
Tripal File Documentation

Release 7.x-1.x

Stephen Ficklin

Mar 27, 2023

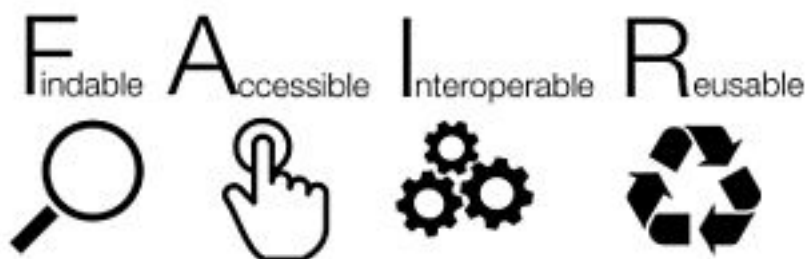
TRIPAL FILE EXTENSION MODULE

1	Overview	1
2	Installation	3
2.1	Step 1: Module Installation	3
2.2	Step 2: Install the EDAM Vocabulary	4
2.3	Step 3: Set Permissions	4
3	Usage	7
3.1	Adding a License	7
3.1.1	Example Licenses	8
3.1.1.1	Public Domain Data	8
3.1.1.2	CC0 1.0 Universal (CC0 1.0)	8
3.1.1.3	Attribution 4.0 International (CC BY 4.0)	9
3.2	Adding a File	9
3.3	Adding File Metadata	12
3.3.1	Manually Adding Metadata	12
3.3.2	Adding Metadata in Bulk	14
3.3.2.1	File Module Chado Tables	15
3.4	Associating a File with Other Content	16
3.5	Accessing Files via Web-Services	18

OVERVIEW

Warning: The Tripal file module is **compatible only with Tripal v3.6 or higher**.

The Tripal File module supports association of data files with content in a Chado database and for integration of those associations with content types on a Tripal website.



These associations are meant to support [FAIR data principles](#) by integrating with Tripal content web services such that:

- Files are findable and accessible via Tripal's content web services.
- Metadata about files use globally unique controlled vocabularies. These metadata can be assigned as properties of each file.
- License and usage details can be assigned to each file.

This module provides two new content types for Tripal sites: File and License.

Development of the Tripal File module was funded by the [National Science Foundation award #1659300](#) and the [National Research Support Program \(NRSP\) 10 project](#) and the McIntire-Stennis project 1019284.

INSTALLATION

2.1 Step 1: Module Installation

The Tripal File module is available as a full Drupal module. Therefore, it can be installed following the typical Drupal module installation either via the GUI or via Drush. We'll do so here using a Drush command:

```
drush pm-enable tripal_file
```

If you do not have the module already downloaded and available in the *sites/all/modules* folder of your Drupal site, then you will be asked if you would like to download and then enable the module.

The module will create two new content types: *File* and *License*. It will also create a variety of tables in your Chado database for associating files to base records. Upon installation you should see the following:

```
$ drush pm-enable tripal_file

The following extensions will be enabled: tripal_file
Do you really want to continue? (y/n): y
INFO (TRIPAL_ENTITIES): Done.
INFO (TRIPAL_ENTITIES): Done.
tripal_file was enabled successfully. [ok]
Custom table, 'file' , created successfully. [status]
Custom table, 'license' , created successfully. [status]
Custom table, 'file_contact' , created successfully. [status]
Custom table, 'fileloc' , created successfully. [status]
Custom table, 'fileprop' , created successfully. [status]
Custom table, 'fileloc' , already exists. Table structure not changed, but [status]
definition array has been saved.
Custom table, 'file_license' , created successfully. [status]
Custom table, 'analysis_file' , created successfully. [status]
Custom table, 'assay_file' , created successfully. [status]
Custom table, 'biomaterial_file' , created successfully. [status]
Custom table, 'cv_file' , created successfully. [status]
Custom table, 'eimage_file' , created successfully. [status]
Custom table, 'feature_file' , created successfully. [status]
Custom table, 'featuremap_file' , created successfully. [status]
Custom table, 'library_file' , created successfully. [status]
Custom table, 'nd_protocol_file' , created successfully. [status]
Custom table, 'organism_file' , created successfully. [status]
Custom table, 'phylotree_file' , created successfully. [status]
Custom table, 'project_file' , created successfully. [status]
```

