applied sciences

MDPI

Article

TmCalculator: An R Package for Genome-Wide Nucleic Acid Melting Temperature Profiling and Multi-Omics Integration

Junhui Li [1,*] and Lihua Julie Zhu [1,2,3,*]

[1] Department of Molecular, Cell, and Cancer Biology, University of Massachusetts Chan Medical School, Worcester, MA 01605, USA
[2] Program in Molecular Medicine, University of Massachusetts Chan Medical School, Worcester, MA 01605, USA
[3] Department of Genomics and Computational Biology, University of Massachusetts Chan Medical School, Worcester, MA 01605, USA
* Correspondence: junhui.li11@umassmed.edu (J.L.); julie.zhu@umassmed.edu (L.J.Z.)

Abstract

Accurate calculation of nucleic acid melting temperature (Tm) underpins many molecular biology applications, while genome-wide Tm profiling provides a means to investigate sequence-intrinsic thermodynamic properties associated with genome function. Existing Tm tools, however, are designed for individual sequences rather than coordinate-aware genome-scale analyses and integrating their output with genomic datasets requires additional processing. We developed TmCalculator, a Bioconductor-compatible R package that extends Tm analysis to coordinate-aware, genome-scale thermodynamic profiling. It accepts nucleotide sequences, FASTA files, installed BSgenome packages, and GRanges objects, and implements 31 nearest-neighbor parameter sets alongside empirical GC-content models with salt and chemical-denaturant corrections. Results are returned as GRanges objects, so that profiles integrate directly with genomic, epigenomic, and other multi-omics data, and built-in functions support their statistical comparison and visualization. Profiling the complete *Escherichia coli* K-12 MG1655 genome reproduced a published map of MutL-associated replication-error hotspots, recovering the reported association between reduced duplex stability and error-prone regions. A compiled C++ core and parallel genome partitioning profiled the hg38 human genome in approximately three minutes on a six-core laptop. TmCalculator therefore provides an efficient, reproducible framework for investigating thermodynamic determinants of genome function and their relationships to disease.

Keywords: melting temperature calculation; genome-wide thermodynamic profiling; Bioconductor; R package; multi-omics integration

Check for updates

Academic Editors: Keun Ho Ryu and Dirk Tischler

Received: 23 July 2026
Revised: 22 September 2026
Accepted: 3 October 2026
Published: 20 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The melting temperature (Tm) of a nucleic acid duplex is defined as the temperature at which half of the duplex molecules dissociate into single strands [1]. As a fundamental thermodynamic measure of duplex stability, Tm is determined by both intrinsic sequence features and extrinsic environmental conditions, including nucleotide composition, sequence length, mismatches, ionic strength, and chemical additives [2–6]. Three main approaches are commonly used for Tm prediction. The Wallace rule provides a rapid estimate based on weighted nucleotide counts and is primarily applicable for short oligonucleotides

of 14–20 bp [7]. For longer sequences, Tm can be estimated using empirical GC content-based models that incorporate corrections for sequence length and ionic strength [1,3]. The nearest-neighbor (NN) thermodynamic model offers greater accuracy by accounting for sequence-context effects through the summed stacking energies of adjacent base-pair doublets. Experimentally derived parameter sets have been established for DNA duplexes [8,9], RNA duplexes [10] and RNA/DNA hybrids [11], and are curated in the nearest-neighbor database [12]. In addition to sequence-dependent thermodynamics, accurate Tm prediction requires consideration of solution conditions, particularly monovalent and divalent cation concentrations [2,13,14] and helix-destabilizing agents such as formamide or dimethyl sulfoxide (DMSO) [5]. Two principal strategies have been developed to account for salt effects, differing in whether ionic dependence is treated as an external correction or incorporated directly into the thermodynamic parameters. Conventionally, a single parameter set is fitted at a reference ionic strength with effects at other ionic conditions incorporated through a closed-form correction [2,13–15]. Alternatively, enthalpy and entropy parameters are fitted directly against melting temperatures measured at multiple sodium concentrations, thereby incorporating sodium dependence into the parameters themselves [16–18]. For such parameter sets, applying an additional salt correction is inappropriate because it would double-count the ionic contribution. Beyond these developments, NN parameterization has continued to evolve along two complementary directions. One focuses on refining the model itself, including replacement of the single terminal-AU penalty in RNA duplexes with end terms that depend on the penultimate base pair [19], and remeasurement of RNA/DNA hybrid stacking parameters under physiological conditions [20]. The other seeks to better approximate the intracellular environment during measurement, either through adding molecular crowding agents [21,22] or through an artificial cytoplasm containing most of the *E. coli* metabolome [23].

Historically, Tm calculation has served primarily as a key design parameter in molecular biology workflows. In PCR and qPCR, accurate Tm estimation promotes specific primer annealing to intended targets [24]. In DNA microarrays, harmonizing probe Tm values is essential for reproducible hybridization signals [25,26]. In CRISPR-based genome editing, the thermodynamic stability of the guide RNA-DNA heteroduplex influences both on-target cleavage and off-target activity [27,28]. Beyond these applications, the thermodynamic properties that govern nucleic acid hybridization in vitro also influence how DNA is accessed and processed within cells. Consequently, Tm is increasingly recognized not only as a sequence design parameter but also as a locus-specific biophysical property that varies across the genome and may have functional implications [29,30]. For example, promoters and replication origins are often enriched in AT-rich, thermodynamically less stable sequences that facilitate transcription initiation and replication [31,32]. Similarly, open chromatin regions identified by ATAC-seq tend to exhibit lower thermodynamic stability, suggesting that local Tm contributes to the physical accessibility of regulatory DNA [33,34]. Replication error hotspots have likewise been associated with regions of reduced duplex stability, which often coincide with elevated densities of microsatellites, cruciform-forming inverted repeats, and GATC mismatch repair sites [35]. Collectively, these observations suggest that thermodynamic stability represents a biologically informative genomic feature that extends beyond oligonucleotide design and warrants systematic investigation at genome scale, particularly in conjunction with other genomic and epigenomic datasets.

Existing tools for Tm estimation can be broadly categorized by their primary design objectives. Web-based calculators such as the NEB Tm Calculator [36] and the ThermoFisher Tm Calculator [37] are designed for primer optimization, often in conjunction with specific commercial polymerase systems. OligoCalc [38] is designed for interactive Tm calculation from user-provided sequence inputs and supports several methods, including

nearest-neighbor thermodynamic models. Programmable frameworks provide scriptable Tm calculations, including MELTING 5 [39,40], which is accessible from R through rmelting [41] and a separate package distributing the MELTING 5 Java archives [42], and the Bio.SeqUtils.MeltingTemp module in Biopython [43].

Other efforts have focused on developing and refining thermodynamic parameters. VarGibbs derives nearest-neighbor enthalpy and entropy parameters directly from experimentally measured melting temperatures and distributes parameter sets fitted at defined sodium concentrations for DNA, RNA, and DNA/RNA hybrid duplexes [16–18]. The companion CompareTm web service enables comparison of Tm predictions across parameter sets and experimental conditions.

Beyond nearest-neighbor approaches, mesoscopic statistical-physics models of the Peyrard-Bishop family describe the full melting transition and can recover microscopic properties such as hydrogen-bonding strength and stacking flexibility [44]. These models are implemented in TfReg, which uses transfer-matrix methods to compute melting temperatures and base-pair opening profiles [45].

Although these tools serve their intended purposes well, two limitations become apparent when the objective is genome-wide thermodynamic profiling. First, they operate on individual sequence strings and lack native support for genomic coordinates. As a result, users must generate genomic windows, retrieve sequences from reference genomes, perform Tm calculations for each interval, and reconstruct the results into coordinate-aware data structures suitable for downstream analysis. Second, the thermodynamic parameter sets available for routine Tm evaluation are generally restricted to those incorporated when the software was developed. Parameter sets optimized at defined sodium concentrations through VarGibbs, as well as more recently reported measurements obtained under cellular-like conditions [20,21], are not available in most commonly used calculators. Consequently, the challenge of genome-wide Tm analysis lies less in the Tm calculation itself than in the surrounding infrastructure required for genome-scale data processing, integration, and access to an up-to-date thermodynamic model library.

To address these challenges, we developed TmCalculator, a Bioconductor-compatible R package that extends melting temperature analysis from individual sequences to genome-wide thermodynamic profiling. TmCalculator supports multiple input formats, including nucleotide sequences, FASTA files, installed BSgenome packages, and GRanges objects, and integrates directly with Bioconductor reference genome resources through BSgenome [46]. The package implements three classes of Tm calculation methods: the Wallace rule, empirical GC-content models, and nearest-neighbor thermodynamic models with salt and chemical corrections.

TmCalculator provides 31 nearest-neighbor parameter sets spanning DNA, RNA, and RNA/DNA hybrid duplexes. These include both classical reference-salt parameterizations and parameter sets fitted under defined ionic or cellular-like conditions, allowing users to select thermodynamic models appropriate for the experimental or biological context under investigation. Results are returned as GRanges objects [47], the standard genomic interval data structure within the Bioconductor ecosystem, enabling direct integration with downstream genomic analyses.

The sequence-scanning engine shared by all three calculation methods is implemented in C++ through Rcpp [48–51]. Combined with built-in parallelization, this enables efficient profiling of whole genomes or user-defined genomic regions with a single function call. TmCalculator further facilitates multi-omic integration by projecting external genomic datasets onto a common Tm reference framework and assigning each thermodynamic window a representative value. This addresses the coordinate mismatch that arises because Tm profiles are generated on regular genomic tilings, whereas peak-based and coverage-

based datasets are defined by dataset-specific intervals. The resulting unified GRanges object supports direct comparison, statistical analysis, and visualization across multiple genomic data layers.

We illustrate the utility of TmCalculator using the *Escherichia coli* K-12 MG1655 genome by revisiting the map of MutL-associated replication error hotspots reported by Hasenauer et al. [35]. This study provides a compelling use case because its melting-temperature analyses were generated using TmCalculator, representing an independent biological application of the package. As such, it enables both reproduction of the published findings and demonstration of the broader genome-wide thermodynamic profiling workflow supported by TmCalculator. We present an end-to-end analysis encompassing Tm calculation, multi-omic integration, statistical comparison, and visualization within a unified coordinate-aware framework.

Reproducing the published analysis also entails adopting its original parameter settings, including the 200 bp window size used by Hasenauer et al. [35]. Because window size can influence both the resolution of a thermodynamic profile and its biological interpretation, we additionally examine its impact on the resulting Tm landscape and on the association between thermodynamic stability and replication error hotspots.

