

Getting the CHEORI Screaton Lab Online

by Charles Malo,

SEG 3901

Coop Session May-August 2010

Coop Session placement:

Children Hospital of Eastern Ontario (CHEO)/

Apoptosis Research Center (ARC)

*By Charles Malo, Children Hospital of Eastern Ontario, Ottawa, Ontario, Canada
With supervision by Dr. Robert Screenshot and Stephen Blair*

May-August 2010

Keywords: RNAi, plate, website, Drupal, apoptosis, cell, genome.

Table of Contents:

A. Abstract

B. Screenshot Lab Website Development

I. Introduction

II. Design

1) Setting up the Drupal based website

2) Designing the layout and theme.

III. Maintenance

C. ARC webpage Analysis tools modification

I. Introduction

II. Tools Modifications

1) Data Output page

2) Analysis tools login restriction

3) RNAi pipeline tool Fix

D. Conclusion

E. Copyrights and Acknowledgements

F. References

List of Figures

Fig 1: People Page.

Fig 2: Administer Menu:

A. Abstract

For my second work term (SEG 3901), completed from May to August 2010, I was employed as a software developer in charge of "getting the Screaton Lab Online". My duties were to create, assisted by both supervisors, a website for the Screaton Lab, led by Dr. Robert Screaton, as well as modifying current programs available on the Apoptosis Research Center (ARC) website.

In particular, I was asked to establish a website for the Screaton team in order for them to publish their projects as well as their data online, making it easier for them to access them at any time. However, my first assignment was to edit some tools available on the ARC website, which the Screaton Lab used to be part of. Some of those tools were a Platemap Data Output page, the RNAi analysis pipeline, and other minor tools used daily in the lab members' daily routine. More detailed information about my task is included in the following.

B. Setting up the Drupal based website

I. Introduction

The Screaton Lab is a team consisting of mostly of chemical or biochemical students that work under Dr. Robert Screaton's supervision. The main task of the group is to conduct various researches focusing mainly on studying cell nuclei and monitor how various substances affect the cellular activity over a period of time.

The group already had a website to display its work; however it was out-dated and needed to attain more functions. From that point, I was asked to build a new website that offers specific "tools" for users to help them post their current data, statistics, results, conclusions and possible theories based on their current experiments. On a side note, the website should provide an employee database alongside with a contact list in case someone needs to communicate with a specific member.

II. Design

1) Setting up the Drupal based website:

In order to avoid building the Screaton Lab website from scratch, my supervisor and I agreed to use Drupal in order to create the overall website without the hassle of manually writing every required aspect of the site.

Drupal is a Content Management Software, (CMS), used particularly in creating websites published online. Similar to Joomla, and other commonly used CMS programs, a user can establish a website database on a server and continually build an online website template simply by installing the program. Such program was a good idea to use, since it only needs maintenance from a person with experience in using Drupal, not in programming. Most of my colleagues accessing the website fall in the previously mentioned category.

When I started working on the website, my supervisor had already prepared a domain as well as a database to host the website. My initial work was to search online for plug-ins and modules to install into the Drupal directory.

Modules resemble small "applications" that can be downloaded from the Drupal website (www.drupal.org) and installed into the website's database in order to unlock a specific feature. For about a month of my work term, my main task was to research about modules as well as downloading the required ones for the website to be set up properly to meet the requests of my supervisors.

I was particularly asked to download and test a decent text editor that would allow the user to publish content online that can include text, as well as images. Drupal does have a variety of free modules available online, however finding the right module that does not cause a conflict with currently installed modules was somewhat a problem.

Details aside, some modules I had to find online required some "fine tuning" which translates into modifying some of the code files downloaded to make the applications compatible and ready to use.

