IGV IonTorrent Features 3.8

Table of Contents

Ion Torrent Menu Items

In the alignment track, right click on a base to view the available Ion Torrent functionality. These include:

- View flow confidence distribution
- View ionogram alignment
- Shade bases by model-data confidence

Copy the confidence distribution for this base to the clipboard

View flow confidence distrubtion for

View ionogram alignment

Group alignments by

Sort alignments by

Color alignments by

Shade bases by...

Ion Torrent Tool Tip Texts

When hovering over a base in the alignment track, you can see details about the called base (depending on what data is available and what has been loaded):

The raw flow value is the measured flow value without any adjustments or corrections.

The predicted flow value is what the raw measurement value would be in theory if everything was perfect based on the predictive model.

The model deviation is the difference between measurement and prediction, and the confidence is how confident we are that the call is correct (compared to the next/previous value of the model prediction).

Sample = E94135-047a02-12-L5841
Read group = N1UOG

Read name = N1UOG:01740:01046
Alignment start = 6529090 (+)
Cigar = 90M
Mapped = yes
Mapping quality = 93

Base = T
Base phred quality = 25
ZM = -1C,[67T],3A,17C
Flow position = 32
Base called = **T**
Raw flow value = **67**
Predicted value = 95
Model deivation = -28
Conf = 68
Model deviation = -4C,[-28T],-4A,5C

Shading by confidence

This option will shade the bases by confidence. The lower the model-data confidence, the more clearly the base is visible, and the more likely it is that the HP length of this base is wrong. In the example below, there is a HP stretch of 4Gs with a low confidence.
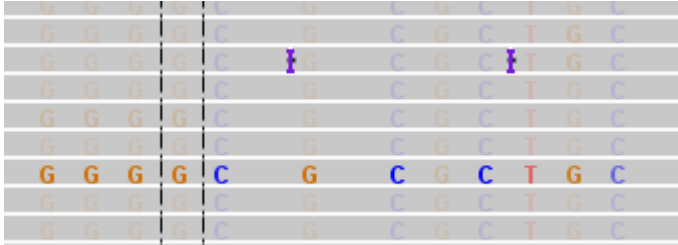
Model-Data Confidence Distribution

For a given genome location, you can view the statistics of the model-data confidence for all the HPs at that position for the different reads and the different bases (if there is more than one possible base call at this position):
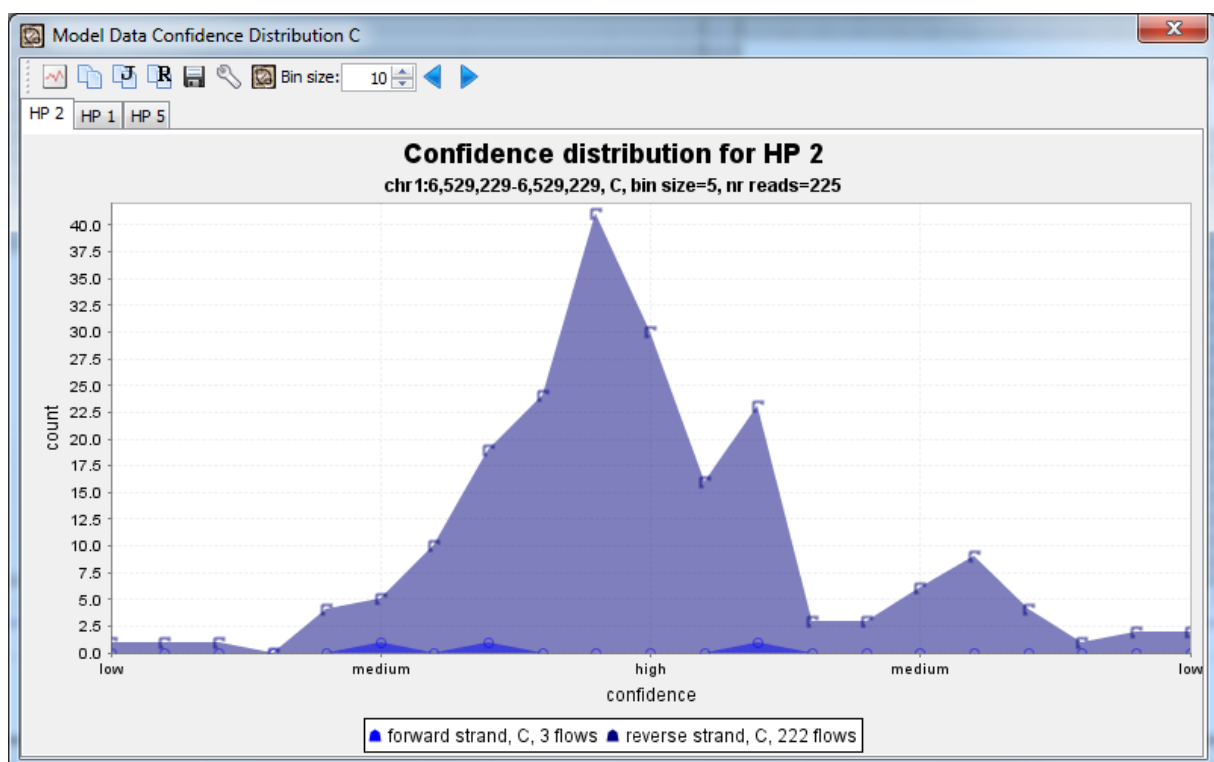

