Package 'ProbeDeveloper'

October 12, 2022

Type Package

Title Develop Hybridization Probes

Version 1.1.0

Date 2022-01-30

Author Junhui Li

Maintainer Junhui Li <junhuili@cau.edu.cn>

Description Hybridization probes for target sequences can be made based on melting temperature value calculated by R package 'TmCalculator' <https://CRAN.R-project.org/package=TmCalculator> and methods extended from Beliveau, B. J.,(2018) <doi:10.1073/pnas.1714530115>, and those hybridization probes can be used to capture specific target regions in fluorescence in situ hybridization and next generation sequence experiments.

License GPL (>= 2)

Imports TmCalculator (>= 1.0.2),Biostrings(>= 2.12.0)

Depends R (>= 2.10)

RoxygenNote 7.1.2

NeedsCompilation no

Repository CRAN

Date/Publication 2022-01-31 00:10:02 UTC

R topics documented:

ProbeMake *Make probes*

Description

Probes are made with a FASTA-formatted input file containing the target sequence. User can specify the allowable ranges of probe length, percent GC content, and adjust melting temperature calculated using nearest neighbor thermodynamics or empirical formulas based on GC content. Candidate probe sequences passing all checks output in BED format.

Usage

```
ProbeMake(
  fafile,
  LN = 90,
  ln = 60,
  TM = 80,
  tm = 60,
  CG = 70,
  cg = 30,
  gap = 0,
  method = c("S2L", "L2S"),
  direction = c("3to5", "5to3"),
  prohibitseq = NULL,
  TmMethod = c("Tm_GC", "Tm_NN"),
  variant = c("Primer3Plus", "Chester1993", "QuikChange", "Schildkraut1965",
   "Wetmur1991_MELTING", "Wetmur1991_RNA", "Wetmur1991_RNA/DNA", "vonAhsen2001"),
  nn_table = c("DNA_NN4", "DNA_NN1", "DNA_NN2", "DNA_NN3", "RNA_NN1", "RNA_NN2",
    "RNA_NN3", "R_DNA_NN1"),
  tmm_table = "DNA_TMM1",
  imm_table = "DNA_IMM1",
  de_table = c("DNA_DE1", "RNA_DE1"),
  dnac1 = 25,
  dnac2 = 25,
  Na = 0,
  K = 0,
  Tris = 0,
  Mg = 0,
  dNTPs = 0,
  saltcorr = c("Schildkraut2010", "Wetmur1991", "SantaLucia1996", "SantaLucia1998-1",
    "SantaLucia1998-2", "Owczarzy2004", "Owczarzy2008"),
  DMSO = 0,
  fmd = 0,
  DMSOfactor = 0.75,
  fmdfactor = 0.65,
  fmdmethod = c("concentration", "molar")
)
```

Arguments

fafile	Input file with a FASTA format read by function readDNAStringSet in R package 'Biostrings'
LN	The maximum allowed probe length, default is 90
ln	The minimum allowed probe length, default is 60
TM	The maximum allowed melting temperature, default is 80
tm	The minimum allowed melting temperature, default is 60
CG	The maximum allowed percent GC content, default is 70
cg	The minimum allowed percent GC content, default is 30
gap	The minimum gap between adjacent probes, default is 0
method	'S2L' is used to design probe extending from minimal length to the maximum until passing all checks, conversely 'L2S' make probe from maximal length to the minimum. Default is 'S2L'
direction	Design probes from 3 to 5 end or from 5 to 3 end of target sequence, default is '3to5'
prohibitseq	Prohibited sequence list, e.g prohibitseq=c("GGGGG","CCCCC"), default is NULL
TmMethod	The method used to calculate Tm, 'Tm_NN' and 'Tm_GC' can be seleted
variant	Empirical constants coefficient with 8 variant for 'Tm_GC' method: Chester1993, QuikChange, Schildkraut1965, Wetmur1991_MELTING, Wetmur1991_RNA, Wetmur1991_RNA/DNA, Primer3Plus and vonAhsen2001
nn_table	Thermodynamic NN values, eight tables are implemented. For DNA/DNA hybridizations: DNA_NN1,DNA_NN2,DNA_NN3,DNA_NN4 For RNA/RNA hybridizations: RNA_NN1,RNA_NN2,RNA_NN3 For RNA/DNA hybridizations: R_DNA_NN1
tmm_table	Thermodynamic values for terminal mismatches. Default: DNA_TMM1
imm_table	Thermodynamic values for internal mismatches, may include insosine mismatches. Default: DNA_IMM1
de_table	Thermodynamic values for dangling ends: DNA_DE1(default),RNA_DE1
dnac1	Concentration of the higher concentrated strand [nM]. Typically this will be the primer (for PCR) or the probe. Default: 25
dnac2	Concentration of the lower concentrated strand [nM]. Default: 25
Na	Millimolar concentration of Na, default is 0
K	Millimolar concentration of K, default is 0
Tris	Millimolar concentration of Tris, default is 0
Mg	Millimolar concentration of Mg, default is 0
dNTPs	Millimolar concentration of dNTPs, default is 0
saltcorr	Salt correction method. Options are "Schildkraut2010", "Wetmur1991","SantaLucia1996","SantaLucia19 1","Owczarzy2004","Owczarzy2008". Note that "SantaLucia1998-2" is not available for this function.

