

# Package ‘vermeulen’

July 22, 2025

**Type** Package

**Title** Biomarker Data Set by Vermeulen et al. (2009)

**Version** 0.1.2

**Description** The biomarker data set by Vermeulen et al. (2009)

<[doi:10.1016/S1470-2045\(09\)70154-8](https://doi.org/10.1016/S1470-2045(09)70154-8)> is provided. The data source, however, is by Ruijter et al. (2013)

<[doi:10.1016/j.ymeth.2012.08.011](https://doi.org/10.1016/j.ymeth.2012.08.011)>. The original data set may be downloaded from

<<https://medischebiologie.nl/wp-content/uploads/2019/02/qpcrdatamethods.zip>>.

This data set is for a real-time quantitative polymerase chain reaction (PCR) experiment that comprises the raw fluorescence data of 24,576 amplification curves. This data set comprises 59 genes of interest and 5 reference genes. Each gene was assessed on 366 neuroblastoma complementary DNA (cDNA) samples and on 18 standard dilution series samples (10-fold 5-point dilution series x 3 replicates + no template controls (NTC) x 3 replicates).

**License** CC BY 4.0

**Encoding** UTF-8

**RoxygenNote** 7.3.1

**Depends** R (>= 2.10)

**Imports** memoise

**Suggests** dplyr, ggplot2, rmarkdown, spelling, testthat (>= 3.0.0),  
tibble

**Language** en-US

**URL** <https://github.com/ramiromagno/vermeulen>,  
<https://rmagno.eu/vermeulen/>

**BugReports** <https://github.com/ramiromagno/vermeulen/issues>

**Config/testthat/edition** 3

**NeedsCompilation** no

**Author** Ramiro Magno [aut, cre] (ORCID:  
    <<https://orcid.org/0000-0001-5226-3441>>),  
    Pattern Institute [cph, fnd]  
**Maintainer** Ramiro Magno <[rmagno@pattern.institute](mailto:rmagno@pattern.institute)>  
**Repository** CRAN  
**Date/Publication** 2024-04-22 22:40:03 UTC

Contents

get_biomarker_dataset . . . . .	2
Index	4

---

get_biomarker_dataset	<i>Import the Biomarker data set</i>
-----------------------	--------------------------------------

---

Description

This function retrieves the Biomarker data set, a data set containing raw fluorescence amplification data: 24,576 amplification curves, of 50 cycles each.

Usage

```
get_biomarker_dataset()
```

Details

Data was gathered from Ruijter et al. (2013), [doi:10.1016/j.ymeth.2012.08.011](https://doi.org/10.1016/j.ymeth.2012.08.011) but original source is by Vermeulen et al. (2009), [doi:10.1016/S14702045\(09\)701548](https://doi.org/10.1016/S14702045(09)701548). The tidy version of the data is kept at the repository of the source of {vermeulen} package. This function fetches such data and thus requires internet connection. It takes a few seconds to run.

The Biomarker data set comprises a set of 59 targets previously identified as a 59-mRNA gene expression signature, that has been developed and validated for improved outcome prediction of children with neuroblastoma. In short, 59 biomarkers and 5 reference genes were measured in 8 µl reactions in a 384-well plate using the LightCycler480 SYBR Green Master (Roche) in a sample maximization experiment design. The 59 genes were carefully selected as being previously reported as prognostic genes in neuroblastoma in at least two independent studies. Each plate contained 366 cDNA samples (n = 1) from primary tumor biopsies, a 5-point 10-fold serial dilution series based on an external oligonucleotide standard (n = 3, from 150,000 to 15 copies), and a no template control (NTC, n = 3). Raw (baseline uncorrected) fluorescent data were exported from the LightCycler480 instrument software.

**Value**

A data frame with 24,576 amplification curves, of 50 cycles each:

`plate` Plate identifier. Because one plate was used per gene, the name of the plate is the same as the values in `target`.

`well` Well identifier.

`dye` In all reactions the SYBR Green I master mix (Roche) was used, so the value is always "SYBR".

`target` Target identifier, in almost all cases the name of a gene.

`target_type` Target type: either target of interest ("toi") or reference target ("ref").

`sample` Sample identifier.

`sample_type` Sample type.

`copies` Standard copy number.

`dilution` Dilution factor. Higher number means greater dilution.

`cycle` PCR cycle.

`fluor` Raw fluorescence values.

**Source**

- <https://medischebiologie.nl/wp-content/uploads/2019/02/qpcrdatamethods.zip>
- <https://github.com/ramiromagno/vermeulen/tree/main/data-raw>

**Examples**

```
# Takes ~ 10-30 sec  
head(get_biomarker_dataset())
```

# Index

`get_biomarker_dataset`, [2](#)