It shows the most likely homo polymer length first (here a C with HP length 2), and shows the confidence for all HP's of that same size. Depending on what option was picked, the forward and reverse distribution may be displayed as a separate data set (as is the case here).

A high confidence means it is likely that the HP length is correct. A low confidence means the HP length may not be correct. If the confidence is low on the left side, it indicates that the HP length might be shorter, and if it is on the right side, it means that the HP length might be longer. The

confidence is determined based on how far away the next or previous predicted value would be based on the model.

To see if it might be possible that the HP length is actually 1 and not 2, we can select the second tab, which is the second most likely HP length:
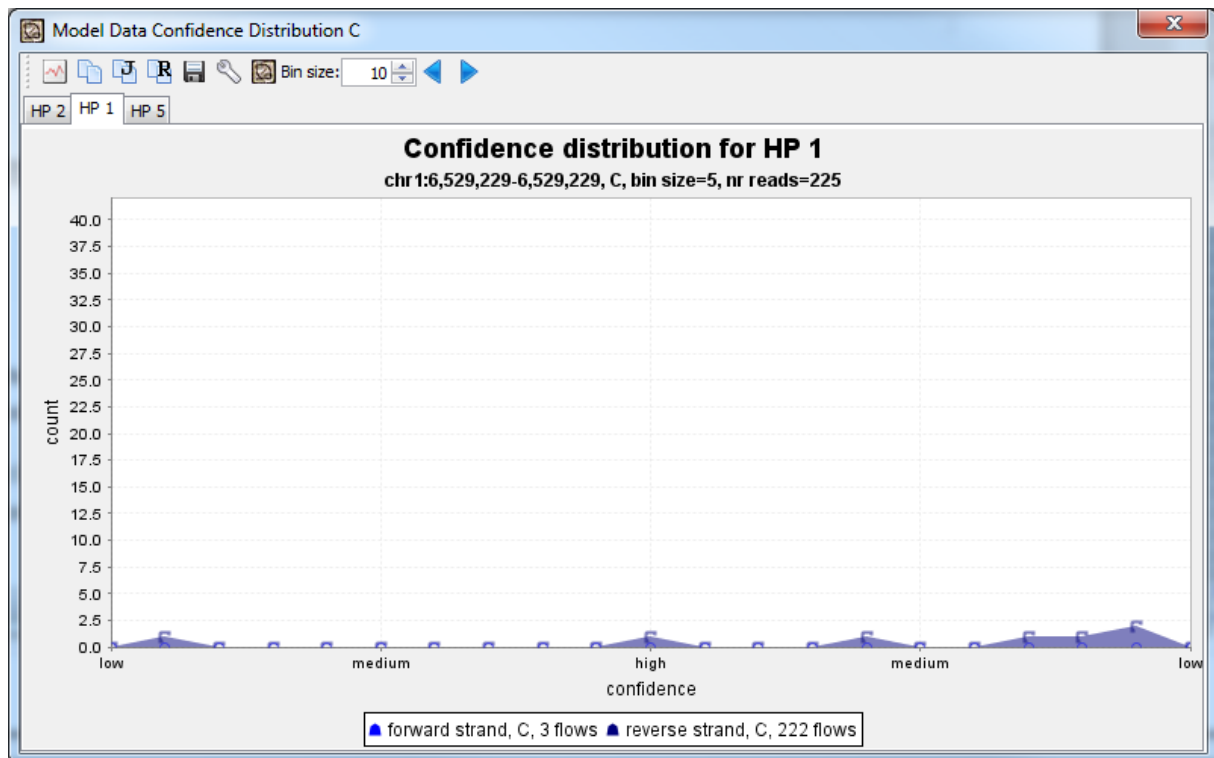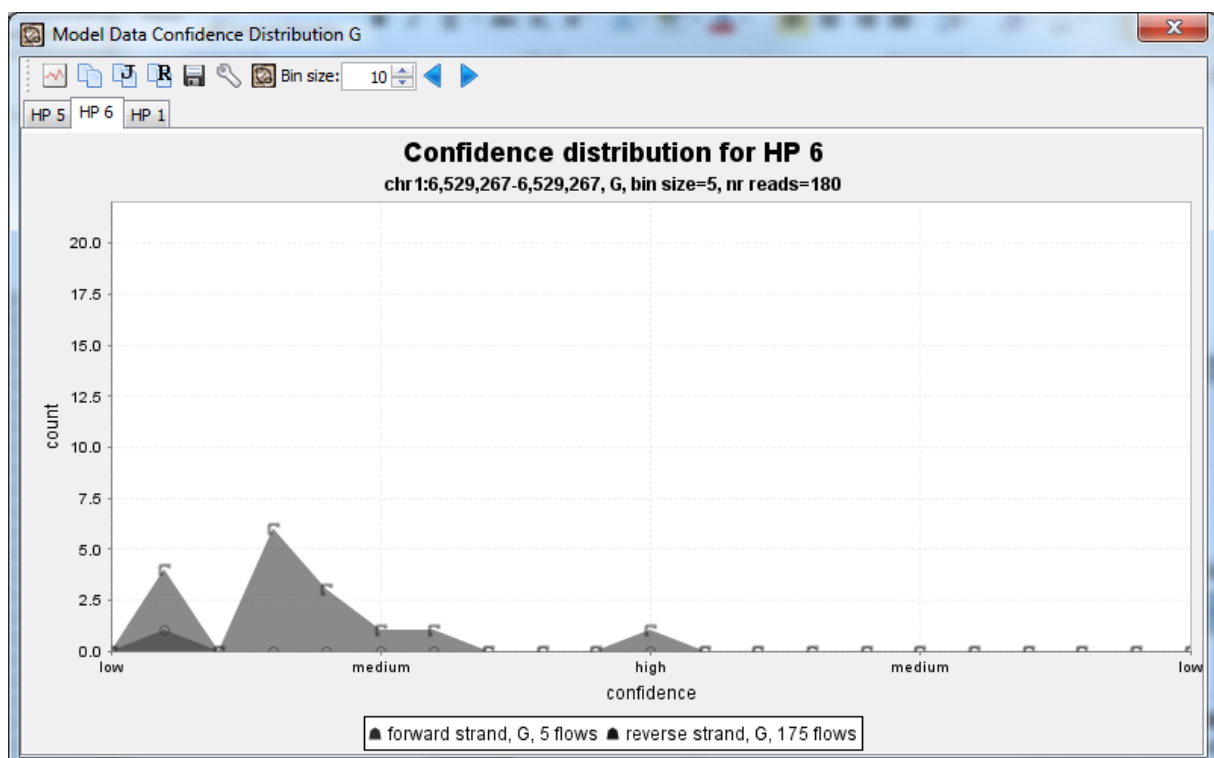

There are only very few base calls with HP of size 1, and the confidence tends to be low, meaning, most likely the HP size is actually 2.