(continues on next page)

(continued from previous page)

```

Custom table, 'pub_file' , created successfully. [status]
Custom table, 'stock_file' , created successfully. [status]
Custom table, 'stockcollection_file' , created successfully. [status]
Custom table, 'study_file' , created successfully. [status]
NOTICE: Added the EDAM ontology. Please navigate to Tripal >> Data Loaders >> [status]
Chado Vocabularies >> OBO Vocabulary Loader and submit a job to import the
EDAM vocabulary.

```

2.2 Step 2: Install the EDAM Vocabulary

The [EDAM vocabulary](#) is needed for the File module because it provides many of the file types (e.g. FASTA, GFF3, VCF, etc). Any file that is managed by the Tripal File module requires a file type.

The EDAM vocabulary details are automatically added by the Tripal File module, but you will need to install the vocabulary. To do this navigate to the **Administration > Tripal > Data Loaders > Chado Vocabularies > OBO Vocabulary Loader** page via the administrative menu. Select the **EDAM** vocabulary from the **Ontology OBO File Reference** drop down and click the **Import OBO File**.

2.3 Step 3: Set Permissions

Now that the Tripal File module is installed, we must set permissions so that users can view and or create File and License content types. To set permissions, navigate to **People** page and click the **Permissions** tab in the top right. Look for permissions that begin with prefix *File:* and *License:* and set the according to your needs.

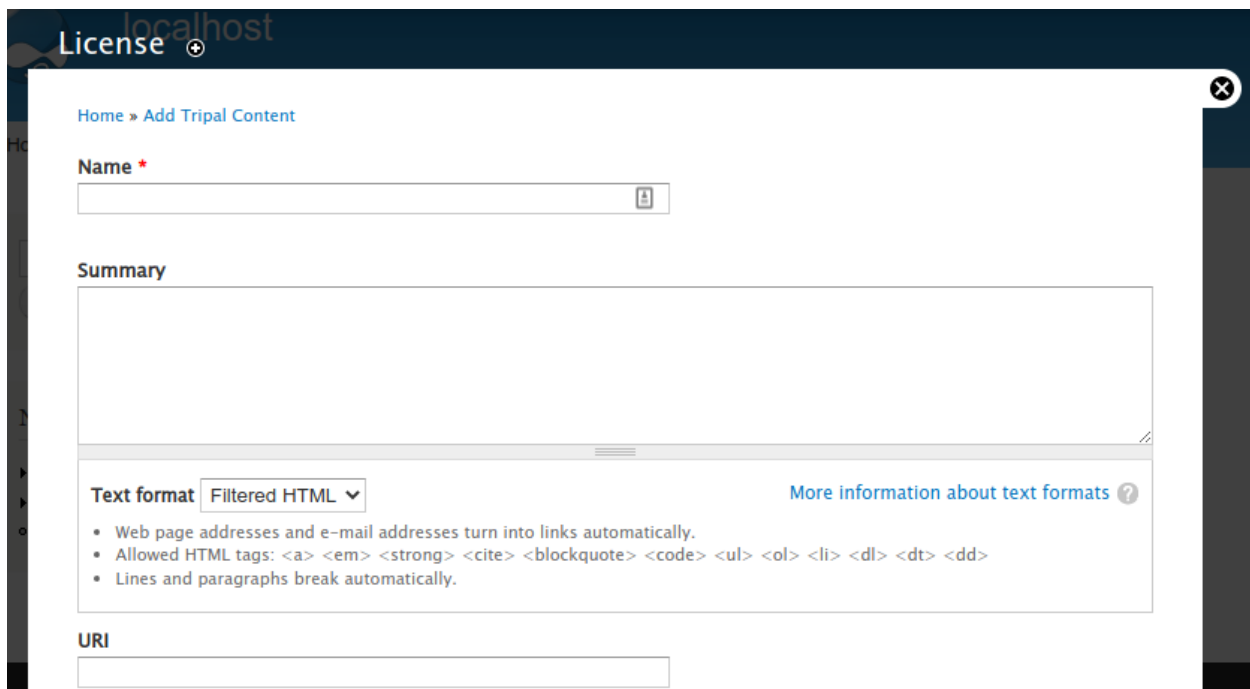
PERMISSION	ANONYMOUS USER	AUTHENTICATED USER	ADMINISTRATOR
File: View Content Allow the user to view <i>File</i> content	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
File: Create Content Allow the user to create <i>File</i> content <i>Warning: Give to trusted roles only; this permission has security implications.</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
File: Edit Content Allow the user to edit <i>File</i> content <i>Warning: Give to trusted roles only; this permission has security implications.</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
File: Unpublish Content Allow the user to unpublish <i>File</i> content. Unpublishing of content removes it from visibility on the site but does not delete the record in the underlying database.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
File: Delete Content Allow the user to delete <i>File</i> content. When content is deleted it is first unpublished and then deleted from the database. <i>Warning: Give to trusted roles only; this permission has security implications.</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
License: View Content Allow the user to view <i>License</i> content	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
License: Create Content Allow the user to create <i>License</i> content <i>Warning: Give to trusted roles only; this permission has security implications.</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
License: Edit Content Allow the user to edit <i>License</i> content <i>Warning: Give to trusted roles only; this permission has security implications.</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
License: Unpublish Content Allow the user to unpublish <i>License</i> content. Unpublishing of content removes it from visibility on the site but does not delete the record in the underlying database.	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
License: Delete Content Allow the user to delete <i>License</i> content. When content is deleted it is first unpublished and then deleted from the database. <i>Warning: Give to trusted roles only; this permission has security implications.</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Warning: You should not give the anonymous user any permissions other than 'view' permission.

