KLab Tiling Program_Fast

The efficacy of the KLab interspecies tiling program was demonstrated by interrogating the endogenous and human protamine loci in a transgenic mouse line (1). Species specific probe sets were designed by adapting a previously reported maskless array synthesis strategy (2).

User defined parameters:

- Length: Select probe length (bp). This setting may be automatically relaxed to allow for Tm considerations.
- Stp: Set step size between probe start sites.
- Na: Salt concentration of hybridization conditions. Be mindful of hybridization buffers.
- Tm: Acceptable range is restricted to 60 - 80 C.
- Simple: default calculations require this box to be checked.
- GC clamp : Use default. Changes the degree of local GC content accepted in order to limit intramolecular bonds (GC clamp).
- Palindrome : Use default. Relaxes or restricts the amount of palindromic sequence allowed within a probe subsequence.
- Keep the 'fast' box checked as the default setting for all analyses.

Installation

DesignTilingFast.exe (Windows executable) requires formatted genome library to examine probe sequences. Presently, libraries for Human (build 36) and Mouse (build 36) are available. Zip files (HS-36-1.zip, MM-36-1.zip) containing desired libraries should be downloaded and extracted to same folder where DesignTilingFast.exe resides.

Instructions

"Check probes"

This feature allows for the evaluation of already existing probe sets against the above parameters. Designate the column containing the probe sequence in the tab-delimited input txt file in "F" box. A prompt will direct the user to select the input file. This is followed by a prompt to choose the name and save location of the output file. Detection of a probe which does not

meet the specified criteria (Tm, GC, etc) terminates analysis of the probe set. The list of high quality probes preceding the low quality probe are saved in the output file.

"Create Probes"

Ensure that the Uniqueness/Complete (slow) box is checked. Check desired parameters w/ respect to species specificity. “P and unique” requires probe sequence to be present and unique within the specified genome(s). Checking "A" requires that the resulting probes sequences be absent from the specified genome(s). In example, for probes targeting mouse sequence check P and unique in Mm and A in Hs. Check “unique in supplied sequence”. This ensures that probes are unique within the input file before comparing the sequence to genome/chromosome files. Select input file. All input must be in FASTA format saved as txt files. There should be no headers or carriage returns. Select output name and location. Once running check output file to ensure the program is running. All possible probe sequences from the supplied input file are listed regardless of their quality.

The tested probe sequence is followed by the local GC/hairpin score. Probes that pass are listed as "OK" and probes that fail are listed as "Bad". The higher the score associated with this value the greater the length of the hairpin sequence. In example "Bad:4" signifies that the probe was rejected because 4 bases are predicted to form a hairpin structure.

The Tm score indicates the predicted Tm (in Celsius) and whether the probe was rejected or accepted on this parameter.

The complete probe sequences which pass the hairpin and Tm criteria are checked for their uniqueness within the input file and within the human and/or mouse genomes.

Complete probe sequences found to be unique within the input and genome sequences are then assigned a quality score. 14 bp subsequences of the probe using a single bp step are checked for their uniqueness against the human and mouse genomes to assign a probe quality score.

Probes which fulfill all criteria are listed as "Good" and those that fail for any reason are listed as "Failed"

1. Johnson G.D., P. A. E., Lalancette C., Goodrich R., Krawetz S.A. (*In Press*) *Systems Biology in Reproductive Medicine*
2. Graf, S., Nielsen, F. G., Kurtz, S., Huynen, M. A., Birney, E., Stunnenberg, H., and Flicek, P. (2007; open soucre code is available at http://www.ebi.ac.uk/) *Bioinformatics* **23**, i195-204