Finally, we extend the analysis from a bacterial chromosome to the human genome and evaluate the computational resources required for genome-scale profiling. This assessment demonstrates that whole-genome thermodynamic analysis can be performed efficiently on standard laboratory computing hardware, supporting routine integration of Tm profiling into modern genomics workflows.

2. Materials and Methods

2.1. Package Implementation and Availability

TmCalculator is implemented in R [52] within the Bioconductor ecosystem and is organized around four components: input processing, Tm calculation, genomic data representation, and multi-omics integration. Sequences and genomic intervals are represented using GenomicRanges, IRanges, and GenomeInfoDb [47], and all outputs are returned as GRanges objects, providing a coordinate-aware framework that supports downstream analysis, visualization, and integration with other genomic datasets. FASTA inputs are parsed with Biostrings [53], whereas coordinate-based inputs retrieve sequences directly from reference genomes through BSgenome [46] and Biostrings. Visualization functions leverage karyoploteR [54] for genome-wide displays and ggplot2 [55], ggridges [56] and ggforce [57] for comparative analyses.

TmCalculator implements three classes of Tm calculation methods: the Wallace rule, empirical GC-content models, and nearest-neighbor thermodynamic models with support for salt and chemical corrections. The package provides 31 nearest-neighbor parameter sets spanning DNA, RNA, and RNA/DNA hybrid duplexes, including both classical reference-salt parameterizations and parameter sets derived under defined ionic or cellular-like conditions. Algorithmic details of sequence preprocessing, nearest-neighbor parameter lookup, thermodynamic calculations, and performance optimizations are provided in Supplementary Methods S1 and S2.

Genome-wide thermodynamic profiling is performed through a single tm_calculate() function that accepts nucleotide sequences, FASTA files, installed BSgenome packages, or GRanges objects as its source, with a regions argument selecting what to cover. User-defined window and slide parameters control genomic tiling, and parallel execution is requested by passing a BiocParallel [58] backend as the BPPARAM argument. Thermodynamic profiles are stored as GRanges objects and can be integrated directly with external genomic and epigenomic datasets through the package's coordinate-aware multi-omics

framework (Supplementary Methods S3 and S4). Runtime performance and memory usage are evaluated in Sections 3.1 and 3.5.

The package passes R CMD check without errors or warnings. Correctness is validated using a testthat [59] test suite that includes regression tests against published melting-temperature values as well as tests covering input parsing, coordinate-based calculations, and multi-omics integration functions. TmCalculator is available on CRAN. All analyses reported here were performed with TmCalculator 1.1.2 under R 4.4.1 and Bioconductor 3.19, with BSgenome.Ecoli.NCBI.ASM584v2 1.0.0, Rcpp 1.1.2 and ggplot2 3.5.2.

2.2. Benchmarking Setup for Accuracy and Computational Cost Against Existing Tools

TmCalculator was benchmarked against two widely used nearest-neighbor Tm implementations: rmelting 1.20.0 [41], an R interface to the MELTING 5 Java engine [39,40], and the Bio.SeqUtils.MeltingTemp module in Biopython 1.85 [43] under Python 3.9.12. Comparisons were restricted to oligonucleotide Tm calculation, the functionality shared by all three tools. Features unique to TmCalculator, including coordinate-aware genome-scale analysis and multi-omics integration, were evaluated separately and are therefore not reflected in runtime comparisons.

To ensure comparability, all tools were configured to use the identical nearest-neighbor parameters [6], a sodium concentration of 50 mM, equal strand concentrations of 25 nM, and the same sodium correction model [15]. Default settings in MELTING 5 were overridden where necessary to ensure that all implementations used equivalent thermodynamic assumptions. Detailed benchmark configuration and parameter harmonization procedures are described in Supplementary Methods S5.1 and S5.2.

Benchmark datasets consisted of 200 bp windows randomly sampled from the *Escherichia coli* K-12 MG1655 genome, matching the window size used in the case study presented below. Dataset sizes ranged from 10^2 to 3×10^5 sequences with logarithmic spacing across multiple scales. Runtime, peak memory consumption, and agreement of computed Tm values were assessed for each implementation. Numerical consistency was quantified using the mean and maximum absolute deviation relative to TmCalculator. Detailed benchmark datasets, execution procedures, and evaluation metrics are described in Supplementary Methods S5.3–S5.5.

All benchmarks were repeated three times, and median values are reported. The benchmark script (bench_crosstool.R) is distributed with the package and can be located after installation using system.file("scripts", package = "TmCalculator").

2.3. Workflow for Genome-Wide Thermodynamic Profiling and Multi-Layer Integration in E. coli K-12 MG1655

We illustrate the utility of TmCalculator by revisiting the genome-wide analysis of spontaneous DNA replication error hotspots in rapidly proliferating *Escherichia coli* K-12 MG1655 reported by Hasenauer et al. [35]. In that study, MutL-associated regions (MutL-ARs), identified by ChIP-seq of the mismatch repair protein MutL, were associated with reduced thermodynamic stability and several genomic features linked to replication errors. Because the published melting-temperature analyses were generated using TmCalculator, this dataset provides an independent biological application through which the complete genome-wide thermodynamic profiling workflow can be reproduced and extended.

The analysis integrates seven data sources: the *E. coli* K-12 MG1655 reference genome, genome-wide Tm profiles, MutL-AR coordinates, microsatellite density, cruciform-forming palindrome density, single-stranded DNA (ssDNA-seq), and GATC methylation site density. These data are combined within a unified genomic framework to demonstrate Tm calculation, coordinate-aware feature integration, statistical comparison, and multi-layer

visualization. Data sources, preprocessing procedures, and integration workflows are described in Supplementary Methods S6.

Genome-wide Tm profiles were generated using non-overlapping 200 bp windows to match the original study [35], and thermodynamic properties were compared between MutL-associated regions and background genomic intervals. Statistical analyses and visualization were performed using functions provided by TmCalculator. Complete reproducible workflows are distributed through the package vignettes, including genome_wide_tm_ecoli.

Because the 200 bp window size was inherited from the original analysis, we additionally evaluated the impact of window size on genome-wide Tm profiles and downstream biological interpretation. Analyses were repeated using 50, 100, 200, and 500 bp windows while holding all other parameters constant. The complete workflow is available in the package vignette window_size_sensitivity, and methodological details are provided in Supplementary Methods S7.

2.4. Benchmarking Setup for Runtime, Memory, and Parallel Scaling on the Human Genome

To evaluate computational performance beyond the bacterial-scale analyses described above, we profiled the human reference genome (UCSC hg38, equivalent to GRCh38; approximately 3.1 Gb), obtained from the BSgenome.Hsapiens.UCSC.hg38 package using the same thermodynamic model and windowing strategy employed in the *E. coli* case study. Genome-wide Tm profiles were generated using non-overlapping 200 bp windows and the nearest-neighbor method with the DNA_NN_Breslauer_1986 parameter set at 50 mM sodium concentration.

Performance was further characterized by varying the number of worker processes from one to six. Runtime and memory consumption were measured on a six-core laptop and a high-performance computing (HPC) cluster node. Measurements included total execution time and peak memory requirements during genome-wide profiling. Detailed descriptions of the computational environments, task-partitioning strategy, benchmarking procedures, and hardware configurations are provided in Supplementary Methods S8.

To facilitate reproducibility, all benchmarking scripts and benchmark outputs used in these analyses are distributed with TmCalculator and can be accessed through the package installation directory.

3. Results

3.1. TmCalculator: An Integrated Framework for Genome-Wide Thermodynamic Profiling

TmCalculator is organized into four functional layers: input, computation, data representation, and integration (Figure 1). The input layer supports nucleotide sequences, FASTA files, reference genomes provided through BSgenome, and GRanges objects, allowing analyses to be initiated from either sequence-based or coordinate-based genomic resources. User-defined genomic regions and tiling parameters enable thermodynamic profiling at arbitrary genomic resolutions.

The computation layer implements three classes of melting-temperature calculation methods: the Wallace rule, empirical GC-content models, and nearest-neighbor thermodynamic models. Nearest-neighbor calculations support multiple parameter sets and optional salt and chemical corrections, enabling analyses under a range of experimental and biological conditions.

Results are returned as TmCalculator objects whose core data component is a GRanges object, providing a standardized genomic interval representation fully compatible with the Bioconductor ecosystem. In addition to thermodynamic measurements, the object records all analysis parameters required for reproducibility.

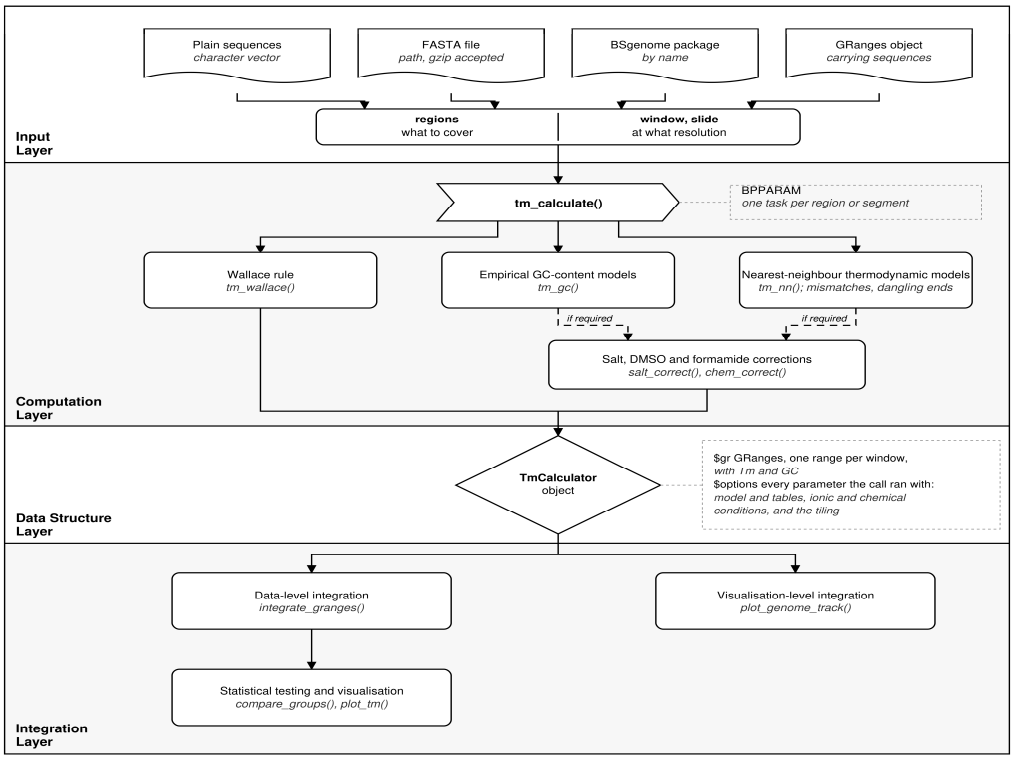


Figure 1. Architecture and workflow of TmCalculator, showing the input, computation, data representation, and multi-omics integration layers.