3.1 Adding a License

To ensure data files offered by your Tripal site meet [FAIR data principles](#), all files must be associated with a license. Therefore, before adding files, you must first create new **License** content types. You can create as many **License** files as needed for the data on your site.

To create a new license page, navigate to **Admin > Add Tripal Content Types** and scroll to the bottom of the page where the **Other** category is found. There you will see the **License** content type. Click the link to create a new license page.



The screenshot shows the 'Add Tripal Content' form for a 'License' content type. The form is titled 'License' with a plus icon. The breadcrumb trail is 'Home > Add Tripal Content'. The form fields include: 'Name' (required, with a red asterisk and a small icon), 'Summary' (a large text area), 'Text format' (set to 'Filtered HTML' with a dropdown arrow), and 'URI' (a text field). A 'More information about text formats' link with a question mark icon is located next to the 'Text format' dropdown. Below the 'Text format' dropdown, there is a list of allowed HTML tags: <a>, , , <cite>, <blockquote>, <code>, , , , <dl>, <dt>, <dd>. The list also includes a note: 'Web page addresses and e-mail addresses turn into links automatically.' and 'Lines and paragraphs break automatically.'

To create your license page you can:

1. Option 1: Fully define the license by giving it a name and provide the full description
2. Option 2: Summarize the license by giving a name, providing a brief summary (or no summary) and providing the URL to the full license online.

Note: It is best practice to provide a human readable summary of the user's rights in the Summary field and to provide

a link to the full legal text of the license via the URI field.

3.1.1 Example Licenses

A few example licenses are provided below along with the values that may be appropriate for the **License** page fields.

3.1.1.1 Public Domain Data

Data that is public, where no one retains rights to it. Therefore, it needs a license page even though “public domain” is not an actual license. You can use the following values when creating a public domain license page:

Field	Value
Name	Public Domain
Summary	The data has no known copyrights or restrictions and can be freely used by anyone.

3.1.1.2 CC0 1.0 Universal (CC0 1.0)

The text for the summary below came from <https://creativecommons.org/publicdomain/zero/1.0/>.

