

Software Description

Cafebr - Citation Amender/Formatter for Biological Research

Daisuke Tsugama ‡

‡ Hokkaido University, Sapporo-shi, Japan

Corresponding author: Daisuke Tsugama (tsugamedaisaku@gmail.com)

Reviewable v1

Received: 14 Sep 2018 | Published: 26 Sep 2018

Citation: Tsugama D (2018) Cafebr - Citation Amender/Formatter for Biological Research. Research Ideas and Outcomes 4: e29773. <https://doi.org/10.3897/rio.4.e29773>

Abstract

Background

A reference list is an essential part of a manuscript for an academic article. It is often necessary to reformat (i.e., change orders of pieces of article information such as authors, article title, publication year, and journal where the article was published) of a preformatted reference list when a manuscript is declined by one journal and submitted to another. EndNote, Zotero and Mendeley are examples of sophisticated reference management programs that help generate a reference list with less errors. However, their multifunctionality seems to have complicated the process of generating a reference list: they require many selection steps to obtain a final output, and also require to manually edit a file outside the execution program even to make small changes in the output.

New information

The author developed a program, Cafebr (Citation Amender/Formatter for Biological Research), to more simply generate a reference list for an article of biological research. It is written in HTML/JavaScript, and as such works on a web browser on any platform. On Cafebr, articles for a final reference list can be either given by a user or collected from PubMed. Pieces of article information are then extracted according to the format of these articles or to the delimiters designated by a user. Preset formats for output are currently

only four, but all of them can be directly edited on the user interface, allowing to change output formats quickly and flexibly. All of the functions of Cafebr is available on its website (either <http://stdtgm.itigo.jp/cafebr/cafebr.html> (main) or <http://studtsugama.s1006.xrea.com/cafebr/cafebr.xhtml> (backup)) with aids of a CGI program. A stand-alone version of Cafebr is available at these websites or Zenodo ([10.5281/zenodo.1404887](https://doi.org/10.5281/zenodo.1404887)).

Keywords

Biology, HTML, JavaScript, Matching and replacement of characters, PubMed, Reference management software

Introduction

A reference list is an essential part of a manuscript for an academic article. Different journals require different formats for a reference list. Due to this situation, it is often necessary to reformat a preformatted reference list if a manuscript is declined by a journal for publication and submitted to another. It is error-prone and takes time to manually format a reference list. To quickly generate a reference list with less errors, reference management software packages have been developed. EndNote, which was developed by Clarivate Analytics (previously Thomson Reuters), 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 (<https://endnote.com/product-details>). However, it is a commercial, non-free program, and incompatible with Linux operating systems. Zotero (developed by Center for History and New Media at George Mason University in 2006; <https://www.zotero.org>) and Mendeley (developed by Mendeley Ltd., which was purchased by Elsevier; <https://www.mendeley.com>) are examples of free reference management programs that can implement the functions similar to those of EndNote, and can run on 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 selection steps (i.e., cursor movements and clicks) to obtain a final output; even making small changes in the output requires users to look into and edit a file outside a main execution file. To simplify the process of generating a reference list for biological research, the author developed a new program, Cafebr (Citation Amender/Formatter for Biological Research).

Project description

Title: Cafebr - Citation Amender/Formatter for Biological Research

Study area description: Cafebr was developed to more simply generate a reference list. It uses only PubMed (managed by The United States National Library of Medicine (NLM) at the National Institutes of Health (NIH), <https://www.ncbi.nlm.nih.gov/pubmed>) as a search engine to find articles of interest, and thus primarily for biological research. However, once a list (or database) of articles is provided, Cafebr can work on it regardless of whether the articles are those of biological research or not.

Design description: Cafebr was originally written in Perl, and most of the functions were copied to a GUI version written in HTML/JavaScript. This GUI version, as an HTML file, works on a web browser on any platform. To obtain an output reference list simply and flexibly, Cafebr focuses its functions on finding and replacing certain patterns of words rather than on managing (i.e., storing and displaying) articles.