The integration layer supports both data-level and visualization-level integration of genomic features. Data-level integration is performed through integrate_granges(), which reconciles coordinate differences between Tm profiles and external genomic datasets using overlap-, nearest-, window-, or bin-based strategies. Visualization-level integration uses a unified track interface that renders multiple datasets as aligned genomic tracks. Supported visualizations include circular genome plots, genome-wide karyotype views, and locus-specific multi-track displays, facilitating comparison of thermodynamic profiles with complementary genomic and epigenomic features.

3.1.1. Input Layer: Sequence Sources and Region Selection

TmCalculator accepts four types of sequence sources through a unified tm_calculate() interface: (i) nucleotide sequences supplied as character vectors, (ii) FASTA files (compressed or uncompressed), (iii) installed BSgenome packages specified by package name, and (iv) GRanges objects containing sequence information. Custom reference genomes generated with BSgenomeForge are supported alongside BSgenome packages distributed through Bioconductor [46,60].

A common regions argument specifies the sequences or genomic intervals to be analyzed regardless of input source. Regions may be identified by sequence name, numerical index, coordinate interval (for example, chr1:1–1000), a mixture of these formats, or directly as a GRanges object. This unified syntax allows the same workflow to be applied to individual oligonucleotides, selected genomic intervals, chromosomes, or complete genomes without modifying downstream analysis steps. For example, chr1 refers to a chromosome in a BSgenome package, a record in a FASTA file, or a sequence name in a GRanges object, whereas 1:1–200 selects the first 200 bases of the first sequence when the input consists of unnamed character vectors.

The resolution of thermodynamic profiling is controlled by the window and slide parameters. When slide equals window, non-overlapping windows are generated; when

slide is smaller than window, overlapping windows are produced. By adjusting these parameters, users can generate Tm profiles at arbitrary genomic resolutions. If window is not specified, each selected region is treated as a single sequence and a single Tm value is returned, supporting applications such as primer, probe, or amplicon analysis.

3.1.2. Computation Layer: Tm Calculation Methods

The unified interface described above enables analyses ranging from individual oligonucleotides to whole genomes through a single function call. Genome-scale analyses are automatically partitioned into independent tasks that can be executed serially or in parallel while producing identical thermodynamic profiles. Details of the compiled implementation, parallel execution architecture, thermodynamic parameter handling, and environmental correction models are provided in Supplementary Methods S2.

TmCalculator implements three classes of Tm calculation methods that differ in complexity and intended application: the Wallace rule, empirical GC-content models, and nearest-neighbor thermodynamic models. The Wallace rule provides a simple composition-based estimate suitable for short oligonucleotides [7,61]. Empirical GC-content models estimate Tm from GC percentage, sequence length, and ionic conditions, offering efficient calculations for longer sequences while disregarding sequence order. Eight published GC-content formulas covering DNA, RNA, and RNA/DNA hybrid duplexes are implemented (Table 1) [1–3,5,62–64].

Nearest-neighbor thermodynamic models estimate Tm from the sequence-dependent stacking enthalpies and entropies of adjacent nucleotide pairs, thereby capturing duplex-stability effects that composition-based methods cannot resolve. The implementation supports perfectly matched duplexes as well as terminal mismatches, internal mismatches, and dangling ends through dedicated parameter tables (Table 2). TmCalculator provides 31 nearest-neighbor parameter sets spanning classical reference-condition models, parameter sets optimized for defined ionic conditions, and parameterizations derived under cell-mimicking environments (Table 2).

In addition to the built-in parameter library, users may incorporate custom thermodynamic parameter tables, enabling evaluation of newly published models or modified nucleotides without altering the package code. Custom parameter tables are validated automatically and can be combined with existing mismatch and dangling-end models through the same analysis framework. Both empirical GC-content and nearest-neighbor methods support salt, DMSO, and formamide corrections [5,65], with reaction conditions supplied by the user rather than fixed by the software. All calculation parameters, thermodynamic models, and environmental settings are recorded in the returned object to ensure reproducibility.

Table 1. GC content models implemented in TmCalculator.

Method	Type	Formula
Primer3Plus [62]	DNA	$Tm = 81.5 + 0.41(\%GC) - 600/N + 16.6\log_{10}[Mon]$
Chester1993 [63]	DNA	$Tm = 69.3 + 0.41(\%GC) - 650/N$
QuikChange [64]	DNA	$Tm = 81.5 + 0.41(\%GC) - 675/N - \%mismatch$
Schildkraut1965 [2]	DNA	$Tm = 81.5 + 0.41(\%GC) - 675/N + 16.6\log_{10}[Mon]$
Wetmur1991_MELTING [3]	DNA	$Tm = 81.5 + 0.41(\%GC) - 500/N + 16.6\log_{10}([Mon]/(1.0 + 0.7[Mon])) - \%mismatch$
Wetmur1991_RNA [3]	RNA	$Tm = 78 + 0.7(\%GC) - 500/N + 16.6\log_{10}([Mon]/(1 + 0.7[Mon])) - \%mismatch$

Table 1. *Cont.*

Method	Type	Formula
Wetmur1991_RNA/DNA [3]	RNA/DNA	$Tm = 67 + 0.8(\%GC) - 500/N + 16.6\log_{10}([Mon]/(1 + 0.7[Mon])) - \%mismatch$
vonAhsen2001 [5]	DNA	$Tm = 77.1 + 0.41(\%GC) - 528/N + 11.7\log_{10}[Mon]$

Note. %GC denotes GC content (%), *N* denotes sequence length (nt), [Mon] denotes the monovalent ion concentration (mM), and %mismatch denotes the percentage of mismatched bases.

Table 2. Nearest-neighbor thermodynamic parameter sets implemented in TmCalculator.

Category	Parameter [1]	Type
Perfect match: reference-condition set	DNA_NN_SantaLucia_2004 [6]	DNA/DNA
	DNA_NN_Breslauer_1986 [8]	DNA/DNA
	DNA_NN_Sugimoto_1996 [66]	DNA/DNA
	DNA_NN_Allawi_1997 [4]	DNA/DNA
	RNA_NN_Freier_1986 [67]	RNA/RNA
	RNA_NN_Xia_1998 [10]	RNA/RNA
	RNA_NN_Chen_2012 [68]	RNA/RNA
	RNA_DNA_NN_Sugimoto_1995 [11]	RNA/DNA
	RNA_NN_Zuber_2022 [19]	RNA/RNA
Perfect match: condition-specific set	DNA_NN_Weber_2015 [18]	DNA/DNA
	DNA_NN_Weber_OW04_69/_119/_220/_621/_1020 [18]	DNA/DNA
	RNA_NN_Weber_VIF_71/_121/_221/_621/_1021 [17]	RNA/RNA
	RNA_NN_Weber_FIF_71/_121/_221/_621/_1021 [17]	RNA/RNA
	RNA_DNA_NN_Weber_2019_FT/_VH [16]	RNA/DNA
	RNA_DNA_NN_Weber_2019_LS [16]	RNA/DNA
	RNA_DNA_NN_Banerjee_2020 [20]	RNA/DNA
Perfect match: cell-mimicking set	DNA_NN_Ghosh_2020_PEG200 [22]	DNA/DNA
	RNA_NN_Ghosh_2023_PEG200 [21]	RNA/RNA
Terminal mismatch	DNA_TMM_SantaLucia_2001 [69]	DNA/DNA
Internal mismatch	DNA_IMM_Peyret_1999 [70]	DNA/DNA
Dangling end	DNA_DE_Bommarito_2000 [71]	DNA/DNA
	RNA_DE_Turner_2010 [12]	RNA/RNA

Note. [1] In addition to the parameter sets listed here, user-supplied parameter tables may be provided as matrices. Supplied tables must contain the complete key set of a compatible built-in reference table and are validated automatically before use. Additional keys, such as those introduced for modified nucleotides, are retained, whereas missing reference keys are rejected during validation.

3.1.3. Data Structure Layer: Granges Output for Bioconductor Ecosystem Integration

Calculated Tm results are returned as a TmCalculator object whose gr element is a GRanges object [47] containing genomic coordinates together with computed Tm values, GC content, and associated metadata. By adopting GRanges as its native data structure, TmCalculator produces thermodynamic profiles that are immediately compatible with the Bioconductor ecosystem.

This representation enables direct integration with a wide range of genomic data types, including ATAC-seq peaks, ChIP-seq binding sites, gene annotations, replication origins, chromatin accessibility profiles, and genetic variants, using standard GenomicRanges operations such as findOverlaps() [47]. GRanges-based outputs can also be incorporated into higher-level Bioconductor data containers, including SummarizedExperiment [72] and MultiAssayExperiment [73], and used together with packages such as VariantAnnotation [74] for variant-centric analyses. Together, these capabilities allow Tm to function

as a standardized genomic annotation that can be analyzed alongside existing biological measurements within established Bioconductor workflows.

In addition to the thermodynamic profile itself, each TmCalculator object records the complete set of analysis parameters used to generate the results. These include the calculation method, thermodynamic parameter sets, ionic and chemical conditions, environmental correction models, and, when applicable, tiling and parallelization settings. By preserving both results and execution parameters within a single object, TmCalculator supports reproducible thermodynamic profiling and downstream analyses.

3.1.4. Integration Layer: Multi-Omics Data Integration and Visualization

The integration layer enables genome-wide thermodynamic profiles to be combined with external genomic and epigenomic datasets spanning diverse resolutions and feature types. The integrate_granges() function merges Tm profiles with one or more external GRanges objects and returns a unified GRanges object in which integrated features are stored as metadata columns for downstream analysis and visualization.

