# Guide to Transcribing Label Data in the CCH2 Portal

last updated by Katie Pearson on 20 January, 2020

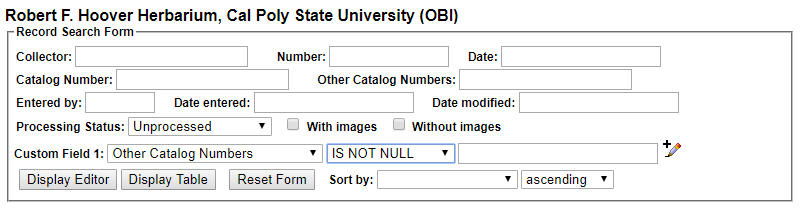
**Goal:** This protocol describes how to transcribe specimen data from images of skeletal records in the CCH2 portal.

## I. Navigating to records to transcribe

1. Log in to your account in CCH2 (cch2.org).



1. Click My Profile.
2. Click the Specimen Management tab.
3. Click the name of the collection for which you will transcribe specimens (there will likely only be one).
4. In the Data Editor Control Panel, click Edit Exiting Occurrence Records.
5. In the resulting Record Search Form, select “Unprocessed” from the dropdown list in the Processing Status field.
6. For Custom Field 1, select “Other Catalog Numbers” from the first dropdown list and “IS NOT NULL” from the second dropdown list.
   * Note: this step applies to institutions that associate barcode numbers with accession numbers in a separate step; if you are not sure whether this is true, ask a supervisor or omit this step

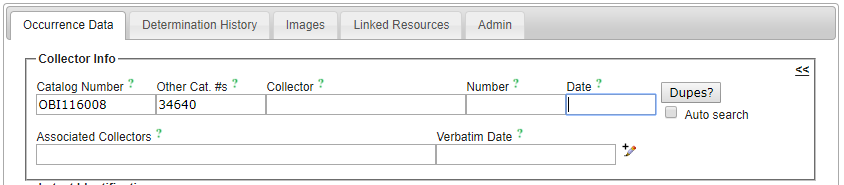


1. Click the Display Editor button.

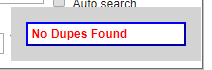
## II. Transcribing records

If you find that you are unable to transcribe part of a specimen label, set the processing status of the specimen to Expert Required, explain the reason you had trouble transcribing the record in the Notes (Occurrence Remarks) field, click the Save Edits button, and move to another specimen record.

1. When you have a specimen record in front of you, scroll down to the bottom of the page to the Status Auto-Set field, which is right above the Save Edits button. Select Pending Review from the dropdown menu (an explanation for the Processing Status field is provided in Appendix 2).
2. Scroll back up to the top of the specimen record. Zoom in to the specimen label on the image on the right by clicking and holding on the image, holding the Tab key on your keyboard, and moving the mouse up. Alternatively, you can hold the Control (or Command) button and click somewhere on the specimen image to zoom in.
   * The specimen label is usually on the bottom right of the specimen.
   * If the label is too blurry to read, click the High Res. button above the image.
3. Type the name of the primary collector into the Collector field.
   * Type the name exactly as it appears on the label
   * If there are multiple collectors, use the first collector name or the collector name that is associated with the collector number.



1. Type the collector number in the Number field.
   * A collector number is a consecutive number that collectors assign as they collect. This is not the same as the catalog number, barcode number, or accession number.
   * This number will almost always be typed or written on the label rather than stamped onto the herbarium sheet.
   * Collector numbers will occasionally include letters, dashes, or other characters. Enter these as they appear on the label. **\*\*\*When in doubt, type it out!\*\*\***
2. Type the collection date in the following format: YYYY-MM-DD. For example, if the label says 5 August 2017, the correctly formatted date would be 2017-08-05.
   * For single-digit months and days, include a preceding zero (e.g., 08, 09, 01).
   * If the date does not include a year, month, or date value, enter zeros for that value (e.g., if the label says “August 1987” you would enter 1987-08-00)
3. Click the Dupes? button (or, if Auto search is checked, the next step will happen automatically).
   * If you see the “No Dupes Found” message, continue to step 7.