To find articles, Cafebr can search the MEDLINE database with PubMed search engine (managed by The United States National Library of Medicine (NLM) at the National Institutes of Health (NIH), <https://www.ncbi.nlm.nih.gov/pubmed>), 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 (Fig. 1, upper part). Alternatively, a user can provide a list (or any texts) of articles for Cafebr. Cafebr can precisely extract pieces of information from texts 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.)). Texts in none of these formats can also be separated into pieces (fields) of information if delimiters are designated. For example, the reference "Tsunami D (2018) Cafebr development." can be separated into the three fields "Tsunami D", "2018" and "Cafebr development" if the delimiting pattern "F1 (F2) F3." is designated. 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 program (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 (Fig. 2, upper part).

The extracted pieces of information or fields are formatted to generate a final reference list. Currently only four preset formats are available. However, all of them can be directly edited on the user interface, allowing to change output formats quickly and flexibly (Fig. 1 and Fig. 2, lower parts). The article ordering options include not only those to order by author names, publication years and PubMed ID but also the "As in manuscript" option. This option outputs references as they are cited in the manuscript, and converts citations such as "(Tsunami et al., 2018)", "(Mike et al., 2014)" to "[Ref1]", "[Ref2]" in the manuscript. Some of the preset formats add HTML tags to specific fields of article information to display them in the italic, bold, or superscript style on a browser. Microsoft Word, which would be usually used for the final manual formatting of a whole manuscript, can maintain italic, bold and superscript styles if words in these styles are copied and pasted with the "Match

Destination Formatting" option. All of these functions of Cafebr are available on its website (either <http://stdtgm.itigo.jp/cafebr/cafebr.html> (main) or <http://studtsugama.s1006.xrea.com/cafebr/cafebr.xhtml> (backup) with the aid of a CGI program (cafebr.cgi or cafebr.xcg, respectively). A stand-alone version of Cafebr is available at these websites and Zenodo ([10.5281/zenodo.1404887](https://doi.org/10.5281/zenodo.1404887)).

Cafebr
-- Citation Amender/Formatter for Biological Research

(c) 2018 Daisuke Tsugama (MIT License applies)
The 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> -<Hello, Cafebrman>

1. Choose how to collect and/or handle input references (help)

PubMed searching DB browsing Delimiter
 PubMed XML correction PubMed abstract correction

Keywords:
tsugama+2013 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

Keywords: Input 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.
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Kanase J, Tsugama D, Liu S, Takano T	The Arabidopsis adaptor protein AP-3 β interacts with the G-protein β subunit AGB1 and is involved in abscisic acid regulation of germination and post-germination development. <i>J. Exp. Bot.</i> 64, 5611-5621	2013 <i>J. Exp. Bot.</i> 64, 16 5611-5621	PubMed: 24098050 PMC: F36/2871816 DOI: 10.1093/jxb/er3247	Open
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Tsugama D, Liu S, Takano T	Metal-binding ability of VIP1: a bZIP protein in Arabidopsis thaliana	2013 <i>Protein J.</i> 32, 7 1026-532	PubMed: 24557918 PMC: NA DOI: 10.1007/s10038-013-0613-3	Open
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Tsugama D, Liu S, Takano T	A bZIP protein, VIP1, interacts with Arabidopsis heterotrimeric G protein β subunit, AGB1	2013 <i>Plant Physiol. Biochem.</i> 71, 240-246	PubMed: 23974356 PMC: NA DOI: 10.1016/j.plaphy.2013.07.024	Open
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Tsugama D, Liu S, Takano T	Arabidopsis heterotrimeric G protein β subunit, AGB1, regulates brassinosteroid signalling independently of BZR1	2013 <i>J. Exp. Bot.</i> 64, 11 3213-3223	PubMed: 23814275 PMC: F36/2321166 DOI: 10.1093/jxb/er1159	Open

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:
[FATH, (FEBS) IFTITLE, (J) JOURNAL (J), IPVAL, IPAGE]

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

Recorded Clear output Download as a file View in a new window

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

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

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

Figure 1. [doi](https://doi.org/10.21956/2023.01.001)

Example of generation of a reference list from a PubMed search on Cafebr.