Four integration strategies are supported: overlap, nearest, window, and bin. These approaches allow Tm profiles to be associated with discrete genomic features such as ATAC-seq and ChIP-seq peaks, gene annotations, genetic variants, and quantitative genomic signals including RNA-seq coverage. Flexible aggregation options enable datasets with differing genomic resolutions to be integrated within a common coordinate framework. Details of the integration algorithms, overlap-weighting schemes, aggregation procedures, and coverage calculations are provided in Supplementary Methods S4.

Statistical comparisons between genomic region classes are implemented through compare_groups(), which supports both the Wilcoxon rank-sum test and Student's t-test. Multiple-testing correction methods, including Bonferroni and Benjamini–Hochberg procedures, are available for controlling family-wise error rates and false discovery rates.

In addition to data-level integration, TmCalculator supports visualization-level integration through a track_list framework that allows each dataset to retain its native coordinates while being displayed along a shared genomic axis. The plot_genome_track() function generates both circular whole-genome views and linear genome-browser-style visualizations, supporting multiple track types and display configurations. Distributional comparisons between classes of genomic regions are provided by plot_tm(), which supports boxplots, violin plots, density plots, ridgeline plots, rank plots, and sina plots, with optional annotation of multiple-testing-corrected p values.

Together, these capabilities enable thermodynamic profiles to be integrated, analyzed, and visualized alongside diverse genomic and multi-omics datasets within a unified and reproducible workflow.

3.2. Comparison with Existing Tm Calculation Tools

TmCalculator was compared with rmelting, an R interface to the MELTING 5 Java engine, and Bio.SeqUtils.MeltingTemp from Biopython, with all three tools configured to use identical nearest-neighbor parameters, salt corrections, and reaction conditions (Section 2.2; Supplementary Methods S5).

TmCalculator and Biopython produced numerically identical melting temperatures, agreeing to the limits of floating-point precision (Table 3). In contrast, MELTING 5 exhibited deviations of up to 1.4 °C under the same thermodynamic model and reaction conditions, indicating differences in implementation despite nominally equivalent parameterization. Detailed thermodynamic consistency analyses are provided in Supplementary Methods S5.6.

Table 3. Agreement in melting temperature between the three tools. Deviations are relative to TmCalculator on identical input.

Tool	Sequences	Mean ΔT_m (°C)	Max $\lvert \Delta T_m \rvert$ (°C)	Pearson *r*	Max ΔΔH (kcal mol^{-1}) [1]	Max ΔΔS (cal mol^{-1} K^{-1}) [1]
Bio.SeqUtils.MeltingTemp	1000	6.5×10^{-11}	5.0×10^{-9}	1.0000	0.000	0.000
rmelting (MELTING 5)	100	0.5	1.4	0.9975	0.307	1.954

Note. [1] Enthalpy and entropy are not exposed by any of the three interfaces. They were recovered from T_m measured at two strand concentrations differing tenfold, which cancels the concentration convention and leaves ΔH determined by the ratio alone. The ΔΔH offset is near-constant across sequences, indicating a difference in the duplex initiation term rather than in the stacking sum; the ΔΔS deviation is not constant.

For computational performance, both TmCalculator and Biopython exhibited linear scaling across the tested input sizes (Figure 2A; Table 4). At larger workloads (>10^4 sequences), TmCalculator achieved approximately 4.6-fold higher computational throughput than Biopython, reflecting the use of a compiled C++ implementation. For smaller workloads, this advantage was offset by the startup overhead associated with loading the R and Bioconductor environment. As a result, total elapsed time favored Biopython for small inputs, whereas TmCalculator became faster at approximately 1.5×10^4 sequences and above (Figure 2B).

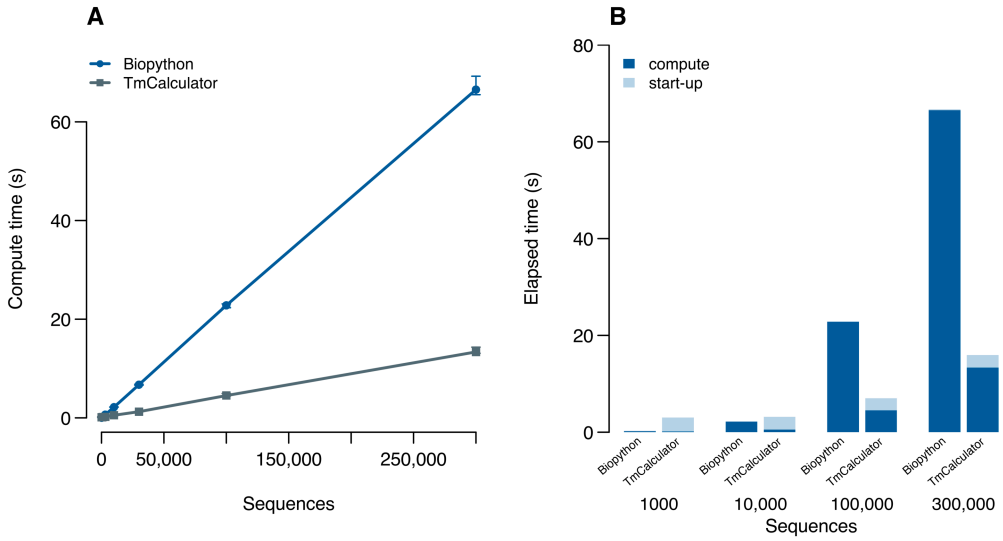


Figure 2. Computational performance of TmCalculator and Bio.SeqUtils.MeltingTemp on identical 200 bp windows derived from the *Escherichia coli* K-12 MG1655 genome. (**A**) Compute time as a function of input size. Points represent median values from three benchmark repetitions, with bars indicating the observed range. (**B**) Elapsed time for single benchmark runs at four input sizes, partitioned into startup and computation components. Startup includes process launch and package initialization. File input/output is excluded because it was introduced by the benchmarking framework rather than by the thermodynamic calculations. MELTING 5 is omitted because benchmarking was restricted to input sizes up to 10^3 sequences (Table 4).

Memory requirements showed the opposite trend. TmCalculator carried a larger fixed memory footprint because of the R session and associated Bioconductor dependencies, whereas per-sequence memory consumption remained modest for both implementations (Table 4). Runtime decomposition, memory analyses, and startup-overhead measurements are described in Supplementary Methods S5.7.

Table 4. Runtime and memory of the three tools on identical 200 bp windows. Medians of three repetitions, range in brackets.

Tool	Sequences	Compute Time (s) [1]	Throughput (s^{-1})	Start-Up (s)	Peak RSS (GB)
TmCalculator	100	0.1 [0.1–0.1]	1042	2.6	0.34
	300	0.1 [0.1–0.1]	2857	2.5	0.34
	1000	0.1 [0.1–0.1]	7519	2.4	0.34
	3000	0.2 [0.2–0.2]	14,019	2.4	0.34
	10,000	0.5 [0.5–0.5]	20,243	2.4	0.35
	30,000	1.3 [1.3–1.3]	22,624	2.4	0.35
	100,000	4.5 [4.5–4.5]	22,080	2.5	0.39
	300,000	13.7 [13.3–13.8]	21,983	2.3	0.58
Biopython	100	0.0 [0.0–0.0]	4762	0.1	0.01
	300	0.1 [0.1–0.1]	4800	0.1	0.01
	1000	0.2 [0.2–0.2]	4682	0.1	0.01
	3000	0.6 [0.6–0.6]	4833	0.1	0.01
	10,000	2.1 [2.1–2.1]	4831	0.1	0.02
	30,000	6.6 [6.3–7.2]	4520	0.2	0.02
	100,000	21.2 [20.7–21.3]	4710	0.1	0.06
	300,000	62.8 [61.7–63.3]	4777	0.2	0.15
rmelting [2]	100	5.3 [5.3–5.4]	18.8	0.6	0.18
	300	28.5 [28.4–28.5]	10.5	0.5	0.20
	1000	263.2 [254.5–267.9]	3.8	0.7	0.28

Note. [1] Each tool was executed as an independent process so that peak resident-set memory (RSS) could be attributed to the implementation being evaluated. Input and output operations were introduced by the benchmarking framework rather than by the thermodynamic calculations themselves; these operations were therefore measured separately and excluded from the reported runtimes. [2] melting() processes a single sequence per invocation. Consequently, benchmarking n sequences required n independent invocations of the Java engine. Because practical execution times increased substantially with dataset size, MELTING 5 was evaluated only up to 10^3 sequences.

MELTING 5 was benchmarked only up to 10^3 sequences because its runtime increased substantially with input size (Table 4). In addition, neither MELTING 5 nor Biopython provides coordinate-aware inputs, parallel execution support, or native GRanges outputs. Constructing genome-wide thermodynamic profiles with these tools therefore requires external workflows for genomic tiling, sequence retrieval, and result aggregation, whereas these operations are integrated directly within TmCalculator.

3.3. Application: Genome-Wide Thermodynamic Profiling and Multi-Layer Integration in E. coli K-12 MG1655

To illustrate the genome-wide thermodynamic profiling capabilities of TmCalculator, we applied the framework to the complete *E. coli* K-12 MG1655 genome. Analysis using non-overlapping 200 bp windows yielded 23,208 genomic intervals with Tm values ranging from 81.1 to 111.4 °C and a genome-wide mean of 99.0 ± 3.8 °C (median 99.7 °C). GC content across windows ranged from 20.5% to 75.0%, with a mean of 50.8%. Window generation, sequence retrieval and Tm calculation together completed in approximately 1.1 s on a single core of a standard desktop computer (2.6 GHz Intel Core i7, 16 GB RAM, R 4.4). Thermodynamic profiling at the bacterial scale is therefore computationally inexpensive and runs on commodity hardware.

To demonstrate TmCalculator's native support for multi-layer genomic integration and visualization, the genome-wide Tm profile was then combined with the five additional feature layers of the ecoli_rep_hotspots dataset distributed with the package. All layers were assembled into a track_list object and visualized using plot_genome_track(), producing the seven-track circular genome representation of Figure 3A, with two hotspot-

enriched regions shown at higher resolution in Figure 3B. Figure 3C resolves a 20 kb locus (U00096.3:212,801–232,800) at the window level. Three pairs of 200 bp windows there share GC content yet differ in Tm by 4.3, 4.1 and 4.1 °C; a GC-content formula assigns one temperature to each pair, so the difference lies entirely in base arrangement.

