

June 9, 2010  
Robert R. Kowal

## The Use of MicrosoftWord's Merge Feature for Extracting Data from an EXCEL File to Create Labels for Herbarium Specimens

MicrosoftWord can be used for extracting data from an EXCEL file to create labels for herbarium specimens. The template for the labels contains "merge codes", each of which corresponds to a column heading in the EXCEL file. On merging, each merge code is replaced by the data within the corresponding cell of the row corresponding to a collection ("recipient"; MicrosoftWord's terminology reflects the original purpose, which is to create documents related to mailings). Within a template, the merge codes are arranged in an appropriate way, separated by lines, punctuation, and strings (e.g., "Collection number: " and "Date: ") so as to form an acceptable herbarium label. The names of the column headings in the EXCEL file must correspond to the names used for merge codes in the label template, so each user must create his/her own customized template (or set of templates).

The following are the column headings that I use: NUMBER, COLLECTOR 1, COLLECTOR 2, COLLECTOR 3, TAXON, ? TAXON, ? SSP., ? VAR., DATE, COUNTRY, STATE, COUNTY, CITY, LOCATION 1, LOCATION 2, ALT., HABITAT 1, HABITAT 2, HABITAT 3, HABITAT 4, NOTES 1, NOTES 2, NOTES 3, NOTES 4, HERBARIUM, FL., FR., VEG., FIELD TRIP, T&R, T&R2, Lat&Long, LatDeg, LatMin, LatSec, LatNS, LongDeg, LongMin, LongSec, LongEW, LatDec, LongDec, # DUPS., STATUS. Note that in order for the table to accommodate many different kinds of information (e.g., different systems of coordinates giving locations), it has a plethora of column headings. For any given collection ("recipient"), many cells are empty.

The following is my most-used template for a label. "Merge codes" are represented by "<<xxx>>".

### WISCONSIN STATE HERBARIUM (WIS) University of Wisconsin-Madison

«TAXON»

«COUNTRY», «STATE», «COUNTY» Co., «CITY», «LOCATION\_1»  
«LOCATION\_2» «ALT» «TR» «LatLong»

«HABITAT\_1» «HABITAT\_2» «HABITAT\_3» «HABITAT\_4»

«NOTES\_1» «NOTES\_2» «NOTES\_3» «NOTES\_4»

«COLLECTOR\_1»: «NUMBER»

«DATE»

Note that the template uses different font sizes for different entries. Unfortunately, different fonts, like italic and bold, do not carry over from the source EXCEL file. One

could include them within the template for certain entries, the taxon name for example. However, because I include the author citation as part of the taxon name, I do not do this. If one were to have the taxon name, per se, and the author citation in different columns, italicizing and bolding the former would be feasible. Also note that the date merge code is included at the right-hand margin but on a separate line from the collector and the collection number. One can get them on the same line at either of the two opportunities to edit individual labels after the merge is completed, before printing.

The following is one of my labels, to give an idea of what the final product looks like.

WISCONSIN STATE HERBARIUM (WIS)

University of Wisconsin-Madison

*Packera paupercula* (Michx.) Á. Löve & D. Löve var.  
*savannarum* R.R. Kowal

U.S.A., WI, Jefferson Co., Red Cedar Lake, ca 4 mi E of Cambridge, on  
Hwy 12. T6N R13E S17&20

Savanna; abundant on hillsides facing lake. Grassy slopes beneath  
*Quercus macrocarpa* and depression between the hills.

Earliest anthesis. Individual 2 (normal sized). Voucher for chromosome  
count:  $2n = 22 II$ ; 3 cells counted (1972 Nov 22). On Spring Flora  
(Botany 401) examination of Kurt Schmidt & Vallerie Mellerop.

Kowal, Robert R.: 1583

1972 May 25

I have found it convenient to use, in general, a template corresponding to six labels per page (2 columns of 3 labels each). I print on sheets that are scored for easy tearing into six labels [University Products, Holyoke, MA: PERMA/DUR BOND LASER PRINT LABEL, 6-UP]. For a few humongous labels, I use a template corresponding to eight labels per page (2 columns of 4 labels each) but split the sheets into only four labels. For smaller labels I provide a template corresponding to eight labels per page (2 columns of 4 labels each). Although one can print on normal paper and cut it into labels, use of pre-scored sheets saves much time.

