Programmable Chromosome Engineering (PCE) system: Designing pegRNAs for Various Large-Scale DNA Manipulations

Given the complexity of designing pegRNAs for the PCE and RePCE systems, we developed a web-based tool, **PCE Targets Designer v1.0** (http://www.engineeringchromosome.net/), to streamline and simplify the design of target sites for diverse chromosome editing applications. These applications include integration, deletion, inversion, replacement, translocation, and scarless editing. This tool aims to provide researchers with a convenient and efficient platform for utilizing our systems.

When designing target sites, the tool considers several critical factors:

- The PAM types associated with the target site,
- The distance between cleavage sites for the dual-pegRNA strategy,
- The secondary structure formed between the target site sequence and the RT sequence, and
- The potential off-target effects of the target site sequence.

The single pegRNA designer part of **PCE Targets Designer** was developed by modifying CRISPR-GE (Xie *et al.*, 2017). The output includes a user-friendly list of target sites based on the input genomic DNA sequences, ranked according to the factors mentioned above. For subsequent pegRNA primer design, users can refer to the the protocol of prime editing (Jin *et al.*, 2023) and the **PlantPegDesigner** tool (http://www.plantgenomeediting.net/) (Lin *et al.*, 2021).

PCEDesign_multiple.py (Design pegRNA for integration, deletion, inversion, replace ment and translocation manipulations)

Usage: python PCEDesign_multiple.py ./example/test_PCE.fa

Input file: ./example/test_PCE.fa

Output file: ./result_PCE/results/*.TwinPair.txt

PCEDesignScarless_multiple.py (Design pegRNA for removing scar)

Usage: python PCEDesignScarless_multiple.py ./example/test_scarless.fa

Input file: ./example/test_scarless.fa

Output files: ./ result_scarless/results/*.TwinPair.txt

Supplementary Note 1. The format of the input sequence

The input sequence should be the reference sequence of rice genome with length of 100-400 bp, while with 100-200 bp for designing target sites for scarless editing. The sequence format could be FASTA or not, and uppercase or lowercase (eg, >rice gene 1

GATTGAGATATGGCAGTTAATTTTCCCAGATCTGAGTACTTTTTTTCGCA
TGTTTAGTTGCTGAATTATGCTGTTTGCTACTGCATATGGCTTTGATTTTCGTG
TGGTGTGCTCACGATTGATGAATTCTATGTTGGTCGTGTGATTTGCAGACCTA
CATGTTCAAGTACGATACCGTGCACGGCCAATGGAAGCACAGCGACATCAAG
ATCAAGGACTCCAAGACTCTGCTCTTGGGCGAGAAGCCGGTCACCGTTTTCG
GCATCAGGTAACTTGATATTGATATTAC or
GATTGAGATATGGCAGTTAATTTTCCCAGATCTGAGTACTTTTTTTCGCATGTT
TAGTTGCTGAATTATGCTGTTTGCTACTGCATATGGCTTTGATTTTCGTGTGGT
GTGCTCACGATTGATGAATTCTATGTTGGTCGTGTGATTTGCAGACCTACATG
TTCAAGTACGATACCGTGCACGGCCAATGGAAGCACAGCGACATCAAGATCA
AGGACTCCAAGACTCTGCTCTTGGGCGAGAAGCCGGTCACCGTTTTCGGCAT
CAGGTAACTTGATATTGATATTAC or

>rice gene 1

gattgagatatggcagttaattttcccagatctgagtacttttttttcgcatgtttagttgctgaattatgctgtttgctactgcatat
ggctttgattttcgtgtggtgtgctcacgattgatgaattctatgttggtcgtgtgatttgcagacctacatgttcaagtacgataccgt
gcacggccaatggaagcacagcgacatcaagatcaaggactccaagactctgctcttgggcgagaagccggtcaccgttttcg
gcatcaggtaacttgatattgatattac or
gattgagatatggcagttaattttcccagatctgagtacttttttttcgcatgtttagttgctgaattatgctgtttgctactgcatatggctt
tgattttcgtgtggtgtgctcacgattgatgaattctatgttggtcgtgtgatttgcagacctacatgttcaagtacgataccgtgcacg
gccaatggaagcacagcgacatcaagatcaaggactccaagactctgctcttgggcgagaagccggtcaccgttttcggcatc
aggtaacttgatattgatattac).

Supplementary Note 2. The selection of the output target sites

We simply sorted the output results based on the cutting distance between two pegRNAs, and the results also listed the GC ratio and off-target score. From the perspective of efficiency, we generally recommend that researchers choose NGG PAMs with a cleavage distance of 30-50 bp. Users can also choose targets with relatively low efficiency but high specificity according to their own needs.

Supplementary Note 3. Primer design and vector construction

After completing target site selection, we added notes indicating the RT sequence of the pegRNA used in each module (integration, deletion, inversion, replacement, and translocation) below the page. Users can design and construct vectors and primers based

on this sequence. As for the scarless editing module, researchers can design the desired sequence for replacing the Lox site sequence according to their needs. This sequence can be the original genome sequence or other sequence desired.

References

Xie, X., Ma, X., Zhu, Q., Zeng, D., Li, G., Liu, YG., (2017). CRISPR-GE: A Convenient Software Toolkit for CRISPR-Based Genome Editing. Mol. Plant 10, 1246–1249.

Lin, Q., Jin, S., Zong, Y., Yu, H., Zhu, Z., Liu, G., Kou, L., Wang, Y., Qiu, J. L., Li, J., & Gao, C. (2021). High-efficiency prime editing with optimized, paired pegRNAs in plants. Nature biotechnology, 39, 923–927.

Jin, S., Lin, Q., Gao, Q., & Gao, C. (2023). Optimized prime editing in monocot plants using PlantPegDesigner and engineered plant prime editors (ePPEs). Nature protocols, 18, 831–853.