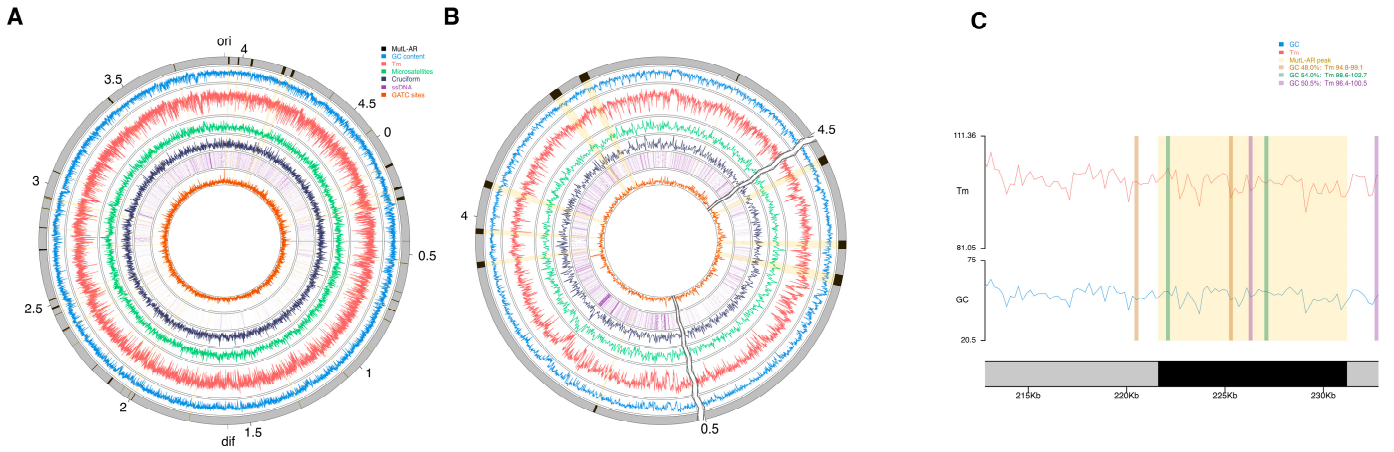


Figure 3. Genome-wide thermodynamic profiling and multi-layer integration in *Escherichia coli* K-12 MG1655. (**A**) Seven-track circular genome map generated by TmCalculator. Concentric rings from outermost to innermost show MutL-AR ChIP-seq peaks (grey rectangles), GC content per 200 bp window (%), Tm per 200 bp window (°C), microsatellite density, cruciform-prone sequence density, single-stranded DNA regions, and GATC methylation-site density. The replication origin (*ori*, position 3,925,804) and terminus (*dif*, position 1,590,777) are annotated on the outermost ring. (**B**) Enlarged circular view of two genomic regions (0.1–0.5 Mb and 3.6–4.5 Mb) displayed at higher resolution. (**C**) High-resolution linear view of U00096.3:212,801–232,800, a 20 kb interval centered on a MutL-AR peak, showing GC-content and Tm tracks. Yellow shading indicates MutL-AR peaks. Three pairs of 200 bp windows are highlighted in amber, green, and purple. Within each pair, windows have identical GC content (48.0%, 54.0%, and 50.5%, respectively) and identical length but differ in Tm by 4.3, 4.1, and 4.1 °C.

Each of the 23,208 windows was then annotated for MutL-AR overlap using integrate_granges(), classifying 806 windows as MutL-AR-associated and 22,402 as background. compare_groups() showed that MutL-AR windows have significantly lower Tm values than background (Wilcoxon rank-sum test, $p = 1.2 \times 10^{-44}$; median Tm 98.0 °C versus 99.7 °C) and significantly lower GC content (Wilcoxon rank-sum test, $p = 8.5 \times 10^{-22}$; median GC 50.0% vs. 52.0%). The reduced Tm and GC content of these hotspots, together with the elevated microsatellite and cruciform-prone sequence density visible in Figure 3, reproduce the thermodynamic landscape previously reported for this dataset.

Together, these analyses demonstrate the utility of TmCalculator for genome-wide thermodynamic profiling, multi-omics integration, statistical analysis, and visualization within a unified workflow.

3.4. Sensitivity of the Thermodynamic Profile to Window Size

Recomputing the profile at 50, 100, 200 and 500 bp showed that window size affects the scale of the Tm distribution but not its spatial structure (Table 5). Mean Tm rises with window length, because duplex stability increases with length in the nearest-neighbor model, so absolute Tm values are comparable only within a single window size; conversely, shorter windows widen the distribution, the standard deviation at 50 bp being 1.6 times that at 500 bp. Mean GC content is identical to two decimal places at every window size, confirming that the composition landscape itself does not depend on the tiling. Aggregating each profile onto a common 1 kb grid, which all four window sizes divide exactly, leaves

them highly concordant, and the four profiles resolve the same local features at different degrees of smoothing (Figure 4).

Table 5. Effect of window size on the genome-wide Tm profile and on the MutL-AR comparison in *Escherichia coli* K-12 MG1655.

Window-Size (bp)	Windows Number	Tm Mean ± SD (°C)	GC Mean (%)	r vs. 1 kb Grid [1]	MutL-AR Windows [2]	Tm Peak − Background (°C) [3]	p (Tm) [4]
50	92,833	84.2 ± 5.3	50.8	0.998	3103	−2.0	5.3×10^{-90}
100	46,416	94.0 ± 4.4	50.8	0.999	1571	−1.7	1.2×10^{-65}
200	23,208	99.0 ± 3.8	50.8	1.000	806	−1.8	1.2×10^{-44}
500	9283	102.0 ± 3.2	50.8	reference	343	−1.7	6.1×10^{-25}

Note. [1] Pearson correlation between each profile aggregated onto a common 1 kb grid and the 500 bp profile aggregated onto the same grid. Because all four window sizes divide 1 kb exactly (20, 10, 5, and 2 windows per 1 kb interval for 50, 100, 200, and 500 bp windows, respectively), aggregation introduces no remainder. The 500 bp profile serves as the reference. [2] Number of windows overlapping a MutL-AR ChIP-seq peak; all remaining windows constitute the background.[3] Difference between the median Tm of MutL-AR windows and the median Tm of background windows. Negative values indicate lower Tm in MutL-AR windows. [4] Wilcoxon rank-sum test comparing Tm between MutL-AR and background windows. Corresponding GC-content comparisons were significant at all window sizes ($p \leq 1.4 \times 10^{-12}$) and are omitted for brevity.

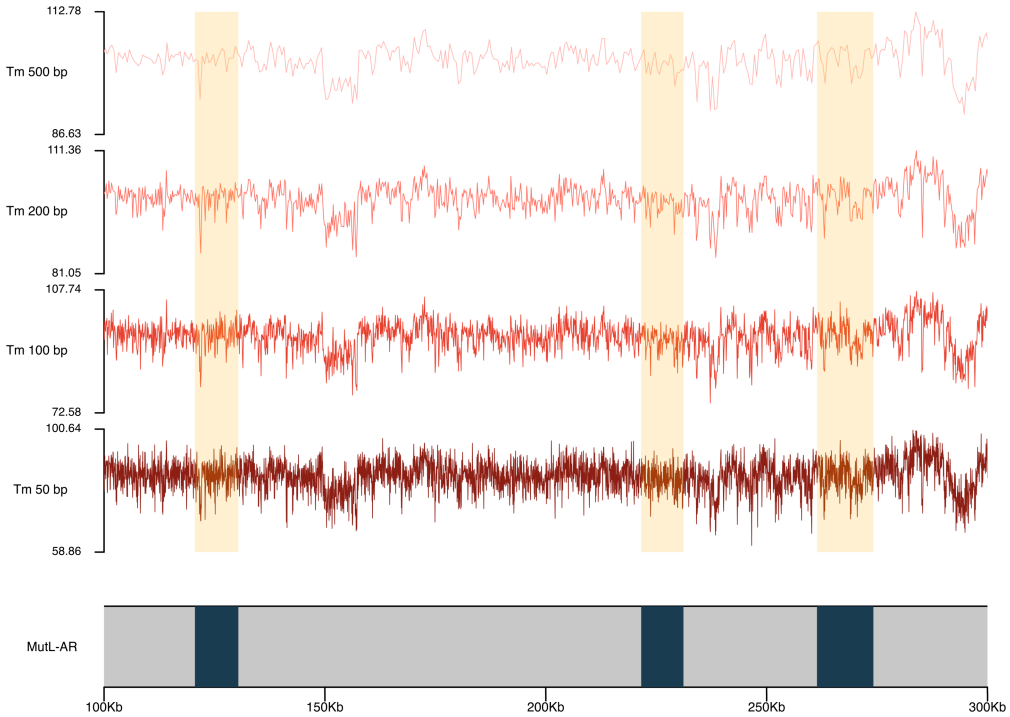


Figure 4. Effect of window size on genome-wide thermodynamic profiles. Tm profiles of the same genomic region computed using non-overlapping windows of 50, 100, 200, and 500 bp. The plotted interval corresponds to positions 0.1–0.3 Mb of the *Escherichia coli* K-12 MG1655 chromosome, within the region shown in Figure 3B. Profiles are displayed from darkest to lightest shading in order of increasing window size. MutL-AR ChIP-seq peaks are shown as the upper track and highlighted in yellow across all profiles.

MutL-AR windows have lower median Tm and lower GC content than the genomic background at all four window sizes, with the effect size essentially constant and every Wilcoxon rank-sum test significant. Across the 50 to 500 bp range examined, the spatial thermodynamic landscape and the MutL-AR association remained highly consistent, indi-

cating that the analysis is robust to window-size selection within this range. The complete analysis, including the multi-scale integrated visualization, is provided in the package vignette: vignette("window_size_sensitivity", package = "TmCalculator").

3.5. Runtime and Memory at Mammalian Genome Scale

Genome-wide thermodynamic profiling of hg38 using non-overlapping 200 bp windows generated 14,687,330 analyzable windows after exclusion of windows containing ambiguous bases. A single tm_calculate() invocation completed the analysis in 590.3 s on the laptop and 494.5 s on the compute node (Table 6, Figure 5). Increasing the number of worker processes reduced runtime in both environments, although scalability differed between the two systems.

Table 6. Runtime and memory requirements for genome-wide thermodynamic profiling of hg38 (14,687,330 non-overlapping 200 bp windows) across worker counts and computational environments.

