clones** and **sgRNA oligos** will generate three plasmid constructs with the **AtU6-26:sgRNA:guideRNA** sequence, ready for use in the **LR Cloning** reaction. Shown above is a schematic image of the first entry clone.

Using **Multisite Gateway Cloning**, the **pDe-Cas9** destination vector and the three **entry clones** carrying **sgRNAs** can be combined in a single LR Cloning reaction during which homologous recombination will occur between all four vectors to finally generate a single **Cas9 expression vector carrying all three sgRNA sequences**, each with its own promoter (not shown).

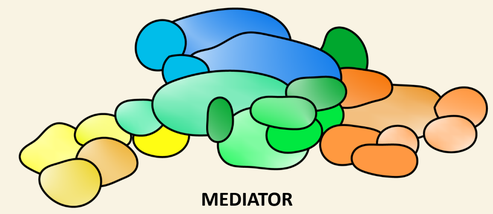
This vector can be subsequently used for stable ***Arabidopsis thaliana*** transformation employing the *Agrobacterium*-mediated transformation protocol, followed by transgenic plant selection and verification.

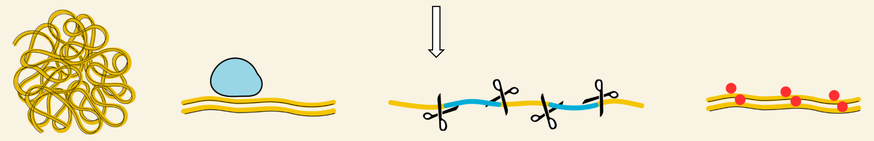
Fauser et al (2014) PLANT J 79(2): 348 - 359
Sung et al (2001) PHYSIOL PLANT 113: 443 - 451

LITERATURE BACKGROUND

In eukaryotes, the highly conserved 70-kDa heat shock protein (HSP70) family performs protein folding, chaperoning and protection from adverse effects of stress, such as preventing aggregation and refolding of denatured proteins.

Members of the cytosolic/nuclear subclass of HSP70 family, such as HSP70-4, were also shown to be involved in the Mediator complex, a multiprotein complex involved in processes such as transcription regulation at the level of chromatin architecture, RNA pol II assembly, RNA processing and epigenetic regulation.





More recently, other members of the HSP70 family have been shown to form protein chaperone complexes along with members of other HSP families, with a probable function in regulation of DNA methylation.

Association of heat-shock proteins in regulation of gene expression at the level of epigenetic regulation opens up novel pathways of studying the different and highly interconnected modes of action through which plants adapt to environmental changes.

Buendia-Monreal et al (2016) DEV BIOL 419 (1): 7 – 18
Feng et al (2021) J INTEGR PLANT BIOL 63 (8): 1451 - 1461

STSM OUTCOMES

- **research**

This project resulted in assessment of the cytosolic/nuclear subclass of HSP70 family and generation of CRISPR/Cas gene editing protocols for obtaining highly specific *A. thaliana* mutants. In the future, these mutant plants will be used to study the interplay of **heat stress** and changes at the level of **gene transcription** and **epigenetic regulation**, with a hope of better understanding how plants might cope with adverse effects of climate change in years to come.

- **idea sharing**

The laboratory for Plant Molecular Biology of Zagreb's Faculty of Science does not routinely employ the CRISPR/Cas technology as a genome-editing method. This project was one of the first steps toward the advancement and modernization of current methodology within the laboratory, but also within the broader network of Croatia's plant molecular biologists.

- **collaboration**

This STSM was the first collaboration between the Molecular Biology department of Zagreb's Faculty of Science and the Plant Breeding department of Wageningen University & Research, with both sides open to additional collaborations in the future.

ACKNOWLEDGEMENTS

This STSM was carried out online in collaboration with Dr. **Martina Juranić** of Wageningen University & Research, The Netherlands, under COST Action CA18111 *Genome editing in plants - a technology with transformative potential*.

Funded by the Horizon 2020 Framework Programme of the European Union