Computation of melting temperature of nucleic acid duplexes with `rmelting`

Aravind, J.[1] *and Krishna, G. K.*[2]

2018-12-24

1. Division of Germplasm Conservation, ICAR-National Bureau of Plant Genetic Resources, New Delhi.

2. Division of Crop Physiology, ICAR-Indian Agricultural Research Institute, New Delhi.

Contents

1 Introduction

The `R` package `rmelting` is an interface to the MELTING 5 program (Le Novère, 2001; Dumousseau et al., 2012) to compute melting temperatures of nucleic acid duplexes (DNA/DNA, DNA/RNA, RNA/RNA or 2'-O-MeRNA/RNA) along with other thermodynamic parameters such as hybridisation enthalpy and entropy.

Melting temperatures are computed by Nearest-neighbour methods for short sequences or approximative estimation formulae for long sequences. Apart from these, several corrections are availabel to take into account the presence of Cations (Na, Tris, K and Mg) or denaturing agents (DMSO and formamide).

1.1 Installation

The package can be installed using the following functions:

```
# Install development version from Github
devtools::install_github("aravind-j/rmelting")
```

Then the package can be loaded using the function

```
library(rmelting)
```

2 Basic usage

Melting temperatures are computed in `rmelting` throught the core function `melting` which takes a number of arguments (see `?melting`). The following are the essential arguments which are mandatory for computation.

- `sequence`
  - 5' to 3' sequence of one strand of the nucleic acid duplex as a character string. Recognises A, C, G, T, U, I, X_C, X_T, A*, AL, TL, GL and CL (**Table 1**). U and T are not considered identical.

Table 1: Recognized sequences

Code	Type
A	Adenine
C	Cytosine
G	Guanine
T	Thymine
U	Uracil
I	Inosine
X_C	Trans azobenzenes
X_T	Cis azobenzenes
A*	Hydroxyadenine
AL	Locked nucleic acid
TL	”
GL	”
CL	”

- `Comp.sequence`
  - Mandatory if there are mismatches, inosine(s) or hydroxyadenine(s) between the two strands. If not specified, it is computed as the complement of `sequence`. Self-complementarity in `sequence` is detected even though there may be (are) dangling end(s) and `comp.sequence` is computed.
- `nucleic.acid.conc`
  - In molar concentration (M or mol L^{-1}).
- `Na.conc, Mg.conc, Tris.conc, K.conc`
  - At least one cation (Na, Mg, Tris, K) concentration is mandatory, the other agents(dNTP, DMSO, formamide) are optional.
- `hybridisation.type`
  - The possible options for hybridisation type are as follows (**Table 2**).

Table 2: Hybridisation type options

Option	Sequence	Complementary sequence
`dnadna`	DNA	DNA
`rnarna`	RNA	RNA
`dnarna`	DNA	RNA
`rnadna`	RNA	DNA
`mrnarna`	2-o-methyl RNA	RNA
`rnamrna`	RNA	2-o-methyl RNA

With these arguments, the melting temperature can be computed as follows.

```
melting(sequence = "CAGTGAGACAGCAATGGTCG", nucleic.acid.conc = 2e-06,
        hybridisation.type = "dnadna", Na.conc = 1)
```

```
## [1] 73.35168
```

