

# Cafebr: Citation Amender/Formatter for Biological Research

Daisuke Tsugama<sup>1</sup>

<sup>1</sup> Research Faculty of Agriculture, Hokkaido University

DOI: [10.21105/joss.00912](https://doi.org/10.21105/joss.00912)

## Software

- [Review ↗](#)
- [Repository ↗](#)
- [Archive ↗](#)

Submitted: 23 August 2018

Published: 24 August 2018

## License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

## Summary

A reference list is an essential part of a manuscript for an academic article. Generating a reference list consists of three steps: (1) Finding and/or collecting appropriate articles to cite, (2) extracting necessary pieces of information from them, and (3) outputting these pieces of information in a style specific to a journal to which the manuscript is submitted. The steps 2 and 3 are error-prone if manually performed. To perform these steps with less errors, reference management software packages have been developed. EndNote (Thomson Reuters 1995) is an example of such packages. It allows users to search for articles of interest, to deposit them in a database, and to output information about selected articles in a style chosen from more than 2000 preset styles. However, it is a commercial, non-free program, and incompatible with Linux operating systems. Zotero (Center for History and New Media at George Mason University 2006) and Mendeley (Mendeley Ltd. 2008) are examples of popular free reference manager packages that can implement the functions similar to those of EndNote, and can run on many platforms such as Windows, Mac OS X, and Linux. All of these programs have many functions and user-friendly interfaces, and have still been improving. However, such multifunctionality seems to have been counteracting with simplicity: It requires many clicks to obtain a final output, and even making small changes in the output requires users to look into and edit a file outside a main execution file. In reality, it is often necessary to reformat a formatted reference list (for example, when a manuscript is declined by a journal and it is submitted to another journal), but such reformatting with the above packages is also not very simple. Here, the author introduces Cafebr (Citation Amender/Formatter for Biological Research), which was developed to simplify the process of generating a reference list. Cafebr was originally written in Perl, and has been converted to a GUI version, cafebr.html, using HTML and JavaScript. As such, it can run on web browsers (at least Firefox, Chrome, and Microsoft Edge) in any platforms.

To achieve the above step 1 (i.e., for articles), Cafebr can search the MEDLINE database with the PubMed search engine (The United States National Library of Medicine (NLM) at the National Institutes of Health (NIH) 1996), which covers articles with a wide variety of biology fields and is therefore most commonly used for biological research. Cafebr then extracts pieces of information on articles from search results, and displays them in a table for selecting articles for further processing, achieving the above step 2. It is also possible for Cafebr to extract such pieces of information from texts provided by users. Extracted pieces of information should be precise if the provided texts are in the PubMed XML format, the PubMed abstract (text) format, or the Cafebr database format, in which one line corresponds to one article record consisting of 12 tab-delimited data fields (Authors, Article Title, Publication Year, Journal name, Volume, Issue, Pages, PubMed ID (PMID), PubMed Central ID (PMCID), DOI, Attributes, Author Information (affiliation etc.)). Information can be extracted even from texts in none of these formats if delimiters are

provided on Cafebr. For example, the reference “Tsugama D (2018) Cafebr development.” can be delimited by the three fields “Tsugama D”, “2018” and “Cafebr development” if the delimiting pattern “F1 (F2) F3.” is provided. In this “Delimiter” option, it is also possible to use a journal name as a delimiter if the journal name is available in PubMed. All journal names are listed in the cafebr.html file itself. The journal names are major contributors to the large (~2-MB) file size of cafebr.html, but allows it to be stand-alone.

The extracted pieces of information are organized to generate a final list, achieving the above steps 2 and 3. For this, a template for field arrangement and an order of references can be designated. Some of the preset templates add HTML tags to specific fields to display them in the italic, bold, or superscript style on a web browser. Microsoft Word, which would be usually used for the final manual formatting of a manuscript, can maintain such a style if such words are copied and pasted with the “Match Destination Formatting” option. The number of preset templates for field arrangement is only five thus far (will be increased in the future). However, it is possible to edit them directly on the interface of Cafebr, and to get flexible results. As an ordering option, “As in manuscript” is available. This option outputs references as they are cited in the manuscript, and converts citations such as “(Tsugama et al., 2018)”, “(Mike et al., 2014)” to “[Ref1]”, “[Ref2]”. All of these functions except the PubMed search can be locally executed with the single file cafebr.html. Its online version can execute all of the functions with the aid of a CGI program (cafebr.cgi), and is available at either <http://stdtgm.itigo.jp/cafebr/cafebr.html> (Tsugama 2018a) or <http://studtsugama.s1006.xrea.com/cafebr/cafebr.xhtm> (Tsugama 2018b).

