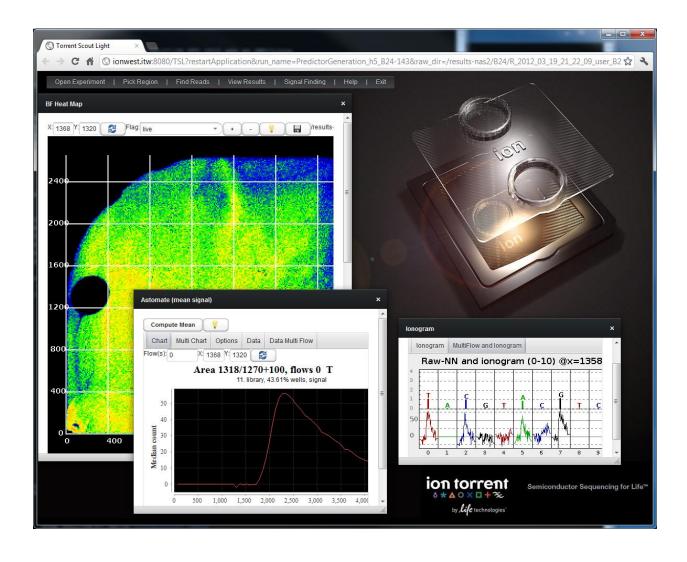
# Torrent Scout Light

V2.2.12, Chantal Roth June 8th, 2012





# **Table of Contents**

DESCRIPTION	3
STARTING	3
OPENING AN EXPERIMENT	4
SEARCHING FOR RUNS	
PICKING A REGION ON THE CHIP	
BFMASK HEAT MAP	8
RAW CHIP VIEW	10
PROTON VIEW	11
VIEWING RESULTS	13
Well Table	13
RAW DATA VIEW	15
IONOGRAM	17
VIEWING RAW SIGNALS IN IONOGRAM VIEW	17
GENOME TO READ	18
Masks	19
WORKING WITH RAW DATA	22
Process	22
FIT COMPONENT	23
AUTOMATE (COMPUTE MEAN SIGNAL)	25
FIND READS	28
Save your search result	31
CREATE A MASK FROM YOUR SEARCH RESULT	32
REGIONAL CHARTS	34
REGIONAL EMPTY TRACES	34
REGIONAL PER FRAME CHARTS	36
REGIONAL PER FLOW CHARTS	38
XYZ CHARTS	39
Computing missing data	41
PICKING DATA POINTS	43



EXPORTING DATA					
ALIGNMENT SEARCHES VIA PLUGIN	47				
VIEWING RESULTS IN THE REPORT PAGE	47				
VIEWING RESULTS INTERACTIVELY					
TORRENT SCOUT LIGHT URL	50				
HELP	51				
TROUBLESHOOTING AND FAQ	52				
I CAN'T MOVE THE SCROLL BAR	52				
I see a wait symbol and the GUI doesn't seem to respond	52				
I SEE A RED EXCLAMATION POINT	52				
THE APPLICATION STOPPED WORKING, AND THERE IS AN ERROR MESSAGE SAYING "OUTOFHEAPSPACE"	52				
HOW CAN I CHECK THE LOG FILE FOR ERRORS?	52				
How do I restart tomcat?	52				
HOW DO I DEPLOY THE APP?	52				
HOW CAN I USE THE TOMCAT WEB ADMIN INTERFACE?	52				
HOW DO LACCESS THE TOMOAT WER ADMIN INTEREACE?	52				

# **Description**

Torrent Scout Light is a pure web application written using Vaadin. The server side is written in Java and is the same code as used in the full Torrent Scout. Vaadin is a software framework that allows the developer to write 100% Java code, and then translates the gui for the web pages into Java Script, and uses the Google toolkit.

The goal of Torrent Scout Light is to provide quick access to data to get a first glance as to what the data looks like, with no installation overhead. As it is purely web based, the functionality is limited. To use the full functionality, you might want to try the full Torrent Scout client.

# **Starting**

The most convenient way is to enable the TorrentScout plugin, and to view the results directly from the Default Report page. The plugin computes all the heat maps index files, so that you don't have to wait



when you want to look at the data:



Just click on the View results in Torrent Scout Light link. This will automatically load this run and show you a heat map.

Alternatively, you can open a browser and type in the URL of your Torrent Server, followed by :8080/TSL. For instance:

http://myserver.com:8080/TSL

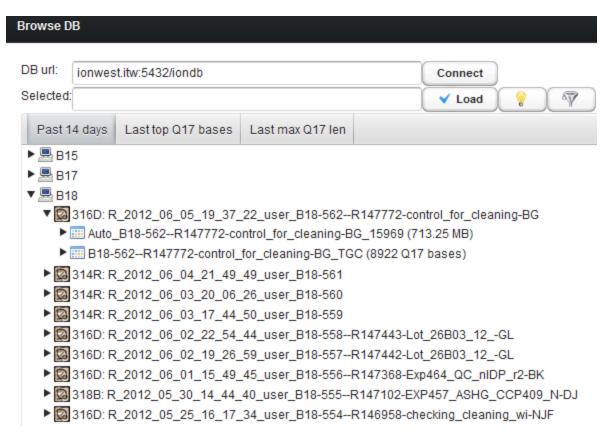
In most cases you can now also use the URL without the port:

http://myserver.com/TSL

# **Opening an Experiment**

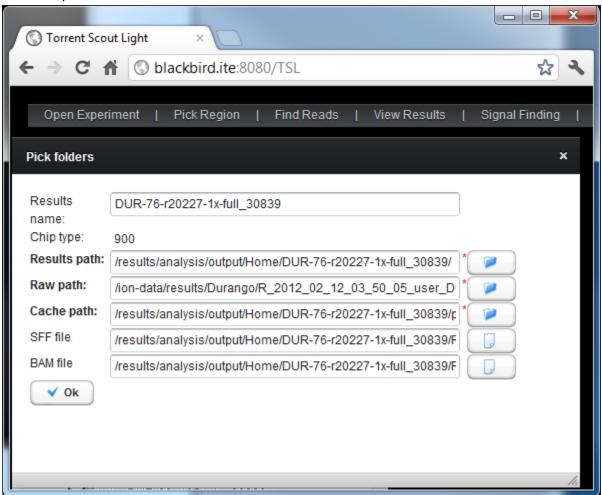
You can select an experiment by browsing in the database. Open a PGM node, then an experiment node, and select an analysis result as shown in the image.



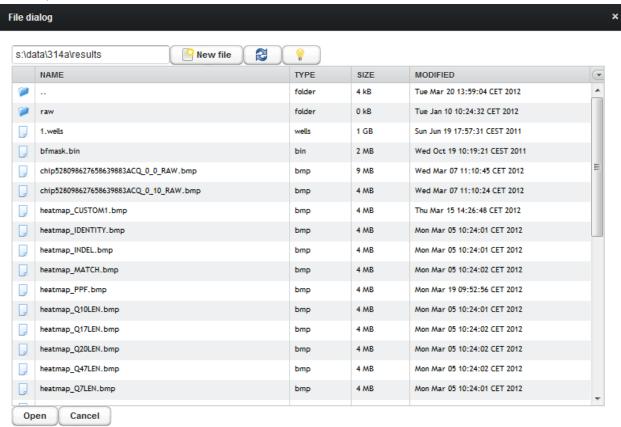


- Click on an experiment or run node to see a description popup.
- To load a run, click on a run node (such as Auto\_xyz) and click on Load.
- You can specify a different db in the DB url but keep in mind that you can only view the data if the server has access to the files. One scenario that makes sense is if you have an rnd machine where you run TSL, but that machine is on the same network as the head node has access to the exact same files using the same paths as the head node.

After you click OK (the button is on top), it will show you the paths it has found from the database and will show you the details.



If the paths are not correct, you can change it here manually or by clicking the folder/document icon, which opens a file browser:

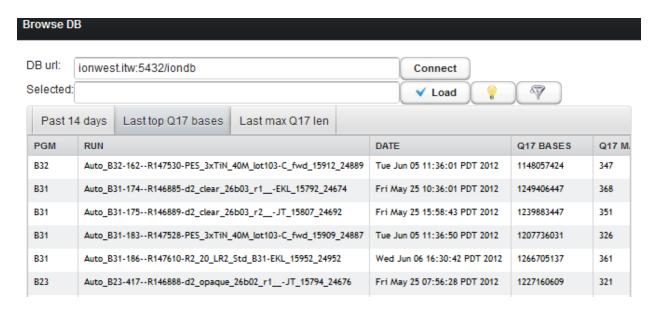


You can also open other experiments that way that are not in the database, by selecting "Open Experiment / Pick Folders".

