

Bioconductor \LaTeX Style 2.0

Andrzej Oleś¹, Martin Morgan², and Wolfgang Huber¹

¹European Molecular Biology Laboratory, Heidelberg, Germany

²Roswell Park Cancer Institute, Buffalo, NY

August 3, 2016

Abstract

This vignette describes the new Bioconductor \LaTeX document style. It aims to be at once a demonstration of the features of the style and a guide to their use. The vignette focuses on the use with documents typeset as markup `.Rnw` files and processed using *Sweave*[1] or *knitr*[2]. The same style is also applied to *R* markdown `.Rmd` files which are rendered to PDF output format.

Package

BiocStyle 2.0.3

Report issues on <https://github.com/Bioconductor/BiocStyle/issues>

Contents

1	Authoring package vignettes	3
1.1	Getting started	3
1.1.1	<i>Sweave</i>	3
1.1.2	<i>knitr</i>	3
1.1.3	R markdown.	3
1.2	Workflow	4
2	Structuring the document	5
2.1	Title page	5
2.1.1	Title and running headers	5
2.1.2	Authors and affiliations	5
2.1.3	Abstract, package version, and table of contents	6
2.2	Style macros	6
2.3	Code chunks	7
2.4	Figures.	7
2.4.1	<code>\incfig</code> convenience macro	10
2.5	Equations	11
2.6	Footnotes	12
2.7	Bibliography	12
A	Session info	13
B	Attached L ^A T _E X packages.	13
C	Known issues	13
C.1	Use with L ^A T _E X package <code>float</code>	13
C.2	Compatibility with the <i>xtable</i> R package	13

1 Authoring package vignettes

1.1 Getting started

BiocStyle provides standard formatting of documents rendered to PDF output format. It consists of a L^AT_EX document style definition which can be used with either plain *Sweave*, *knitr*, or *rmarkdown*. To enable *BiocStyle* in your package vignette follow the instructions below depending on the input format.

1.1.1 *Sweave*

To use with *Sweave*, add the following to your package 'DESCRIPTION' file:

```
Suggests: BiocStyle
```

and add this code chunk to the preamble¹ of your .Rnw file:

```
<<style-Sweave, eval=TRUE, echo=FALSE, results=tex>>=  
BiocStyle::latex2()  
@
```

¹before the `\begin{document}` command, and preferably right after the `\documentclass` definition

See `?latex2` for additional options. It is not necessary to include `\usepackage{Sweave}` in your document preamble, as it is already inserted by *BiocStyle*. See Appendix B for a list of other automatically attached L^AT_EX packages.

1.1.2 *knitr*

To use with *knitr*, add the following to the 'DESCRIPTION' file:

```
VignetteBuilder: knitr  
Suggests: BiocStyle, knitr
```

this to the top of the .Rnw file:

```
%\VignetteEngine{knitr::knitr}
```

and this to the preamble:

```
<<style-knitr, eval=TRUE, echo=FALSE, results="asis">>=  
BiocStyle::latex2()  
@
```

See `?latex2` for additional options, and Appendix B for a list of automatically attached L^AT_EX packages.

1.1.3 R markdown

To enable *BiocStyle* in your R markdown vignette you need to:

1. Edit to the 'DESCRIPTION' file by adding

Bioconductor L^AT_EX Style 2.0

```
VignetteBuilder: knitr
Suggests: BiocStyle, knitr, rmarkdown
```

2. Specify `BiocStyle::pdf_document2` as the output format and add vignette metadata to the document header:

```
---
title: "Vignette Title"
author: "Vignette Author"
output:
  BiocStyle::pdf_document2
vignette: >
  %\VignetteIndexEntry{Vignette Title}
  %\VignetteEngine{knitr::rmarkdown}
  %\VignetteEncoding{UTF-8}
---
```

The vignette section is required in order to instruct *R* how to build the vignette. Note that `\VignetteIndexEntry` should match the title of your vignette.