Only the melting temperature is given as a console output. However, the output can be assigned to an object which contains the details of the environment, Options and the thermodynamics results as a list.

```
out <- melting(sequence = "CAGTGAGACAGCAATGGTCG", nucleic.acid.conc = 2e-06,
               hybridisation.type = "dnadna", Na.conc = 1)
# Environment output
out$Environment
```

```
## $Sequence
## [1] "CAGTGAGACAGCAATGGTCG"
##
## $`Complementary sequence`
## [1] "GTCACTCTGTCGTTACCAGC"
##
## $`Nucleic acid concentration (M)`
## [1] 2e-06
##
## $`Hybridization type`
## [1] "dnadna"
##
## $`Na concentration (M)`
## [1] 1
##
## $`Mg concentration (M)`
## [1] 0
##
## $`Tris concentration (M)`
## [1] 0
##
## $`K concentration (M)`
## [1] 0
##
## $`dNTP concentration (M)`
## [1] 0
##
## $`DMSO concentration (%)`
## [1] 0
##
```

```
## $`Formamide concentration (M or %)`
## [1] 0
##
## $`Self complementarity`
## [1] FALSE
##
## $`Correction factor`
## [1] 4
```

```
# Options used
out$Options
```

```
## $`Approximative formula`
## [1] NA
##
## $`Nearest neighbour model`
## [1] "all97"
##
## $`GU model`
## [1] NA
##
## $`Single mismatch model`
## [1] "allsanpey"
##
## $`Tandem mismatch model`
## [1] "allsanpey"
##
## $`Single dangling end model`
## [1] "bom00"
##
## $`Double dangling end model`
## [1] "sugdna02"
##
## $`Long dangling end model`
## [1] "sugdna02"
##
## $`Long dangling end model`
## [1] "sugdna02"
##
## $`Internal loop model`
## [1] "san04"
##
## $`Single bulge loop model`
## [1] "tan04"
##
## $`Long bulge loop model`
## [1] "san04"
##
## $`CNG repeats model`
## [1] NA
##
## $`Inosine bases model`
## [1] "san05"
##
## $`Hydroxyadenine bases model`
```

```
## [1] "sug01"
##
## $`Azobenzenes model`
## [1] "asa05"
##
## $`Locked nucleic acids model`
## [1] "mct04"
##
## $`Ion correction method`
## [1] NA
##
## $`Na equivalence correction method`
## [1] "ahs01"
##
## $`DMSO correction method`
## [1] "ahs01"
##
## $`Formamide correction method`
## [1] "bla96"
##
## $Mode
## [1] "def"
```

```
# Thermodynamics results
out$Results
```

```
## $`Enthalpy (cal)`
## [1] -159000
##
## $`Entropy (cal)`
## [1] -430
##
## $`Enthalpy (J)`
## [1] -664620
##
## $`Entropy (J)`
## [1] -1797.4
##
## $`Melting temperature (C)`
## [1] 73.35168
```

```
# Command for MELTING 5
attributes(out)$command
```

```
## [1] "-S CAGTGAGACAGCAATGGTCG -H dnadna -P 2e-06 -E Na=1 -T 60"
```

3 Melting temperature computation

Melting temperature is computed by either approximative or nearest neighbour methods according to the length of the oligonucleotide sequences. For longer sequences (longer than the threshold value, the threshold value set by `size.threshold` with the default value 60) approximative method is used, while for others, nearest neighbour method is used.

3.1 Approximative methods

The approximative method for computation can be specified by the argument `am.method`. The available methods are given in **Table 3**.

Table 3: Details of approximative methods

Formula	Type	Limits/Remarks	Reference
`ahs01`	DNA	No mismatch	Ahsen et al. (2001)
`che93`	DNA	No mismatch; Na=0, Mg=0.0015, Tris=0.01, K=0.05	Marmur and Doty (1962)
`che93corr`	DNA	No mismatch; Na=0, Mg=0.0015, Tris=0.01, K=0.05	Marmur and Doty (1962)
`marschdot`	DNA	No mismatch	Wetmur (1991), Marmur and Doty (1962), Chester and Marshak (1993), Schildkraut and Lifson (1965), Wahl et al. (1987), Britten et al. (1974), Hall et al. (1980)
`owe69`	DNA	No mismatch	Owen et al. (1969), Frank-Kamenetskii (1971), Blake (1996), Blake and Delcourt (1998)
`san98`	DNA	No mismatch	SantaLucia (1998), Ahsen et al. (2001)
`wetdna91`*	DNA		Wetmur (1991)
`wetrna91`*	RNA		Wetmur (1991)
`wetdnarna91`*	DNA/RNA		Wetmur (1991)