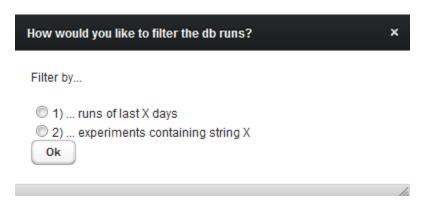
# **Searching for Runs**

The database window has two additional tabs where you can see the latest runs sorted by top Q17 bases and the longest Q17 reads:





In case you are looking for an older run, or a specific run where you know part of the name, you can also use the filter icon:

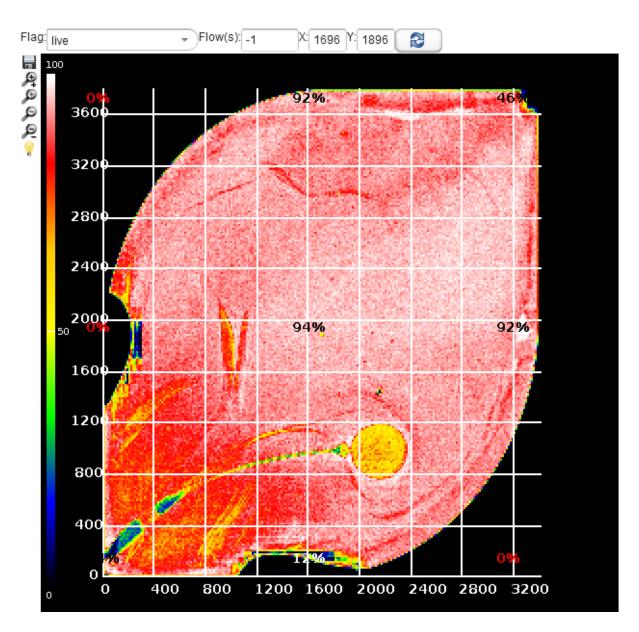


# Picking a region on the chip

### **BfMask Heat Map**

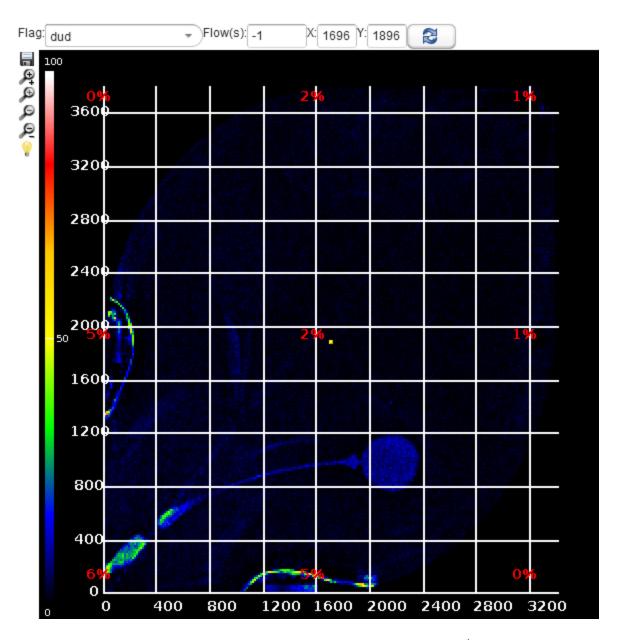
If there is a bfmkask.bin file, it will open the BfMask Heat Map view that shows you the various flags such as bead, live, dud, ambiguous etc in a heat map:





- To pick an area just click into the heat map
- You can also enter the coordinates manually
- The +- buttons (in all components) allow you to zoom in and out
- The save icon lets you save the image
- The light bulb icon shows some info on what you can do (available in most components)
- You can pick any flag from the dropdown box here the dud mask:





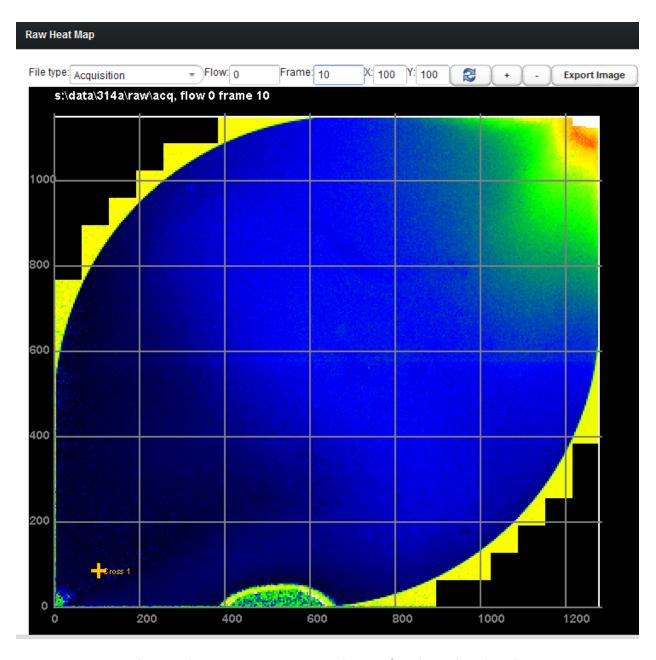
To pick a coordinate, move the orange cursor, or enter a coordinate in the X/Y text boxes.

In the flag drop down box you can chose different flags (such as bead, keypasse, dud etc) to view.

## **Raw Chip View**

If there is no bfmask.bin file, or if you select "Pick region/ Raw heat map" it will open the WholeChip view that displays one frame of a raw .dat file:





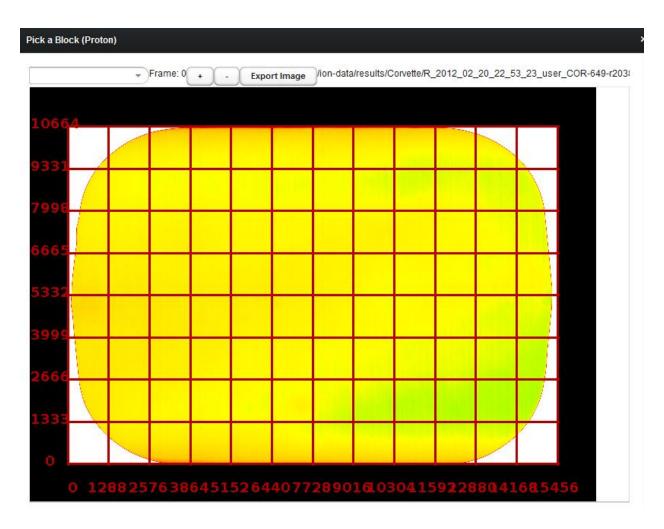
You can change the file type (acquisition, prerun bead find pre/post), the flow (any flow number starting from 0) and any frame (default is frame 10).

To pick a coordinate, move the orange cursor or enter the coordinates as in the bf mask view, and you can also zoom in and out with the +/- buttons.

#### **Proton View**

For proton chips, where the result is composed of multiple sub-blocks, you can pick a block for viewing with the "Pick Region/ Pick a block" menu:

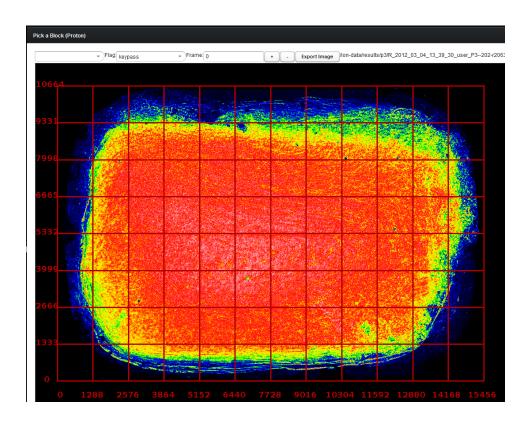




The Proton view also lets you pick other heat maps as in the BF heat map component.

Note that if the heatmap has not been computed by the plugin, it will have to do it on the fly, which will take several minutes and use a lot of memory/CPU on the server...