**Figure 1. Example of use of the PubMed searching option.**

**Figure 2. Example of use of the Delimiter searching option.**

## References

- Center for History and New Media at George Mason University. 2006. “Zotero.” 2006. <https://www.zotero.org>.
- Mendeley Ltd. 2008. “Mendeley.” 2008. <https://www.mendeley.com>.
- The United States National Library of Medicine (NLM) at the National Institutes of Health (NIH). 1996. “PubMed.” 1996. <https://www.ncbi.nlm.nih.gov/pubmed>.
- Thomson Reuters. 1995. “EndNote.” 1995. <https://endnote.com>.
- Tsugama, Daisuke. 2018a. “Cafebr – Citation Amender/Formatter for Biological Research.” 2018. <http://stdtgm.itigo.jp/cafebr/cafebr.html>.
- . 2018b. “Cafebr – Citation Amender/Formatter for Biological Research.” 2018. <http://studtsugama.s1006.xrea.com/cafebr/cafebr.xhtm>.

### Cafebr

-- Citation Amender/Formatter for Biological Research

(c) 2018 Daisuke Tsugama ([MIT License](#) applies)  
 This program uses the PubMed search engine (<https://www.ncbi.nlm.nih.gov/pubmed/>),  
 managed by The United States National Library of Medicine (NLM) at the National Institutes of Health (NIH)

[Download a stand-alone version](#) > [Download a command-line version](#) > [Help. Cafebrman](#)

**1. Choose how to collect and/or handle input references** ([help](#))

PubMed searching    DB browsing    Delimiter  
 PubMed XML correction    PubMed abstract correction

Keywords:  X

View in: [PubMed](#) [PubMed XML](#) [Cafebr \(up to 20 records\)](#) [Cafebr \(> 20 records\)](#)

**2. Give and edit input references** ([help](#))

Select a file  No file selected    Add new input to old   [Download/Add to the database](#)

Kansup J, Tsugama D, Liu S, Takano T. The Arabidopsis adaptor protein AP-3 $\mu$  interacts with the G-protein  $\beta$  subunit AGB1 and is involved in abscisic acid regulation of germination and post-germination development. 2013. *J. Exp. Bot.* 64, 18, 5611-5621  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3871816/> <https://doi.org/10.1093/jxb/ert327>  
 Journal of Experimental Botany. U.S. Gov't Author. Information. (1)Asian Natural Environmental Science Center, The University of Tokyo, Nishitokyo, Tokyo 188-0002, Japan.  
 Heterotrimeric G-proteins (G-proteins) have been implicated in ubiquitous signalling mechanisms in

Keywords:  X   [Use the above for step 3](#) [Clear input](#)

**3. Choose references for formatted output** (click references to select/unselect)

Shown is Page 1 (records 1-4) of  
 Pages: 1

Records per page:  5    20    50    200    All

[Select all](#) [Clear selection](#) [Add selected references to output](#) [Hide the table](#)

Author	Title	Published in	Links	Abstract etc.
Kansup J, Tsugama D, Liu S, Takano T	The Arabidopsis adaptor protein AP-3 $\mu$ interacts with the G-protein $\beta$ subunit AGB1 and is involved in abscisic acid regulation of germination and post-germination development	2013 <i>J. Exp. Bot.</i> 64, 18 5611-5621	PubMed: <a href="#">24098050</a> PMC: <a href="#">PMC3871816</a> DOI: <a href="#">10.1093/jxb/ert327</a>	<a href="#">Open</a>
Tsugama D, Liu S, Takano T	Metal-binding ability of VIP1: a bZIP protein in <i>Arabidopsis thaliana</i>	2013 <i>Protein J</i> 32, 7 526-532	PubMed: <a href="#">24057918</a> PMC: NA DOI: <a href="#">10.1007/s10930-013-9512-3</a>	<a href="#">Open</a>
Tsugama D, Liu S, Takano T	A bZIP protein, VIP1, interacts with Arabidopsis heterotrimeric G protein $\beta$ subunit, AGB1	2013 <i>Plant Physiol. Biochem.</i> 71 240-246	PubMed: <a href="#">23974356</a> PMC: NA DOI: <a href="#">10.1016/j.plaphy.2013.07.024</a>	<a href="#">Open</a>
Tsugama D, Liu S, Takano T	Arabidopsis heterotrimeric G protein $\beta$ subunit, AGB1, regulates brassinosteroid signalling independently of BZR1	2013 <i>J. Exp. Bot.</i> 64, 11 3213-3223	PubMed: <a href="#">23814276</a> PMC: <a href="#">PMC3733148</a> DOI: <a href="#">10.1093/jxb/ert159</a>	<a href="#">Open</a>