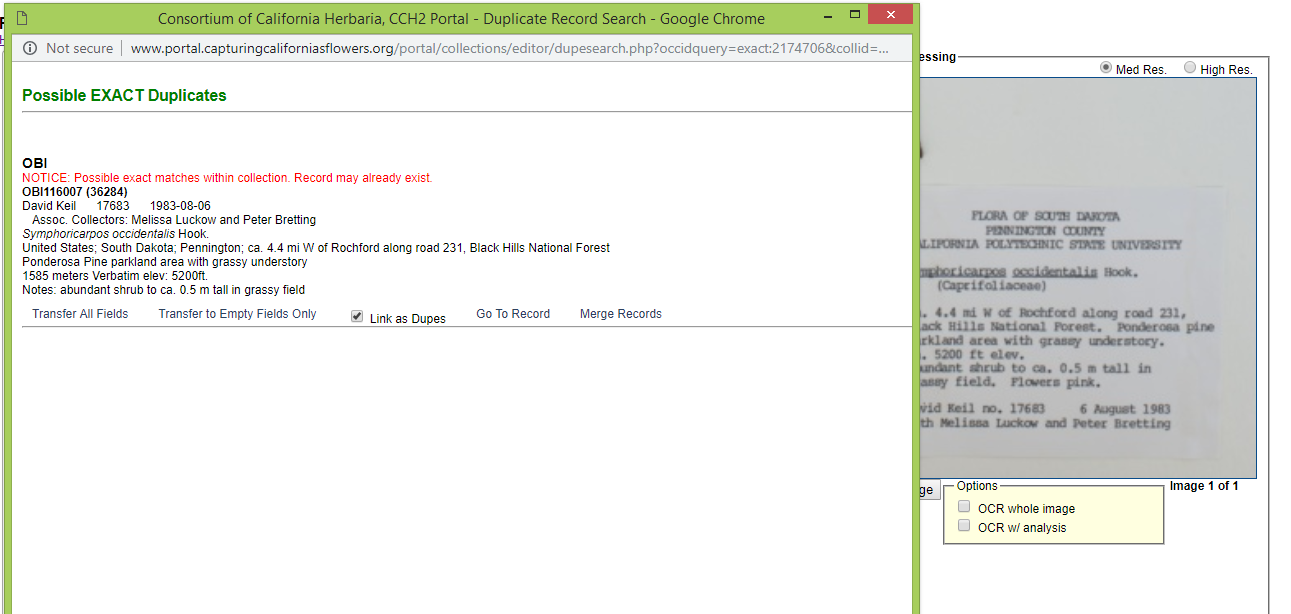
* + If a duplicate is found, a pop-up window will show up. (Make sure that your pop-up blocker is disabled to allow this). Examine the results of this pop-up window.
    - If you see the “Possible Matching Duplicate EVENTS” message, exit out of the window and proceed to step 7.



* + - If you see the “Possible EXACT Duplicates” message, look at the record listed in the pop-up window.



Check to see whether the scientific name, collector, date, and general details about the potential duplicate record(s) are the same as those on the label of the specimen you are transcribing. It may help to view the label and the potential duplicate record side-by-side as shown below.



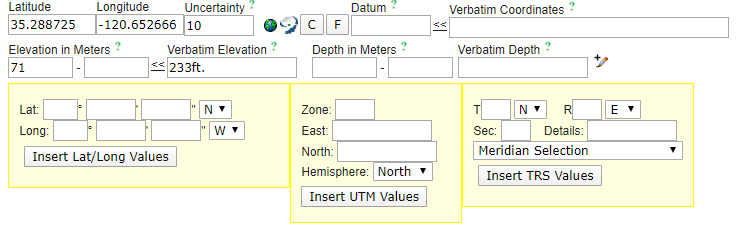
* + - If the records are the same, make sure the “Link as Dupes” box is checked in the pop-up window and click “Transfer to Empty Fields Only.” Check that the information imported into the data entry form matches the label exactly and that the data have been imported into the correct fields. You may have to re-type or move information around so it is stored in the correct fields (see step 11).
      * *If there are multiple potential duplicates, select the one that matches the specimen you are transcribing AND has georeferenced (latitude/longitude) coordinates, if possible.*
    - If the records are not the same, double check that you have correctly entered the collector name, number, and date in the original occurrence editor form. If you are sure these values are correct, notify a supervisor that you have found a potential data quality issue. Make sure you record the catalog numbers of the potentially problematic records.

1. On the record editor form, enter any collectors other than the primary collector in the Associated Collectors field. Separate names in this field by a comma and do not include words like “with” or “and.”
2. If any part of the date information was not clear in the Date field above (e.g., if the label said “Spring 2019”), type in the date EXACTLY as it is shown on the label into the Verbatim Date field.
3. Begin to type the ORIGINAL scientific name of the specimen into the Scientific Name field. If there are annotations (i.e., a more recent name for the specimen), you will add these in part III. Once the field begins to auto-populate, **select the appropriate scientific name from the dropdown menu**.
   * The ORIGINAL scientific name is the one that is on the original label. If there is no scientific name on the original label, leave this field blank.
   * Note that you do not need to enter authorship (e.g., the last names and abbreviations after the genus and specific epithet); these will be auto-populated.
   * You also do not need to enter the family name, since it will also be auto-populated.
4. Start typing the Country, State, and County into the appropriate fields and **select the appropriate values from the dropdown menus** as they appear.
   * Always select from the dropdown menus whenever possible.
   * Enter any reference to the USA as “United States”
   * The State/Province field primarily applies to states or provinces of the United States, Canada, and Mexico. If you are unsure about what to enter here, leave it blank.
5. Look at the label to see what further information is provided on the label. The most common fields for which there are label data are described below. See section IV for examples of specimen labels and how they would be transcribed into CCH2.
   * **Locality** – geographic description of where the plant was collected
     + This should not include country, state, or county information, since these data are found in alternate fields, unless the locality includes a description of an area within the county or state (e.g., “southwest corner of Jefferson county” or “northeast Arizona”).
     + This field should include directions to the collection location, city names, road names, names of parks, etc.
     + Some habitat information may be included in this field if it would help someone be able to re-locate the specimen’s locality. This information can be entered into this field AND the habitat field.
     + Note that you can check the “Locality Security” box if the locality information should be protected (e.g., protected species location).