Here is another example for a G, with the most likely HP size 5:
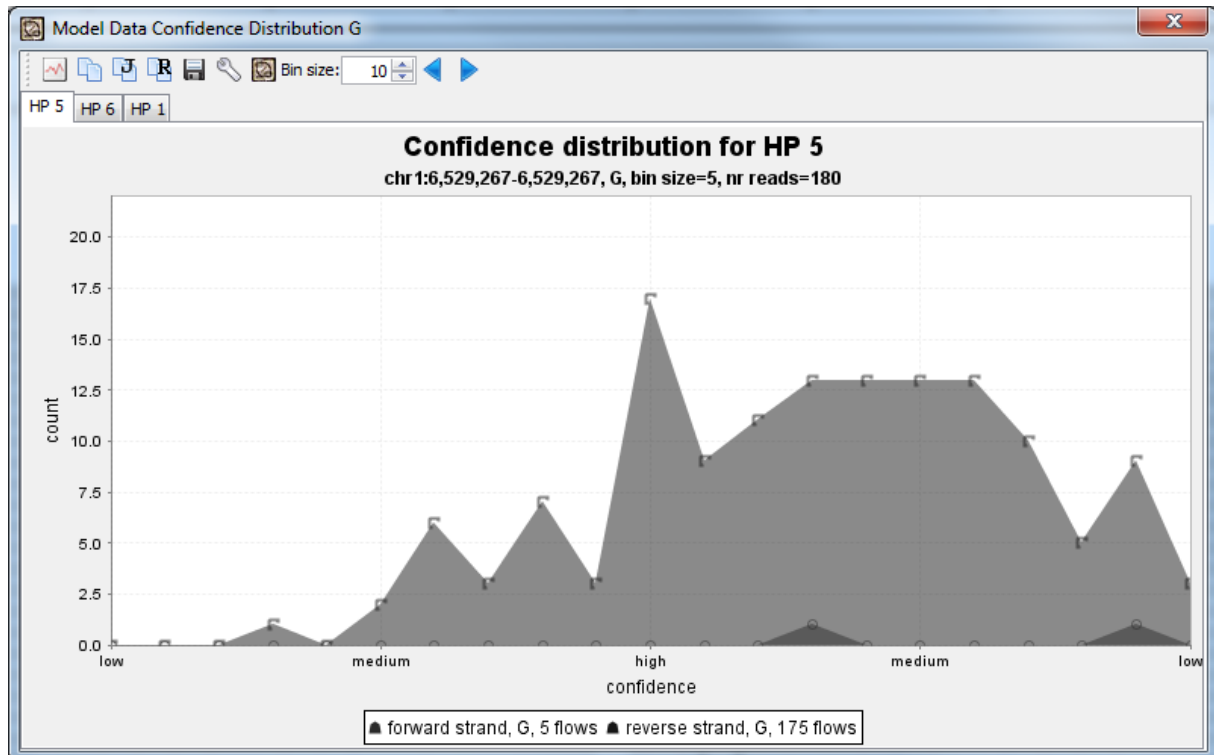
The confidence is not as high as in the case above, and the next most likely HP length is 6:

But even in this case, where the confidence is lower and is indicating a possibly longer HP length, once we look at the HP size 6 tab, it is clear that it is not very likely that the HP length is 6 (there are only a few reads with HP length 6, and the confidence of those is mostly low, and on the left side).
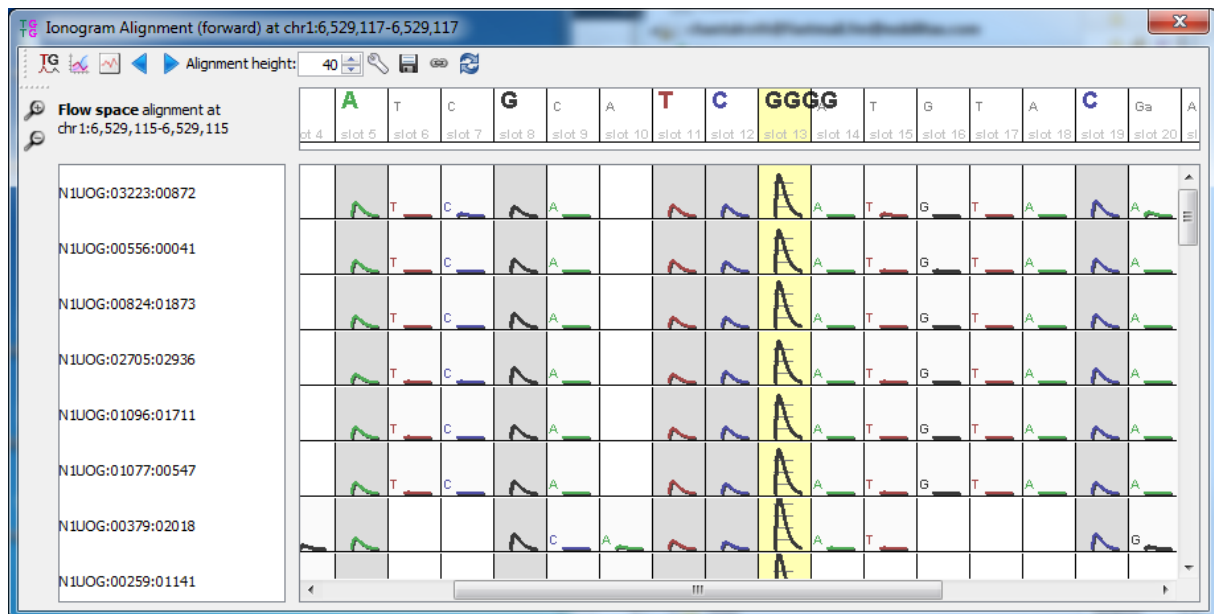
Ionogram Alignment