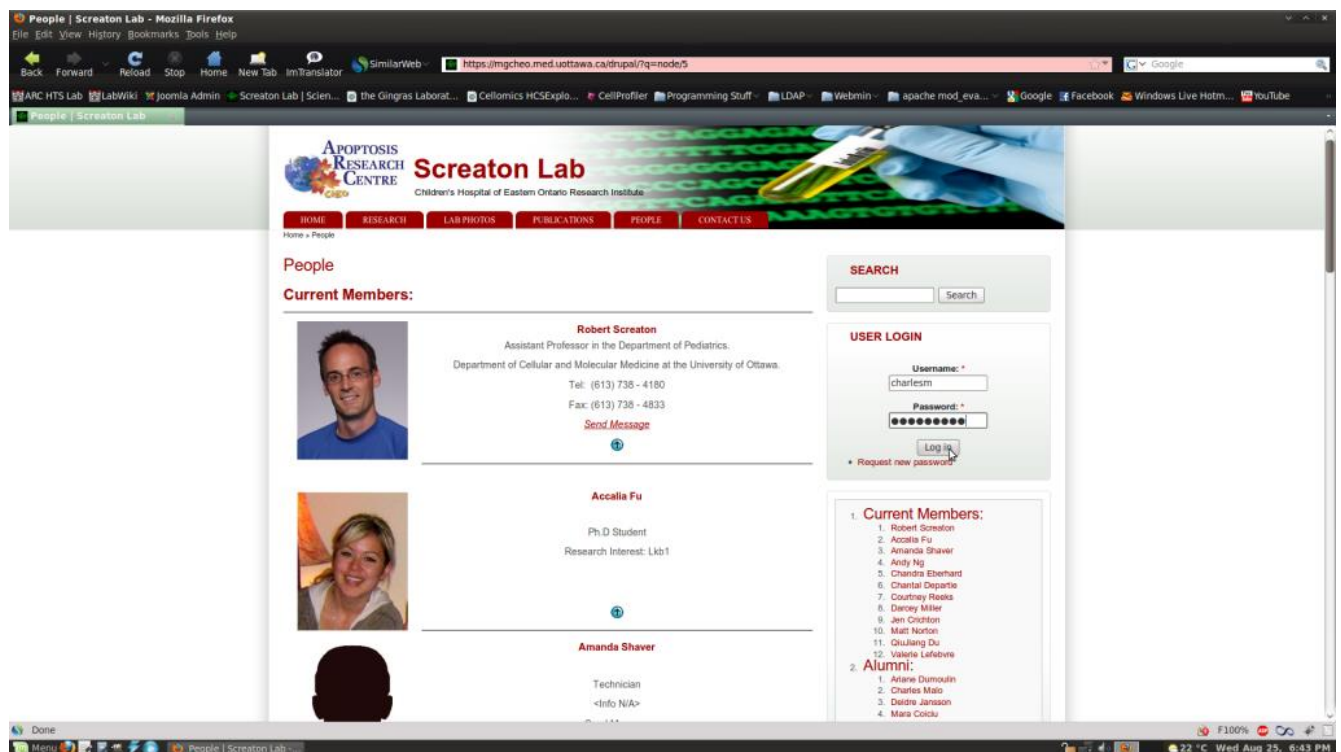


Fig 1: People Page: A section of the website showing the members of the ScreatonLab. Page still under construction.

Sometimes, a module would become outdated, or incompatible with other modules or with the current version of the CMS used. This incident had occurred many times, and it was my biggest challenge when creating the website, as the modules would appear to be the right ones, yet end up creating more errors in the website design steps. This required a lot of maintenance, since modules get updated on a weekly basis, whenever new glitches appear.

This was somewhat an issue, since sometimes we would be searching for a specific module that is yet to be developed, and the request for it would have been postponed for a later time when such module would be created.

Other than Drupal for the CMS used, we could have used Joomla since the ARC webpage was built with that software. However, Drupal offers more flexibility, as its online database supplies the user with more applications that are up to date and action specific than Joomla. The downside is that Drupal is by far more complicated than Joomla in terms of functionality: the learning curve on Drupal is much steeper than the one on Joomla.

Although it did require us great deal of time and effort in order to understand the complexity behind the CMS that is Drupal, such skill can only be beneficial in the future, since the Drupal software is universal when it comes to creating websites.

In the following image, we can clearly see the main administrator menu from which a user can create content for the webpage, or edit the modules, themes and current menus in order to change the appearance as well as functionality of the website itself. The other categories that are not circled aren't really put into emphasis in the design process.