Environment [1]	Workers	Wall Time (s) [2]	Speedup [3]	Peak RSS per Worker (GB) [3]
Laptop	1	590.3 [586.1–590.7]	1.0	3.47
	2	336.8 [335.4–337.6]	1.8	2.50
	3	239.6 [233.1–243.9]	2.5	2.63
	4	197.4 [196.4–199.8]	3.0	2.18
	5	179.7 [178.0–184.0]	3.3	2.00
	6	179.8 [176.4–214.9]	3.3	2.81
Compute node	1	494.5 [485.8–529.0]	1.0	3.55
	2	277.4 [276.8–278.1]	1.8	3.53
	3	197.0 [196.7–197.3]	2.5	2.88
	4	154.3 [154.1–156.3]	3.2	2.95
	5	129.9 [128.8–130.3]	3.8	3.10
	6	114.4 [114.0–116.0]	4.3	3.59

Note. [1] Laptop: six-core 2.6 GHz Intel Core i7, 16 GB memory, macOS, R 4.4. Compute node: two 20-core Intel Xeon Gold 6230 processors at 2.10 GHz, hyper-threading disabled, 376 GB memory, R 4.4.1, evaluated using six allocated slots and a 16 GB memory limit on a shared node. [2] Median wall-clock time from three benchmark repetitions; observed ranges are shown in brackets. Reported times include worker startup. [3] Speedup is defined as the one-worker wall time divided by the wall time of the corresponding configuration within the same environment. Peak RSS represents the resident-set size of the worker with the highest memory consumption.

On the laptop, runtime decreased from 590.3 s with a single worker to 179.7 s with five workers, corresponding to a 3.3-fold speedup. Increasing the worker count to six provided no additional benefit (179.8 s) and was associated with substantially greater variability across repetitions. On the compute node, runtime decreased monotonically to 114.4 s at six workers, corresponding to a 4.3-fold speedup. Runtime variation between repetitions remained small across all parallel configurations.

The different scaling behavior reflects processor availability rather than memory capacity. On the laptop, five workers plus the coordinating process already occupy all six physical cores, leaving no dedicated core for a sixth worker. In contrast, the six allocated workers on the compute node run on a system with forty physical cores. Peak resident memory ranged from 2.0 to 3.6 GB per worker across all configurations and did not limit performance in either environment, because each worker processes no more than a single 50 Mb genomic segment at a time.

Fitting the wall times to Amdahl's law yielded $T(n) = 79.4 + 509.2/n$ for the laptop and $T(n) = 41.6 + 456.4/n$ for the compute node, with $R^2 \geq 0.997$ in both cases. The parallelizable component was similar across environments, consistent with the identical computational workload, whereas the serial component differed substantially between the two systems.

These results indicate that near-optimal performance is generally achieved when using one worker per physical core minus one; however, the optimal worker count should be determined empirically for the target computational environment. The benchmarking workflows used in this study are distributed with the package.

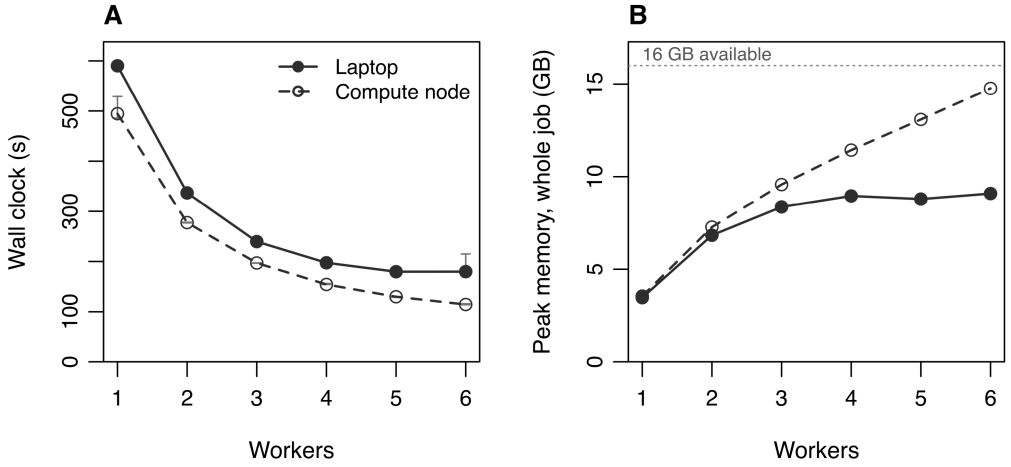


Figure 5. Runtime and memory requirements for genome-wide thermodynamic profiling of the human genome. Genome-wide Tm profiling of hg38 using non-overlapping 200 bp windows (14,687,330 windows) evaluated across one to six worker processes in two computational environments. Solid lines with filled symbols represent the six-core laptop; dashed lines with open symbols represent the compute node. Points denote the median of three benchmark repetitions and bars indicate the observed range. (**A**) Wall-clock time, including worker startup. (**B**) Peak resident-set size (RSS) of the heaviest worker process.

4. Discussion

Existing Tm calculators are primarily designed for individual sequences or primer optimization and do not support coordinate-aware, genome-scale profiling within the Bioconductor ecosystem. TmCalculator complements these tools by changing the unit of analysis from individual sequences to genomic intervals. Its native GRanges representation allows Tm to be treated as a quantitative genomic annotation that can be integrated, statistically compared, and visualized alongside coverage tracks, genomic peaks, gene annotations, variants, and other interval-based measurements. The same framework scales to mammalian genomes: a complete human-genome profile was generated in approximately three minutes on a six-core laptop, demonstrating that genome-wide thermodynamic profiling can be performed on commonly available computing hardware.

Several limitations should be considered when interpreting genome-wide Tm profiles generated by TmCalculator. First, all three calculation approaches, including the Wallace rule, empirical GC-content formulas, and nearest-neighbor thermodynamic models, are derived from measurements of purified nucleic acids in solution. They do not directly model the cellular environment, where chromatin organization, DNA supercoiling, molecular crowding, and interactions with DNA-binding proteins can influence duplex accessibility and stability. Genome-wide Tm profiles should therefore be interpreted as sequence-intrinsic thermodynamic baselines against which functional genomic measurements can be compared, rather than as direct predictions of duplex unwinding in living cells.

Second, nearest-neighbor parameters are generally derived from short oligonucleotides. Their application to longer genomic windows assumes that local stacking contributions remain additive and transferable across sequence lengths. Accuracy may decrease for sequences with compositions or motifs poorly represented in the experimental data used for parameter estimation. In addition, absolute Tm values depend on window size,

thermodynamic parameterization, ionic conditions, and correction model, and should therefore be compared only among analyses performed under consistent settings. Salt corrections introduce an additional source of uncertainty when the requested ionic condition differs substantially from the condition under which a parameter set was derived. To reduce this limitation, TmCalculator includes parameter sets optimized directly at defined sodium concentrations [16–18] and suppresses additional salt correction when the selected set already represents the requested condition.

Third, the built-in thermodynamic library supports canonical nucleotides, perfectly matched duplexes, internal and terminal mismatches, and dangling ends, but does not currently distribute parameter sets for chemically modified nucleotides, such as 5-methylcytosine or 5-hydroxymethylcytosine. User-supplied parameter tables can be incorporated without modifying the package, if they contain the complete key set required by a compatible built-in reference table. Thus, the software infrastructure can accommodate modified-base models, but its accuracy will depend on the availability and quality of experimentally derived thermodynamic parameters.

Fourth, tm_calculate() models a linear duplex and sums local stacking contributions along the aligned strands. It does not calculate competing intramolecular folding energies. Consequently, Tm estimates for sequences capable of forming hairpins, G-quadruplexes, cruciforms, or other secondary structures do not account for competition between duplex formation and alternative conformations. Analyses of such regions should combine TmCalculator profiles with results from dedicated nucleic acid secondary-structure models.

Several developments could extend the biological and analytical scope of TmCalculator. Parameter sets derived under molecular-crowding or intracellular-like conditions may help narrow the gap between dilute-solution measurements and cellular environments [21,22]. Complementary cellular variables, including chromatin accessibility, nucleosome occupancy, transcription-factor binding, and transcriptional activity, can be incorporated as genomic tracks aligned to the same Tm windows. This approach preserves the distinction between sequence-intrinsic thermodynamic properties and experimentally measured cellular context while allowing their relationships to be evaluated directly.

The transferability of nearest-neighbor parameters to long genomic windows could also be examined through calibration against melting measurements for longer duplexes. Such studies could define the sequence lengths and nucleotide-composition ranges over which current models remain reliable. Future versions could additionally report whether a genomic window lies outside the sequence-composition or experimental range represented in the data used to fit the selected thermodynamic model.

Support for modified nucleotides represents another important direction. The existing parameter-table interface already permits additional stacking keys, allowing experimentally derived modified-base parameter sets to be incorporated without changing the computational framework. Published observations concerning the effects of cytosine methylation and molecular crowding [75] may inform initial model development, but generalized additive treatment should be validated across sequence contexts and experimental conditions before being adopted as a default. As comprehensive stack-level measurements become available, methylation-aware Tm profiles could be constructed by integrating modified-base thermodynamic parameters with genomic methylation measurements.

Secondary-structure analysis could similarly be incorporated as a complementary genomic layer. Duplex stability and single-stranded folding free energy could be represented as separate, coordinate-aligned tracks, enabling identification of regions in which alternative structures may compete strongly with duplex formation. This design would retain the modularity of the current framework while extending genome-wide analysis beyond a fully paired linear-duplex model.

The implemented RNA/DNA hybrid parameter sets, including RNA_DNA_NN_Sugimoto_1995 [11] and parameter sets optimized at defined sodium concentrations [16–18], also provide a foundation for applications involving guide RNA-DNA heteroduplexes and R-loops. However, comprehensive support for these applications will require additional parameterization and correction models that address mismatches, guide modifications, buffer composition, and cellular conditions. These requirements are particularly important for CRISPR applications, in which thermodynamic stability represents only one component of on-target and off-target activity.

More broadly, systematic integration of genome-wide Tm profiles with chromatin accessibility, transcription, protein-DNA binding, repetitive elements, secondary-structure propensity, and disease-associated variation may help clarify how sequence-intrinsic thermodynamic properties contribute to genome function and disease-associated genomic variation. As increasingly comprehensive genomic, epigenomic, and single-cell datasets become available, extending thermodynamic models through experimentally validated parameter sets and improved representations of cellular environments will expand the biological questions that can be addressed through genome-wide thermodynamic profiling.