Field	Value
Name	CC0 1.0 Universal (CC0 1.0)
Summary	<p>The person who associated a work with this deed has dedicated the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law. You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission.</p> <ul style="list-style-type: none"> • In no way are the patent or trademark rights of any person affected by CC0, nor are the rights that other persons may have in the work or in how the work is used, such as publicity or privacy rights. • Unless expressly stated otherwise, the person who associated a work with this deed makes no warranties about the work, and disclaims liability for all uses of the work, to the fullest extent permitted by applicable law. • When using or citing the work, you should not imply endorsement by the author or the affirmer.
URI	https://creativecommons.org/publicdomain/zero/1.0/legalcode

3.1.1.3 Attribution 4.0 International (CC BY 4.0)

The text for the summary below came from <https://creativecommons.org/licenses/by/4.0/>.

Field	Value
Name	Attribution 4.0 International (CC BY 4.0)
Summary	<p>You are free to:</p> <ul style="list-style-type: none"> • Share: copy and redistribute the material in any medium or format • Adapt: remix, transform and build upon the material for any purpose, even commercially. <p>The licensor cannot revoke these freedoms as long as you follow the following license terms:</p> <ul style="list-style-type: none"> • Attribution You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. • No additional restrictions You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits <p>Notices:</p> <ul style="list-style-type: none"> • You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation. • No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.
URI	https://creativecommons.org/licenses/by/4.0/legalcode

3.2 Adding a File

To demonstrate adding a file, we will provide an example using the genome files that are provided in the [Tripal User's Guide](#) in the section titled [Setup of an Example Site](#). Those instructions direct the reader to create an **Analysis** page to represent the whole genome assembly of the *Citrus sinensis* v1.0 genome. We will associate the files to that analysis. If you have followed along with the Tripal User's Guide you will already have those files uploaded.

If you have not followed the Tripal User's Guide but would like to follow along here, you must follow [these instructions to create an Analysis page](#). And obtain the files here:

- [Citrus sinensis-orange1.1g015632m.g.gff3](#)
- [Citrus sinensis-scaffold00001.fasta](#)
- [Citrus sinensis-orange1.1g015632m.g.fasta](#)

To create a new file page, navigate to **Admin > Add Tripal Content Types** and scroll to the bottom of the page where the **Other** category is found. There you will see the **File** content type. Click the link to create a new file page. The following page appears.

File

My

[Home](#) » [Add Tripal Content](#)

Name *

Description

Source

✂

📄

📁

📅

📅

ABC

↶

↷

🔍

↺

📊

🖼

🕒

📅

☺

Ω

🔄

📄

Format

B

I

U

~~S~~

x₂

x²

I_x

≡

≡

≡

≡

”

≡

≡

≡

≡

🔗

🔗

🚩

[Switch to plain text editor](#)

Text format

Filtered HTML

- Web page addresses and e-mail addresses turn into links automatically.
- Allowed HTML tags: <a> <cite> <blockquote> <code> <dl> <dt> <dd>

On this page you can:

1. Provide a **name** (required) and **description** for the file.
2. Indicate a file **type** (required).
3. Indicate the file **source** (or contact person) who was rights to the data.
4. Indicate the file **location** (either locally or via remote URL).
5. And set the **license** (required).

First, we will create the whole genome FASTA file page. Enter the following in the File page fields:

Field	Value
Name	<i>Citrus sinensis</i> Whole Genome Assembly v1.0
Description	The whole genome assembly, v1.0, of <i>Citrus sinensis</i> .
Type	FASTA
File Source	<i>Leave blank or provide any contact you may have already</i>
Location	Citrus sinensis-scaffold000001.fasta
File License	<i>Select whatever license you may have already added</i>

If you already have the file uploaded to Tripal then it will automatically appear in the **Tripal Managed File** section:

DOWNLOAD *

Remote URL

+

If the file is not managed by Drupal or Tripal then enter the full URL to the file. Use this field if the file is shared publicly on this same web server but not i

Tripal Managed File

Citrus_sinensis-scaffold00001.fasta

Please select a file.

Remote URL

+

If the file is not managed by Drupal or Tripal then enter the full URL to the file. Use this field if the file is shared publicly on this same web server but not i

Tripal Managed File

-- Select a File --

Please select a file.

The download location of the file.