### **Viewing results**

Once you have chosen an experiment and a location (and a block in Proton experiments), you can now view results.

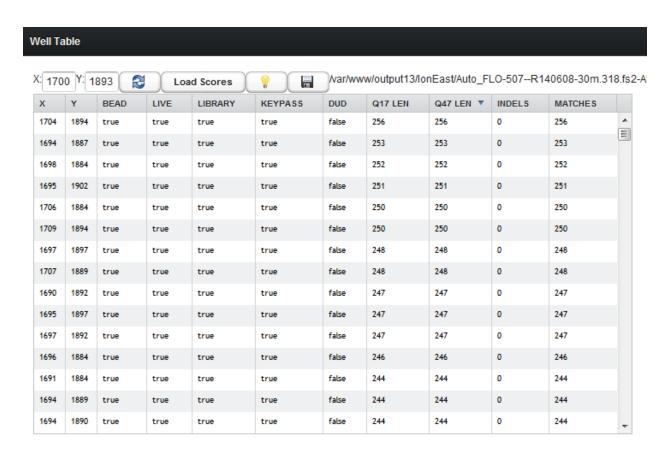
#### **Well Table**

The well table shows a few thousand wells surrounding the currently selected well. It shows a few flags and scores such as Q17 lengths, indels and so on. The scores might not always load if the index files have not been created yet. To force them to load, click on Load Scores (it might take a few minutes if there are no index files yet).

You can sort the table by clicking on the table header.

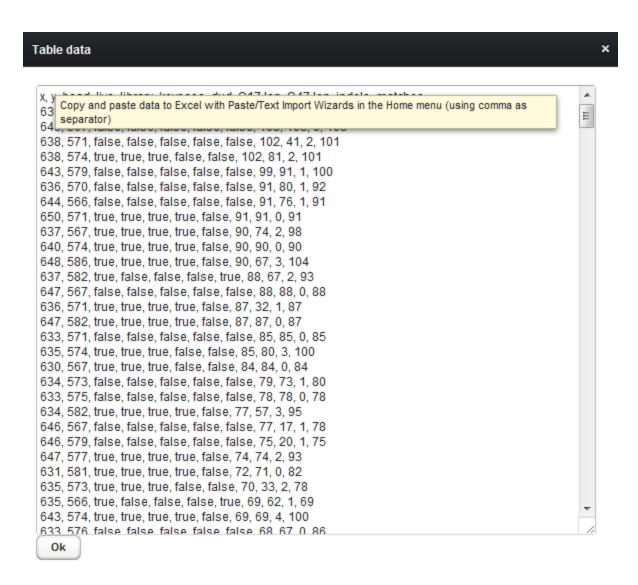
You can select a well by either clicking on a row, or by entering the coordinates on top. To view the results, select any of the viewers from the menu "View Results" (such as lonogram, Alignment, Raw View etc)





The export button lets you export the table data (from all tables):

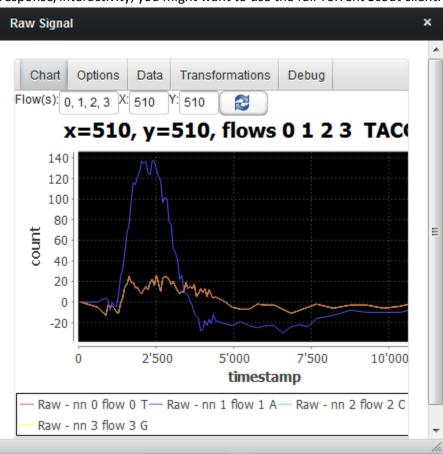




#### **Raw Data View**

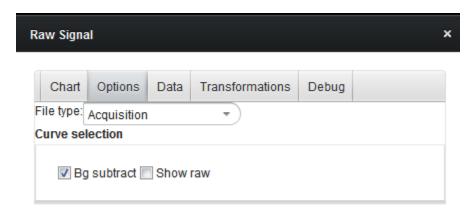
You can enter one of more flow numbers (separated by comma) into the flows text field. You can also enter a coordinate right there. Note that sometimes it takes time to load the raw data, in particular as this all goes to the torrent server first (which might be busy). For better performance and better





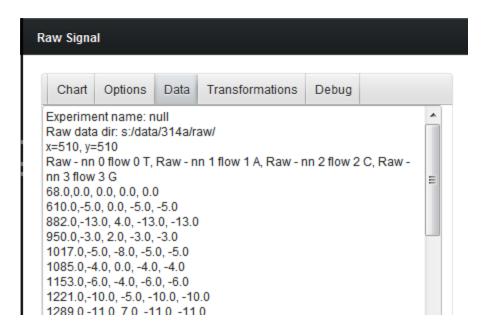
response/interactivity, you might want to use the full Torrent Scout client.

In the options tab you can pick to view the raw data and/or the bg subtracted data and the file type. The bg subtraction simply subtracts the average signal of the surrounding wells that are considered empty.



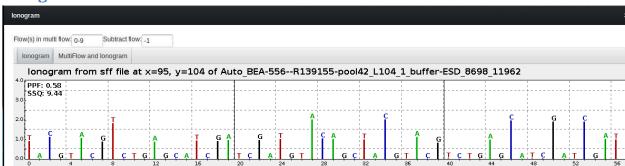
The data tab shows the currently loaded data:





You can copy and paste this to Excel (Use Paste/Text import wizard, comma as separator) to create your own charts.

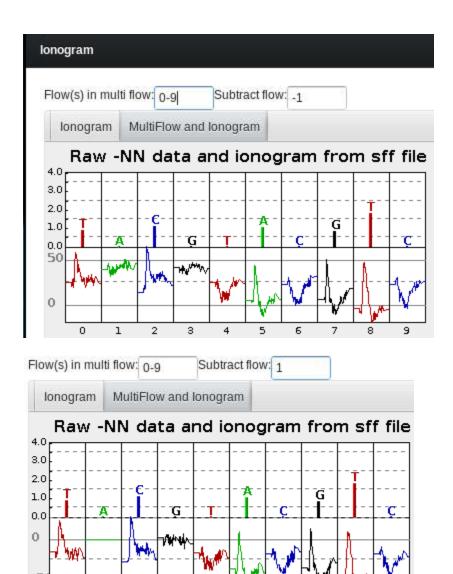
#### **Ionogram**



#### Viewing raw signals in ionogram view

The second tab shows any number of nearest neighbor subtracted raw signals for this well: just enter the flow numbers (starting at 0) into the text field and hit enter:



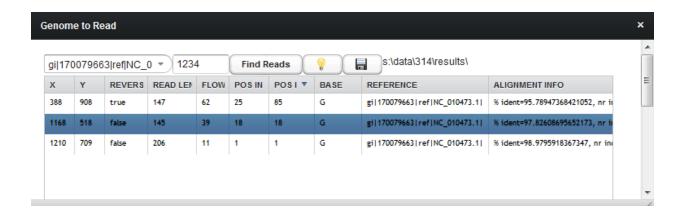


#### **Genome To Read**

In the component Find Reads/ Find reads by genome position you can locate any read that maps to a certain genome position.

Pick the reference sequence in the drop down box and enter a genome location in the text box. This will return a list of all reads (including the flow number) that map to that position. To view the ionogram or alignment of any read you can just open that component and select the read in the table.





#### **Masks**

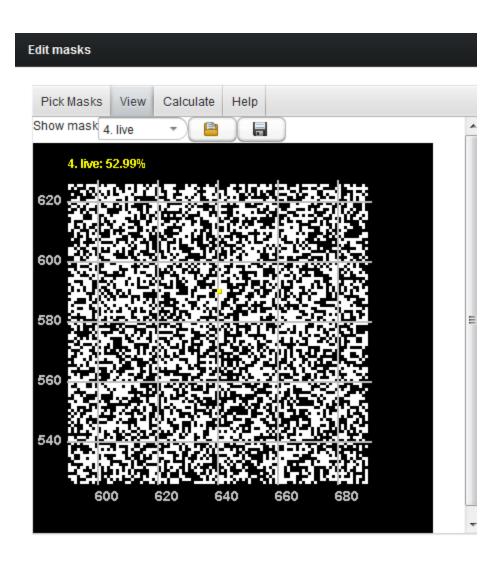
The Signal Finding/ Edit Masks component lets you view and do calculations with masks of a specific area on the chip.

The default is a 100x100 area, but you can zoom in and out of this with the Process component (upon request I can of course make this more flexible :-).