DMSO	Percent of DMSO
fmd	Formamide concentration in percentage (fmdmethod="concentration") or molar (fmdmethod="molar")
DMSOfactor	Coeffecient of Tm decreases per percent DMSO. Default=0.75 von Ahsen N (2001) <PMID:11673362>. Other published values are 0.5, 0.6 and 0.675.
fmdfactor	Coeffecient of Tm decrease per percent formamide. Default=0.65. Several papers report factors between 0.6 and 0.72.
fmdmethod	"concentration" method for formamide concentration in percentage and "molar" for formamide concentration in molar

Value

Returns a bed file in the format TargetID <tab> Chr <tab> Start <tab> End <tab> Sequence <tab> Tm <tab> GC

Author(s)

Junhui Li

References

Beliveau B J, Kishi J Y, Nir G, et al. (2017). OligoMiner: A rapid, flexible environment for the design of genome-scale oligonucleotide in situ hybridization probes. bioRxiv.

Breslauer K J , Frank R , Blocker H , et al. Predicting DNA duplex stability from the base sequence.[J]. Proceedings of the National Academy of Sciences, 1986, 83(11):3746-3750.

Sugimoto N , Nakano S , Yoneyama M , et al. Improved Thermodynamic Parameters and Helix Initiation Factor to Predict Stability of DNA Duplexes[J]. Nucleic Acids Research, 1996, 24(22):4501-5.

Allawi, H. Thermodynamics of internal C.T mismatches in DNA[J]. Nucleic Acids Research, 1998, 26(11):2694-2701.

Hicks L D , Santalucia J . The thermodynamics of DNA structural motifs.[J]. Annual Review of Biophysics & Biomolecular Structure, 2004, 33(1):415-440.

Freier S M , Kierzek R , Jaeger J A , et al. Improved free-energy parameters for predictions of RNA duplex stability.[J]. Proceedings of the National Academy of Sciences, 1986, 83(24):9373-9377.

Xia T , Santalucia , J , Burkard M E , et al. Thermodynamic Parameters for an Expanded Nearest-Neighbor Model for Formation of RNA Duplexes with Watson-Crick Base Pairs,[J]. Biochemistry, 1998, 37(42):14719-14735.

Chen J L , Dishler A L , Kennedy S D , et al. Testing the Nearest Neighbor Model for Canonical RNA Base Pairs: Revision of GU Parameters[J]. Biochemistry, 2012, 51(16):3508-3522.

Bommarito S, Peyret N, Jr S L. Thermodynamic parameters for DNA sequences with dangling ends[J]. Nucleic Acids Research, 2000, 28(9):1929-1934.

Turner D H , Mathews D H . NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure[J]. Nucleic Acids Research, 2010, 38(Database issue):D280-D282.

Sugimoto N , Nakano S I , Katoh M , et al. Thermodynamic Parameters To Predict Stability of RNA/DNA Hybrid Duplexes[J]. Biochemistry, 1995, 34(35):11211-11216.

Allawi H, SantaLucia J: Thermodynamics and NMR of internal G-T mismatches in DNA. Biochemistry 1997, 36:10581-10594.

Santalucia N E W J . Nearest-neighbor thermodynamics of deoxyinosine pairs in DNA duplexes[J]. Nucleic Acids Research, 2005, 33(19):6258-67.

Peyret N , Seneviratne P A , Allawi H T , et al. Nearest-Neighbor Thermodynamics and NMR of DNA Sequences with Internal A-A, C-C, G-G, and T-T Mismatches, [J]. Biochemistry, 1999, 38(12):3468-3477.

Marmur J , Doty P . Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature.[J]. Journal of Molecular Biology, 1962, 5(1):109-118.

Schildkraut C . Dependence of the melting temperature of DNA on salt concentration[J]. Biopolymers, 2010, 3(2):195-208.

Wetmur J G . DNA Probes: Applications of the Principles of Nucleic Acid Hybridization[J]. CRC Critical Reviews in Biochemistry, 1991, 26(3-4):33.

Untergasser A , Cutcutache I , Koressaar T , et al. Primer3–new capabilities and interfaces[J]. Nucleic Acids Research, 2012, 40(15):e115-e115.

von Ahsen N, Wittwer CT, Schutz E , et al. Oligonucleotide melting temperatures under PCR conditions: deoxynucleotide Triphosphate and Dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. Clin Chem 2001, 47:1956-1961.

Examples

```
data(samplefa)
ProbeMake(samplefa,LN=90,ln=60,TM=80,tm=70,CG=80,cg=20,TmMethod="Tm_NN",Na=50)
```

`samplefa` *sample data for target sequence region with class 'DNAStringSet'*

Description

sample data read by function readDNAStringSet in R package 'Biostrings' from fasta format file, there are two target sequence resion in this data

Usage

```
data("samplefa")
```

Format

Formal class 'DNAStringSet' [package "Biostrings"] with 5 slots

sample data read by function readDNAStringSet in R package 'Biostrings' from fasta format file, which is from ncbiRefSeq database for Homo Sapiens with referece genome version hg19

Examples

```
data(samplefa)
```

Index