See the *Bioconductor R markdown style* package `vignette`² for more details on authoring markdown documents with *BiocStyle*. ²in preparation

1.2 Workflow

Provided that *BiocStyle* has been installed, a convenient way to build the vignette as it is being written is with the command

```
R CMD Sweave --pdf vignette.Rnw
```

A short-cut useful for checking the L^AT_EX portion of vignettes is to toggle evaluation of code chunks to `FALSE`.

```
SWEAVE_OPTIONS="eval=FALSE" R CMD Sweave --pdf vignette.Rnw
```

When using *knitr*, the command to process the vignette is

```
R CMD Sweave --engine=knitr::knitr --pdf vignette.Rnw
```

For *R* markdown vignettes, set the engine to `knitr::rmarkdown`.

```
R CMD Sweave --engine=knitr::rmarkdown --pdf vignette.Rmd
```

By default, *knitr* automatically caches results of vignette chunks, greatly accelerating the turnaround time required for edits. Both the default and *knitr* incantations create PDF files using `texi2dvi --pdf`; many versions of this software incorrectly display non-breaking spaces as a tilde, `~`. This can be remedied (as on the *Bioconductor* build system) with a final run of

```
R CMD texi2dvi --pdf vignette.tex
R CMD pdflatex vignette.tex
```

2 Structuring the document

2.1 Title page

Important: most of the methods described here only work once BiocStyle is loaded; therefore, specify title and authors after the code chunk calling `BiocStyle::latex2()`.

2.1.1 Title and running headers

Create a title and running headers by defining `\biocTitle` and `\author` in the preamble

```
\biocTitle[Short title for headers]{Full title for title page}
%% also: \biocTitle{Title used for both header and title page}
%% or... \title{Title used for both header and title page}
\author{Vignette Author\thanks{\email{user@domain.com}}}
```

and use `\maketitle` at the beginning of the document to output the title and author information.

2.1.2 Authors and affiliations

As illustrated above, use the `\email` command to enter hyperlinked email addresses typeset in monospace font. Multiple author affiliations can be specified with the help of the L^AT_EX package `authblk` which is automatically loaded by *BiocStyle*. See the following examples for typical use cases.

If your vignette has just a single author, use standard L^AT_EX syntax to enter the author and affiliation information separated by a new line. You can provide the email address in `\thanks`, or in another new line after the affiliation.

```
\author{Vignette Author\thanks{\email{user@domain.com}}\
      Author's Affiliation}
```

In case of multiple authors, you can specify the author and affiliation information in different manners depending on the number of affiliations. In the basic case, when you don't provide authors' affiliation, or all the authors share the same affiliation, you can just use the regular `\author` command:

```
\author{
  First Author\thanks{\email{first@author.com}},
  Second Author\thanks{\email{second@author.com}}\
  Shared Affiliation
}
```

To provide different affiliations, some of which are potentially shared, enter each author separately and use `\affil` to specify affiliation. The authors will appear all in one (possibly continued) line with automatic footnotes, and the affiliations are displayed in separate lines below.³

³as on this vignette's title page

Bioconductor L^AT_EX Style 2.0

```
\author{First Author}
\author{Second Author}
\affil{First and Second authors' shared affiliation}
\author{Third Author\thanks{\email{corresponding@author.com}}}
\affil{Third author's affiliation}
```

In the case when some authors have more than one affiliation, or the authors with a shared affiliation do not come one after another, you need to manually define footnote markers as optional arguments to `\author` and `\affil` commands, as in the following example.

```
\author[1,2]{First Author\thanks{\email{user@domain.com}}}
\author[1]{Second Author}
\author[2]{Third Author}
\affil[1]{First and Second authors' affiliation}
\affil[2]{First and Third authors' affiliation}
```

2.1.3 Abstract, package version, and table of contents

Abstract can be entered using the standard abstract environment:

```
\begin{abstract}
  Short summary of the vignette's content.
\end{abstract}
```

It is recommended to include information on the specific package version described in the vignette. The following line⁴ inserts this information automatically and in a properly formatted manner.

```
\packageVersion{\Sexpr{BiocStyle::pkg_ver("packageName")}}
```

⁴substitute `packageName` by the name of your package

Use `\tableofcontents` to include a hyperlinked table of contents (TOC), and `\section`, `\subsection`, `\subsubsection` for structuring your document. It is a good practice to start the first section on a new page by calling `\newpage` after the TOC:

```
\tableofcontents
\newpage
\section{My First Section}
```

2.2 Style macros

BiocStyle introduces the following additional markup styling commands useful in typical *Bioconductor* vignettes.

