

Ed Ryder Department of Genetics

Flannotator installation manual and getting started guide

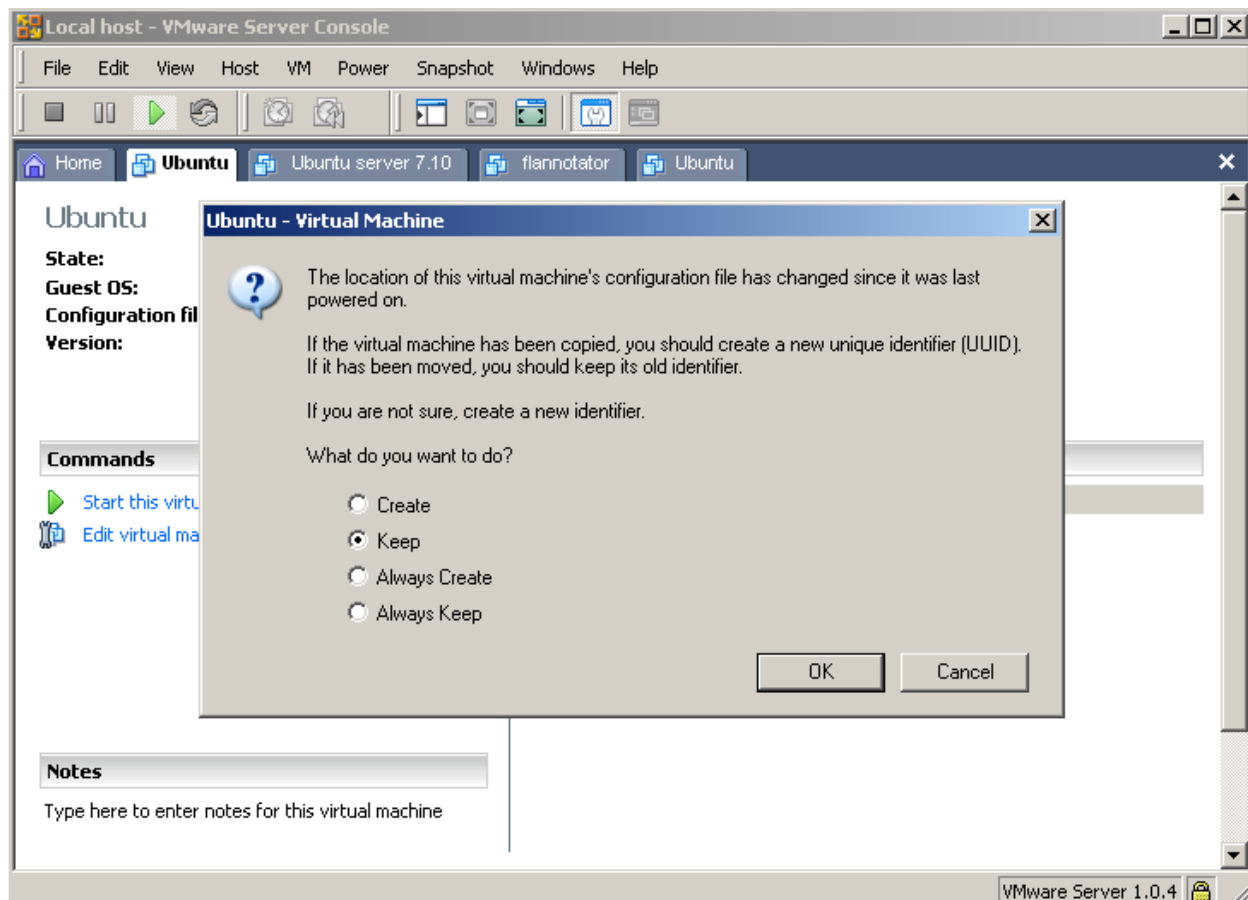
Software installation

The Flannotator VMware image comes as an already configured linux operating system running an apache2 web server and mysql 5 database, along with the Flannotator and Gbrowse 1.68 software. It requires either VMware player or server, both of which are free and can be downloaded from www.vmware.com. The Flannotator also comes preloaded with the *Drosophila* release5.3 gene annotation (FlyBase), anatomy cv terms (FlyBase), GO Slim cellular_component (GO), and selected spatial descriptors (FlyBase) and developmental vocabularies (FlyBase).