The View tab lets you view the masks and what % is flagged:

In the View tab you can save and load masks. It stores the coordinate in addition to all masks, so next time you load a masks file, it will automatically move the view to the coordinate of the mask. Note that you have to be in the correct experiment or block. So for Proton experiment currently you have to load the block for which you created the mask.



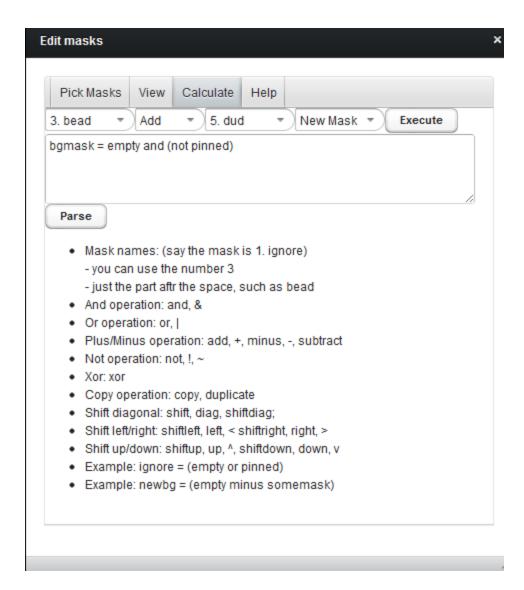


The calculate tab has 2 calculators:

- a) A simple drop down calculator
- b) A parser

The drop down calculator is really simple: pick mask A as first operand, an operation such as add, subtract, xor etc, and if it makes sense, a second operator mask B, and finally either an existing mask as result (Which will be overwritten) or else a new mask.



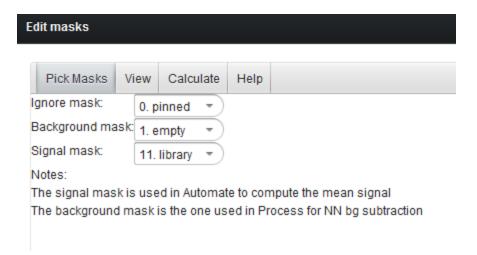


After the computation, it will print the command line string into the text box.

The parser allows you to copy/paste frequent operations directly into the text box, so you don't have to pick the masks in the drop down box :-)

The Pick masks tab is important for the Process and Automate component Shown later )





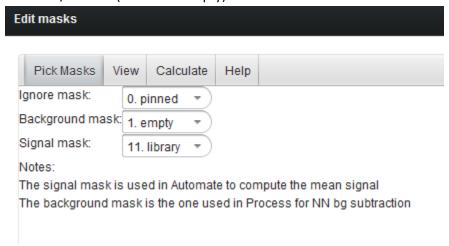
### Working with Raw data

#### **Process**

Basically it lets you view an area of say 100x100 wells, pick a few wells in there at the same time (the cursors), and view NN subtracted signal. This allows you to compare the signal of different wells, maybe look at problems, and to see where approximately the signal appears.

You can snap the cursors to any of the masks (also your own!).

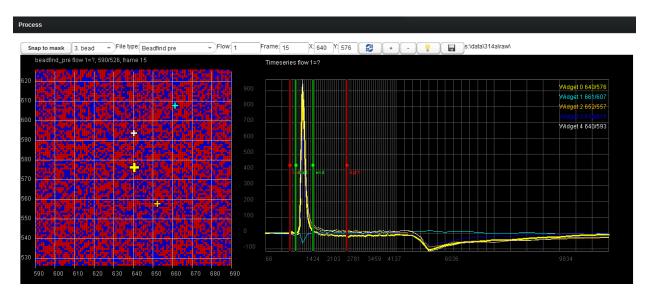
It automatically computes the NN bg subtraction, so make sure you have a reasonable mask selected in the Mask/Pick tab (default tis empty):



The zoom in magnifies the area, and reduces its size be a factor of 2 each time. **Note this currently reloads the data and masks**. So you have created your own masks, they would be lost (will think of a save features soon, but in the meantime, you might want to use the **full Torrent Scout client** to do more professional mask editing as it allows you to save one or all masks!)



You can pick the type of data and flow – here the beadfine pre flow 1 is shown with a very clear bead find signal:



Note the vertical red and green lines. They can be moved around with the mouse.

Those positions are used in **the Fit component** below:

#### **Fit Component**

The fit component is a bit difficult to understand at first I think: basically it helps you to classify wells. That way you can for instance create a mask such as "all live wells with a very high signal", or maybe "duds that don't look like duds", or "bead find with a low signal" etc.

The histogram is calculated using all wells from the selected histogram mask (or all wells if none is selected). For all those wells, it computes the values using the green (and red) frames.

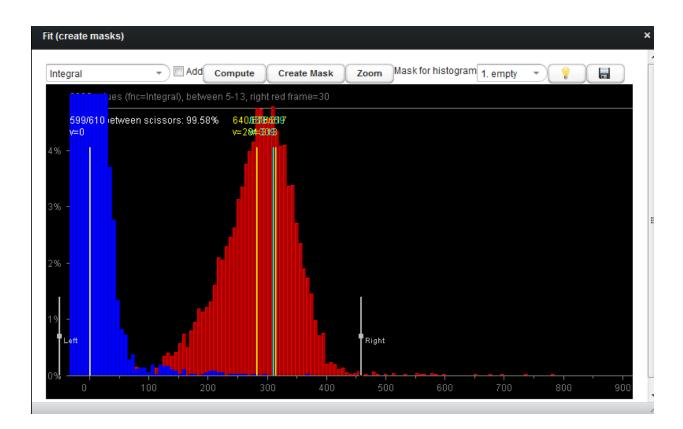
The integral computes the area under the curve between the green frames.

The max-end height computes the max between the green frames, and subtracts the value at the right red frame.

The add check box let's you add a second histogram the next time you compute it. So you might want to compare the data of the empty wells (blue) versus the bead wells (red) as in the example below.

The zoom button zooms into the area where the scissors are (it only shows the last data set currently...)

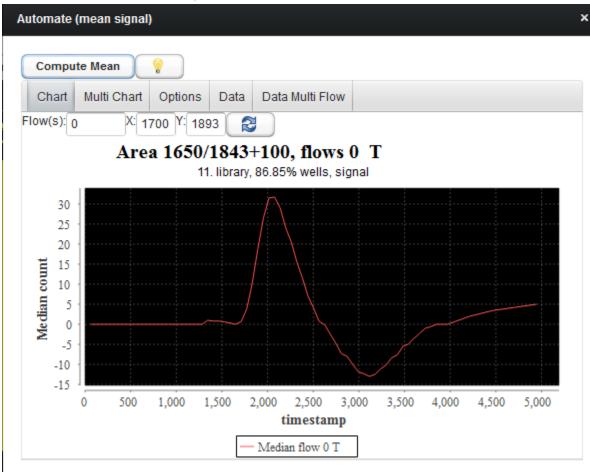






### **Automate (Compute Mean Signal)**

Maybe you want to compare the signal shape of different areas the chip, maybe using different masks. This is where the automate component comes in:



It takes all wells from the Process component (such as a 100x100 area, but it can be larger/smaller depending on the zoom level), and computes the NN subtraction and mean signal for the entire area. Make sure you pick the signal mask you want in the Mask/Pick tab:

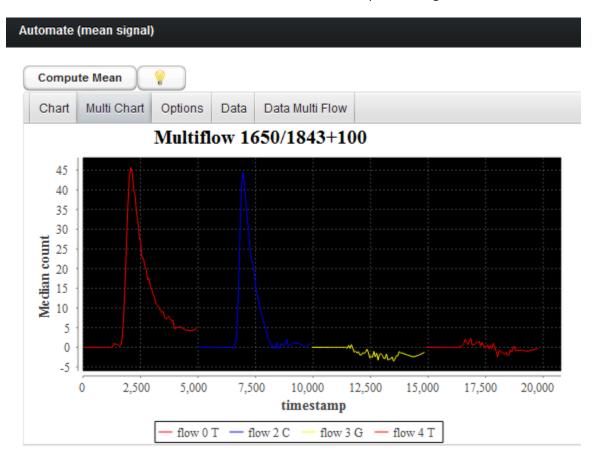


E	dit masks								
	Pick Masks	View		Calcu	ılate	Help			
	Ignore mask:		0. p	inned	<b>v</b> )				
	Background mas	sk:	1. empty 🔻						
Signal mask:			11. library ▼						
	Notes:								
	The signal mask is used in Automate to compute the mean signal								
	The background mask is the one used in Process for NN bg subtraction								

You can enter multiple flows in the text box (such as 1-4 or 1,2,3,4 or any other flow). Select the file type for beadfind in the Options tab (there you can also decide to subtract the result from another (empty) flow).