**Verbatim Coordinates** – the latitude and longitude for the collection location of the specimen

* + - Enter the latitude and longitude coordinates and click the double arrows (<<) .
    - Degree/minutes/seconds lat/longs can be entered by clicking “Tools,” entering Lat/Long in degrees, minutes, seconds boxes, and clicking “Insert Lat/Long Values” to convert to decimal format.
    - If coordinates were provided on the sheet, enter “collector” in the Georeference Source field.

**NOTES:** If coordinates are provided in degrees minutes seconds, decimal degrees, township-range-section (e.g., T4N, R12W), or UTMs (e.g., 3572309E 1340983N), these can be entered by clicking the F button to the right of the Latitude and Longitude fields. Clicking this button will reveal fields where you can enter these data (in yellow boxes below).



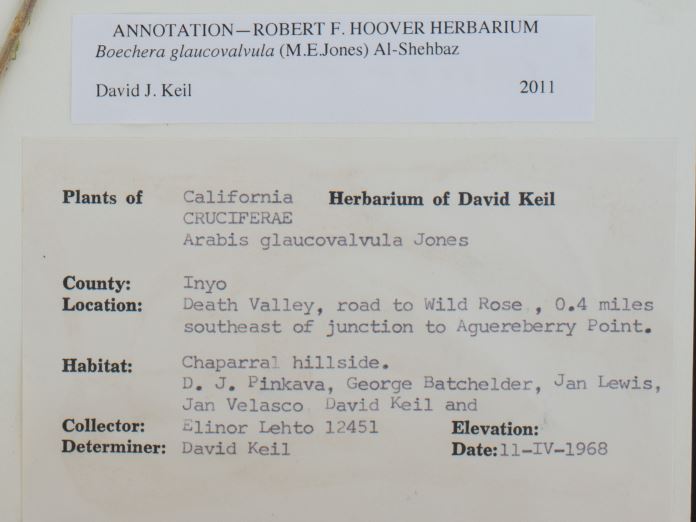
The C button will open the “coordinate cloning tool”, which will look for other specimens that have locality text similar to that which you entered in the locality field. If there are good-quality (i.e., accurate and precise) coordinates that match your locality, you can import these data into your data form by clicking the georeferenced point on the pop-up map and selecting Use Coordinates.

* + **Verbatim Elevation** – elevation or altitude of the collection location of the specimen
    - Enter the elevation value and units provided on the specimen label and click the double arrows (<<) to standardize to meters.
  + **Habitat** – environmental conditions in which the plant was found (e.g., marsh, grassy field, roadside)
  + **Substrate** – technical definitions given for substrate (e.g., soil type/series or rock formation)
  + **Associated Taxa** – other plant taxa listed as growing with the collected specimen.
    - Begin to type scientific names into this field and a dropdown list will appear. Make sure to select the appropriate name from this dropdown list.
    - Separate multiple associated taxa with a comma. Once you have placed a comma, the field will “understand” that you are about to type in a new associated taxon and will clear the dropdown menu.
  + **Description** – information specific to the individual plant (e.g., condition, flower color, leaf shape, height of the plant)
  + **Notes (Occurrence Remarks)** – population characteristics (rare, common, frequent, etc.) and any additional data on the label that does not fit easily into the other data fields.
    - This field can also be used as temporary holder to write notes for records that require further review.
  + **Phenology** – information about whether the plant is flowering or fruiting; only enter information into this field if it is written on the label or if you are instructed otherwise.

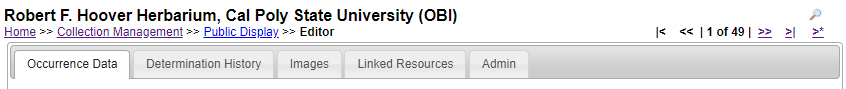
1. Once you have filled out the above fields as appropriate, check to see that any additional information is visible on the specimen label.
   * If the specimen is noted to be cultivated or planted, check the Cultivated/Captive box.
   * You do not need to include label headers (e.g., “Flora of California”) or the name of the herbarium.
   * If there are additional notes about the specimen having been collected for a specific study, enter this information in the Notes field.
   * For additional explanations of the Occurrence Data Fields, see Appendix 1.
2. Click the Save Edits button or press Enter on your keyboard.
3. If your specimen has an annotation label (see part III), move on to part III. Otherwise, return to the beginning of part II.