Since hosts may be set up in a variety of ways please consult with your IT department on how to get VMware player / server and the OS image talking to other machines on your network. As default it is set up to work on a NAT system and accessed locally, but you may wish to assign it its own IP address on a networked server especially if you are going to have multiple users.

For this example we'll assume it's being run under NAT on a user's own machine.

1. Load VMware player and click *file ->open* and browse to where your vmware image is.
2. Click on the green play arrow to start the machine
3. The following dialogue box may appear



4. If you are running just one instance of the Flannotator, selecting *keep* should mean it just works. You can also create a new UUID if you wish (especially if you are running multiple copies) although this may require some working knowledge of vmware to set up the networking properly.
5. If all goes well the login prompt will appear. You will want to change the defaults to keep things more secure using standard linux methods
 - a. Username is *flanuser*
 - b. Password is *flanpass*
6. You should now be logged on to the shell, and the web server and databases will now be running in the background
7. Find out what your vmware ip is by using the command *ifconfig*
 - a. In this case it's 192.168.198.131

```

flanuser@flannotator:~$
flanuser@flannotator:~$ ifconfig
eth0      Link encap:Ethernet  HWaddr 00:0c:29:0e:2b:84
          inet addr:192.168.198.131  Bcast:192.168.198.255  Mask:255.255.255.0
          inet6 addr: fe80::20c:29ff:fe0e:2b84/64  Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:15  errors:0  dropped:0  overruns:0  frame:0
          TX packets:17  errors:0  dropped:0  overruns:0  carrier:0
          collisions:0  txqueuelen:1000
          RX bytes:1920 (1.8 KB)  TX bytes:1774 (1.7 KB)
          Interrupt:16  Base address:0x1400

lo        Link encap:Local Loopback
          inet addr:127.0.0.1  Mask:255.0.0.0
          inet6 addr: ::1/128  Scope:Host
          UP LOOPBACK RUNNING  MTU:16436  Metric:1
          RX packets:0  errors:0  dropped:0  overruns:0  frame:0
          TX packets:0  errors:0  dropped:0  overruns:0  carrier:0
          collisions:0  txqueuelen:0
          RX bytes:0 (0.0 B)  TX bytes:0 (0.0 B)

```

8. The vmware console is very slow and has an annoying habit sometimes of sending two carriage returns instead of one, which can make it very difficult to enter passwords. If you plan to use the console a lot we recommend installing an ssh client (eg Putty) and accessing it that way. The flannotator comes with sshd already installed.
9. Load up your favourite web browser. We recommend FireFox for maximum compatibility across all platforms (windows, mac os, linux) but Internet Explorer 6+ will work just as well. Please note Safari is NOT supported at this time, as it tends to do strange things and anyway FF is a lot better.
10. Enter the IP address of your web server in the address line and add /flannotator on the end



11. The Flannotator homepage should then appear.



[Annotation](#) [Stocks](#) [Browser](#) [Start a query](#) [Docs](#) [Account](#)

Please log on to continue: Username Password [log on](#)

The Flannotator allows annotation of gene expression at all stages of development and tissue types (including sub cellular location) using standard *Drosophila* anatomy ontology. All methods of input use a controlled vocabulary to ensure data integrity.

The web-based input and retrieval system allows multiple groups to work in collaboration and share images and annotation easily, whilst still protecting the original data. If you have any questions about the project or need an access code please contact Ed Ryder at e.ryder[at]gen.cam.ac.uk

Latest news:

[random pic](#)

[random pic](#)

[random pic](#)

[random pic](#)

[random pic](#)

[random pic](#)

[random pic](#)

[random pic](#)

[random pic](#)

[FlyPROT](#)



Flannotator v2.2

Ed Ryder, Department of Genetics, University of Cambridge, UK. 2008

12. Log on to the system with username: *admin*, password: *admin*
13. Now might be a good time to change the admin password
 - a. Select *Account* -> *Admin*
 - b. Click on *edit users*
 - c. Choose 'admin'
 - d. Click on *reset* password and enter a new one
 - e. Don't forget your new password!
14. For more general administration tasks (eg adding new users) please refer to the administration manual
15. You should now be ready to add stocks