5. Conclusions

TmCalculator extends melting-temperature analysis from individual sequences to coordinate-aware, genome-scale thermodynamic profiling. By representing Tm as a native GRanges-based genomic annotation, the package enables thermodynamic information to be integrated directly with genomic, epigenomic, and multi-omics datasets within established Bioconductor workflows. The framework combines multiple Tm calculation methods, comprehensive thermodynamic parameter libraries, genome-scale parallel processing, statistical analysis, and visualization within a unified environment. Together, these capabilities provide a standardized and extensible foundation for incorporating thermodynamic information into computational genomics and multi-omics studies.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app162010400/s1, Supplementary Methods S1: Nearest-Neighbor Thermodynamic Calculation; Supplementary Methods S2: Computational Implementation; Supplementary Methods S3: Genome-Wide Thermodynamic Profiling; Supplementary Methods S4: Multi-Omics Integration Framework; Supplementary Methods S5: Benchmarking Accuracy and Computational Cost; Supplementary Methods S6: Reproduction of the *E. coli* Replication Error Hotspot Analysis; Supplementary Methods S7: Window-Size Sensitivity Analysis; Supplementary Methods S8: Genome-Scale Runtime and Memory Benchmarking.

Author Contributions: J.L. conceived and developed the TmCalculator package, designed and implemented all software components, performed all analyses, and drafted the manuscript. L.J.Z. supervised the project, contributed to study design, and critically revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The TmCalculator R package is available from CRAN at https://cran.r-project.org/web/packages/TmCalculator/index.html (accessed on 2 October 2026). The ecoli_rep_hotspots dataset distributed with TmCalculator contains all feature layers required to reproduce the *Escherichia coli* application analyses presented in this study. The benchmark datasets underlying Tables 3, 4 and 6 and Figures 2 and 5 are provided as CSV files in the package extdata directory. The scripts used to generate these benchmarks are included in the package scripts directory,

enabling full reproduction of the reported performance evaluations. Three annotated vignettes reproduce the principal analyses described in this manuscript: Genome-wide thermodynamic profiling and multi-layer integration in *E. coli* (https://cran.r-project.org/web/packages/TmCalculator/vignettes/genome_wide_tm_ecoli.html, accessed on 2 October 2026). Window-size sensitivity analysis (https://cran.r-project.org/web/packages/TmCalculator/vignettes/window_size_sensitivity.html, accessed on 2 October 2026). Genome-scale profiling and parallel performance evaluation (https://cran.r-project.org/web/packages/TmCalculator/vignettes/hg38_performance_parallel.html, accessed on 2 October 2026). Together, the package, bundled datasets, benchmark files, analysis scripts, and vignettes provide a fully reproducible framework for all analyses presented in this study.

Acknowledgments: The authors thank Kai Hu, Haibo Liu, and Yaroslav Markov for helpful discussions and valuable suggestions. The authors are also grateful to Gerald Weber (Universidade Federal de Minas Gerais, Brazil), whose comments on the preprint prompted the inclusion of melting-temperature-optimized nearest-neighbor parameter sets in TmCalculator and revisions to the description of nearest-neighbor models.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ATAC-seq	Assay for transposase-accessible chromatin with sequencing
ChIP-seq	Chromatin immunoprecipitation sequencing
CRISPR	Clustered regularly interspaced short palindromic repeats
DMSO	Dimethyl sulfoxide
FASTA	Text-based sequence format
FDR	False discovery rate
FWER	Family wise error rate
GC	Guanine-cytosine
GWAS	Genome-wide association study
IGV	Integrative Genomics Viewer
MutL-AR	MutL-associated region
NN	Nearest neighbor
PEG200	Polyethylene Glycol 200
PCR	Polymerase chain reaction
RNA-seq	RNA sequencing
ssDNA-seq	Single-stranded DNA sequencing
Tm	Melting Temperature

References

1. Marmur, J.; Doty, P. Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *J. Mol. Biol.* **1962**, *5*, 109–118. [CrossRef] [PubMed]
2. Schildkraut, C. Dependence of the melting temperature of DNA on salt concentration. *Biopolymers* **1965**, *3*, 195–208. [CrossRef] [PubMed]
3. Wetmur, J.G. DNA probes: Applications of the principles of nucleic acid hybridization. *Crit. Rev. Biochem. Mol. Biol.* **1991**, *26*, 227–259. [CrossRef] [PubMed]
4. Allawi, H.T.; SantaLucia, J., Jr. Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry* **1997**, *36*, 10581–10594. [CrossRef] [PubMed]
5. von Ahsen, N.; Wittwer, C.T.; Schutz, E. Oligonucleotide melting temperatures under PCR conditions: Nearest-neighbor corrections for Mg^{2+}, deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. *Clin. Chem.* **2001**, *47*, 1956–1961. [CrossRef]
6. SantaLucia, J., Jr.; Hicks, D. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 415–440. [CrossRef] [PubMed]
7. Wallace, R.B.; Shaffer, J.; Murphy, R.F.; Bonner, J.; Hirose, T.; Itakura, K. Hybridization of synthetic oligodeoxyribonucleotides to phi chi 174 DNA: The effect of single base pair mismatch. *Nucleic Acids Res.* **1979**, *6*, 3543–3557. [CrossRef] [PubMed]

8. Breslauer, K.J.; Frank, R.; Blocker, H.; Marky, L.A. Predicting DNA duplex stability from the base sequence. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 3746–3750. [CrossRef] [PubMed]
9. SantaLucia, J., Jr. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 1460–1465. [CrossRef] [PubMed]
10. Xia, T.; SantaLucia, J., Jr.; Burkard, M.E.; Kierzek, R.; Schroeder, S.J.; Jiao, X.; Cox, C.; Turner, D.H. Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. *Biochemistry* **1998**, *37*, 14719–14735. [CrossRef] [PubMed]
11. Sugimoto, N.; Nakano, S.; Katoh, M.; Matsumura, A.; Nakamuta, H.; Ohmichi, T.; Yoneyama, M.; Sasaki, M. Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry* **1995**, *34*, 11211–11216. [CrossRef] [PubMed]
12. Turner, D.H.; Mathews, D.H. NNDB: The nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Res.* **2010**, *38*, D280-282. [CrossRef] [PubMed]
13. Owczarzy, R.; You, Y.; Moreira, B.G.; Manthey, J.A.; Huang, L.; Behlke, M.A.; Walder, J.A. Effects of sodium ions on DNA duplex oligomers: Improved predictions of melting temperatures. *Biochemistry* **2004**, *43*, 3537–3554. [CrossRef] [PubMed]
14. Owczarzy, R.; Moreira, B.G.; You, Y.; Behlke, M.A.; Walder, J.A. Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry* **2008**, *47*, 5336–5353. [CrossRef] [PubMed]
15. SantaLucia, J., Jr.; Allawi, H.T.; Seneviratne, P.A. Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry* **1996**, *35*, 3555–3562. [CrossRef] [PubMed]
16. Basilio Barbosa, V.; de Oliveira Martins, E.; Weber, G. Nearest-neighbour parameters optimized for melting temperature prediction of DNA/RNA hybrids at high and low salt concentrations. *Biophys. Chem.* **2019**, *251*, 106189. [CrossRef] [PubMed]
17. Ferreira, I.; Jolley, E.A.; Znosko, B.M.; Weber, G. Replacing salt correction factors with optimized RNA nearest-neighbour enthalpy and entropy parameters. *Chem. Phys.* **2019**, *521*, 69–76. [CrossRef] [PubMed]
18. Weber, G. Optimization method for obtaining nearest-neighbour DNA entropies and enthalpies directly from melting temperatures. *Bioinformatics* **2015**, *31*, 871–877. [CrossRef] [PubMed]
19. Zuber, J.; Schroeder, S.J.; Sun, H.; Turner, D.H.; Mathews, D.H. Nearest neighbor rules for RNA helix folding thermodynamics: Improved end effects. *Nucleic Acids Res.* **2022**, *50*, 5251–5262. [CrossRef] [PubMed]
20. Banerjee, D.; Tateishi-Karimata, H.; Ohyama, T.; Ghosh, S.; Endoh, T.; Takahashi, S.; Sugimoto, N. Improved nearest-neighbor parameters for the stability of RNA/DNA hybrids under a physiological condition. *Nucleic Acids Res.* **2020**, *18*, 12042–12054. Correction in *Nucleic Acids Res.* **2021**, *49*, 10796–10799. https://doi.org/10.1093/nar/gkab780. [CrossRef] [PubMed]
21. Ghosh, S.; Takahashi, S.; Banerjee, D.; Ohyama, T.; Endoh, T.; Tateishi-Karimata, H.; Sugimoto, N. Nearest-neighbor parameters for the prediction of RNA duplex stability in diverse in vitro and cellular-like crowding conditions. *Nucleic Acids Res.* **2023**, *51*, 4101–4111. [CrossRef] [PubMed]
22. Ghosh, S.; Takahashi, S.; Ohyama, T.; Endoh, T.; Tateishi-Karimata, H.; Sugimoto, N. Nearest-neighbor parameters for predicting DNA duplex stability in diverse molecular crowding conditions. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 14194–14201. [CrossRef] [PubMed]
23. Sieg, J.P.; Jolley, E.A.; Huot, M.J.; Babitzke, P.; Bevilacqua, P.C. In vivo-like nearest neighbor parameters improve prediction of fractional RNA base-pairing in cells. *Nucleic Acids Res.* **2023**, *51*, 11298–11317. [CrossRef] [PubMed]
24. Rychlik, W.; Spencer, W.J.; Rhoads, R.E. Optimization of the annealing temperature for DNA amplification in vitro. *Nucleic Acids Res.* **1990**, *18*, 6409–6412. [CrossRef] [PubMed]
25. Burden, C.J.; Pittelkow, Y.E.; Wilson, S.R. Statistical analysis of adsorption models for oligonucleotide microarrays. *Stat. Appl. Genet. Mol. Biol.* **2004**, *3*, 35. [CrossRef] [PubMed]
26. Pozhitkov, A.E.; Tautz, D.; Noble, P.A. Oligonucleotide microarrays: Widely applied--poorly understood. *Brief. Funct. Genom. Proteomic* **2007**, *6*, 141–148. [CrossRef] [PubMed]
27. Kuscu, C.; Arslan, S.; Singh, R.; Thorpe, J.; Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* **2014**, *32*, 677–683. [CrossRef] [PubMed]
28. Alkan, F.; Wenzel, A.; Anthon, C.; Havgaard, J.H.; Gorodkin, J. CRISPR-Cas9 off-targeting assessment with nucleic acid duplex energy parameters. *Genome Biol.* **2018**, *19*, 177. [CrossRef] [PubMed]
29. Bernardi, G. Isochores and the evolutionary genomics of vertebrates. *Gene* **2000**, *241*, 3–17. [CrossRef] [PubMed]
30. Liu, F.; Tostesen, E.; Sundet, J.K.; Jenssen, T.K.; Bock, C.; Jerstad, G.I.; Thilly, W.G.; Hovig, E. The human genomic melting map. *PLoS Comput. Biol.* **2007**, *3*, e93. [CrossRef] [PubMed]
31. Choi, C.H.; Kalosakas, G.; Rasmussen, K.O.; Hiromura, M.; Bishop, A.R.; Usheva, A. DNA dynamically directs its own transcription initiation. *Nucleic Acids Res.* **2004**, *32*, 1584–1590. [CrossRef] [PubMed]
32. Kanhere, A.; Bansal, M. Structural properties of promoters: Similarities and differences between prokaryotes and eukaryotes. *Nucleic Acids Res.* **2005**, *33*, 3165–3175. [CrossRef] [PubMed]