## III. Annotation label transcription

1. On the image that you just transcribed, check to see if there are any annotation labels on the specimen (see the example image below)
   * Annotation labels are usually above the specimen label but may be beside the label or elsewhere on the specimen. Make sure you view the full specimen for annotations before you move on.

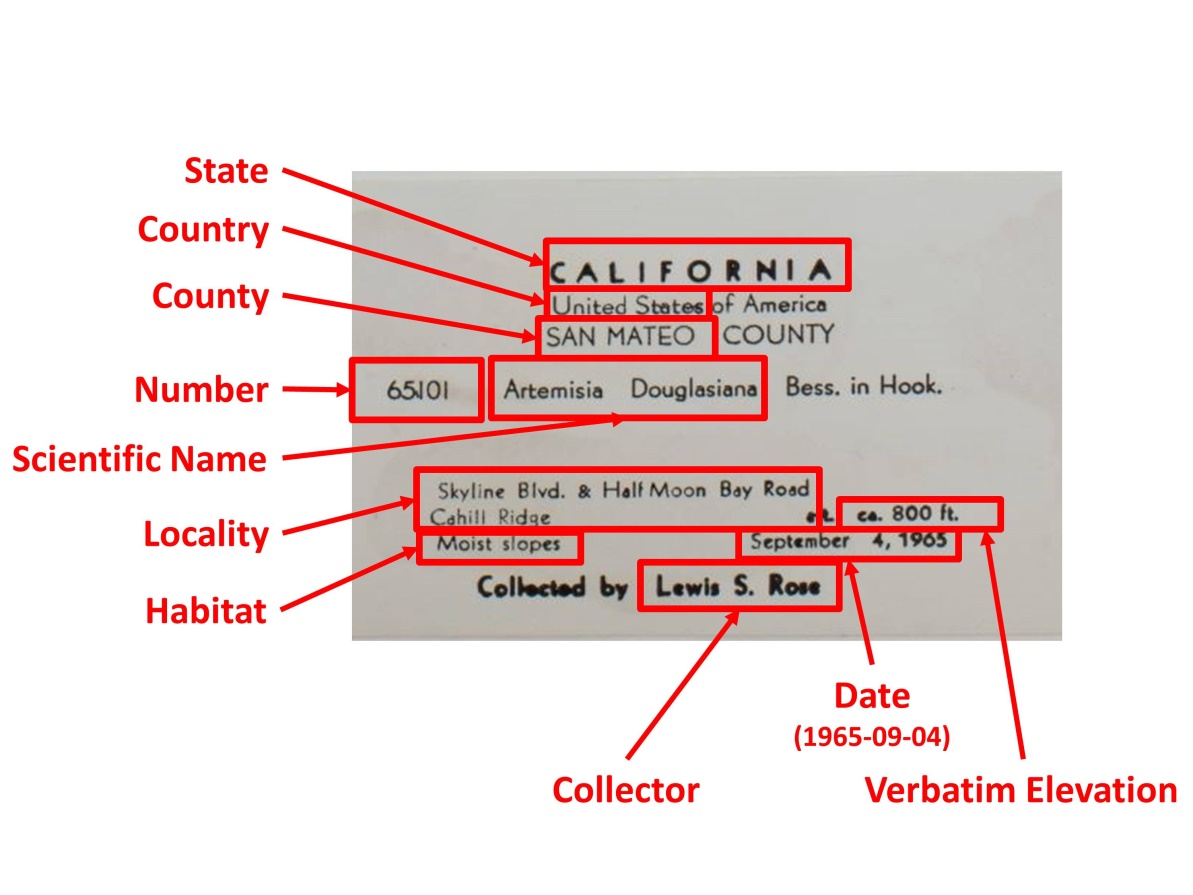


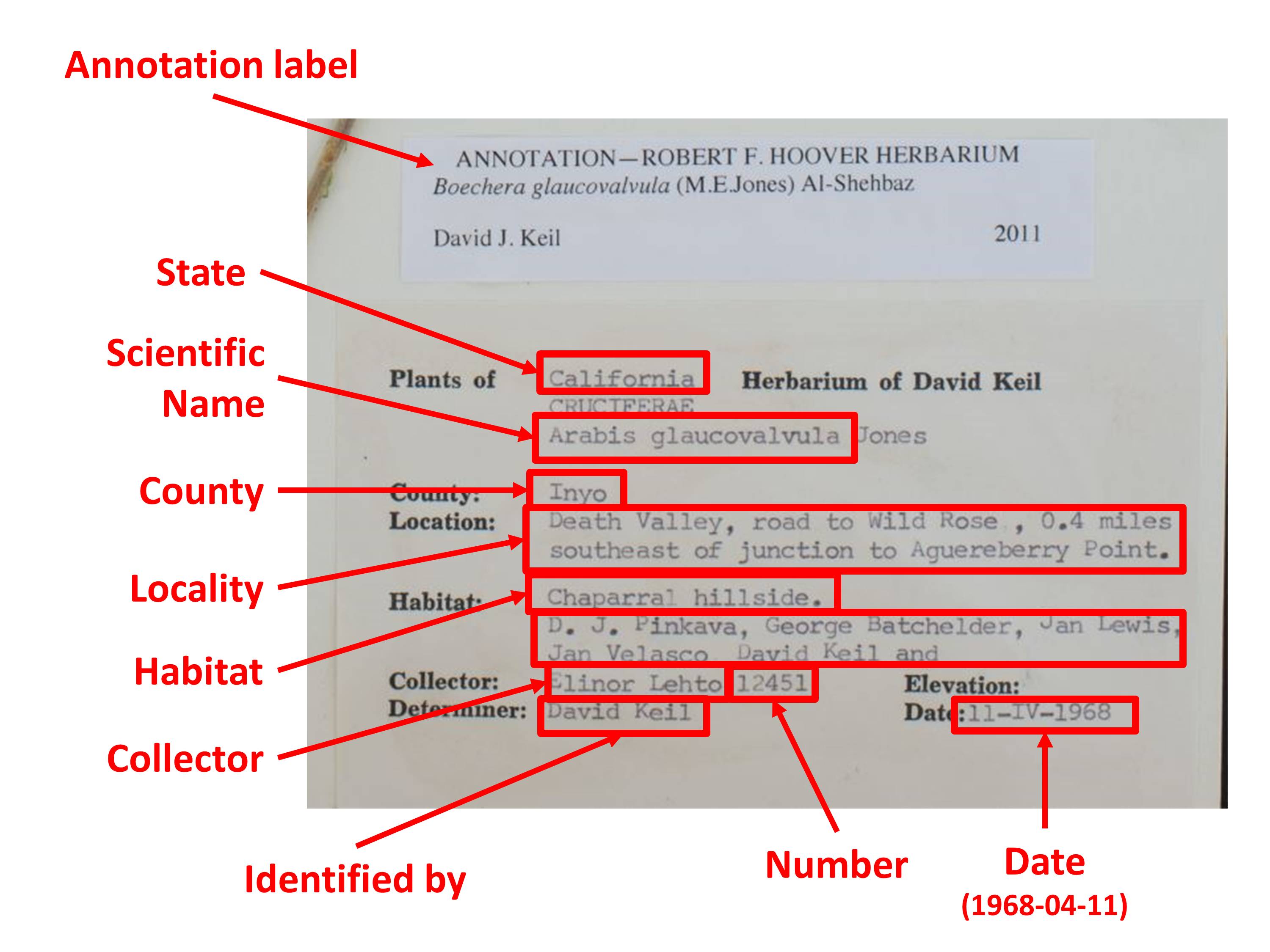
1. If there are annotation labels, click the Determination History tab at the top of the occurrence editor form.