Otherwise you can provide a remote URL for the file. If you have files on your site that are not managed by Tripal, you can use a URL here that points to the full path on your site. In both of these cases file size and md5 checksum cannot be automatically generated. Alternatively, you can use a `public://` prefix for local files, in which case the file size and md5 checksum can be automatically included. As an example, you might upload a file to your local filesystem in the directory `sites/default/files/bulk_data/` and then you could specify a Remote URL of `public://bulk_data/Citrus_sinensis-scaffold00001.fasta`

Note: A file can have more than one download location, and you can combine both local and remote files.

Note: Providing a file source or “contact” is optional, but is recommended. Every file with a license should indicate, via the “file source” field, who retains the license rights (if applicable).

After creation, the file page appears:

Citrus sinensis Whole Genome Assembly v1.0

View

Edit

Reload

[Summary](#)[File Source](#)[License](#)[Location](#)

Summary

Resource Type	File
Name	Citrus sinensis Whole Genome Assembly v1.0
Description	The whole genome assembly, v1.0, of <i>Citrus sinensis</i> .
File Type	FASTA format

License

Name	Attribution 4.0 International (CC BY 4.0)
License Summary	<p>You are free to:</p> <ul style="list-style-type: none"> Share: copy and redistribute the material in any medium or format

3.3 Adding File Metadata

3.3.1 Manually Adding Metadata

You can add additional metadata to a file by adding new fields to the file content type. To do this, navigate to **Admin > Structure > Tripal Content Types** and click the **manage fields** link in the **File** row. Scroll to the section titled **Add new field**. The field Type and field widget should be **Chado Property**:

LABEL	MACHINE NAME	FIELD TYPE	WIDGET	OPERATIONS
+ Download	schema__itemlocation	File Location	File Location	edit delete
+ License	schema__license	License	License	edit delete
+ References	sio__references_file	References	References	edit delete
<div> <div>+ Add new field</div> <div> <input type="text"/> </div> <div> <div>Chado Property</div> <div>Type of data to store.</div> </div> <div> <div>Chado Property</div> <div>Form element to edit the data.</div> </div> </div>				

Name the property according as needed. Suppose we wanted to add a **genus** and **species** property to each file as metadata. We would add each one by first providing the name and clicking the **Save** button at the bottom. Next, Tripal provides a page for more details. Here the most important setting is the **Controlled Vocabulary Term**. To ensure data meets FAIR standards all metadata must be associated with a controlled vocabulary term. Here, we can search for and

select the term **genus** and click the **save settings** button. We can repeat this step for the **species**. Once completed we should see two new fields associated with our File content type.

LABEL	MACHINE NAME	FIELD TYPE	WIDGET	OPERATIONS	TERM	SUPPORTED BY *
File Source	file_contact	Contacts	Contacts	edit delete	contact (local:contact)	CHADO
Download	schema__itemlocation	File Location	File Location	edit delete	itemLocation (schema:itemLocation)	CHADO
License	schema__license	License	License	edit delete	license (schema:license)	CHADO
References	sio__references_file	References	References	edit delete	references (SIO:000631)	CHADO
Genus	field_genus	Chado Property	Chado Property	edit delete	genus (TAXRANK:0000005)	CHADO
Species	field_species	Chado Property	Chado Property	edit delete	species (TAXRANK:0000006)	CHADO

We now have two new fields that can be used for metadata but we must make those visible on the page when data is present for them. You can follow the instructions for [Configuring Page Layout](#) in the Tripal User's Guide to place these new metadata fields where you want on the page. For this example, we will place them in the **Summary** section of the page layout.

If we return to the any File page and edit it then we will now have a new metadata field for storing the genus.

File License *

Attribution 4.0 International (CC BY 4.0) ▼

Please select a license that properly describes to the user their rights, privilege and or usage permissions for this file.

Genus

Citrus

Species

sinensis

Authoring information

Publishing options

Published

Authored by

admin

Leave blank for *Anonymous*.