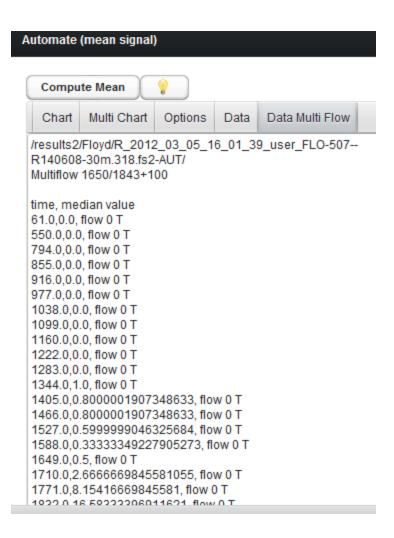


The Multi Chart tab shows the same data, but the flows are spread along the x axis:



You can export the data to excel with the Data tabs. In Excel, use Paste/Text Import Wizard/comma as separator.

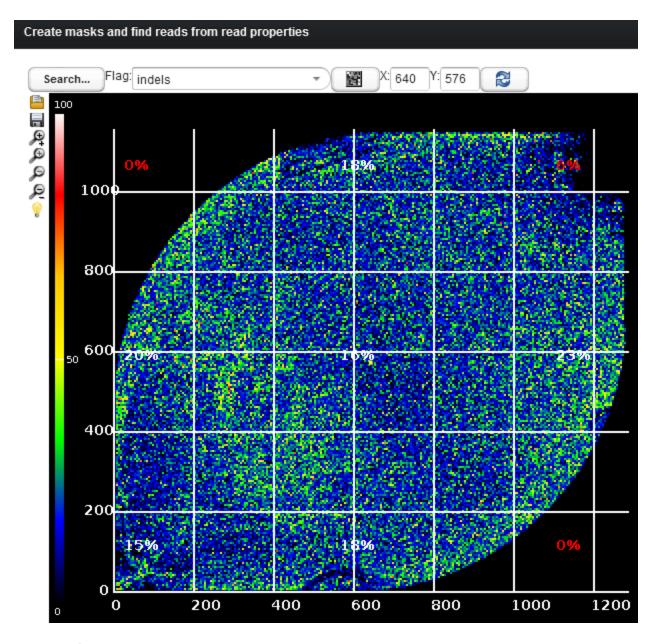




### **Find reads**

The find Reads component shows heat maps of scores such as Q17 length, and lets you filter the reads:

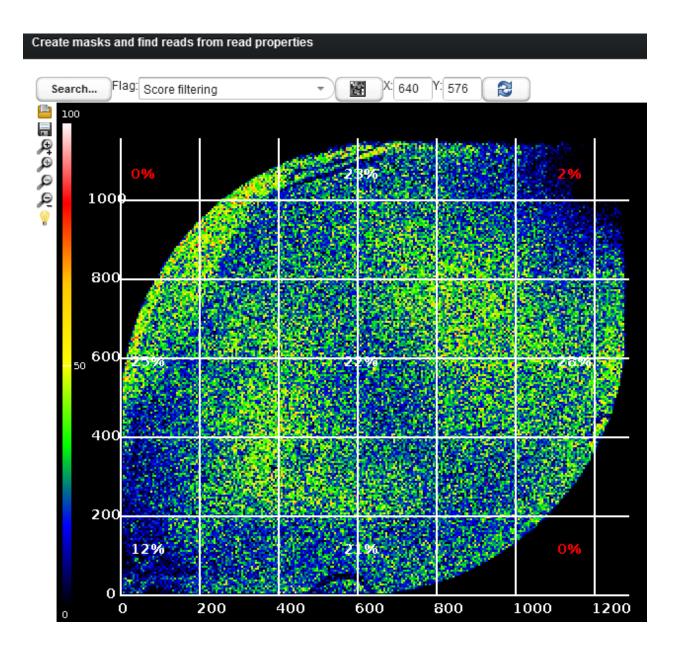




Select the filter button to only show wells that have certain values:

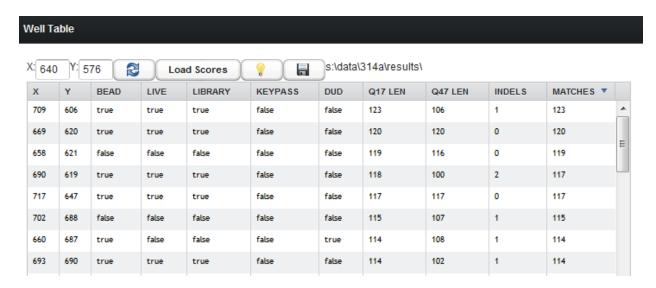






If you now select a well by moving the orange cursor, and open the well table, it will only select wells that fit the search criteria. In this case, only wells with a Q47 length of at least 100:

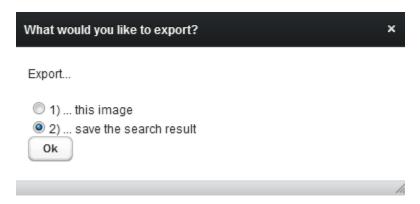




If there are fewer than 5000 wells in the entire chip that fit the result, it will load them all into the table.

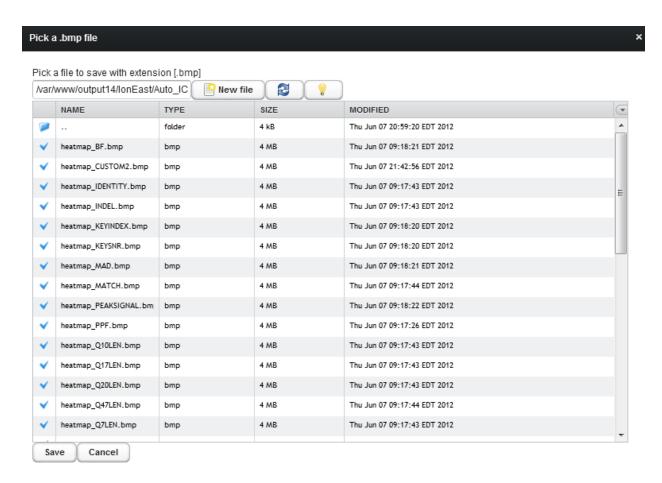
#### Save your search result

You can save your search so that you can go back to it later – just click the disc icon:



And pick/create a .bmp file:





### Create a mask from your search result

You can convert your search result into a mask just like any other mask (like live or empty), which you can then use in any other component.

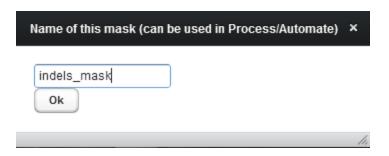
For instance you can use it in the Process component to see the raw data of those wells.

Or you can use it in the Automate component to compute the mean signal of all wells in that mask.

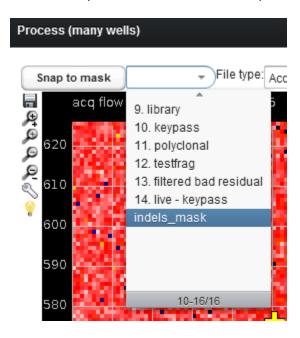
Or you can combine this mask with other Mask Editor.