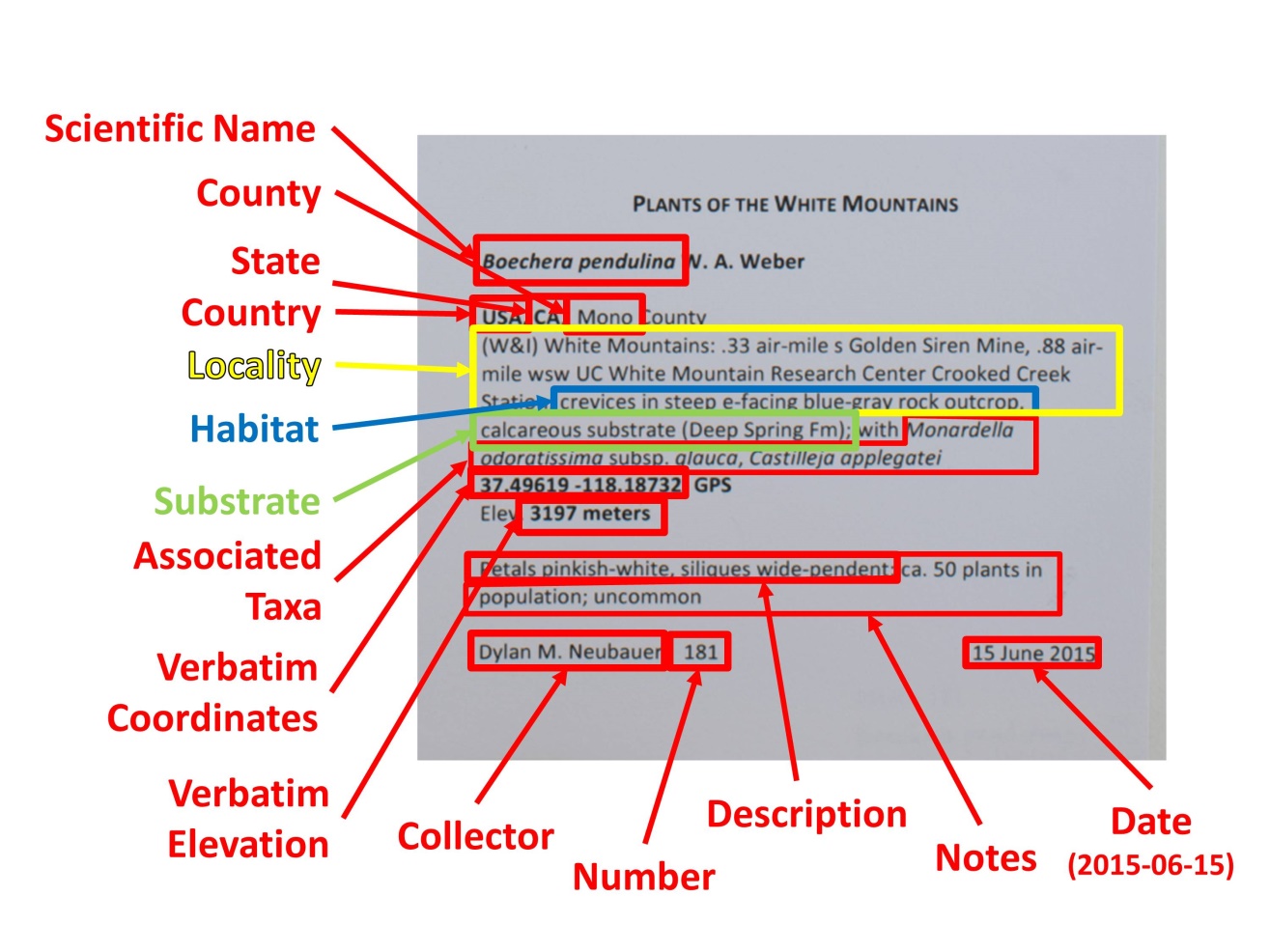
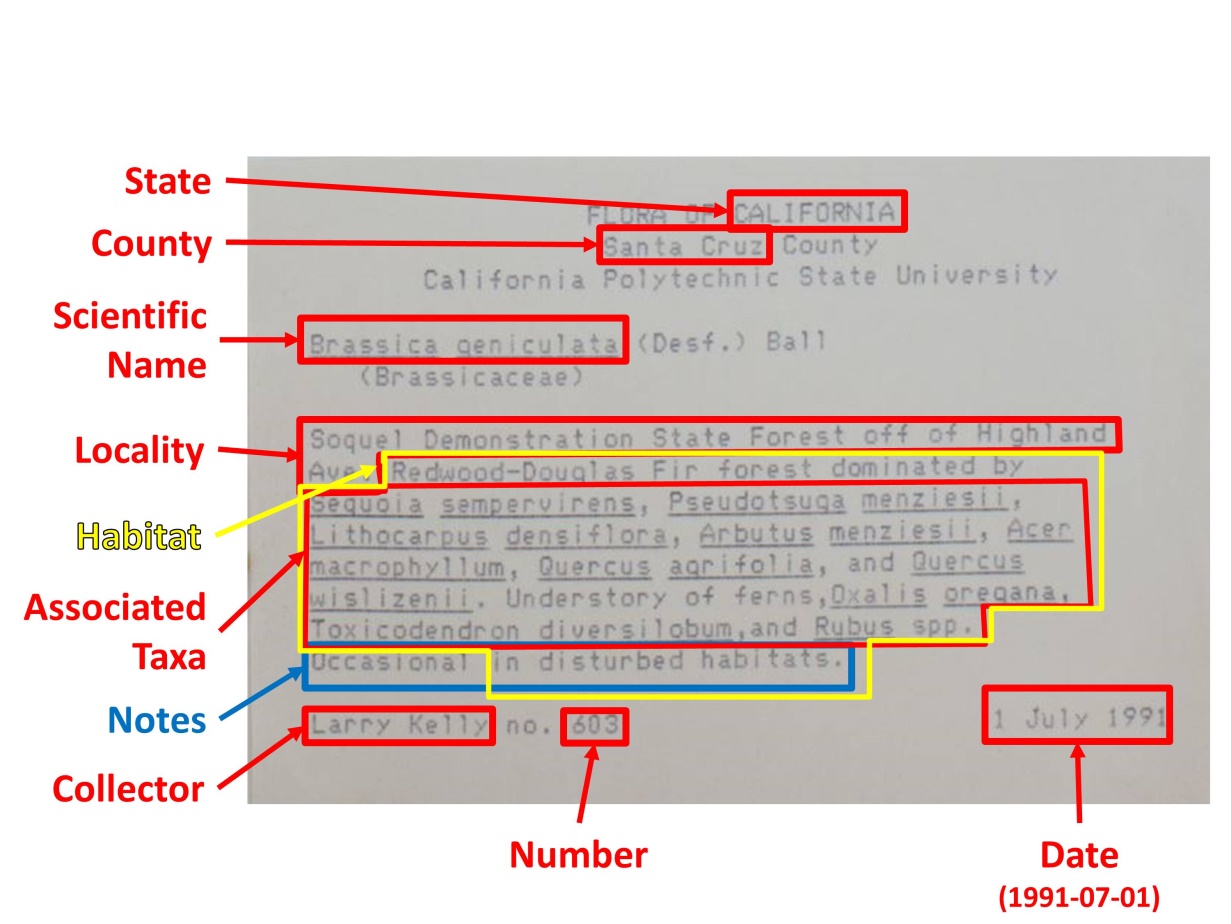