[Select all](#) [Clear selection](#) [Add selected references to output](#) [Hide the table](#)

Records per page:  5    20    50    200    All

Shown is Page 1 (records 1-4) of  
 Pages: 1

**4. Choose a style for output references** ([help](#))

Cafebr default    PSB    FEBS    NPG    Custom

Output style:

!AUTH. !(YEAR) !TITLE, <i>!JOURNAL</i>, !VOL, !PAGE...

[Add selected references to output](#)

**5. Choose how to order output references**

As input    As in a manuscript    Author names

Publication year (ascending)    Publication year (descending)

PubMed ID (ascending)    PubMed ID (descending)

[Add selected references to output](#) [Reorder](#)

**6. Review and edit formatted references**

Add new output to old

Kansup J, Tsugama D, Liu S & Takano T. (2013) The Arabidopsis adaptor protein AP-3 $\mu$  interacts with the G-protein  $\beta$  subunit AGB1 and is involved in abscisic acid regulation of germination and post-germination development. *J. Exp. Bot.* 64, 5611-5621.

Tsugama D, Liu S & Takano T. (2013) A bZIP protein, VIP1, interacts with Arabidopsis heterotrimeric G protein  $\beta$  subunit, AGB1. *Plant Physiol. Biochem.* 71, 240-246.

Tsugama D, Liu S & Takano T. (2013) Arabidopsis heterotrimeric G protein  $\beta$  subunit, AGB1, regulates brassinosteroid signalling independently of BZR1. *J. Exp. Bot.* 64, 3213-3223.

[Reorder](#) [Clear output](#) [Download as a file](#) [View in a new window](#)

Search results

Click to unselect/select references for output

Choose a template and edit it

Designate how to order references

Review and edit output references

Figure 1: Example of use of the PubMed searching option.

Tsugama, (2018). Cafebr: Citation Amender/Formatter for Biological Research. *Journal of Open Source Software*, 3(28), 912. <https://doi.org/10.21105/joss.00912>

3

**1. Choose how to collect and/or handle input references ([help](#))**

PubMed searching  DB browsing  Delimiter  
 PubMed XML correction  PubMed abstract correction

Delimiters for input:  
 X

Processing. Please wait ...

**2. Give and edit input references ([help](#))**

Select a file No file selected  Add new input to collection  
 Kansup J., Tsugama D., Liu S., Takano T. The Arabidopsis AP-3 $\beta$ -subunit interacts with the G-protein  $\beta$  subunit AGB1 and is involved in abscisic acid regulation of germination and post-germination development. *J. Exp. Bot.* 2013; 64: 5611-5621. PMID: 24268050  
 Tsugama D., Liu S., Takano T. A bZIP protein, VIP1, interacts with Arabidopsis heterotrimeric G protein  $\beta$  subunit, AGB1, regulates brassinosteroid signalling independently of BZR1. *J. Exp. Bot.* 2013; 64: 3213-3223. PMID: 23814276

Keywords:  X

**3. Choose references for formatted output (click references to select/unselect)**

Shown is Page 1 (records 1-3) of Pages: 1  
 Records per page:  5  20  50  200  All

<input type="checkbox"/> Select all	<input type="checkbox"/> Clear selection	<input type="checkbox"/> Add selected references to output	<input type="checkbox"/> Hide the table				
Field 1 (F1)	Field 2 (F2)	Field 3 (F3)	Field 4 (F4)	Field 5 (F5)	Field 6 (F6)	Field 7 (F7)	Field 8 (F8)
Kansup J., Tsugama D., Liu S., Takano T	The Arabidopsis adaptor protein AP-3 $\beta$ interacts with the G-protein $\beta$ subunit AGB1 and is involved in abscisic acid regulation of germination and post-germination development	<i>J. Exp. Bot.</i> 2013; 64: 5611-5621. PMID: 24268050					
Tsugama D., Liu S., Takano T	A bZIP protein, VIP1, interacts with Arabidopsis heterotrimeric G protein $\beta$ subunit, AGB1, independently of BZR1	<i>Plant Physiol. Biochem.</i> 2013; 71: 240-246. PMID: 23974356					
Tsugama D., Liu S., Takano T	Arabidopsis heterotrimeric G protein $\beta$ subunit, AGB1, regulates brassinosteroid signalling independently of BZR1	<i>J. Exp. Bot.</i> 2013; 64: 3213-3223. PMID: 23814276					