Authored on

2021-03-01 20:50:12 -0800

Format: 2021-03-01 20:50:12 -0800. The date format is YYYY-MM-DD and -0800 is the time zone

Now when we view the page we can see the genus and species metadata for the file

Citrus sinensis Whole Genome Assembly v1.0

View

Edit

Reload

[Download](#)
[File Source](#)
[License](#)
[Summary](#)

Summary

Resource Type	File
Name	Citrus sinensis Whole Genome Assembly v1.0
Description	The whole genome assembly, v1.0, of <i>Citrus sinensis</i> .
File Type	FASTA format
Genus	Citrus
Species	sinensis

Download

Available at	Size	MD5
Tripal v3 Dev	6.03MB	85934c1cc36308d829dc73e4eb9dcea3

3.3.2 Adding Metadata in Bulk

Because there is no common file format for providing metadata the best way to add metadata in bulk is to use the Tripal Bulk Loader, a module that comes built-in to Tripal v3. You can find a tutorial for learning to use the bulk loader at the [Bulk Loader](#) tutorial of the Tripal User's Guide.

Use of the bulk loader requires a good understanding of Chado, its database tables, their constraints and relationships. If you are comfortable with Chado, you can use the Tripal bulk loader to add metadata about files using a tab-delimited file. To do so you need to

1. Identify controlled vocabulary (CV) terms for each metadata value.
2. Ensure that those CV terms are already in Tripal and if not, you need to add them.
3. Create a tab-delimited file that contains the file name used in Tripal and as many columns of metadata values as desired.
4. Create a bulk loader template to recognize the file and import it.

The following provides the schema for the **file** and **fileprop** tables that were added to Chado by the Tripal File module. The **fileprop** table is where you will add metadata values.

3.3.2.1 File Module Chado Tables

The file Table

The *file* table houses the base file record. The *name* field must be a unique value for each file and thus can be selected on for finding files.

Column	Type	Null	Default Value	Constraint
file_id	integer	No	(auto)	Primary Key
name	text	No		Unique
type_id	integer	No		Foreign Key to cvterm
description	text	Yes		

The fileprop Table

The *fileprop* table holds the properties or metadata about files. The CV term is specified using the *type_id* column and the rank is incremented if multiple values of the same type are stored.

Column	Type	Null	Default Value	Constraint
fileprop_id	integer	No	(auto)	Primary Key
file_id	integer	No		Foreign Key to file
type_id	integer	No		Foreign Key to cvterm
value	text	Yes		
rank	integer	No	0	

The fileloc Table

The *fileloc* table indicates where files can be downloaded. The *uri* column must contain the URI of the file. Even local files have a URI. For example a Drupal URI usually has a *public://* URI prefix. For example: ``public://tripal/users/1/Citrus_sinensis-scaffold0.fasta``. When a file has more than one location to download the can be ordered by setting the *rank* column. The Tripal file module automatically fills in the *size* and *md5checksum* values for local files. If you are adding file locations via the bulk loader you must provide these or the will not be available.

Column	Type	Null	Default Value	Constraint
fileloc_id	integer	No	(auto)	Primary Key
file_id	integer	No		Foreign Key to file
uri	text	No		
rank	text	No	0	
md5checksum	integer	Yes		
size	char(1024)	Yes		

3.4 Associating a File with Other Content

Now that we have a file page we can associate that file with any other Tripal-based content. As stated earlier, we will follow the User's Guide example where an Analysis page for the *Citrus sinensis* whole genome page was already created. We will associate the whole genome FASTA file to the analysis page.

Before we can associate a file with an analysis, we must first add a new field for the file to the Analysis content type. Navigate to **Admin > Structure > Tripal Content Types** and click the **manage fields** link in the **Analysis** row.

Analysis

[Home](#) » [Administration](#) » [Structure](#) » [Tripal Content Types](#) » [Analysis](#)

[+ Check for new fields](#)

LABEL	MACHINE NAME	FIELD TYPE	WIDGET	OPERATIONS
+ Resource Type	<code>rdfs__type</code>	Content Type	Content Type	edit
+ URL redirects	<code>redirect</code>	Redirect module form elements		
+ Name	<code>schema__name</code>	Text	Text field	edit

Click the **Check for new fields** link. Unlike adding fields for Metadata, the Tripal File module can automatically add fields for files. You should see the following message after clicking the link:

Added field: `sio__file`

Once again we must tell Drupal where on the page to display this new field. Follow the instructions on the [Configuring Page Layout](#) of the Tripal User's guide to learn how to do this. For this example, we will use the Tripal Default Display option to automatically place the file field.