Things to do before adding stocks to the database

1. Before you start adding new stocks, several things in the admin section need doing first, so the software can work out what frame insertions are in later on. Please refer to the administration manual on the specifics on how to add / edit features.
 1. *Add a transposon construct*
 - i. This will be the gene trap construct and frame used in the screen.
 - ii. Needed to work out whether an insertion is a likely trap or not
 2. *Add a donor line*
 - i. This will be your ammunition stock from where the transposon construct is jumped from
 - ii. This needs the construct information associated with it.
 - iii. Sequencing information about the donor is optional, but will allow the software to automatically filter out spurious donor sequence from the stock reports later on
 3. *Stock owner*
 - i. The person(s) who look after the stocks. Useful for data tracking purposes.
2. Setting up a COPAS experiment. The flannotator was written to take data from a COPAS select embryo sorter (Union Biometrica) and allows for data tracking for a line right back to the original sort. A perl script is provided (see the flannotator COPAS analysis manual) to analysis this data and for easy upload, with support for Petri-dishes and 24well plates. Even if you don't use a COPAS for generating your lines you need at least one experiment defined in order for the system to work.
 1. Click on *Add copas data*, and select the plate used collected embryos/ larvae.
 2. If you are not using copas data, just click on Petri dish.
 3. Enter the experiment number (must be a number) and date

COPAS sorting data entry

Experiment number:

Date:

Number of plates containing an event:

Donor Line:

Jumps from:

Paste in data from copas_summariser.pl .out file:

```
Experiment: 2008_07_09_430
Axis scales: EXT,TOF,Green,Yellow,Red
Axis scales values: 1024,2048,4096,256,256
Size gate: TOF v EXT
Sort gate: Yellow v Red
Gate region coordinates: 17,68,8,43,30,10,106,63,152,119,63,134
Sort region coordinates: 256,83,117,56,49,32,25,1,256,1
Total events: 77950
Events passed sizing gate: 76553 (98.208%)
Events passed through size and sort gates: 3 (0.004% from sizing gate, 0.004% total events)
Events giving correct collect code: 2 (0.003% from sizing gate, 0.003% total events)
```

Paste in data from copas_summariser.pl .well file:

```
### 2008_07_09_430
PlateCol Row ID Yellow Red Flu-factor
1 A 1 3225 34 6 5.7
1 A 2 49196 69 10 6.9
1 A 3 76370 33 5 6.6
```

Comments:

demo

4. Choose the donor line (which was set up earlier) and whether the jumps were set up in the male or female germ line
5. If you have used the perl analysis scripts, paste in the relevant data in the text boxes. If not, just leave these blank and click on 'add data'
6. The next page shows a plan of the plate (for Petri dishes just consider A1 the whole thing). Enter the number of fluorescent embryos (Y) and non fluorescent (A) embryos found in each well. A box is also provided for comments for each well. If not using the copas system, just enter a 1 in the Y box for well A1
7. COPAS experiments can be edited later if required, and information about each embryo (eg whether it lived or not) can be added if required.
8. Click on *Add data* to commit the data to the database.

This may be a good time to add new users to the system. Please refer to the administration manual on how to do this.

Adding new stocks

1. Click on stocks -> add new stocks.
2. Enter what data you can about the stock – this can be edited later to add / change more information if required.

Enter a new stock

Stock information

Stock Name:

Line Number:

COPAS experiment:

COPAS plate:

Well position:

Donor Line:

Stock Owner:

Isolated fly gender:

Lethality:

Stock kept?

Plated for sequencing?

Phenotype and genotype

YFP:

CFP transposase?

Donor insert present?

Eye colour:

Eye colour variegation?

Balancing

Balanced?

Chromosome:

Balancer:

Other observations

Derived from: [select](#)

Dead?