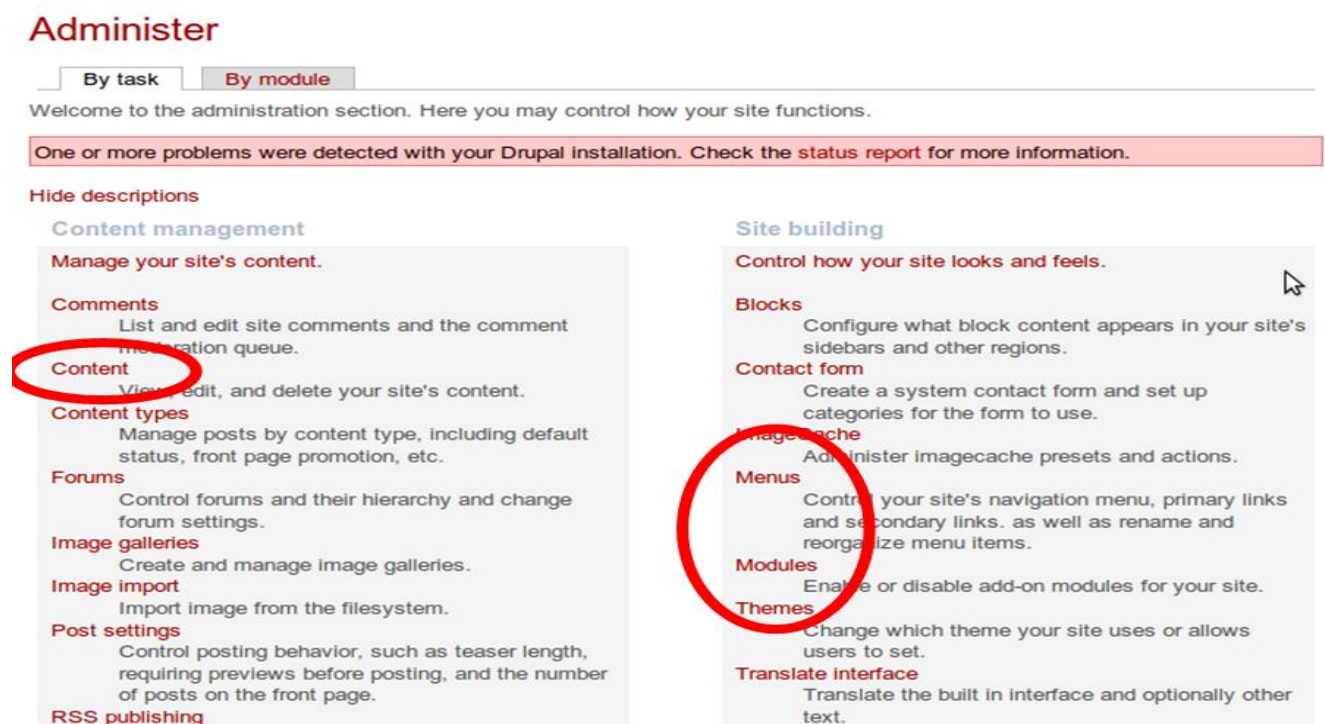


Fig 2: Administer Menu: A section in the website's control panel that enables the user to modify the webpage.

Later on during my work term, I was required to install a few modules that would help add media to the website, as well as create a step by step tutorial for all my colleagues to teach them how to add content to the page while preserving its layout and functionality.

I did face a small challenge in this part of my assignment: a problem occurred that, while working on the webpage itself in order to create a “Contact Us” page containing all the lab members’ info, other students would be working on the same page, and ultimately confusing me since they would make unauthorized changes without consulting me.

At many occasions, I had to personally re-work on the whole page that I worked on for the last couple days. The problem wasn’t simply the fact that content was being lost in such process, but the fact that recreating the content of that page each time had to be done manually and thus it was time consuming.

One possible solution to this problem that was achieved to this problem was discussing the issue with the colleagues and notifying them of future updates of the website prior to any changes done by each user in order for each one to back-up their data.

2) Designing the layout and theme:

Similarly to modules, themes can be downloaded from the same website and installed in order to automatically re-arrange the website and its layout to accommodate the theme change.

The difference between the “themes” section and the “modules” section is that the first one is less complex to modify. Installing a module showed a great deal of complexity when trying to enable a module or disable it, where as a theme is simple. A user can enable a certain theme for the website, disable it, or change the theme for the administer section only of the website to accommodate certain needs, such as viewing the menus differently, and put emphasis on a few links.

A user can download a theme online, and edit certain properties of the theme without the need to edit the themes files on a source code level.

This degree of “easiness” in changing themes would allow the amateur website developer to update his or her webpage without the need of prior knowledge in html and CSS scripting languages. With the fact that most users are of a science background, and not many of them retain experience in programming, it is certain that the users will find this “tool” suitable for their needs. Many Content Management software programs do not even offer that user friendly interface, and rely heavily on the user in programming every change they might require.