Links to six templates are provided: Herb Label Template 4x2.doc; Herb Label Template 3x2.doc; Herb Label Template 3x2 2 collectors.doc; Herb Label Template 2x2.doc; Herb Label Template 2x2 2 collectors.doc; Annotation Label Template with Notes 4x2.doc. None will be directly usable by others, as the "merge codes" correspond to the specific column headings in my EXCEL file. However, they provide models for others to modify to their needs, and, most importantly, they provide templates that are blocked into 1) 2 columns of 4 labels each, 2) 2 columns of 3 labels each, and 3) 2 columns of 2 labels each. These templates are (or attempt to be) compatible with the scored label sheets that are available. They will work most of the time. However, some final editing (at Step 5b in the list at the end of this document) will

commonly be necessary when the merge procedure results in a large amount of information to be fit within the allotted area of an individual label.

### **The Devil is in the Details**

Although the basic concept is simple, the implementation is frustratingly quirky. The following are details that users must contend with.

1. As one would expect, the EXCEL file used as the source of the data must have the first row consisting of column names and the remaining rows each corresponding to a collection (or specimen). As far as I have experienced, the table (including the headings) MUST be a named range, as this is what is selected under Step 3, "Select recipients". I always use the name "Table" for this range. [Note that the individual columns need not be explicitly named; EXCEL does this automatically using the column headings.]
2. Be aware that the arrangement of the columns and their widths in the "recipient list" as viewed within the Merge feature is different from that in the EXCEL file that is its source. [This is a great mystery.]
3. One of my columns (actually my first, "NUMBER" [although NOT the first when viewed within the Merge feature – Ha, Ha, on You!]) contains the collection number of the specimen (or set of specimens); it is the column that I use to select "recipients" for inclusion in the sheet of labels. Because of my sometimes mistakenly using the same collection number twice, some collection numbers have an "A" appended, e.g., "2876A" to distinguish from "2876". If a cell containing such a string is formatted as "General" or "Number" in the EXCEL file, it will appear as "0" in the "recipient list" as viewed within the Merge feature. Explicitly formatting such a column as "Text" solves the problem.
4. Note that the maximum number of characters (including spaces) allowed for one merge code (corresponding to a cell in the EXCEL file) is 255, even though the number of characters allowed within a cell of an EXCEL file is 65,000. All characters above 255 are truncated. This restriction is generally not a problem, but I sometimes use more than 255 characters. For example, I can have a long list of associated species under "Habitat" and comments concerning cytological details connected with a chromosome count under "Notes". My solution is to use two or more cells (e.g., NOTES 1, NOTES 2, NOTES 3, and NOTES 4) such that each contains no more than 255 characters. One can use EXCEL's "=LEN(yyy)" function to count the number of characters in cell "yyy" (where "yyy" is the cell immediately to the right of the cell containing the function, for example).
5. Merge codes referencing empty cells are invisible in the final document. This allows having a template with all the merge codes potentially needed for any collection, even though some of the collections lack the corresponding data. For example, the sequence

<<COUNTRY>>, <<STATE>>, <<COUNTY>>Co., <<CITY>> <<LOCATION 1>>  
<<LOCATION 2>> <<ALT.>> <<T&R>>

can be used for collections that lack data on altitude or "Township & Range", and the sequence

<<HABITAT 1>> <<HABITAT 2>> <<HABITAT 3>> <<HABITAT 4>>

can be used even when only the first cell contains data.

6. Often one has duplicates of a collection and, therefore, needs duplicate labels. The 2003 and more recent versions of Microsoft Excel (according to DoIT) have an option for printing an entire page of the same label; my 2002 version does not. The simplest way to get duplicate labels is to "Copy & Paste" one label into the desired number of label positions at step 5b, "Edit individual labels", the penultimate step in the merge procedure.

