1

Oct 06, 2021

Encoding Probe Design using PaintSHOP, SOP006.v1.1 V.1

Nicole Houchins[1], Rory Kruithoff[1], Douglas Shepherd[1]

[1]Arizona State University

1

protocol .

Human Cell Atlas Method Development Community

Rory Kruithoff
Arizona State University

Document Summary: This document, SOP006 – Encoding Probe Design using PaintSHOP, describes how to construct RNA-targeting, DNA-based encoding probes with the open-source, interactive application known as PaintSHOP. Using this website, a desired target refseq ID (DNA or RNA), or gene name, can be input to generate all of the possible encoding probe target sequences based on your selected criteria and desired stringency. Next, primers and bridge sequences can be appended to the target sequences. These can be pre-determined and input as manual files, or you can choose to use literature-established sets using the drop-down menu. Once sequences are appended, you can download files containing the encoding probe list, order file, reference file and appending file.

Quick Overview:

Part 1 - Generating encoding probe target sequences

Step 1 – Determine gene/genes of interest
Step 2 – Compile your gene RefSeq IDs
Step 3 - Generate encoding probe sequences using the RNA Probe Design tab

Part 2 - Appending primers and readout sequences

Step 1 - Put sequences that will be appended into individual text files
Step 2 - Append sequences from text files using the Append Sequences tab

Part 3: Save completed probe files

Step 1 - Check desired settings for which file will be obtained using the Download tab and save files.

Encoding Probe Design using PaintSHOP.pdf

Nicole Houchins, Rory Kruithoff, Douglas Shepherd 2021. Encoding Probe Design using PaintSHOP, SOP006.v1.1. **protocols.io** https://protocols.io/view/encoding-probe-design-using-paintshop-sop006-v1-1-byc8pszw

PaintSHOP, encoding probe, oligo library, design, bridge, sequence, append, codebook, probe

protocol ,

Sep 19, 2021

Oct 06, 2021

Sep 19, 2021 Julia Rossmanith protocols.io

Sep 28, 2021 Rory Kruithoff Arizona State University

53376

v1.1 revision notes
1. Updated document summary
2. General update and editing of the document

Adapted from: Hershberg et al, 2021; written by Nicole Houchins and Rory Kruithoff, QI2 Laboratory.

References:

Hershberg, E. A., Camplisson, C. K., Close, J. L., Attar, S., Chern, R., Liu, Y., ... & Beliveau, B. J. (2021). PaintSHOP enables the interactive design of transcriptome-and genome-scale oligonucleotide FISH experiments. Nature Methods, 1-8. https://doi.org/10.1038/s41592-021-01187-3

Moffitt, J. R., Hao, J., Bambah-Mukku, D., Lu, T., Dulac, C., & Zhuang, X. (2016). High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing. Proceedings of the National Academy of Sciences, 113(50), 14456-14461. https://doi.org/10.1073/pnas.1617699113

Required Equipment:
This protocol only requires a computer with working internet access.

For hazard information and safety warnings, please refer to the SDS (Safety Data Sheet).

Part 1 - Generating encoding probe sequences - Step 1: Determine gene/genes of interest

1 Determine a gene or list of genes of interest for your specific experimental design. Gene selection may require consideration of expression levels depending on the sample size and distribution of the expression of the selected genes. For example, a lowly expressed gene (<0.01 FPKM) may be expressed so infrequently that it may take several hundred cells to see even the occasional FISH spot. The sensitivity of this method does have limits. Moffitt et al describes FPKM values <0.05 to have substantially reduced correlation with RNAseq data. To reduce uncertainty, we recommend expression values greater than 1 FPKM and, ideally, values >10 FPKM.

Part 1 - Generating encoding probe sequences - Step 2: Compile your gene RefSeq IDs

2 On the UCSC genome browser, https://genome.ucsc.edu/ (or similar), choose the desired species and assembly.

3 Enter the gene name of your gene of interest into the Position/Search Term text field.

4 Use the RefSeqID from the splice variant that most accurately captures the gene for your purposes. The Refseq number is formatted as NM_000XXXX and is also known as an accession number. There may be alternate isoforms of the gene and picking an appropriate isoform accession number is not always simple. You may need to look closely at the exon coverage to select the most comprehensive coverage for all isoforms. There are also curated list references and other resources to help with this selection at https://genome.ucsc.edu/FAQ/FAQgenes.html#whatdo.

> **Note: You can also choose to use an isoform-flattened gene model as a means to target all potential isoforms of the gene.