The downfall to this tool is that the user gets very dependent on downloading and installing themes, without the need of personally modifying them. This usually causes stress for the user when he or she cannot find a theme that fits their requirements for their ideal theme, since the user would avoid desperately the need to modify the themes manually.

On the other hand, such tool can reduce the need for a skilled programmer, and for example, the need of the creator of the theme himself or herself, since their creation can be found online where it can be updated regularly and automatically, without losing the work done on the theme.

Themes

[List](#) [Configure](#)

Select which themes are available to your users and specify the default theme. To configure site-wide display settings, click the "configure" task above. Alternatively, to override these settings in a specific theme, click the "configure" link for that theme. Note that different themes may have different regions available for displaying content; for consistency in presentation, you may wish to enable only one theme.

To change the appearance of your site, a number of **contributed themes** are available.

See the **available updates** page for information on installed modules and themes with new versions released.



Screenshot	Name	Version	Enabled	Default	Operations
	0 Point 0 Point theme is ideal for a wide range of sites. It has been optimized for e-commerce with Ubercart . <i>By Dr. Radut.</i>	6.x-1.10	<input checked="" type="checkbox"/>	<input checked="" type="radio"/>	configure
	A Cold Day A Cold Day Template	6.x-3.0	<input type="checkbox"/>	<input type="radio"/>	

Fig 3: Themes Section : A section in the administer menu that enables the user to install, and edit themes for displaying the website.

A particular theme, called "Zero Point" that was found online, designed by a physician, met most of my supervisor's requirements. After I downloaded that theme, installed it, and modified it, I started working on more specific tasks such as having a different banner image for specific yet different pages on the website.

Although it seemed like a relatively easy task, it turned out to be harder than anticipated, as Drupal is known for its complexity and steep learning curve. This is a task I have yet to have achieved fully, as it required me to create some *.php files from scratch and integrate them in the website's directory. I created some scripts and files that should be able to load an image directory and rotate the pictures based on the current page number in the website, however this theory was yet to be confirmed as I was requested to move on to something more important.

C. ARC Webpage Analysis Tools

Modification

I. Introduction

The ScreatonLab, in addition to other lab groups in the research institute are part of the Apoptosys Research Center. This grants them access to a webpage that allows them to use certain analysis tools in order to study their data. Offering a link in this report to that webpage would be redundant since the page requires the web surfers to authenticate themselves as part of the research group in order to proceed with viewing the page.

My job did not require me to understand fully what the analysis tools do to the experimental data collected; however, I did need to understand that the tools exert some actions/calculations on the data provided by the user, and output a page providing statistical numbers, tables and plots that help the researcher understand the changes that occurred to a specific molecule or element in a cell after being exposed to certain chemicals.

I was asked by my supervisors to edit some of those tools that are available on the ARC website, such as editing the Data Output Page, Restricting Access to the Analysis Tool to people with authentication, and fixing a bug in the RNAither Pipeline tool that would cause a constant fail of the program when specific data is used for a sample study.


II. Tools Modifications

1) Data Output page:

The Data Output page is a web tool that is used by the lab members in order to view the genome platemaps and study different statistical variables in order to estimate the cell activity that had occurred during the experiments done.

Prior to my coop session, that tool used to handle only one data type at a time during a study. In other words, if a certain molecule was tested differently by different chemical substances and under different processes, only one of them is used to create the data and the platemap table in the output.

However, as mentioned above, sometimes a certain plate is analyzed with different markers in order to study another type of data it might hold. Up until recently, the platemap output only displayed the platemap (plate matrix) output of one data type, a default one.



OLD Data Output

Data Output (OLD VERSION)

Choose a barcode from the list to output the data for that plate:

[Save this data in plate format](#)
[Save this data in column format](#)

Data Type: eqEllipseLWR
Date&Time: 2010-03-24 20:19:12

	01	02	03	04	05	06	07	08	09
A									
B		2.3962	2.3872	2.288	2.3278	2.3939	2.3397	2.3631	2.3127
C		2.788	2.3175	2.3317	2.4196	2.303	2.4583	2.406	2.3854

MAIN MENU

- Home
- LabWiki
- Phone Directory
- **Test**
 - Datapilot MySQL Connection
 - Hello World
 - PHP Test
 - ScreenWatcher
 - OLD Data Output
- Database Tools
- Analysis Tools

USER LOGIN

Hi charlesm,

LOG OUT

Fig 4: Old Data Output Page: The previous version of the Data Output page that members used to access in order to display their current data.