To view a multiple sequence alignment of the ionogram, which also includes empty flows, you can select the ionogram alignment menu item in the popup menu: the view shows the selected genome position in the middle (highlighted with yellow background), and a few bases up and downstream of that position. The empty flows are shown with white background, and the incorporation flows are shown with gray backgrounds. The alignment that is shown initially is not recomputed, but taken from the alignment as shown in the BAM file.

Note: The shapes are just cartoons of a raw signal. Further down you can see how to retrieve the actual raw signal.
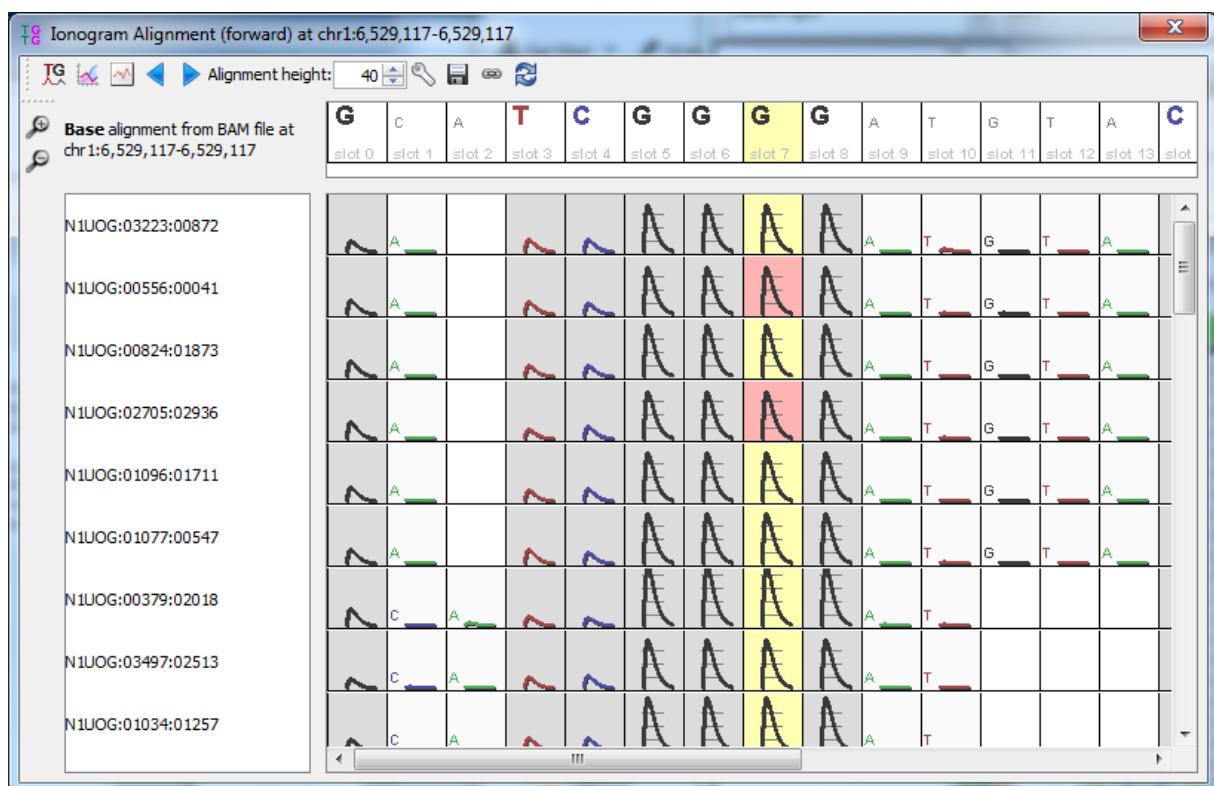
The first button in the tool bar allows you to recompute the multiple sequence alignment, using the empty flows as well. It doesn't use the flow values, as the raw values are not very useful for higher flow numbers, but it uses the length of the homo polymer, and it uses the fact that there are empty flows between incorporation calls, which may lead to a better alignment in some cases.