Notes:

[Cancel](#)

3. Sometimes several lines may produced (eg different balancers are used) from the original stock. Either multiple line numbers can be defined or it may be better

to just call them a different stock number completely and select that they are derived from the original one. This also helps if the other lines are discarded

1. Please note that multiple line numbers were not pursued much when writing other parts of the software and therefore may not display properly in stock reports etc. Please use new stock names and derived status until a new release is out.
4. Most of the options should be pretty self explanatory, but just in case
 1. *YFP*: Is the line confirmed fluorescent? This is usually checked by creating a stock of the putative positive lines and resorting.
 2. *CFP Transposase*: Is the transposase source still present (usually on a marked chromosome. It is a good idea to remove this from the stock as soon as possible to prevent further jumping
 3. *Donor insert present?*: Donors on a marked chromosome balancer can be removed from the stock, which helps sequencing further down the line.
5. If the stock is balanced please choose the chromosome and balancer used
6. If the stock is lost please set the *dead* option to yes
7. Click on '*add new stock*' when you are finished
8. All stock editing is saved and tracked. If you wish to see the life story of a particular stock click on *stocks* -> *stock history* and choose the line of interest to display all of the edits and at what date they occurred.

Stocks and batches

Not all stocks will be fluorescent and not all fluorescent stocks may be in unique genes or introns. To help prevent the people doing the annotation looking at multiple instances of the same thing, or indeed looking at wild type flies by mistake, stocks can be split up into batches, which flags them as lines to look at. Only lines that are part of batches are viewable in the stock report browser.

You can either have lines in one big batch (if you are just one group looking at a set number of lines) or you can split them up into several smaller batches as new lines become available. To create a batch and add lines go to *account* -> *admin* -> *add batches*

Add a batch number to a list of stocks

Entering a batch number that already exists will add more lines to that batch.

Enter batch number:

Paste in CPTI codes, separated by newlines:

```
CPTI-004000  
CPTI-004001  
CPTI-004002  
CPTI-004003  
CPTI-004004
```

Enter a batch number and paste in the line numbers (either from a text file or from Excel) into the text box: lines should be separated by new lines. If you enter a batch number that already exists the new stocks will be added to that batch.

Adding a line which already exists in another batch will overwrite the original information.

Another function of batches is that they allow you to place limits on what data different annotation groups can view. You may not wish annotators to view sequence data or other annotations until they have completed annotation of a particular batch. Setting the 'batch complete' flag on *edit users* in the admin panel will give access to any data for that particular line for that user. Users with a user level of 3 or above (ie power users or admins) will be able to view everything no matter what batches they have annotated.

Please note completing a batch and getting the data released is not done automatically – this has to be set manually by the admin.

Adding sequencing data

1. Please download and view the example plate form from the flannotator website. You will need this to have the data in the correct format for uploading using the web form.
2. Please note our sequencing analysis pipeline is not included in this software package – it's up to you how to get your sequences analysed!
3. Not all samples will have worked and will have to be given the code NULL instead of sequencing results. Any blank cells in the sheet where the plate plan is should be filled with NULL values.
4. Samples also given codes (1 or 0) depending on their fate:

1. PCR – sample PCR worked
 2. Sequenced – sample produced a readable sequence trace
 3. Blast – sample produced a valid mapping blast result
 4. Multi-insert – sample had mixed sequence in the trace file
 5. Multi-blast – sample could not be aligned to a unique place of the genome
 6. Short – sample sequence was too short to produce an alignment
5. All samples must be given a code for each of these outcomes.
6. Remember to save the file when you are finished!