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:
 F1, F2, F3, F4, F5, F6, F7 X

Processing. Please wait ...

2. Give and edit input references (help)

Select a file No file selected

Keywords: X

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

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

Fields resulting from the delimiters

Field 1 (F1)	Field 2 (F2)	Field 3 (F3)	Field 4 (F4)	Field 5 (F5)	Field 6 (F6)	Field 7 (F7)	Field 8 (F8)
Kaneko J, Tsugama D, Liu S, Takano T. The Arabidopsis adaptor protein AP-3u interacts with the G-protein β 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: 24066090							
Tsugama D, Liu S, Takano T. A bZIP protein, VIP1, interacts with Arabidopsis heterotrimeric G protein β subunit, AGB1 in <i>Plant Physiol. Biochem.</i> 2013; 71: 246-246. PMID: 23874206							
Tsugama D, Liu S, Takano T. Arabidopsis heterotrimeric G protein β subunit, AGB1, regulates brassinosteroid signaling independently of BZR1. <i>J. Exp. Bot.</i> 2013; 64: 3213-3223. PMID: 23814276							

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

Output 1 (output style is: 'FAUTH, IFTITLE, <i>IFJOURNAL</i> IFVOL, !FPAGE (IFYEAR).')

4. Choose a style for output references (help)

Cafebr default FBS FBSB NPG Custom

Output style:
 IAUTH, IFTITLE, <i>IFJOURNAL</i>, IFVOL, !FPAGE (IFYEAR)

Authors (FAUTH)	Article Title (IFTITLE)	Year	Journal (IFJOURNAL)	Volume (IFVOL)	Issue (IFISSUE)	Pages (IFPAGE)	DOI (IFDOI)	PubMed ID (IFPMID)	PMID (IFPMID)	Abstract (IFABSTRACT)	Others (IFOTHERS)
1	2	4	3	5	6	8					

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)

6. Review and edit formatted references

Add new output to old

1. Tsugama D, Liu S, Takano T. Arabidopsis heterotrimeric G protein β subunit, AGB1, regulates brassinosteroid signaling independently of BZR1. *J. Exp. Bot.* 64: 3213-3223 (2013).

2. Tsugama D, Liu S, Takano T. A bZIP protein, VIP1, interacts with Arabidopsis heterotrimeric G protein β subunit, AGB1 in *Plant Physiol. Biochem.* 71: 246-246 (2013).

3. Kaneko J, Tsugama D, Liu S, Takano T. The Arabidopsis adaptor protein AP-3u interacts with the G-protein β 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 FBS FBSB NPG Custom

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

Authors (FAUTH)	Article Title (IFTITLE)	Year	Journal (IFJOURNAL)	Volume (IFVOL)	Issue (IFISSUE)	Pages (IFPAGE)	DOI (IFDOI)	PubMed ID (IFPMID)	PMID (IFPMID)	Abstract (IFABSTRACT)	Others (IFOTHERS)
1	2	4	3	5	6	8					

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)

6. Review and edit formatted references

Add new output to old

1. Kaneko J, Tsugama D, Liu S, Takano T worked together to publish a paper (The Arabidopsis adaptor protein AP-3u interacts with the G-protein β 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 β 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 β subunit, AGB1, regulates brassinosteroid signaling independently of BZR1) in *J. Exp. Bot.* in 2013.

Figure 2. doi Example of use of the "Delimiter" option on Cafebr.

Web location (URIs)

Homepage: <http://stdtgm.itigo.jp/cafebr/cafebr.html>

Download page: http://stdtgm.itigo.jp/cafebr/cafebr_gui.zip

Download mirror: http://studtsugama.s1006.xrea.com/cafebr/cafebr_gui.zip

Technical specification

Platform: Web browsers, Command line with Perl enabled

Programming language: HTML/JavaScript, Perl

Operational system: Windows, Mac, Linux

Interface language: HTML, plain text

Service endpoint: N/A

Repository

Type: Git

Browse URI: <https://github.com/Stu-DT/Cafebr>

Usage rights

Use license: Creative Commons Public Domain Waiver (CC-Zero)