1. Start to type the scientific name of the specimen annotation into the Scientific Name field. Select the appropriate scientific name from the dropdown menu once it appears.
   * If there is an identification qualifier (for example: cf., aff., or ?), enter this in the Identification Qualifier field above the Scientific Name field.
2. Enter the name of the determiner (the person making the annotation) exactly as it appears on the annotation label into the Determiner field.
3. Enter the determination date in the YYYY-MM-DD format in the Date field.
4. If any references by which the specimen was identified or other notes occur on the annotation label, enter them in the appropriate fields.
5. If this is the most recent annotation, make sure the “Make this the current determination box” is checked. If there are more recent annotations after the annotation you are entering, uncheck this box.
6. Click Submit Determination.
7. Repeat steps 16-21 for any remaining annotation labels on the specimen.

## IV. Example diagrammed labels



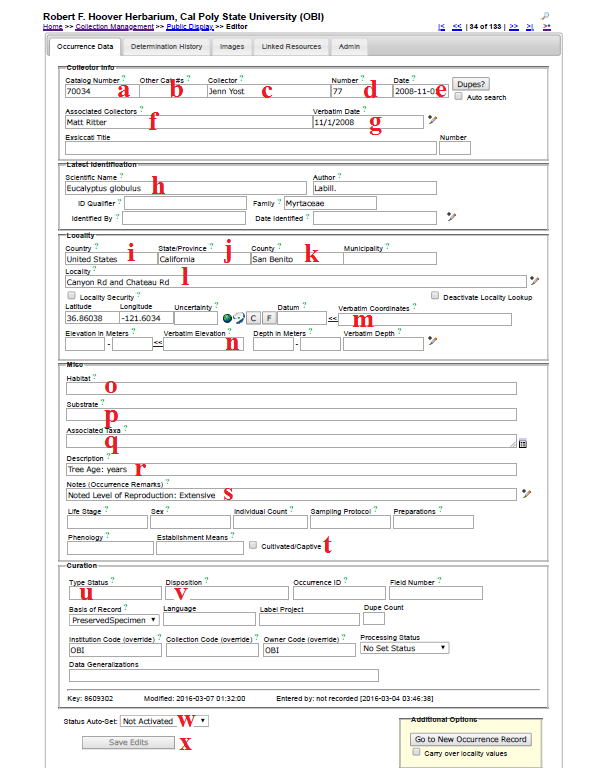




## Appendix 1: Explanation of Specimen Record Data Fields

(adapted from Skema & Barber 2018)

The major fields of interest are described in the following two pages. In the portal, you can click on the green question mark next to any field to get further explanation of the use of that field. Additional fields are viewable by clicking the pencil icon to the right of some fields. For further explanation of all fields, visit http://symbiota.org/docs/data-entry-guides-for-openherbarium/.