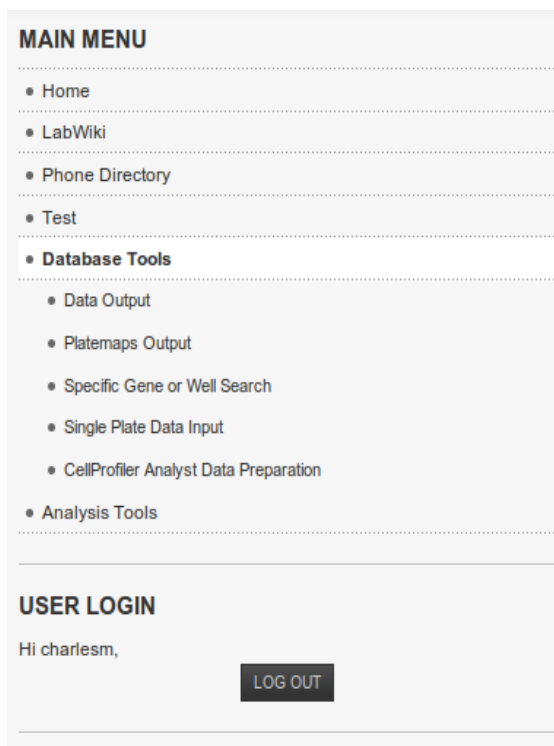
I was required to modify the perl scripts that define this tool so that the user gets the option of viewing a specific data type to display the platemaps in, and not be forced to use the default one at all times. This script needed to be interactive with the user and instantaneous, in such way that, whenever the user changes the data type, the platemap output would automatically be regenerated without the need to re-start the tool process from scratch.

Learning to program in Perl was the hardest part of this task. If someone cannot understand the language, that person cannot correct someone else's grammar or sentence structure! I have personally never seen PERL before, and I had to put my work on hold in order to learn basics of the language.

Once I understood the basic syntax of PERL, using my current knowledge in other programming languages, I modified the PERL and PHP files used to display the platemap output, in order to display an additional drop down menu that allows the user to select and switch in between data types, in order to study the results in different ways.

A common problem I faced in this task is being forced to learn both PERL and PHP simultaneously since the source code files contained a mixture of both languages. Identifying the languages used at each line was a challenge for me at first, but in time, it became an easy task.

The following figure shows the newly added drop down menu, a simple block that would facilitate toggling in between data types instantly, instead of restarting a study all over again.



Data Output

Data Output

Choose a barcode then a data type from the lists to output

[Save this data in plate format](#)

[Save this data in column format](#)

Data Type: eqEllipseLWR
Date&Time: 2010-03-24 20:19:12

	01	02	03	04	05	06	07	08
A								
B		2.3962	2.3872	2.288	2.3278	2.3939	2.3397	2.36
C		2.788	2.3175	2.3317	2.4196	2.303	2.4583	2.4
D		2.0502	2.3793	2.2999	2.3864	2.3934	2.3993	2.27
E		2.3982	2.2912	2.3651	2.3181	2.3163	2.4144	2.37

Fig 5: New Data Output Page: The newly modified Data output page that enables users to change the type of data contained in the tables. This page replaces the page figured in Fig 4.

2) Analysis tools login restriction:

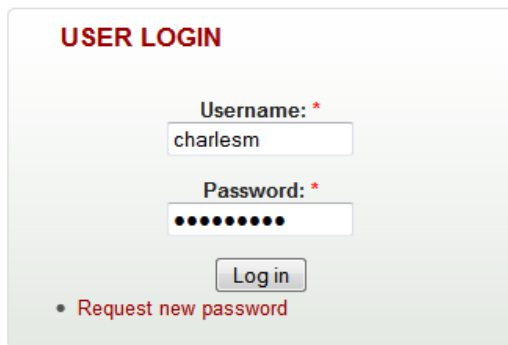
Previously, anyone that could access the ARC website, located on a remote server also had access to use any of the analysis tools without restriction. I was asked to modify that fact in order to restrict people with no authentication of using the tools.

The logs used to show an analysis done "anonymously" by someone who was either too negligent to login or simply does not have enough authentication to use the tool.