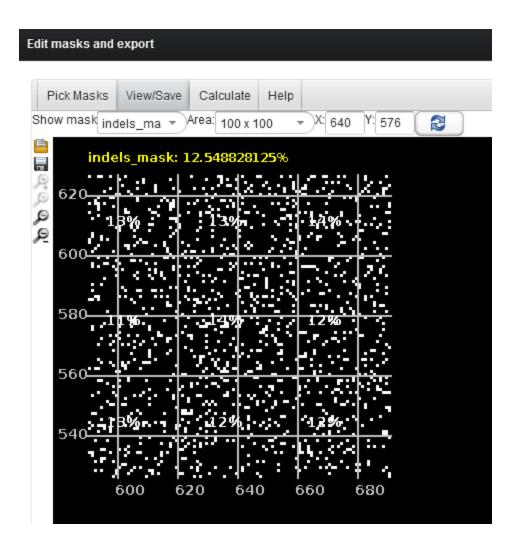


For instance you can now see it in the drop down of the Process component:



You can also work with the new mask in the mask editor:



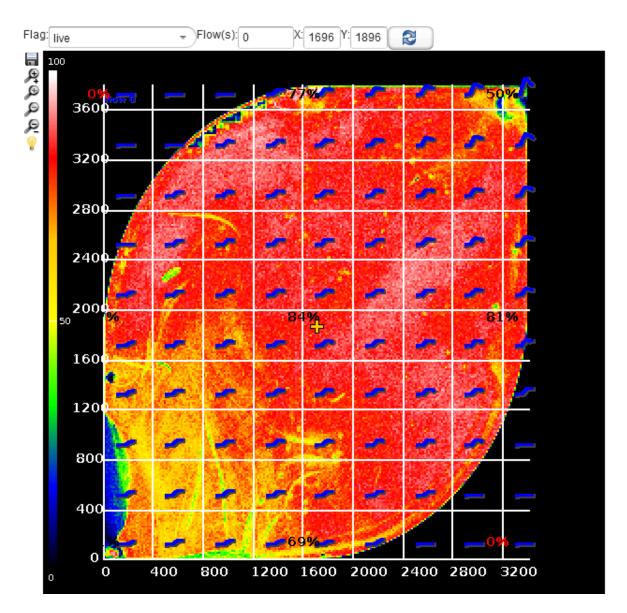


# **Regional Charts**

### **Regional Empty Traces**

After you pick an experiment the bf heat map opens. It also shows the regional empty traces for the flow entered on top (multiple flows can be entered).

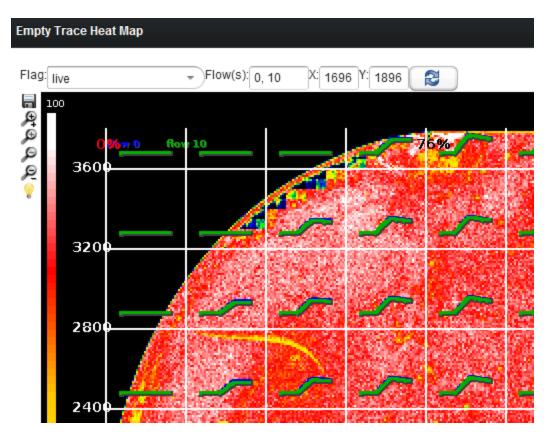




Zoom in to see more details. If you enter more than one flow, it will draw them on top of each other – the idea is to just get a general sense of the shape of the empty trace and relative sizes to each other.

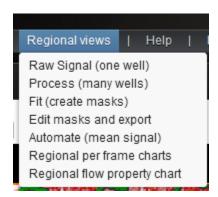
Note: to hide the empty traces, just enter -1 for flow!





#### **Regional per Frame Charts**

To view any particular region in more detail, first double click on an area, and then pick one of the regional views :



Pick the "Regional per frame charts" to see the empty traces (and other data) for a particular region:

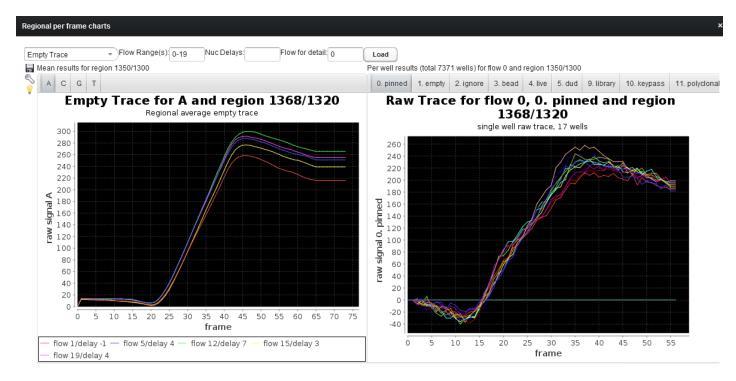
On the top left drop down, you can pick the type of data to view (currently there are is just empty traces, but in the future there might be other types).

The chart on the left is partitioned by base (ACGT). You can pick one or more flow ranges, such as:

0-19,20-30, 34, 50-60



If you wish to further limit what is shown in the charts, you can also enter one or more nuc waits (such as 1, 7)

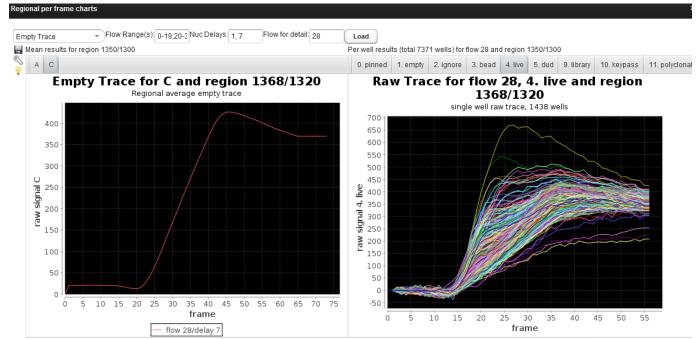


On the right side, it will show the actual raw data from the .dat file for any entered flow (flow for detail) (I will add the ability to pick the flow by clicking on a trace on the left shortly).

The right side is partitioned by bf flag. So in the pinned tab, it shows the raw for all wells that are flagged as pinned and so on.

The disc icon on the very left border allows you to save all charts and all data (in Excel format).





### Example for looking at flow 28 with nuc wait 7

The second new regional view is the "Regional flow property chart", which lets you see properties by flow:

In the left drop down you can pick the type of data to view. Currently it only contains step size, but in the future there might be additional types such as base call error and others.

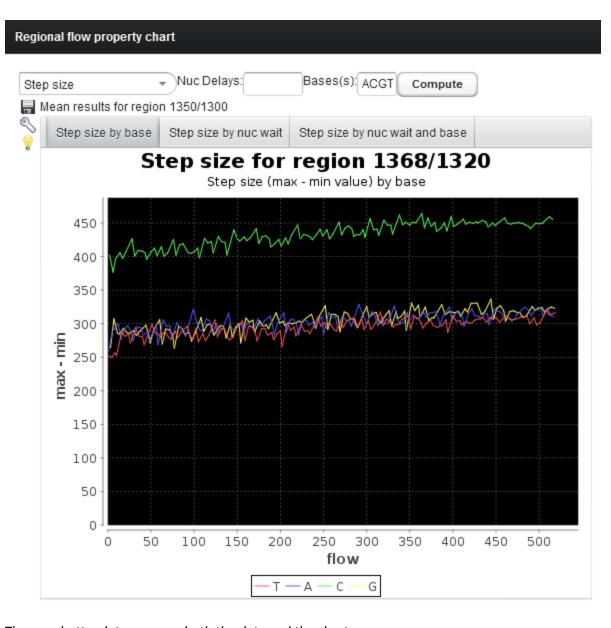
## **Regional per Flow Charts**

Again you can limit the data to see by nuc delay. The same data is shown in 3 charts, which are partitioned in different ways:

- a) Partitioning by base (ACGT)
- b) Partitioning by nuc wait
- c) Partitioning by both nuc wait and base

If you only wish to see certain bases (say AC), you can enter them in the Bases text field.