Software:

- `\R{}` and `\Bioconductor{}` to reference *R* software and the *Bioconductor* project.
- `\software{GATK}` to reference third-party software, e.g., *GATK*.

Packages:

- `\Biocpkg{IRanges}` for *Bioconductor* software, annotation and experiment data packages, including a link to the release landing page or if the package is only in devel, to the devel landing page. *IRanges*.
- `\CRANpkg{data.table}` for *R* packages available on CRAN, including a link to the FHCRC CRAN mirror landing page, *data.table*.
- `\Githubpkg{rstudio/rmarkdown}` for *R* packages available on GitHub, including a link to the package repository, *rmarkdown*.
- `\Rpackage{MyPkg}` for *R* packages that are *not* available on *Bioconductor* or CRAN, *MyPkg*.

Code:

- `\Rfunction{findOverlaps}` for functions `findOverlaps`.
- `\Robject{olaps}` for variables `olaps`.
- `\Rclass{GRanges}` when referring to a formal class *GRanges*.
- `\Rcode{log(x)}` for *R* code, `log(x)`.

Communication:

- `\bioccomment{additional information for the user}` communicates *comment: additional information for the user*.
- `\warning{common pitfalls}` signals *warning: common pitfalls*.
- `\fixme{incomplete functionality}` provides an indication of *fixme: incomplete functionality*.

General:

- `\email{user@domain.com}` to provide a linked email address, `user@domain.com`.
- `\file{script.R}` for file names and file paths 'script.R'.

2.3 Code chunks

BiocStyle automatically adjusts the line length⁵ of output code chunks depending on the document's font size setting. Therefore, it shouldn't be necessary to set it manually. If for some reason you wish to alter it, use the `width` argument in the call to `BiocStyle::latex2`.

⁵`options("width")`

2.4 Figures

In addition to the standard `\figure` environment *BiocStyle* provides special environments for small and wide figures. These three environments differ in size, default aspect ratio⁶, and horizontal placement on the the page. See the following table for details, and Figures 1, 2 and 3 for examples.

⁶device dimensions are automatically set only by *knitr* and *rmarkdown*, and not by *Sweave*

environment	description	output width	device dimensions	aspect ratio
figure	regular figure	100mm (≈ 3.9 in)	7.5 × 5.0 in	1.5
figure*	wide figure	140mm (≈ 5.5 in)	10.0 × 5.0 in	2.0
smallfigure	small figure	65mm (≈ 2.6 in)	5.0 × 5.0 in	1.0

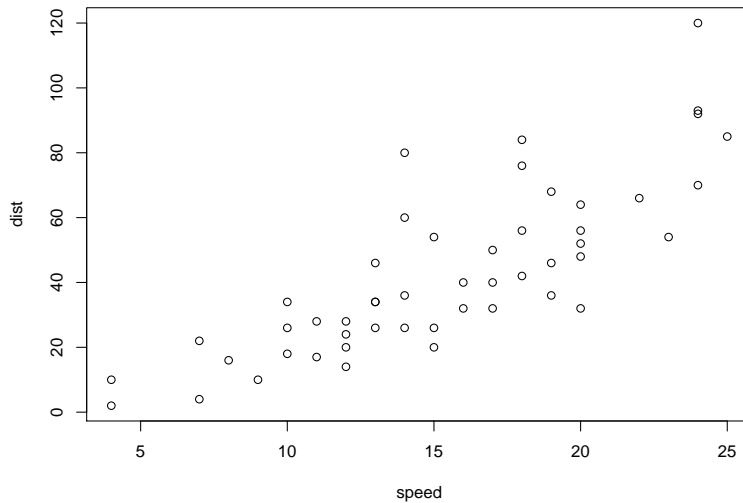


Figure 1: Regular figure. A plot displayed with the `\figure` environment.