Records per page:  5  20  50  200  All  
 Shown is Page 1 (records 1-3) of Pages: 1

**Fields resulting from the delimiters**

**Output 1 (output style is: !FAUTH. !FTITLE. <i>!FJOURNAL</i> <b>!FVOL</b>, !FPAGE (!FYEAR).')**

**4. Choose a style for output references ([help](#))**

Cafebr default  PSB  FEBS  NPG  Custom

Output style:  
 !FAUTH !FTITLE <i>!FJOURNAL</i> <b>!FVOL</b>, !FPAGE  
 (!FYEAR)

Authors	Article Title	Year	Journal	Volume	Issue	Pages	DOI	PubMed ID	PMC ID	Abstract	Others
1	2	4	3	5	6	8					

Add selected references to output  Reorder

**5. Choose how to order output references**

As input  As in a manuscript  Author names  
 Publication year (ascending)  Publication year (descending)  
 PubMed ID (descending)  PubMed ID (descending)

Add selected references to output  Reorder

**6. Review and edit formatted references**

Add new output to old

Tsugama, D., Liu, S., Takano, T. Arabidopsis heterotrimeric G protein  $\beta$  subunit, AGB1, regulates brassinosteroid signalling independently of BZR1. *J. Exp. Bot.* 64, 3213-3223 (2013).  
 Tsugama, D., Liu, S., Takano, T. A bZIP protein, VIP1, interacts with Arabidopsis heterotrimeric G protein  $\beta$  subunit, AGB1. *Plant Physiol. Biochem.* 71, 240-246 (2013).  
 Kansup, J., Tsugama, D., Liu, S., Takano, T. The Arabidopsis adaptor protein AP-3 $\beta$  interacts with the G-protein  $\beta$  subunit AGB1 and is involved in abscisic acid regulation of germination and post-germination development. *J. Exp. Bot.* 64, 5611-5621 (2013).

**Output 2 (output style is: 'F1 worked together to publish a paper (F2) in F3 in F4.'**

**4. Choose a style for output references ([help](#))**

Cafebr default  PSB  FEBS  NPG  Custom

Output style:  
 F1 worked together to publish a paper (F2) in F3 in F4

Authors	Article Title	Year	Journal	Volume	Issue	Pages	DOI	PubMed ID	PMC ID	Abstract	Others
1	2	4	3	5	6	8					

Add selected references to output  Clear Field numbers

**5. Choose how to order output references**

As input  As in a manuscript  Author names  
 Publication year (ascending)  Publication year (descending)  
 PubMed ID (descending)  PubMed ID (descending)

Add selected references to output  Reorder

**6. Review and edit formatted references**

Add new output to old

Kansup J., Tsugama D., Liu S., Takano T. worked together to publish a paper (The Arabidopsis adaptor protein AP-3 $\beta$  interacts with the G-protein  $\beta$  subunit AGB1 and is involved in abscisic acid regulation of germination and post-germination development) in J. Exp. Bot. in 2013.  
 Tsugama D., Liu S., Takano T. worked together to publish a paper (A bZIP protein, VIP1, interacts with Arabidopsis heterotrimeric G protein  $\beta$  subunit, AGB1) in Plant Physiol. Biochem. in 2013.  
 1. Kansup J., Tsugama D., Liu S., Takano T. worked together to publish a paper (The Arabidopsis adaptor protein AP-3 $\beta$  interacts with the G-protein  $\beta$  subunit AGB1 and is involved in abscisic acid regulation of germination and post-germination development) in J. Exp. Bot. in 2013.  
 2. Tsugama D., Liu S., Takano T. worked together to publish a paper (A bZIP protein, VIP1, interacts with Arabidopsis heterotrimeric G protein  $\beta$  subunit, AGB1) in Plant Physiol. Biochem. in 2013.  
 3. Tsugama D., Liu S., Takano T. worked together to publish a paper (Arabidopsis heterotrimeric G protein  $\beta$  subunit, AGB1, regulates brassinosteroid signalling independently of BZR1) in J. Exp. Bot. in 2013.

**Figure 2:** Example of use of the Delimiter searching option.