Once recomputed, the alignment shows the HPs in a single cell (the 4 G are in one cell).

Viewing Raw Signals in IGV

Both the ionogram alignment and confidence distribution has a way to let you retrieve the actual raw signal:

- In the Ionogram Alignment, select one or more of the yellow slots for which you would like to view the raw signal (highlighted in red)
- Select the Icon with the red line to tell it to get the raw data
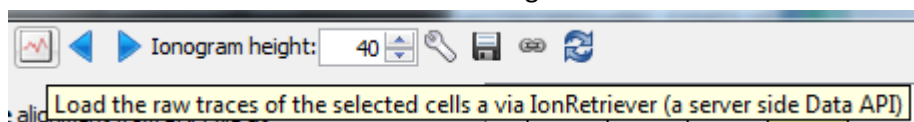
- In the settings panel, verify that IGV is using the correct server. Use the server is where the raw data is located. The PK/Run should either contain the name of the run, or the primary key (the last part of a run name). It will try to figure it out by itself based on the .BAM file that is opened, but in case the BAM file is a custom bam file or uses a different naming scheme, it may not know which run this is.
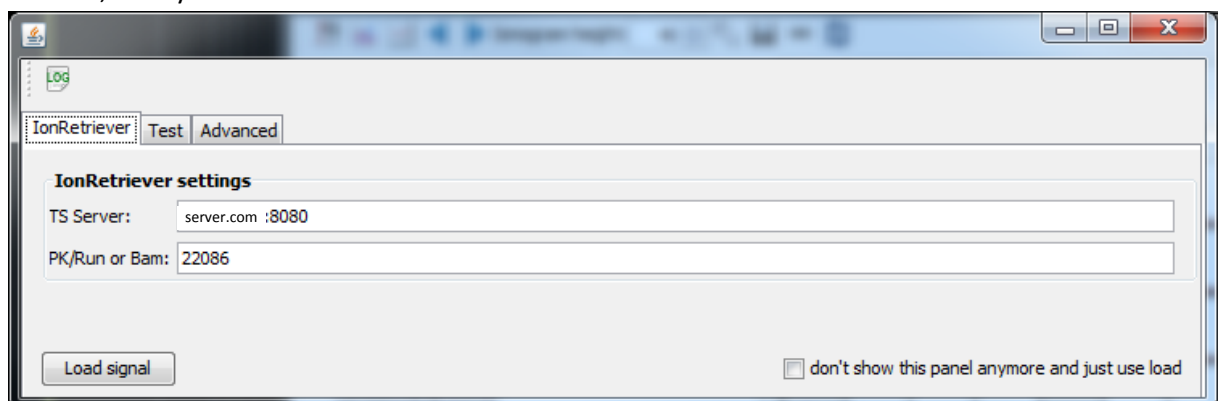
- It will now show you the URL it is planning to use. This view is mainly for me to test so that I can modify it if I have to, but you can actually use this URL and paste it into the Browser if you want to view the raw data in Excel, or you can use it to make any changes if you want to