The `figure` environment produces regular figures which are about 0.8 paragraph wide and right-aligned with the margin, as Figure 1. To utilize the whole available width use the `figure*` environment. It produces figures which are about 1.2 paragraph wide and extend on the right margin (Figure 2). The `smallfigure` environment is meant for possibly rectangular plots which are about half as wide as the

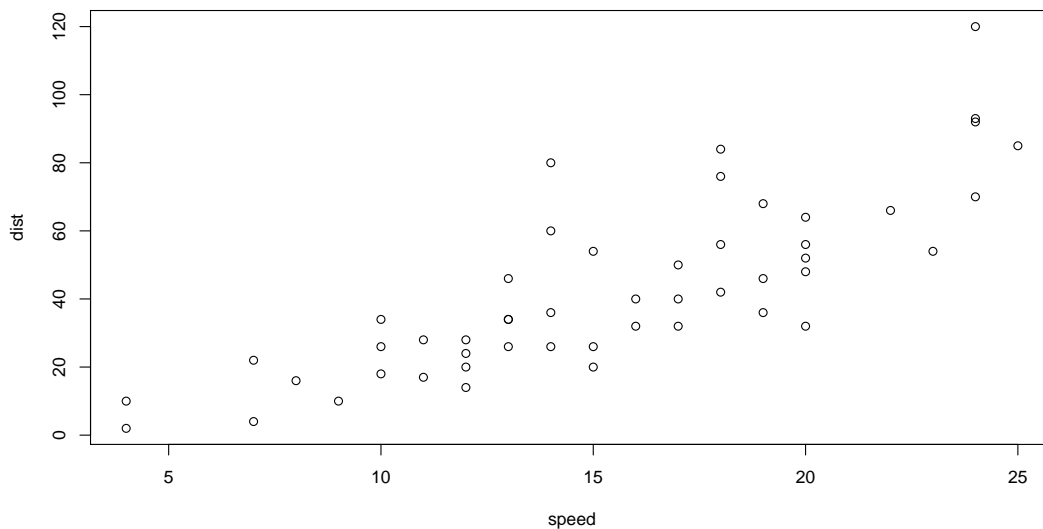


Figure 2: Wide figure. A plot displayed with the `\figure*` environment.

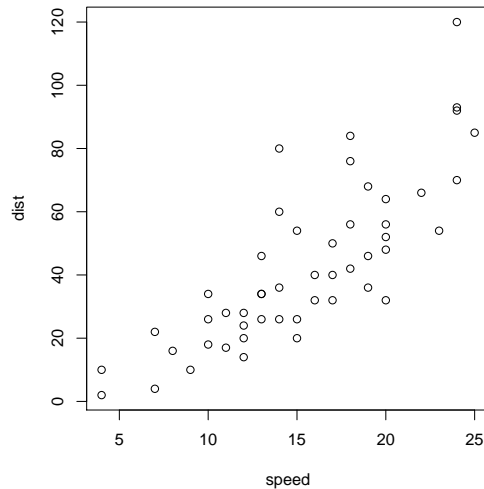


Figure 3: Small figure. A plot displayed with the `\smallfigure` environment.

paragraph (Figure 3). The default placement specifier for *BiocStyle* floats is `htbp`, which typically outputs them in the place where they are defined. The first sentence of figure captions is emphasized to serve as figure title. This feature can be disabled by setting the argument `titlecaps=FALSE` in the call to `BiocStyle::latex2`.

To use figure environments in *Sweave*, write explicit L^AT_EX code which inserts the figure in combination with the *Sweave* option `include=FALSE`. For example, Figure 2 was produced with the following code.

```
<<widefig, echo=FALSE, fig=TRUE, include=FALSE, width=10, height=5>>=
par(mar=c(4,4,0.5,0.5))
plot(cars)
@

\begin{figure*}
  \includegraphics{\jobname-widefig}
  \caption{\label{fig:wide}Wide figure. A plot displayed
    with the {\ttfamily\textbackslash figure*} environment.}
\end{figure*}
```