3.2 Nearest neighbour methods

Table 4: Details of nearest neighbour methods

Model	Type	Limits/Remarks	Reference
`all97`*	DNA		Allawi and SantaLucia (1997)
`bre86`	DNA		Breslauer et al. (1986)
`san04`	DNA		SantaLucia and Hicks (2004)
`san96`	DNA		SantaLucia et al. (1996)
`sug96`	DNA		Sugimoto et al. (1996)
`tan04`	DNA		Tanaka et al. (2004)
`fre86`	RNA		Freier et al. (1986)
`xia98`*	RNA		Xia et al. (1998)
`sug95`*	DNA/ RNA		SantaLucia et al. (1996)
`tur06`*	2'-O-MeRNA/ RNA	A sodium correction (`san04`) is automatically applied to convert the entropy (Na = 0.1M) into the entropy (Na = 1M).	Kierzek et al. (2006)

```
melting(sequence = "CAGTGAGACAGCAATGGTCG", nucleic.acid.conc = 2e-06,
        hybridisation.type = "dnadna", Na.conc = 1, method.nn = "bre86")
```

```
## [1] 83.2203
```

```
melting(sequence = "CAGTGAGACAGCAATGGTCG", nucleic.acid.conc = 2e-06,
        hybridisation.type = "dnadna", Na.conc = 1, method.nn = "san04")
```

```
## [1] 73.30191
```

```
melting(sequence = "CAGTGAGACAGCAATGGTCG", nucleic.acid.conc = 2e-06,
        hybridisation.type = "dnadna", Na.conc = 1, method.nn = "san96")
```

```
## [1] 75.7102
```

```
melting(sequence = "CAGTGAGACAGCAATGGTCG", nucleic.acid.conc = 2e-06,
        hybridisation.type = "dnadna", Na.conc = 1, method.nn = "sug96")
```

```
## [1] 78.17556
```

4 Corrections

4.1 Nucleic acid concentration

4.2 Ion Concentration

4.2.1 Na

4.2.2 Mg

4.2.3 Na Mg

4.2.4 NaEq

4.3 Denaturing agent concentration

4.3.1 DMSO

4.4 Formamide

4.5 Batch

4.6 Citing rmelting

```
##
## To cite the R package 'rmelting' in publications use:
##
##   Aravind, J. and Krishna, G. K. (2018).  rmelting: R Interface to
##   MELTING 5. R package version 0.99.1,
##   https://aravind-j.github.io/rmelting/.
##
## A BibTeX entry for LaTeX users is
##
##   @Manual{,
##     title = {rmelting: R Interface to MELTING 5},
##     author = {J. Aravind and G. K. Krishna},
##     year = {2018},
##     note = {R package version 0.99.1},
##     note = {https://aravind-j.github.io/rmelting/},
```

```
##   }
##
## This free and open-source software implements academic research by
## the authors and co-workers. If you use it, please support the
## project by citing the package.
```

4.7 Session Info

```
sessionInfo()
```

```
## R Under development (unstable) (2018-10-27 r75507)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows >= 8 x64 (build 9200)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_India.1252  LC_CTYPE=English_India.1252
## [3] LC_MONETARY=English_India.1252 LC_NUMERIC=C
## [5] LC_TIME=English_India.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] rmelting_0.99.1 testthat_2.0.1  rJava_0.9-10    readxl_1.1.0
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_1.0.0       cellranger_1.1.0 compiler_3.6.0   pillar_1.3.0
##  [5] highr_0.7        tools_3.6.0      digest_0.6.18    pkgload_1.0.2
##  [9] evaluate_0.12    tibble_1.4.2     rlang_0.3.0.1    bibtex_0.4.2
## [13] rstudioapi_0.8   yaml_2.2.0       xfun_0.4         withr_2.1.2
## [17] stringr_1.3.1    knitr_1.20       desc_1.2.0       gbRd_0.4-11
## [21] rprojroot_1.3-2  R6_2.3.0         Rdpack_0.10-3    rmarkdown_1.10
## [25] pander_0.6.3     magrittr_1.5     backports_1.1.2  htmltools_0.3.6
## [29] assertthat_0.2.0 tinytex_0.9      stringi_1.2.4    crayon_1.3.4
```

References

Ahsen, N. von, Wittwer, C. T., and Schütz, E. (2001). Oligonucleotide melting temperatures under PCR conditions: Nearest-neighbor corrections for Mg2+, deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. *Clinical Chemistry* 47, 1956–1961. Available at: http://clinchem.aaccjnls.org/content/47/11/1956 [Accessed February 2, 2018].

Allawi, H. T., and SantaLucia, J. (1997). Thermodynamics and NMR of internal G · T mismatches in dna. *Biochemistry* 36, 10581–10594. doi:10.1021/bi962590c.

Blake, R. D. (1996). "Denaturation of DNA," in *Encyclopedia of molecular biology and molecular medicine*, ed. R. A. Meyers (Weinheim, Germany: VCH Verlagsgesellschaft), 1–19.

Blake, R. D., and Delcourt, S. G. (1998). Thermal stability of DNA. *Nucleic Acids Research* 26, 3323–3332. doi:10.1093/nar/26.14.3323.