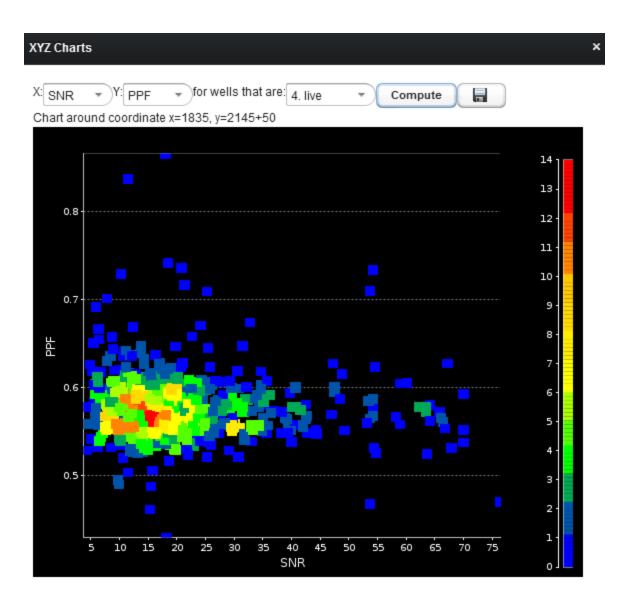


The save button lets you save both the data and the charts

# **XYZ Charts**

To see relationships between different read properties – say to see what kind of effect the signal to noise ratio has on the Q17 length, we can use the XYZ charts:

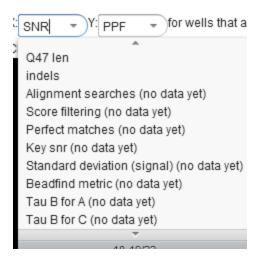




You can pick two properties (here SNR and PPF) and pick a mask. The colors indicate how many reads/wells it found for that particular combination of values. So in the chart above, there were lots of reads with SNR around 15 and PPF around 0.57

Depending on how/if the TS plugin was executed, there may be properties where it says "no data yet". This simply means that if you chose to use it, it first has to compute that particular heat map, which may take up to 10 minutes (depending on the chip size).

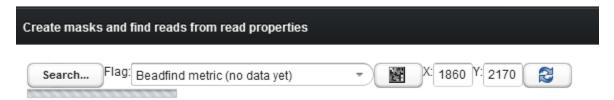




A lot of the listed properties are from the Separator file (beadfind metric, Tau B values and others). So in case that file is missing, you won't be able to use those properties.

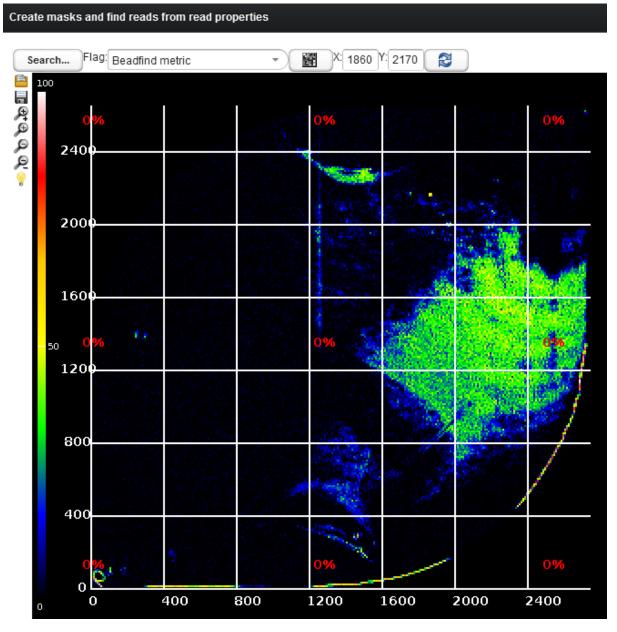
## **Computing missing data**

The Score heat map lets you pick those heat maps and it will show you in a progress bar how far along it is in computing any missing data – in the example below the beadfind metric.



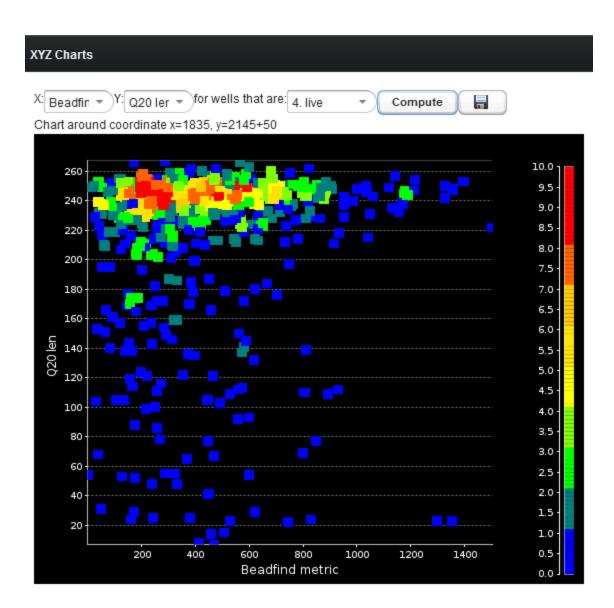


After the map has been computed, it can be used in the XYZ charts.



This chart below shows the relationship between beadfind metric and Q20 length (which is extracted from the .BAM file).

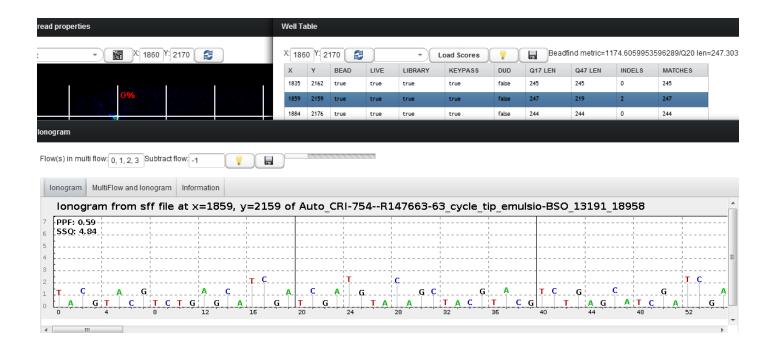




### **Picking data points**

You can **click into the chart**, which will select all wells/read that have this particular set of properties. It will list those reads in the table – from which you can do anything you want, look at ionograms, alignments, look at raw signal etc (in this example I clicked on the top right green spots):

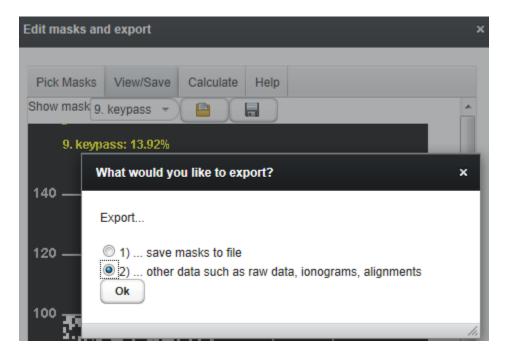




# **Exporting Data**

In both the Process and Mask Edit window you can export raw data, alignments and ionogram. Simply pick the save icon:

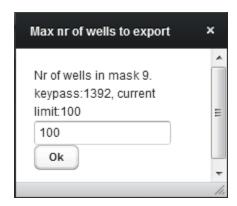
You can save the masks or other data:







You can limit the nr of data points to export not only by mask, but you can also give it an absolute limit



A progress bar indicates how far along the export is



You can pick if you want .csv or .html. If you chose to export raw data, you can pick if you prefer to export the raw data as it is, or NN subtracted.

Example for alignment in .html format:



## Data of area starting at x=50, y=50, size 100 "Only using data for mask 9. keypass, with 1392=13.92% wells"

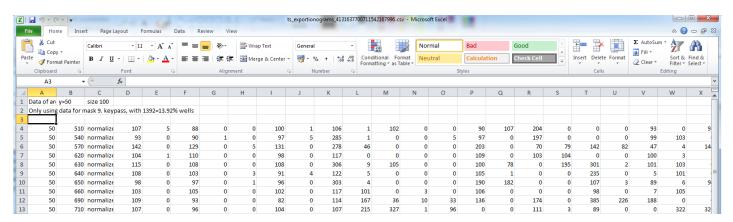
#### Alignment at 500, 510

Flags: 16 (reverse: true)
Genome position: 489343-489418
Alignment in sequencing order:

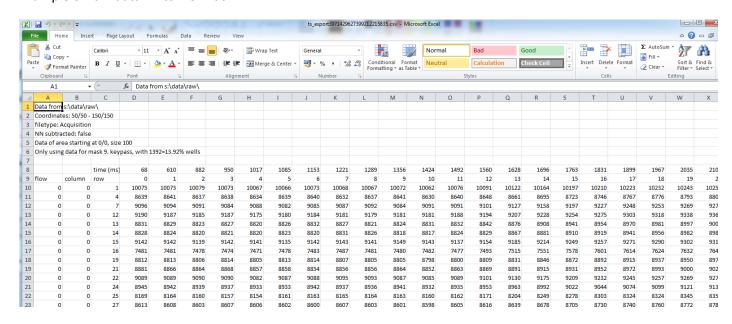
10 20 30 40 50 60 seq: CAGCCGTAAGCCGCTTCCAGTGCCTTTTTTTTTTTGGCTGATAAACCGGACCAATAAAT ref: CAGCCGTAAGCCGCTTCCAGTGCCTCTTTTTTTTTTTGTGCTGATAAACCGGACCAATAAAT