Note: You may be asking, why do I have to add the field then set the display? Why can't the module do it automatically? The reason is that Tripal supports full customization of every page by the site admin and the Tripal File module does not want to make any assumptions about when and where any fields will be displayed. So the site admin must add them manually.

Next, find the Analysis page via **Admin > Find Tripal Content**. If you followed the Tripal User's Guide it will be named "Whole Genome Assembly and Annotation of Citrus Sinensis (JGI)". But if do not have this page, any practice Analysis page will do. Edit the page and scroll to the new section titled **File**. You should see a drop down containing a list of available files.

[Show row weights](#)

FILE

Available Files

+

Citrus sinensis Whole Genome Assembly v1.0 (FASTA format) ▼

Please select a file.

Available Files

+

-- Select a File -- ▼

Please select a file.

A file.

Add another item

Select the file and click the **Save** button at the bottom. We now have a **File** item for the page.

Whole Genome Assembly and Annotation of Citrus Sinensis (JGI)

View

Edit

Reload

[Summary](#)
[Annotations](#)
[File](#)
[Publication](#)
[Relationship](#)

File	
File	Type
Citrus sinensis Whole Genome Assembly v1.0	FASTA format

Summary	
Resource Type	Analysis
Name	Whole Genome Assembly and Annotation of Citrus Sinensis (JGI)
Description	Note: The following text comes from phytozome.org:

Clicking the file link will take the user to the full file page where they can download the file, view the license information, and view metadata about the file.

3.5 Accessing Files via Web-Services

All information about files can be obtained via the built-in Tripal Content web services. You can learn about how to access Tripal web services via the [Web Services](#) page of the Tripal User's Guide. When web services are enabled you can:

1. Retrieve the list of files associated with any content.
2. Retrieve all information about a file, including its name, description, license, source, download locations and metadata.

The following screen shot provides an example of the JSON returned for the Analysis content to which we associated the genome file:

```
{
  "@context": "http://localhost/sites/default/files/tripal/ws/context/content.v0_1.analysis.2.json",
  "@id": "http://localhost/web-services/content/v0.1/Analysis/2",
  "@type": "Analysis",
  "label": "Whole Genome Assembly and Annotation of Citrus Sinensis (JGI)",
  "ItemPage": "http://localhost/bio_data/2",
  "type": "Analysis",
  "name": "Whole Genome Assembly and Annotation of Citrus Sinensis (JGI)",
  "description": "<p><strong><em>Note: </em>The following text comes from phytozome.org:</strong></p>\r\n\r\n<p><em>version (v.1) of the assembly is 319 Mb spread over 12,574 scaffolds. Half the genome is accounted for by 236 sc gene set (orangel.1) integrates 3.8 million ESTs with homology and ab initio-based gene predictions (see below). predicted, each with a primary transcript. An additional 20,771 alternative transcripts have been predicted, ger transcripts. 16,318 primary transcripts have EST support over at least 50% of their length. Two-fifths of the pr support over 100% of their length.</p>\r\n\r\n<p><u>Sequencing Method</u><br />\r\nGenomic sequence was generate approach with 2Gb sequence coming from GS FLX Titanium; 2.4 Gb from FLX Standard; 440 Mb from Sanger paired-end libraries</p>\r\n\r\n<p><u>Assembly Method</u><br />\r\nThe 25.5 million 454 reads and 623k Sanger sequence reac effort by 454 Life Sciences, University of Florida and JGI. The assembly was generated by Brian Desany at 454 Li assembler.</p>\r\n\r\n<p><u>Identification of Repeats</u><br />\r\nA de novo repeat library was made by running Hubley) on the genome to produce a library of repeat sequences. Sequences with Pfam domains associated with non-library of repeat sequences and the library was then used to mask 31% of the genome with RepeatMasker.</p>\r\n\r\n/>\r\nWe aligned the sweet orange EST sequences using Brian Haas's PASA pipeline which aligns ESTs to the best ; filters hits to ensure proper splice boundaries.</p>\r\n",
  "software": "Assembly and Annotation Performed by JGI",
  "version_number": "Phytozome v9",
  "algorithm": null,
  "time_executed": "2021-02-01 17:53:00",
  "source_data": {
    "name": "JGI Citrus sinensis assembly/annotation v1.0 (154)",
    "version_number": null,
    "uri": "http://www.phytozome.net/citrus.php"
  },
  "database_cross_reference": "http://localhost/web-services/content/v0.1/Analysis/2/database+cross+reference",
  "annotation": "http://localhost/web-services/content/v0.1/Analysis/2/annotation",
  "publication": "http://localhost/web-services/content/v0.1/Analysis/2/publication",
  "relationship": "http://localhost/web-services/content/v0.1/Analysis/2/relationship",
  "file": "http://localhost/web-services/content/v0.1/Analysis/2/file"
}
```