Adding data to the Flannnotator

1. The data can now be uploaded to the Flannnotator

Column	Row	element_name	line_id	primer_id	restriction	plate_id	comments	pcr?	sequence?	blast?	multi-insert?	multi-blast?	short?	start	stop	seq_chr_id	strand	expect	q_start	q_stop	identity	pcr_length	sequence				
7	1	F	CPT1-000095	2	dpn	5	157 NULL	1	1	1	0	0	0	0	814953	814903	20	1	4.00E-54	1	106	100	100	AATATATCCGAT			
8	1	G	CPT1-000503	1	dpn	5	157 NULL	0	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
9	1	H	CPT1-000671	1	dpn	5	157 NULL	0	0	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
10	2	A	CPT1-000803	2	dpn	5	157 NULL	1	1	1	0	0	0	0	2547221	2547935	X	-1	0	1	715	100	715	AAGTCTGGGTTG			
11	2	B	CPT1-000930	1	dpn	5	157 NULL	1	1	1	0	0	0	0	5004955	5005374	Z	-1	0	1	978	98	97433809	979	CTGCTGGTGGCT		
12	2	C	CPT1-002174	1	dpn	5	157 NULL	1	1	0	1	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
13	2	D	CPT1-002140	1	dpn	5	157 NULL	0	0	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
14	2	E	CPT1-002487	1	dpn	5	157 NULL	1	1	0	0	0	0	0	1	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
15	2	F	CPT1-002496	1	dpn	5	157 NULL	0	0	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
16	2	G	CPT1-002529	1	dpn	5	157 NULL	0	0	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
17	2	H	CPT1-002482	1	dpn	5	157 NULL	1	1	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
18	3	A	CPT1-002377	1	dpn	5	157 NULL	0	0	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
19	3	B	CPT1-002370	1	dpn	5	157 NULL	0	0	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
20	3	C	CPT1-002523	1	dpn	5	157 NULL	0	0	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
21	3	D	CPT1-002181	1	dpn	5	157 NULL	1	1	1	0	0	0	0	0	7233916	7234592	X	1	0	1	676	98	6763366	676	GCGAACCAAAA	
22	3	E	CPT1-002186	1	dpn	5	157 NULL	1	1	1	0	0	0	0	0	2546430	2547034	X	1	0	1	612	96	6122376	612	ACAATAATGCGC	
23	3	F	CPT1-002391	1	dpn	5	157 NULL	1	1	0	1	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
24	3	G	CPT1-002180	1	dpn	5	157 NULL	1	1	1	0	0	0	0	0	8149537	8149099	20	1	2.00E-81	1	152	100	152	TTAGCGATTITAG		
25	3	H	CPT1-002535	1	dpn	5	157 NULL	1	1	0	0	0	0	0	1	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
26	4	A	CPT1-001254	1	dpn	5	157 NULL	1	1	1	0	0	0	0	0	1594007	1594010	3L	-1	0	1	302	98	3034578	303	GTACCCCTGTA	
27	4	B	CPT1-000586	1	dpn	5	157 NULL	1	1	1	0	0	0	0	0	146236	144706	3R	-1	0	1	531	99	4360326	531	ACAAACGAGCA	
28	4	C	CPT1-000576	1	dpn	5	157 NULL	1	1	1	0	0	0	0	0	960051	948099		4	-1	0	32	997	96	79098027	1014	GTACTACTAAR
29	4	D	CPT1-001962	1	dpn	5	157 NULL	1	1	1	0	0	0	0	0	9806535	9806403	2R	-1	2.00E-63	3	136	96	50746269	136	GTAGTACTAAGC	
30	4	E	CPT1-001933	1	dpn	5	157 NULL	1	1	1	0	0	0	0	0	4026245	4026337	2L	-1	2.00E-46	1	93	100	93	AGTAGTAAACTA		
31	4	F	CPT1-002366	1	dpn	5	157 NULL	1	1	1	0	0	0	0	0	6114670	6114760	2L	-1	2.00E-39	9	96	97	8027979	792	AGATGAAAGTAA	
32	4	G	CPT1-000501	1	dpn	5	157 NULL	1	1	0	1	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
33	4	H	CPT1-000324	1	dpn	5	157 NULL	1	1	0	1	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
34	5	A	CPT1-000400	1	dpn	5	157 NULL	1	1	0	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
35	5	B	CPT1-000578	1	dpn	5	157 NULL	1	1	0	0	0	0	0	0	17368467	17368905	20	-1	0	1	483	99	37889199	483	TAATAATGCGATA	
36	5	C	CPT1-000577	1	dpn	5	157 NULL	0	0	0	0	0	0	0	0	0	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL			
37	5	D	CPT1-001822	1	dpn	5	157 NULL	1	1	1	0	0	0	0	0	6417968	6417518	2S	-1	0	1	452	99	77826106	452	TCTTTGATTGTC	

2. Highlight the area on the plate record from the sample names to the column containing the sequence data
3. Paste the highlighted area into the sequencing upload page on the Flannnotator. Click on 'add data to database' to upload the data.
4. The sequencing data, along with the donor and frame information for the stock should now be able to automatically determine whether any insertion into an intron is in frame or not to produce a fluorescent gene trap.
5. Sequencing data can be added at any time in the future (it may be a while between a stock being established and the sequence data being determined) and multiple sequences can be added to any stock. They can also be deleted if required via the admin menu.
6. Sequence that includes the donor will be automatically screened out of the stock report if the donor location is added in the admin menu for donors.