In *knitr* and *rmarkdown* the output environment for a graphics-producing code chunk can be specified in `fig.env` chunk option, e.g. set `fig.env='smallfigure'` to get `\begin{smallfigure}`. It is also possible to specify the wide and small figure environments by setting `fig.wide` or `fig.small` option to `TRUE`. The following two *knitr* code chunks are equivalent and produce the same output, similar to Figure 3.

```
<<smallfig, fig.cap="Small figure.", fig.env="smallfigure">>=
plot(cars)
@

<<smallfig, fig.cap="Small figure.", fig.small=TRUE>>=
plot(cars)
@
```

When *knitr* is used plots are cropped so that blank margins around them are removed to make better use of the space in the output document⁷. This feature can be switched off by setting the chunk option `knitr::opts_chunk$set(crop=NULL)`.

⁷otherwise one often has to struggle with `par` to set appropriate margins

2.4.1 `\incfig` convenience macro

Besides the usual L^AT_EX capabilities (the figure environment and `\includegraphics` command) *BiocStyle* defines a macro

```
\incfig[placement]{filename}{width}{title}{caption}
```

which expects four arguments:

filename The name of the figure file, also used as the label by which the float can be referred to by `\ref{}`. Some *Sweave* and *knitr* options place figures in a subdirectory; unless `short.fignames=TRUE` is set the full file name, including the subdirectory and any prefixes, should be provided. By default, these are '`<sweavename>`' for *Sweave* and '`figure/`' for *knitr*. Also note the different naming scheme used by *knitr*: figure files are named '`<chunkname>-i`' where *i* is the number of the plot generated in the chunk.

width Figure width.

title A short title, used in the list of figures and printed in bold as the first part of the caption.

caption Extended description of the figure.

The optional **placement** specifier controls where the figure is placed on page; it takes the usual values allowed by L^AT_EX floats, i.e., a list containing t, b, p, or h, where letters enumerate permitted placements⁸.

⁸if no placement specifier is given, the default `htbp` is assumed

For `incfig` with *Sweave* use

```
<<figureexample, fig=TRUE, include=FALSE, width=4.2, height=4.6>>=
par(mar=c(4,4,0.5,0.5))
v = seq(0, 60i, length=1000)
plot(abs(v)*exp(v), type="l", col="Royalblue")
@
\incfig{\jobname-figureexample}{0.5\textwidth}{A curve.}
{The code that creates this figure is shown in the code chunk.}
as shown in Figure-\ref{\jobname-figureexample}.
```

This results in

```
> par(mar=c(4,4,0.5,0.5))
> v = seq(0, 60i, length=1000)
> plot(abs(v)*exp(v), type="l", col="Royalblue")
```

as shown in Figure 4.

When option `short.fignames=TRUE` is set, figure names used by `\incfig` and `\ref` do not contain any prefix and are identical to the corresponding code chunk labels

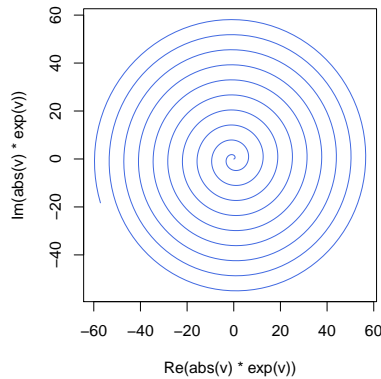


Figure 4: A curve. The code that creates this figure is shown in the code chunk.

(plus figure number in case of *knitr*). For example, in *Sweave* the respective code for the above example would be `\incfig{figureexample}{...}{...}{...}` and `\ref{figureexample}`.

For `\incfig` with *knitr*, use the option `fig.show='hide'` rather than `include=FALSE`. The *knitr*-equivalent code for Figure 4 is:

```
<<figureexample, fig.show='hide', fig.width=4, fig.height=4>>=
par(mar=c(4,4,0.5,0.5))
v = seq(0, 60i, length=1000)
plot(abs(v)*exp(v), type="l", col="Royalblue")
@
```

Note the difference in option names setting the figure width and height compared to *Sweave*. Unless `short.fignames=TRUE` is set, use the default 'figure/' prefix when inserting and referring to figures, e.g.:

```
\incfig{figure/figureexample-1}{0.5\textwidth}{A curve.}
{The code that creates this figure is shown in the code chunk.}
```

A custom prefix for figure file names can be passed to `latex2` in the `fig.path` argument. When `short.fignames=TRUE`, figures can be referred to directly by code chunk labels, i.e., `\incfig{figureexample-1}...` and `\ref{figureexample-1}`.