Identity: 75 = 98.7% Gaps in seq: 1 = 1.3%

### Example for ionogram in .csv format:



### Example of raw data in .csv format:





# Alignment searches via Plugin

Searching for alignment patterns can take 10 or more minutes for large chips. A more convenient way to so this is to run the Torrent Scout plugin:

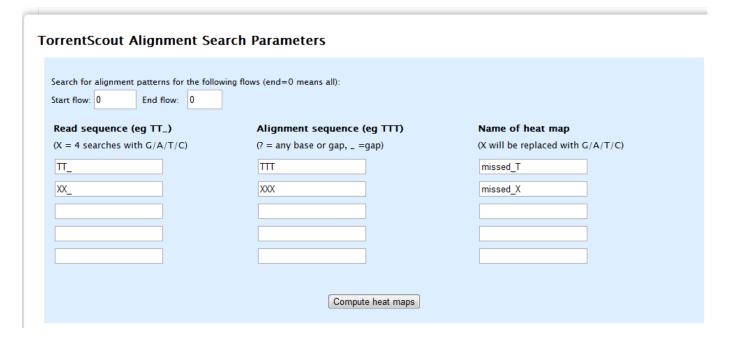
You can enter up to 5 alignment searches and specify the range of flows in which to search the alignment pattern.

Any X will be replaced with 4 different searches, with another base each time. So if you search for

TXTTX, it will search for TATTA, TCTTC, TGTTC and TTTTT.

A question mark? means that any base is accepted. So a TT?T results in just one search, where? can be any base.

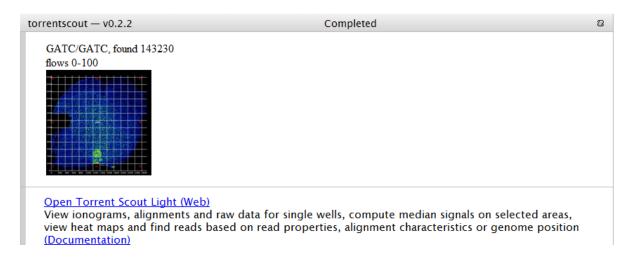
If you were to search for TXTT?, it would run 4 searchesL TATT?, TCTT?, TGTT? and TTTT?



## Viewing results in the report page

Once the alignment search is complete, it is displayed on the results page:





# Viewing results interactively

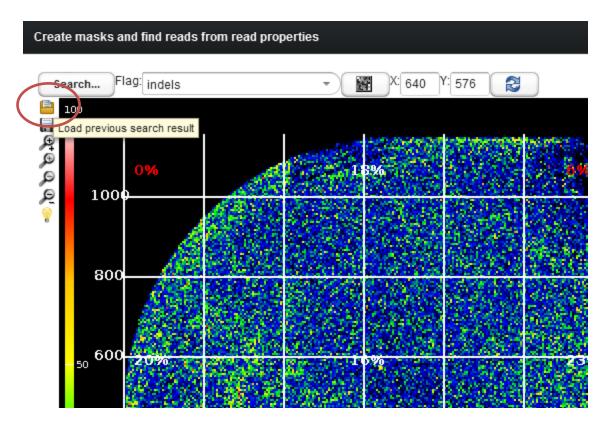
But you can also interactively work with the results by opening it in Torrent Scout with the link from the plugin results page.

Once started, go to the Find/Find reads and create masks window:



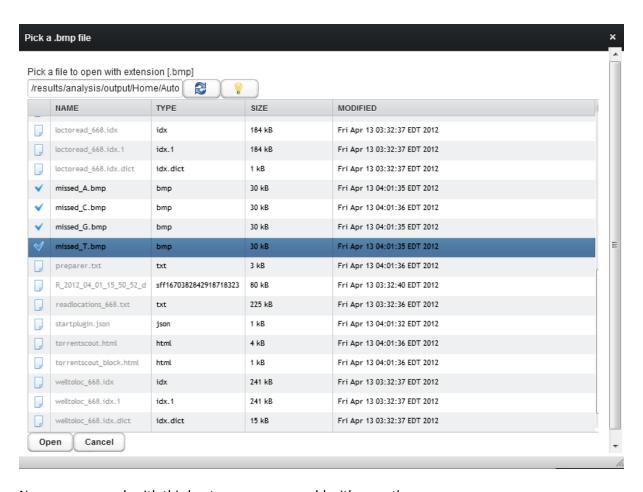
and click the open button:





Your search results will have names as specified on the plugin page, such as missed\_T:





Now you can work with this heat map as you would with any other.

For instance if you click anywhere on the heat map, it will show the wells with this flag set in the table (all of them if < 5000, or else in the area where you clicked), from which you can then view the ionograms, raw data, alignments etc.

# **Torrent Scout Light URL**

The url for Torrent Scout is now simplified without the 8080 port:

http://blackbird.itw/TSL

http://blackbird.ite/TSL

http://blackbird.bev/TSL

http://10.33.106.11/TSL

http://rnd1.ite/TSL

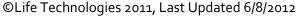
http://ioneast.ite/TSL

http://ionwest.itw/TSL

http://rnd3.itw/TSL

http://rndbev.apg.per.na.ab.applera.net/TSL

http://aruba.apg.per.na.ab.applera.net/TSL

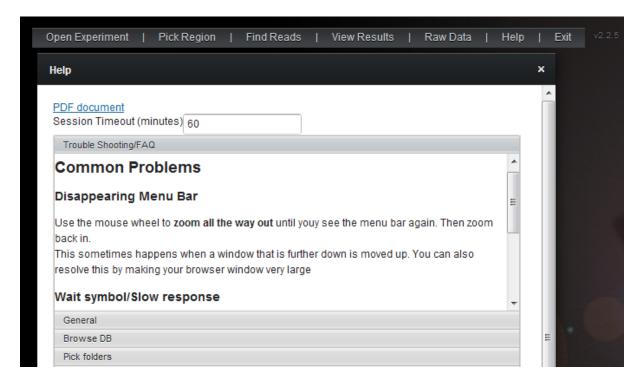




# Help

The help menu item shows a short description for each component (same as the light bulb icon in each component).

The help pages also contains info about common problems (Please let me know so I can add them!), and you can also extend your session time (current limit is 60 minutes):



# **Troubleshooting and FAQ**

#### I can't move the scroll bar

To move a scroll bar of any windows, first activate it by clicking on the title bar

### I see a wait symbol and the GUI doesn't seem to respond

If your server is far away (network wise) then it might take a while to process clicks from the client. It is better to wait until the server response is back, otherwise it might make it worse :-). It could also indicate an error (is there a red exclamation point? If so move your mouse over it and you will see a stack trace)

### I see a red exclamation point

This indicates an error. If you move your mouse over it, you will see the exception. If it is a memory error, see below on how to fix it. Otherwise, if you don't know what the issue is or it seems like a bug, let me know ©.

### The application stopped working, and there is an error message saying "OutOfHeapSpace"

Since the application runs on the server, it can only handle 1-2 users at the same time as all the data is loaded in the **server** memory. You could increase the server memory (if your server has enough) by changing the setting in **/etc/default/tomcat6** and restarting tomcat. It is the line with:

JAVA\_OPTS="-Djava.awt.headless=true -Xms2048m -Xmx6096m -XX:MaxPermSize=512M"

### How can I check the log file for errors?

sudo tail -f /var/logs/tomcat6/catalina.out

### How do I restart tomcat?

sudo /etc/init.d/tomcat6 restart

### How do I deploy the app?

Copy the TSL.war file to /var/lib/tomcat6/webapps (you might have to restart with sudo /etc/init.d/tomcat6 restart)

#### How can Luse the Tomcat web admin interface?

Edit the file /etc/tomcat6/tomcat-users.xml:

```
<tomcat-users>
    <user username="tomcat" password="tomcat" roles="admin, manager, manager-gui"/>
</tomcat-users>
```

#### How do I access the Tomcat web admin interface?

http://yourserver.com:8080/manager/html (enter username/pw as specified in the file tomcatusers.xml)