Also, with the new restrictions, it gives the website administrators more control on who is the owner of what files on the server.

I was afraid that the new restriction would cause a personal outrage from a few senior members who are accommodated the use of this method "on the go", without the hassle of logging in. However, the restriction was mostly accepted without any problems surfacing.

Fortunately, such login restriction was easy to create using the Joomla administrator tools. Joomla is, similarly to Drupal, a CMS which was used to create the ARC page, and thus contains tools and modules useful in cases like these.



USER LOGIN

Username: *
charlesm

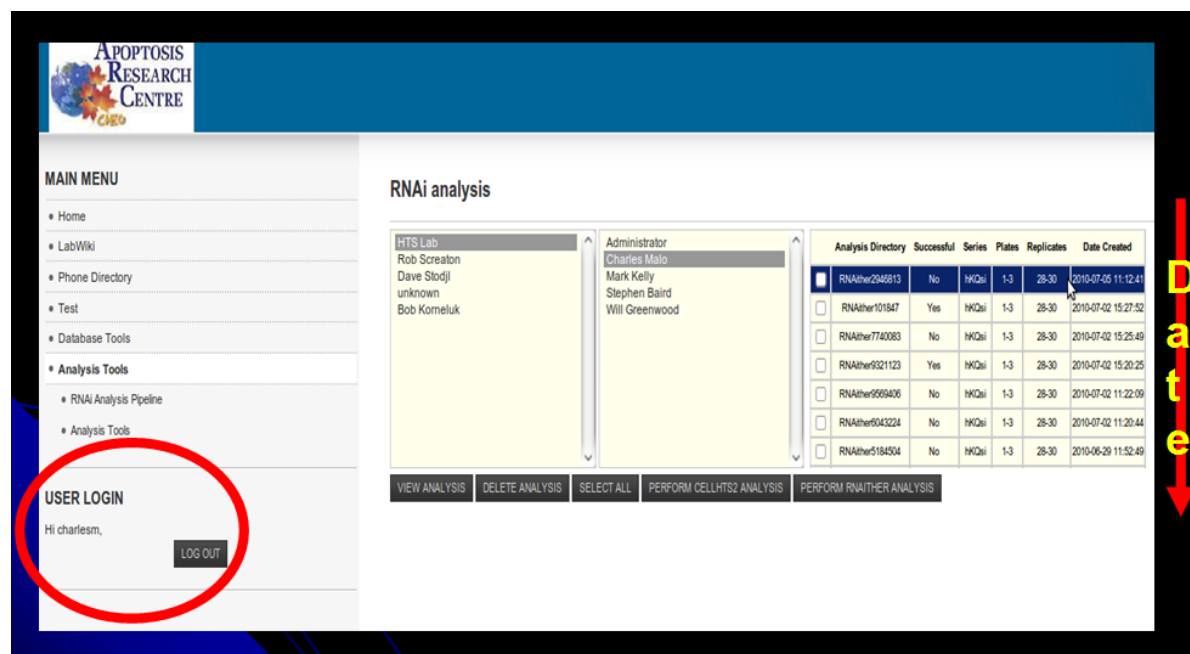
Password: *
●●●●●●●●

Log in

• [Request new password](#)

Fig 6: User login Block: A block on the main page of the website that forces users to login in order to access the tools provided on the page.

On a side note, I was also asked to find a way to sort the recently created analyses by date. This required me to learn about mySQL databases as well as SQL embedded in php code in order to modify the current scripts and databases in order to add a new date stamp field to the SQL database and use it to sort out the analyses displayed in the table by decreasing date stamp.



The screenshot shows the RNAi analysis tool interface. On the left is a 'MAIN MENU' with links to Home, LabWiki, Phone Directory, Test, Database Tools, and Analysis Tools. The 'Analysis Tools' section is expanded, showing 'RNAi Analysis Pipeline' and 'Analysis Tools'. The 'USER LOGIN' block is circled in red, showing 'Hi charlesm,' and a 'LOG OUT' button. The main content area is titled 'RNAi analysis' and contains two lists of users: 'HTS Lab' (Rob Screation, Dave Stodji, unknown, Bob Korneluk) and 'Administrator' (Charles Malo, Mark Kelly, Stephen Baird, Will Greenwood). Below these lists are buttons for 'VIEW ANALYSIS', 'DELETE ANALYSIS', 'SELECT ALL', 'PERFORM CELLHTS2 ANALYSIS', and 'PERFORM RNAITHER ANALYSIS'. To the right is a table of analyses, sorted by 'Date Created' in descending order. A red arrow labeled 'Date' points downwards next to the table.