### **Steps in MicrosoftWord (2002) for Using the Merge Feature**

The following are the steps in MicrosoftWord (2002) for using the Merge feature to create a page (or pages) of labels. Naturally the steps will have to be modified for other versions of MicrosoftWord, but at least they provide a beginning for such modifications.

0. Tools, Letters and Mailings, Mail Merge Wizard
  - a. Labels
  - b. Next: Starting Document
1. Select starting document [i.e., the template with merge codes]
  - a. Start from existing document
  - b. Select an existing document
  - c. Open
  - d. Next: Select recipients
2. Select recipients
  - a. Use an existing list [i.e., the EXCEL file containing collection data]
    - i. Select a different list (OPTIONAL: Use if the current source file is not appropriate)
      - (1) Choose an EXCEL file from any directory
      - (2) Open
      - (3) Select the named range "Table"
        - (a) Insofar as I know, the EXCEL file must have its contents given a range name, e.g., "Table", and this must be selected
    - ii. OK
  - b. Edit recipient list
    - i. "Clear all records" before starting a new page of labels
    - ii. Check the rows ("recipients", each corresponding to a collection) to be inserted into the label template
    - iii. OK

- iv. Next: Arrange your labels
- 3. Arrange your labels (OPTIONAL: Only use if you want to modify the label template, thereby creating a new template. First time users will use this option to modify one of my sample templates, for example, to create their own, containing merge codes corresponding to the headings in their own EXCEL file.)
  - a. More items
    - i. Insert merge codes & characters within a label
  - b. Update all labels
    - i. This automatically repeats the first edited label format for all the labels on a page, if this is desired
  - c. Ctr-S can be used to save a modified label template
    - i. A new name for the modified template should be used, in general
  - d. Next: Preview your labels
- 4. Preview your labels
  - a. To see all your checked labels you may have to move "Recipient: #" up or down. I presume that the first label displayed is the one whose row happens to be highlighted in the recipient list.
  - b. Some editing of individual labels can be done here, although no editing within the string of a merge code is possible at this point
    - i. To some extent, large labels can be gotten to fit by eliminating blank lines
    - ii. Inappropriate characters can be edited out, as necessary (e.g., " Co., " for collections for which there is no county)
  - c. Next: Complete the merge
- 5. Complete the merge
  - a. Print
    - i. If no editing of the document is necessary
    - ii. All/Current record/from xxx to yyy?
      - (1) As I print only one page of labels at a time, I presume the first two options are effectively the same
    - iii. With my set-up I initially get an error message concerning a problem with margins. Continue.
    - iv. My BROTHER printer's "Alarm" red light goes on. Press the RESET button to continue.
  - b. Edit individual labels
    - i. This creates a MicrosoftWord document that will be printed onto the scored herbarium label paper
    - ii. One can edit the document, including the strings derived from merge codes, using all the facilities of MicrosoftWord
      - (1) Bold-face & italics can be added as desired
        - (a) These do not carry over from the source EXCEL file
        - (b) Although I prefer having taxon names in bold face and italics, I do not take the time to do this [Life is full of compromises!]
      - (2) One can "fine-tune" the labels, adding or subtracting material for individual labels as necessary, should the automatic format not be appropriate for a given collection

- (3) One can duplicate a label by “Copying & Pasting” it into the desired number of label positions
  - (a) One can select only one recipient at step 2, “Select recipients” and “Copy & Paste” it into the remaining label positions at this step
- iii. Print the document using the icon on the tool bar, as one cannot go back to the previous page with the “Print” hot button

*Acknowledgments.* My thanks to Stephanie Lyon for introducing me to the use of MicrosoftWord’s merge feature for making labels. To the long-suffering Mark A. Wetter go my eternal gratefulness and a bottle of Tequila for providing the impetus for this project and for putting up with my endless picayune queries concerning herbarium protocol, label formatting, printing, and God-Knows-What-Else, all with politesse, aplomb, and a verisimilitude of interest.