1. **Catalog Number** – the barcode number (and institution code, if included on the barcode label) on the specimen (e.g., OBI00045897)
2. **Other Catalog #s** – the accession number of the specimen, if applicable. If your institution already had data in NANSH or the Consortium of California Herbaria, this may have previously been the number in the Catalog Number field.
3. **Collector** – primary or only collector. The name should be in order of “First name Last name”. On a specimen, this is often the name that follows “col.”, “leg.”, and “legit.” If no collector is present, “Unknown” should be entered in lieu of other institutional standards.
4. **Number** – collector number. If no number is present, “s.n.” (*sine numero* is Latin for ‘without number’) should be entered in lieu of other institutional standards. If letters or symbols are included in the collector number on the label (e.g., KP-004), they should be included.
5. **Date** – when the specimen was collected in the format YYYY-MM-DD. Institutional standards should be followed for entering 0s when a day or month is omitted, or if a range of dates are listed. See also “Verbatim Date” (g).
6. **Associated Collectors** – collectors following the primary collector. Each name should be in order of “First name Last name.” Multiple names should be separated by a semicolon (;).
7. **Verbatim Date** – date as it appears, verbatim, on the label (e.g., July 21st, 2016). This is particularly useful if part of the date is omitted, the date is in another language, or if a range is listed.
8. **Scientific Name** – the scientific name of the specimen. When populating this field, be sure to select an option from the dropdown list rather than typing out the name. If the specimen is filed under a different name than the name on the specimen, this information should be noted in the Notes field (e.g., “filed under...”) in lieu of other institutional standards.
9. **Country** – the country in which the specimen was collected. If the specimen was collected in the U.S., “United States” should be entered in this field.
10. **State/Province** – the state (or parish, in Louisiana) in which the specimen was collected, if in the U.S., or the province if the specimen was collected in Canada. The state/province name should be entered as a full name rather than an abbreviation (e.g., “California” instead of “CA”).
11. **County** – the county in which the specimen was collected. County designations (e.g., “Co.” or “County”) should not be included in this field. This value should be selected from the dropdown list.
12. **Locality** – geographic description of where plant was collected. This should not include country, state, or county information, since these data are found in alternate fields, unless the locality includes a description of an area within the county or state (e.g., “southwest corner of Jefferson county” or “northeast Arizona”). This field should include directions to the collection location, city names, road names, names of parks, etc. Note that you can check the “Locality Security” box if the locality information should be protected (e.g., protected species location).
13. **Verbatim Coordinates** – If latitude and longitude are provided on the sheet, they should be entered in this field and standardized by clicking << . Degree/minutes/seconds lat/longs can be entered by clicking “Tools,” entering Lat/Long in degrees, minutes, seconds boxes, and clicking “Insert Lat/Long Values” to convert to decimal. If coordinates were provided on the sheet, “Coordinates recorded at time of collection” should be entered in the Georeference Remarks field.
14. **Elevation in Meters** – if elevation in meters is provided on the sheet, it should be entered in this field. If provided in other unit, e.g., feet, the value can be entered into the “Verbatim Elevation” field with its unit and converted to metric by clicking the << .
15. **Habitat** – environmental conditions in which the plant was found (e.g., marsh, grassy field)
16. **Substrate** – technical definitions given for substrate, e.g., soil type/series or rock formation
17. **Associated Taxa** – other plant taxa listed as growing with the collected specimen. To promote standardization and easy entry into this field, a dropdown list of scientific names in the portal’s taxon table can be accessed by clicking the small table icon on the small table icon to the right of the data field.
18. **Description** – information specific to the individual plant (e.g., condition, color)
19. **Notes (Occurrence Remarks)** – population characteristics (e.g., frequency: rare, common, etc.) and any additional data on the label that does not fit easily into the other data fields. This field can also be used as temporary holder to write notes for specimens that require further review.
20. **Cultivated/Captive Checkbox** – this should be checked if the label indicates that the sample was from a cultivated plant
21. **Type Status –** the type status (e.g., holotype, isotype) of the specimen. This field should only contain data if the specimen is a type
22. **Disposition –** the location of the specimen in the herbarium or on loan. This field is especially useful if the specimen is being used for a display, research project, or education use, or if the specimen has two different species on it (i.e., it can only be filed with one of the species).
23. **Processing Status** – this field is used to indicate the record’s stage in the quality control workflow (see Appendix 1: Processing Status for instructions)
24. **Save Edits** – click to save record. If the button is unclickable, click anywhere on screen outside of a field, then click “Save Edits” button again.

## Appendix 2: Use of Processing Status Field in CCH2

(adapted from Skema & Barber 2018)

Processing Status is a field in Symbiota that helps keep track of what stage a record is at in the workflow. There is no one standard for how to use the different levels of Processing Status across all users of Symbiota, although there is some standardization within projects. If each user in a particular collection uses the processing status field consistently, both in remembering to change the status as required, and in assigning the appropriate status, it is of great help to the workflow. The following table shows the recommended definitions for the status levels in Processing Status to use for the California Phenology TCN.

|  |  |
| --- | --- |
| **Status** | **Definition** |
| **Unprocessed** | skeletal record (typically with just an image and catalog number) |
| **Pending Review** | volunteer/technician transcribed record and had no problems/questions  **OR**  record imported from a dataset that was previously minimally quality controlled |
| **Expert Required** | volunteer/technician transcribed or reviewed record and had problem(s)/questions (volunteer/technician should leave concise notes in Occurrence Remarks to explain problem/question) |
| **Reviewed** | record transcribed; transcription reviewed and correct  **OR**  record imported from a dataset that was previously quality controlled |
| **Stage 3** | record problematic; to be researched |
| **Closed** | record transcribed, georeferenced, and fully reviewed (or as best as possible given available information) |

## References

Skema C, Barber A. 2018. Mid-Atlantic Megalopolis digitization standards and training manual. Morris Arboretum of University of Pennsylvania. Available from https://docs.wixstatic.com/ugd/6f7156\_800579c353574e59918cfb541172d749.pdf