33. Buenrostro, J.D.; Giresi, P.G.; Zaba, L.C.; Chang, H.Y.; Greenleaf, W.J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **2013**, *10*, 1213–1218. [CrossRef] [PubMed]
34. Klemm, S.L.; Shipony, Z.; Greenleaf, W.J. Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* **2019**, *20*, 207–220. [CrossRef] [PubMed]
35. Hasenauer, F.C.; Barreto, H.C.; Lotton, C.; Matic, I. Genome-wide mapping of spontaneous DNA replication error-hotspots using mismatch repair proteins in rapidly proliferating *Escherichia coli*. *Nucleic Acids Res.* **2025**, *53*, gkae1196. [CrossRef] [PubMed]
36. Biolabs, N.E. NEB Tm Calculator. Available online: https://tmcalculator.neb.com/#!/main (accessed on 13 July 2026).
37. Scientific, T. Tm Calculator for Primers. Available online: https://www.thermofisher.com/us/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html (accessed on 13 July 2026).
38. Kibbe, W.A. OligoCalc: An online oligonucleotide properties calculator. *Nucleic Acids Res.* **2007**, *35*, W43–W46. [CrossRef] [PubMed]
39. Le Novere, N. MELTING, computing the melting temperature of nucleic acid duplex. *Bioinformatics* **2001**, *17*, 1226–1227. [CrossRef] [PubMed]
40. Dumousseau, M.; Rodriguez, N.; Juty, N.; Le Novere, N. MELTING, a flexible platform to predict the melting temperatures of nucleic acids. *BMC Bioinform.* **2012**, *13*, 101. [CrossRef] [PubMed]
41. Aravind, J.; Krishna, K.G. rmelting: R Interface to MELTING 5. R Package Version 1.20.0. Available online: https://aravind-j.github.io/rmelting/ (accessed on 13 July 2026).
42. Rudis, B. *melting5jars: An R Package Wrapper for MELTING 5 Jars*; 2017. Available online: https://github.com/hrbrmstr/melting5jars (accessed on 13 July 2026).
43. Cock, P.J.; Antao, T.; Chang, J.T.; Chapman, B.A.; Cox, C.J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; et al. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **2009**, *25*, 1422–1423. [CrossRef] [PubMed]
44. Weber, G.; Essex, J.W.; Neylon, C. Probing the microscopic flexibility of DNA from melting temperatures. *Nat. Phys.* **2009**, *5*, 769–773. [CrossRef]
45. Weber, G. TfReg: Calculating DNA and RNA melting temperatures and opening profiles with mesoscopic models. *Bioinformatics* **2013**, *29*, 1345–1347. [CrossRef] [PubMed]
46. Pagès, H. *BSgenome: Software Infrastructure for Efficient Representation of Full Genomes and Their SNPs*, R package version 1.80.0; 2026. Available online: https://bioconductor.org/packages/BSgenome (accessed on 13 July 2026).
47. Lawrence, M.; Huber, W.; Pages, H.; Aboyoun, P.; Carlson, M.; Gentleman, R.; Morgan, M.T.; Carey, V.J. Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **2013**, *9*, e1003118. [CrossRef] [PubMed]
48. Eddelbuettel, D.; Francois, R.; Allaire, J.; Ushey, K.; Kou, Q.; Russell, N.; Ucar, I.; Bates, D.; Chambers, J. *Rcpp: Seamless R and C++ Integration*; R package version 1.1.2; 2026. Available online: https://CRAN.R-project.org/package=Rcpp (accessed on 13 July 2026).
49. Eddelbuettel, D. *Seamless R and C++ Integration with Rcpp*; Springer: New York, NY, USA, 2013.
50. Eddelbuettel, D.; Joseph, J. Extending R with C++: A Brief Introduction to Rcpp. *Am. Stat.* **2018**, *72*, 28–36. [CrossRef]
51. Eddelbuettel, D.; François, R. Rcpp: Seamless R and C++ Integration. *J. Stat. Softw.* **2011**, *40*, 1–18. [CrossRef]
52. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2026.
53. Pagès, H.A.; Aboyoun, P.; Gentleman, R.; DebRoy, S. *Biostrings: Efficient Manipulation of Biological Strings*, R package version 2.80.1; 2026. Available online: https://bioconductor.org/packages/Biostrings (accessed on 13 July 2026).
54. Gel, B.; Serra, E. karyoploteR: An R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* **2017**, *33*, 3088–3090. [CrossRef] [PubMed]
55. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag: New York, NY, USA, 2016.
56. Wilke, C.O. *ggridges: Ridgeline Plots in 'ggplot2'*, version 0.5.7; 2025. Available online: https://wilkelab.org/ggridges/ (accessed on 13 July 2026).
57. Pedersen, T.L. *ggforce: Accelerating 'ggplot2'*, version 0.5.0.9000; 2025. Available online: https://ggforce.data-imaginist.com (accessed on 13 July 2026).
58. Wang, J.; Morgan, M.; Obenchain, V.; Lang, M.; Thompson, R.; Turaga, N. *BiocParallel: Bioconductor Facilities for Parallel Evaluation*; 2026. Available online: https://bioconductor.org/packages/BiocParallel (accessed on 13 July 2026).
59. Wickham, H. testthat: Get Started with Testing. *R J.* **2011**, *3*, 5–10. [CrossRef]
60. Pagès, H.K. Atuhurira Kirabo *BSgenomeForge: Forge Your Own BSgenome Data Package*, R package version 1.12.0; 2026. Available online: https://bioconductor.org/packages/BSgenomeForge (accessed on 13 July 2026).

61. Wallace, R.B.; Johnson, M.J.; Hirose, T.; Miyake, T.; Kawashima, E.H.; Itakura, K. The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequence to rabbit beta-globin DNA. *Nucleic Acids Res.* **1981**, *9*, 879–894. [CrossRef] [PubMed]
62. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3--new capabilities and interfaces. *Nucleic Acids Res.* **2012**, *40*, e115. [CrossRef] [PubMed]
63. Chester, N.; Marshak, D.R. Dimethyl sulfoxide-mediated primer Tm reduction: A method for analyzing the role of renaturation temperature in the polymerase chain reaction. *Anal. Biochem.* **1993**, *209*, 284–290. [CrossRef] [PubMed]
64. Heydenreich, F.M.; Miljus, T.; Jaussi, R.; Benoit, R.; Milic, D.; Veprintsev, D.B. High-throughput mutagenesis using a two-fragment PCR approach. *Sci. Rep.* **2017**, *7*, 6787. [CrossRef] [PubMed]
65. Blake, R.D.; Delcourt, S.G. Thermodynamic effects of formamide on DNA stability. *Nucleic Acids Res.* **1996**, *24*, 2095–2103. [CrossRef] [PubMed]
66. Sugimoto, N.; Nakano, S.; Yoneyama, M.; Honda, K. Improved thermodynamic parameters and helix initiation factor to predict stability of DNA duplexes. *Nucleic Acids Res.* **1996**, *24*, 4501–4505. [CrossRef] [PubMed]
67. Freier, S.M.; Kierzek, R.; Jaeger, J.A.; Sugimoto, N.; Caruthers, M.H.; Neilson, T.; Turner, D.H. Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 9373–9377. [CrossRef] [PubMed]
68. Chen, J.L.; Dishler, A.L.; Kennedy, S.D.; Yildirim, I.; Liu, B.; Turner, D.H.; Serra, M.J. Testing the nearest neighbor model for canonical RNA base pairs: Revision of GU parameters. *Biochemistry* **2012**, *51*, 3508–3522. [CrossRef] [PubMed]
69. SantaLucia, J., Jr.; Peyret, N. Method and System for Predicting Nucleic Acid Hybridization Thermodynamics and Computer-Readable Storage Medium for Use Therein. Google Patent WO2001094611A2, 13 December 2001.
70. Peyret, N.; Seneviratne, P.A.; Allawi, H.T.; SantaLucia, J., Jr. Nearest-neighbor thermodynamics and NMR of DNA sequences with internal A.A, C.C, G.G, and T.T mismatches. *Biochemistry* **1999**, *38*, 3468–3477. [CrossRef] [PubMed]
71. Bommarito, S.; Peyret, N.; SantaLucia, J., Jr. Thermodynamic parameters for DNA sequences with dangling ends. *Nucleic Acids Res.* **2000**, *28*, 1929–1934. [CrossRef] [PubMed]
72. Morgan, M.; Obenchain, V.; Hester, J.; Pagès, H. *Summarized Experiment: A Container (S4 Class) for Matrix-Like Assays*, R package version 1.42.0; Bioconductor: Seattle, WA, USA, 2026.
73. Ramos, M.; Schiffer, L.; Re, A.; Azhar, R.; Basunia, A.; Rodriguez, C.; Chan, T.; Chapman, P.; Davis, S.R.; Gomez-Cabrero, D.; et al. Software for the Integration of Multiomics Experiments in Bioconductor. *Cancer Res.* **2017**, *77*, e39–e42. [CrossRef] [PubMed]
74. Obenchain, V.; Lawrence, M.; Carey, V.; Gogarten, S.; Shannon, P.; Morgan, M. VariantAnnotation: A Bioconductor package for exploration and annotation of genetic variants. *Bioinformatics* **2014**, *30*, 2076–2078. [CrossRef] [PubMed]
75. Tsuruta, M.; Sugitani, Y.; Sugimoto, N.; Miyoshi, D. Combined Effects of Methylated Cytosine and Molecular Crowding on the Thermodynamic Stability of DNA Duplexes. *Int. J. Mol. Sci.* **2021**, *22*, 947. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.