Note, the file key at the bottom of the JSON output? Following this URL provides information about the file. If more than one file is associated with the content then an array of files is provided.

Below is a screenshot of the JSON result for the file:

```

"@context": "http://localhost/sites/default/files/tripal/ws/context/content.v0_1.file.27.json",
"@id": "http://localhost/web-services/content/v0.1/File/27",
"@type": "file",
"label": "Citrus sinensis Whole Genome Assembly v1.0",
"ItemPage": "http://localhost/bio_data/27",
"type": "File",
"name": "Citrus sinensis Whole Genome Assembly v1.0",
"description": "<p>The whole genome assembly, v1.0, of <em>Citrus sinensis</em>.</p>\r\n",
"additionaltype": "FASTA format",
"contact": "http://localhost/web-services/content/v0.1/File/27/contact",
"itemlocation": {
  "@id": "http://localhost/web-services/content/v0.1/File/27/itemLocation",
  "@type": "Collection",
  "label": "Download",
  "totalItems": 1,
  "member": [
    {
      "uri": "http://localhost/sites/default/files/tripal/users/1/Citrus_sinensis-scaffold00001.fas",
      "rank_order": "0",
      "sequence_checksum": "85934c1cc36308d829dc73e4eb9dcea3",
      "filesize": "6.03MB"
    }
  ]
},
"license": {
  "name": "Attribution 4.0 International (CC BY 4.0)",
  "disambiguatingdescription": "<p>You are free to:</p>\r\n\r\n<ul>\r\n\r\n<li>\r\n\r\n<p><strong>Share</strong> or format</p>\r\n\r\n</li>\r\n\r\n<li>\r\n\r\n<p><strong>Adapt</strong>: remix, transform and build upon t</p>\r\n\r\n</li>\r\n\r\n</ul>\r\n\r\n<p>The licensor cannot revoke these freedoms as long as you follow t</p>\r\n\r\n<ul>\r\n\r\n<li>\r\n\r\n<p><strong>Attribution</strong> You must give appropriate credit, pr</p>\r\n\r\n</li>\r\n\r\n</ul>\r\n\r\n<p>were made. You may do so in any reasonable manner, but not in any way that suggests the licensor end</p>\r\n\r\n<strong>No additional restrictions</strong> You may not apply legal terms or technological measures</p>\r\n\r\n</li>\r\n\r\n</ul>\r\n\r\n<p><strong>Notices:</strong></p>\r\n\r\n<ul>\r\n\r\n<li>\r\n\r\n<p>elements of the material in the public domain or where your use is permitted by an applicable except</p>\r\n\r\n</li>\r\n\r\n</ul>\r\n\r\n<p>warranties are given. The license may not give you all of the permissions necessary for your intende</p>\r\n\r\n</li>\r\n\r\n</ul>\r\n\r\n",
  "uri": "https://creativecommons.org/licenses/by/4.0/legalcode",
  "@id": "http://localhost/web-services/content/v0.1/License/26",
  "@type": "license"
},
"references": "http://localhost/web-services/content/v0.1/File/27/references",
"genus": "Citrus",
"species": "sinensis"

```