5 Repeat this process for all of the gene targets to compile a list of corresponding RefSeq IDs for your study.

Part 1 - Generating encoding probe sequences - Step 3: Generate encoding probe target sequences using the RNA Probe Design tab

6 Generate your PaintSHOP input files (skip to Part 2 for more information). Because PaintSHOP's server will time out in 30 minutes, you may find it more useful to establish files with your appending sequences (Part 2) prior to generating your encoding probe target

sequences. This way, your appending files can be quickly applied to the target sequences in a single session. Below is an example list of the plain text files you should generate for entering to PaintSHOP.

7 Pull up the PaintSHOP website, and open the 'RNA Probe Design' tab.

8 Choose the probe set depending on the desired species, then enter all of the RefSeq IDs separated by a comma and a space (e.g. NM_020309, NM_015267, NM_018008, …) into the "Enter RefSeq IDs manually" text field. A file can also be uploaded instead of manually entering the RefSeq IDs. In this text file, the RefSeq IDs should each be placed on their own line and you should then choose 'custom' rather than 'manual' entry

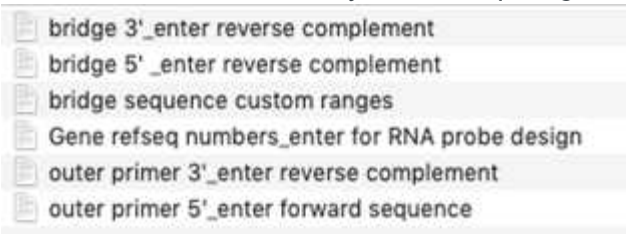
9 Choose the appropriate settings under the Advanced Probe Settings to adjust for off target score and K-mer count.

10 Should you desire the same number of probes per gene of interest, use the 'Balance Set' section, select the number of probes desired and select the 'balance' option.

11 Once complete, select 'submit' to generate the results. This will provide you a list of the target sequences as well as a summary for review.

Part 2 - Appending Primers and Readout Sequences - Step 1: Put sequences that will be appended into individual text files

12 For a typical encoding probe set, you will need a set of plain text files (example below), which include the bridge sequences or primers to be used on either the 5' or the 3' end of the encoding probe. Additionally, you should include a list of the ranges for your bridge sequences as well as a file with a list of your RefSeq IDs generated from Part 1 above.



13 Put each of the sequences (e.g. primers, readout sequences, etc.) that will be appended into text files. If multiple sequences will be appended for one section of the probe (e.g. multiple readout sequences for the 5' and 3' bridge sequences), separate each sequence by line in the text file. The sequences that go into these files are the exact sequences of the primers or readouts, etc. In PaintSHOP, you will append the correct sequence for the encoding probes by

entering the file as either the forward orientation or reverse complement. To ensure this is easy to double-check, you should include the input method on the file name as seen in the example above.

Part 2 - Appending Primers and Readout Sequences - Step 2: Append sequences from text files using the Append Sequences tab

14 With the Appending files above and Part 1 of this SOP complete, open the Append Sequences tab and select the type of probe to be used. Choose the RNA Probe Design option under the Design Scheme section for probes designed to visualize RNA.

15 Click 'Append' for each of the sections that will have sequences appended to them to reveal more options. Append the sequence files with the appropriate settings. A few notes on appending:

15.1 Note that the 5' and 3' Outer Primer are used for our PCR primers, and the 5' and 3' Bridge Sequences are used for readout sequences.

15.2 Within these revealed sections, choose the appropriate settings. These settings include orientation, format, custom ranges (optional), number per target (optional), select sequence set, and upload custom set (optional).

15.3 Generally, primers will all be the same for all probes while bridge sequences will be ordered according to the 'bridge sequence custom ranges' file.

15.4 For the '3' Appending Scheme' section specifically, select 'Primer/Bridge/Universal' option to reveal more options regarding sections on the 3' side of the probe. Once selected, you can choose to append a 3' primer and bridge sequence.

16 Once all of the sequences have been appended, double-check all of the entries and selections and then click 'Append' at the bottom of the page.

Part 3 - Save completed probe files - Step 1: Check desired settings for which kind of file will be obtained in the Download tab and save files.

17 Open the Download tab.

18 Choose the appropriate settings (Appended sequences and RNA Probe Design) regarding the completed file that will be generated. Such settings include used appending, design scheme, and file type (order, full probe, appending, or citation). Save the following files.

a. Appending File
b. Order File
c. Full Probe File
d. Reference File (optional)