2.5 Equations

When referencing equations, e.g. **1**, use `\eqref` to ensure proper label formatting.

1 $\sin^2 \theta + \cos^2 \theta \equiv 1$

2.6 Footnotes

One of the distinctive features of the style is an asymmetric page layout with a wide margin on the right. This provides additional space for ancillary information in side notes. These can be entered in footnotes⁹ typeset as margin notes, which has the advantage that the notes appear close to the place where they are defined.

⁹this is a side note
entered using the
`\footnote` command

2.7 Bibliography

`BiocStyle::latex2` has default argument `use.unsrurl=TRUE` to automatically format bibliographies using *natbib*'s `unsrurl` style. There is no need to explicitly include *natbib*, and it is an error to use a `\bibliographystyle` command. The *unsrurl.bst* format, e.g., [1, 2], supports hyperlinks to DOI and PubMed IDs but not `\citet` or `\citep`.

To use a bibliography style different from `unsrurl`, set `use.unsrurl=FALSE` and follow normal L^AT_EX conventions.

References

- [1] Friedrich Leisch. *Compstat: Proceedings in Computational Statistics*, chapter Sweave: Dynamic Generation of Statistical Reports Using Literate Data Analysis, pages 575–580. Physica-Verlag HD, Heidelberg, 2002. URL: <http://leisch.userweb.mwn.de/Sweave/>, doi:10.1007/978-3-642-57489-4_89.
- [2] Yihui Xie. knitr: A comprehensive tool for reproducible research in R. In Victoria Stodden, Friedrich Leisch, and Roger D. Peng, editors, *Implementing Reproducible Computational Research*. Chapman and Hall/CRC, 2014. ISBN 978-1466561595. URL: <http://www.crcpress.com/product/isbn/9781466561595>.

A Session info

Here is the output of `sessionInfo()` on the system on which this document was compiled:

- R version 3.3.1 (2016-06-21), x86_64-pc-linux-gnu
- Locale: LC_CTYPE=en_US.UTF-8, LC_NUMERIC=C, LC_TIME=en_US.UTF-8, LC_COLLATE=C, LC_MONETARY=en_US.UTF-8, LC_MESSAGES=en_US.UTF-8, LC_PAPER=en_US.UTF-8, LC_NAME=C, LC_ADDRESS=C, LC_TELEPHONE=C, LC_MEASUREMENT=en_US.UTF-8, LC_IDENTIFICATION=C
- Base packages: base, datasets, grDevices, graphics, methods, stats, utils
- Other packages: BiocStyle 2.0.3
- Loaded via a namespace (and not attached): Rcpp 0.12.6, digest 0.6.10, evaluate 0.9, formatR 1.4, htmltools 0.3.5, knitr 1.13, magrittr 1.5, rmarkdown 1.0, stringi 1.1.1, stringr 1.0.0, tools 3.3.1, yaml 2.1.13

B Attached L^AT_EX packages

BiocStyle loads the following L^AT_EX packages:

authblk, beramono, caption, changepage, color, enumitem, fancyhdr, fontenc, footmisc, framed, geometry, graphicx, hyperref, ifthen, mathtools, parnotes, parskip, placeins, ragged2e, soul, titlesec, titletoc, xstring.

C Known issues

C.1 Use with L^AT_EX package `float`

The current implementation of *BiocStyle* is not compatible with the L^AT_EX package `float`. Please refrain from using `float`, in particular the `[H]` placement specifier.

C.2 Compatibility with the *xtable* R package

BiocStyle does not support tables produced by the R package *xtable* in plain *Sweave* documents. This limitation does not apply to documents compiled with *knitr*. If you would like to use *xtable* in your *BiocStyle*-enabled document, please consider using *knitr*¹⁰ instead of *Sweave*.

¹⁰see Sections 1.1.2 and 2.4 for details on using *BiocStyle* with *knitr*