Paste in sequence results from Excel form:

CPTI-002695	2	2	dpm	5	157	NULL	1	1	1	0	0	0	8149583	8149688	2R
1	4.00E-54	1	106	100	106										
AATATATCGCATCGCTTTCTACGTGGCCAATTTGTTTCAGTGATTCAACCAACGCTGCTAACTTGGCCAGAAAGCAAC															
CPTI-002603	1	2	dpm	5	157	NULL	0	0	0	0	0	0	NULL	NULL	
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
CPTI-002671	1	2	dpm	5	157	NULL	0	0	0	0	0	0	NULL	NULL	
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
CPTI-002603	2	2	dpm	5	157	NULL	1	1	1	0	0	0	2547221	2547935	X
1	0	1	715	100	715										
AAGTCTGGGTTGAAGGCAGTTCGCTGGCTTACTTTTGCATATTTTTCTTTGTTTGTTCAGGTTTCTTGGCTCGAC															
CPTI-002200	1	2	dpm	5	157	NULL	1	1	1	0	0	0	5004355	5003374	2L
-1	0	1	978	98.57433809	979										
CTGCTTGGTCCTTAACCTGCACTAAGTCTTGTGCACTGAGCATTGTTTTGCACGCCATCTCCTATCCCTCCTGTTTTCA															
CPTI-002174	1	2	dpm	5	157	NULL	1	1	0	1	0	0	NULL	NULL	
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
CPTI-002140	1	2	dpm	5	157	NULL	0	0	0	0	0	0	NULL	NULL	
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
CPTI-002487	1	2	dpm	5	157	NULL	1	1	0	0	0	1	NULL	NULL	
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
CPTI-002496	1	2	dpm	5	157	NULL	0	0	0	0	0	0	NULL	NULL	
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
CPTI-002529	1	2	dpm	5	157	NULL	0	0	0	0	0	0	NULL	NULL	
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	
CPTI-002482	1	2	dpm	5	157	NULL	1	1	0	0	1	0	NULL	NULL	

Add data to database
Reset Form

Congratulations you have now hopefully added your first group of stocks and maybe even some sequencing data.

What to do next

Before users can start annotating there are a few more things that need doing in the admin panel.

Microscopes

Adding microscope information will help annotators better define their experiments and images. Click on *account* -> *admin* -> *add microscope* and enter the name (eg Leica MZ16F) and type (fluorescence).

Tissue types

The Flannotator comes preconfigured with a number of common tissue types to use as a base line for annotating. Occasionally it may be necessary to add new tissue types for

annotators to use. You can do this by selecting a cv anatomy term in the left box and clicking 'add' to add them to the list on the right. Click on 'update tissues' when you are finished.

The following form allows you to customise the tissue categories

2. Choose from a list of tissue types.

Term list		Tissue categories
mushroom body	Add >>	adult tracheal system
mushroom body primordium		cell
myoblast		central nervous system
myofibril		circulatory system
myotube		digestive system
mystery cell		embryonic adipose system
Nebenkern		embryonic brain
Nebenkern derivative		embryonic central nervous system
neck		embryonic circulatory system
nerve		embryonic digestive system
		embryonic endocrine system
		Remove
		Update tissues

Select Matching Options: Search

You can also remove tissue categories by highlighting one and clicking on 'remove'. If you try to remove a tissue type which has any annotations associated with it an error will be displayed and it will not be removed.

Random images

The front page of the Flannotator can display random images from the database which are linked to their respective stock reports. To do this simply copy thumbnails from the various annotation directories in the *files* sub directory (thumbnails have the prefix 'small_') to the *random* directory. This will have to be done via the console or ssh session.

1. `cd /var/www/flannotator/files`
2. Navigate through the various user ID sub dirs and experiment groups until you get to a thumbnails directory
3. `cp small_file.jpg /var/www/flannotator/random`