Analysis Directory	Successful	Series	Plates	Replicates	Date Created
<input checked="" type="checkbox"/> RNAither2945913	No	HKQai	1-3	28-30	2010-07-05 11:12:41
<input type="checkbox"/> RNAither101847	Yes	HKQai	1-3	28-30	2010-07-02 15:27:52
<input type="checkbox"/> RNAither7740383	No	HKQai	1-3	28-30	2010-07-02 15:25:49
<input type="checkbox"/> RNAither9321123	Yes	HKQai	1-3	28-30	2010-07-02 15:20:25
<input type="checkbox"/> RNAither5509406	No	HKQai	1-3	28-30	2010-07-02 11:22:09
<input type="checkbox"/> RNAither043224	No	HKQai	1-3	28-30	2010-07-02 11:20:44
<input type="checkbox"/> RNAither5184504	No	HKQai	1-3	28-30	2010-06-29 11:52:49

Fig 7: Login Confirmation: A snapshot confirming the need to be logged as an authenticated user in order to use the RNAi analysis tools. Note: This figure shows the database to the right sorted by order of last modified. No earlier image of the page before the sorted table was available.

3) RNAither pipeline tool Fix:

The RNAither pipeline is, details aside, a tool provided on the ARC website. The main functionality of this tool is to analyze genome data based on some variables provided by the user, after which it recreates a raw data file of the analysis done, some curve plots that help understand certain chemical

activity on the cellular level, and finally print out a stats page containing various information about the analysis made.

However, sometimes the RNAither pipeline analysis tool used to fail and not output all the above mentioned data, stopping before creating the curve plots. After debugging the tool for weeks, and checking all its code files, which are written in php , python and R programming languages, I still couldn't locate the glitch/bug in our program that might cause such an error.

Below is a sample of an RNAither user-defined platemap that used to cause the tool to crash at some point.

Analysis Tools

Data Selection

Series
hKQsi

Barcodes
Select: All, None
☒ hKQsi001W001
☒ hKQsi002W001
☒ hKQsi003W001

Replicates
Select: All, None
☐ Replicate 024
☐ Replicate 025
☐ Replicate 026
☐ Replicate 027
☒ Replicate 028
☒ Replicate 029
☒ Replicate 030
☐ Replicate 031

Data Type
High_ROIA_T2_AveInten

Analysis Type
☐ cellHTS2 ☒ RNAither
 Analyze Platemaps

Well Type Selection

☐ Positive Control
 ☒ Negative Control
 ☐ Sample
 ☐ Empty Well
 ☐ Other

	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	A01	A02	A03	A04	A05	A06	A07	A08	A09	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A23	A24
B	B01	B02	B03	B04	B05	B06	B07	B08	B09	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19	B20	B21	B22	B23	B24
C	C01	C02	C03	C04	C05	C06	C07	C08	C09	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24
D	D01	D02	D03	D04	D05	D06	D07	D08	D09	D10	D11	D12	D13	D14	D15	D16	D17	D18	D19	D20	D21	D22	D23	D24
E	E01	E02	E03	E04	E05	E06	E07	E08	E09	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20	E21	E22	E23	E24
F	F01	F02	F03	F04	F05	F06	F07	F08	F09	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24
G	G01	G02	G03	G04	G05	G06	G07	G08	G09	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24
H	H01	H02	H03	H04	H05	H06	H07	H08	H09	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	H21	H22	H23	H24
I	I01	I02	I03	I04	I05	I06	I07	I08	I09	I10	I11	I12	I13	I14	I15	I16	I17	I18	I19	I20	I21	I22	I23	I24
J	J01	J02	J03	J04	J05	J06	J07	J08	J09	J10	J11	J12	J13	J14	J15	J16	J17	J18	J19	J20	J21	J22	J23	J24
K	K01	K02	K03	K04	K05	K06	K07	K08	K09	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20	K21	K22	K23	K24
L	L01	L02	L03	L04	L05	L06	L07	L08	L09	L10	L11	L12	L13	L14	L15	L16	L17	L18	L19	L20	L21	L22	L23	L24
M	M01	M02	M03	M04	M05	M06	M07	M08	M09	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24
N	N01	N02	N03	N04	N05	N06	N07	N08	N09	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20	N21	N22	N23	N24
O	O01	O02	O03	O04	O05	O06	O07	O08	O09	O10	O11	O12	O13	O14	O15	O16	O17	O18	O19	O20	O21	O22	O23	O24
P	P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24

Fig 8: Analysis Tool Page: Figure representing a sample Analysis tool execution on sample data (no data in the figure is confidential.)