Breslauer, K. J., Frank, R., Blöcker, H., and Marky, L. A. (1986). Predicting DNA duplex stability from the base sequence. *Proceedings of the National Academy of Sciences* 83, 3746. doi:10.1073/pnas.83.11.3746.

Britten, R. J., Graham, D. E., and Neufeld, B. R. (1974). Analysis of repeating DNA sequences by reassociation. *Methods in Enzymology* 29, 363–418. doi:10.1016/0076-6879(74)29033-5.

Chester, N., and Marshak, D. (1993). Dimethyl sulfoxide-mediated primer Tm reduction: A method for analyzing the role of renaturation temperature in the polymerase chain reaction. *Analytical Biochemistry* 209, 284–290. doi:10.1006/abio.1993.1121.

Dumousseau, M., Rodriguez, N., Juty, N., and Le Novère, N. (2012). MELTING, a flexible platform to predict the melting temperatures of nucleic acids. *BMC Bioinformatics* 13, 101. doi:10.1186/1471-2105-13-101.

Frank-Kamenetskii, M. D. (1971). Simplification of the empirical relationship between melting temperature of DNA, its GC content and concentration of sodium ions in solution. *Biopolymers* 10, 2623–2624. doi:10.1002/bip.360101223.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., et al. (1986). Improved free-energy parameters for predictions of RNA duplex stability. *Proceedings of the National Academy of Sciences* 83, 9373. doi:10.1073/pnas.83.24.9373.

Hall, T. J., Grula, J. W., Davidson, E. H., and Britten, R. J. (1980). Evolution of sea urchin non-repetitive DNA. *Journal of Molecular Evolution* 16, 95–110. doi:10.1007/BF01731580.

Kierzek, E., Mathews, D. H., Ciesielska, A., Turner, D. H., and Kierzek, R. (2006). Nearest neighbor parameters for Watson-Crick complementary heteroduplexes formed between 2'-O-methyl RNA and RNA oligonucleotides. *Nucleic Acids Research* 34, 3609–3614. doi:10.1093/nar/gkl232.

Le Novère, N. (2001). MELTING, computing the melting temperature of nucleic acid duplex. *Bioinformatics* 17, 1226–1227. doi:10.1093/bioinformatics/17.12.1226.

Marmur, J., and Doty, P. (1962). Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *Journal of Molecular Biology* 5, 109–118. doi:10.1016/S0022-2836(62)80066-7.

Owen, R., Hill, L., and Lapage, S. (1969). Determination of DNA base compositions from melting profiles in dilute buffers. *Biopolymers* 7, 503–516. doi:10.1002/bip.1969.360070408.

SantaLucia, J. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of the National Academy of Sciences* 95, 1460. doi:10.1073/pnas.95.4.1460.

SantaLucia, J., and Hicks, D. (2004). The thermodynamics of DNA structural motifs. *Annual Review of Biophysics and Biomolecular Structure* 33, 415–440. doi:10.1146/annurev.biophys.32.110601.141800.

SantaLucia, John, Allawi, H. T., and Seneviratne, P. A. (1996). Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry* 35, 3555–3562. doi:10.1021/bi951907q.

Schildkraut, C., and Lifson, S. (1965). Dependence of the melting temperature of DNA on salt concentration. *Biopolymers* 3, 195–208. doi:10.1002/bip.360030207.

Sugimoto, N., Nakano, S., Yoneyama, M., and Honda, K. (1996). Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Research* 24, 4501–4505. doi:10.1093/nar/24.22.4501.

Tanaka, F., Kameda, A., Yamamoto, M., and Ohuchi, A. (2004). Thermodynamic parameters based on a nearest-neighbor model for DNA sequences with a single-bulge loop. *Biochemistry* 43, 7143–7150. doi:10.1021/bi036188r.

Wahl, G. M., Barger, S. L., and Kimmel, A. R. (1987). Molecular hybridization of immobilized nucleic acids: Theoretical concepts and practical considerations. *Methods in Enzymology* 152, 399–407. doi:10.1016/0076-6879(87)52046-8.

Wetmur, J. G. (1991). DNA probes: Applications of the principles of nucleic acid hybridization. *Critical Reviews in Biochemistry and Molecular Biology* 26, 227–259. doi:10.3109/10409239109114069.

Xia, T., SantaLucia, J., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., et al. (1998). Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* 37, 14719–14735. doi:10.1021/bi9809425.