It was particularly hard to debug the parts written in R since they need special tools which are not accessible in my working environment. Also, any debugging at all was relatively harder than expected since the files were running on a remote UNIX server, which added another variable to consider when running the program: the possibility that information exchange between the remote server and the local operating system is failing.

After weeks of research, I reached a dead-end: I could not find any glitch in the programs developed by myself or my predecessors. This led to fallout between my supervisor and I, as he assumed that I simply did not look hard enough. Yet, I decided to go further an extra step, and contacted the creator of the RNAither pipeline tool herself, and asked her for assistance since I sent sample data that used to cause the program to crash.

A few days later, I received a message from the creator in Germany, Miss Nora, confirming that my program is working properly and that the glitch lies in her own program, and it was very difficult to

locate it herself even, since the use case that I provided was so rare that it was not accounted for when developing her RNAither tool which was embedded into our own website.

The following shows a passage of a conversation between myself and Miss Nora Rieber, the creator of the RNAither tool.

“ From: **Nora Rieber** (RNAither@gmx.de)

Sent: July 8, 2010 10:45:38 AM

To: charles.malo@hotmail.com

Hi Charles,

Ok, so I checked with your dataset and if you run the commands directly in R you can actually see that R throws an error.

For some reason the density function R computes for the control distributions has a (slightly) negative tail, which is why the histogram cannot be plotted since it assumes positive values (roughly speaking).

Your dataset's NA distribution is unusual and it made me discover a couple of bugs in the package. I fixed all that and it works on your "doesn't work" dataset now. I uploaded the new version of RNAither to Bioconductor (should be 1.8.1), however it will take 1-2 days for it to show up on the homepage (will get checked by the BioC people). I can send you the corrected R code files if you need them sooner, but as far as I understood they will not help you since you are not using the R gui. Let me know.

Hope I could help!

Wishing you good luck for your studies,

Nora” (4)

After a few correspondences, it was discovered that at some point, her program tries to plot negative values for the curve in places where the values should be positive. She took my use case into consideration and modified her program and requested us to download the updates in order to fix our issues. This was left until the next opportune update to test if the solution provided works or not. Theoretically, the update should have fixed the issue since no more students mentioned the program failing after that.

D. Conclusion

During my COOP work term, I had the chance to learn various programming languages which should be very useful to my studies in the future, as well as for future employments. The learning curve was relatively very steep, however it was worth the time and effort. By the end of the COOP work term, I had assisted in countless tiny tasks, yet my pride is in the website that I built for Dr Robert Screaton's Lab, since I was requested to duplicate that website for other lab groups with minor differences.

E. Copyrights and Acknowledgements:

During the COOP summer work term from May to August 2010, and the continuous development and update of ScreatonLab website, I was assisted by Mr. Stephen Baird and supervised by Mr. Robert Screaton, to insure proper programming as well as time and quality efficiency. I do not in any way take credit for any prior work done on the analysis tools I helped modify: those tools has been developed for years by fellow colleagues and I do not take credit for any of their work. I acknowledge accessing their work to satisfy the main objectives sought, and any changes were confirmed by Mr. Stephen Baird before merging my small modifications with their work.

"I acknowledge that this report as well as its contents is purely my own work. Any modification or editing to this report must be approved by myself and myself only. Also this report requires no confidentiality release form by my supervisors nor the Research Institute since it deals solely with my own work and contains no confidential material, yet should not be made public in total or in part of any form, shape or content without consent from my behalf. Any breach of the requests above will be considered as a violation to this copyright."

Stated by Charles Malo, September 2010

F. References

(1): <http://drupal.org/about>

(2): <https://mgcheo.med.uottawa.ca/drupal/>

(4): (Unfortunately, this conversation is located on a private email account, and there are no valid sources that can confirm the above conversation other than a copy of the conversation that was archived for future references and is not accessible by members outside of the Research Institutes)