change any parameters.
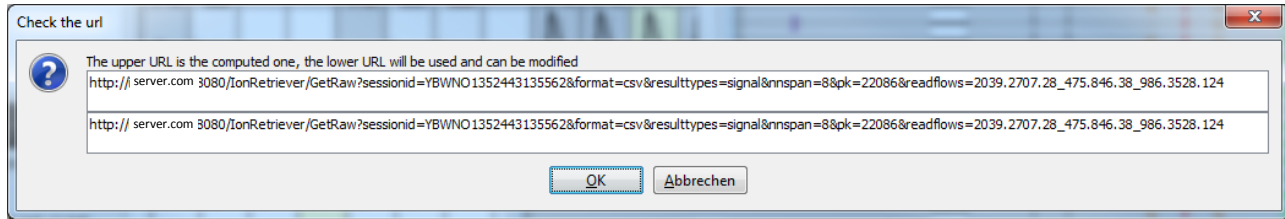
- Depending on whether the run already has regional information avaiable or not, it can will take about 10 seconds per signal to return the data and display it:
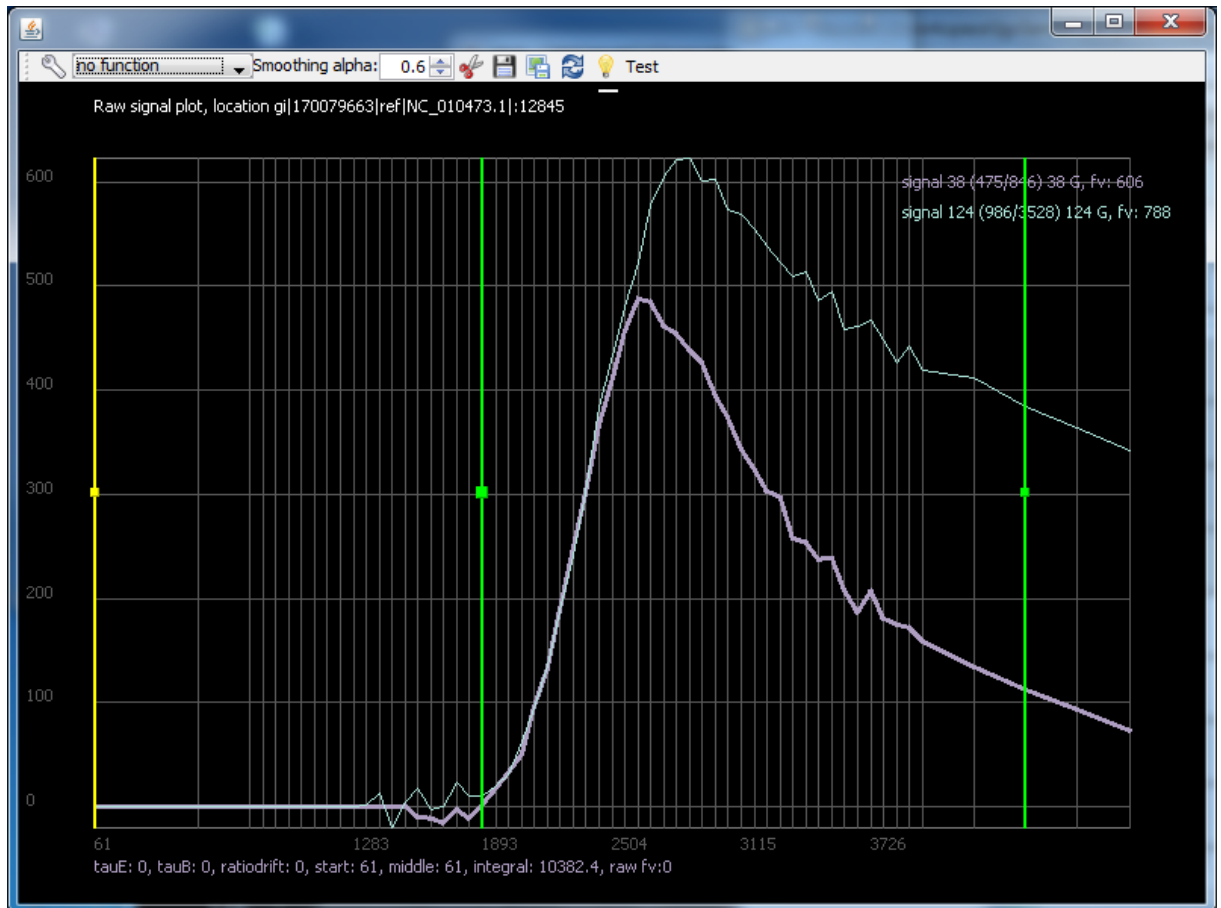
- Note how it computes the integral of the data between the green bars
- It will also display tauE and tauB, if the regional files are there (in this case the file were not available)

- You can smoothe the function by usingthe smooting alpha spinner to apply exponential smoothing (alpha is the weight used)
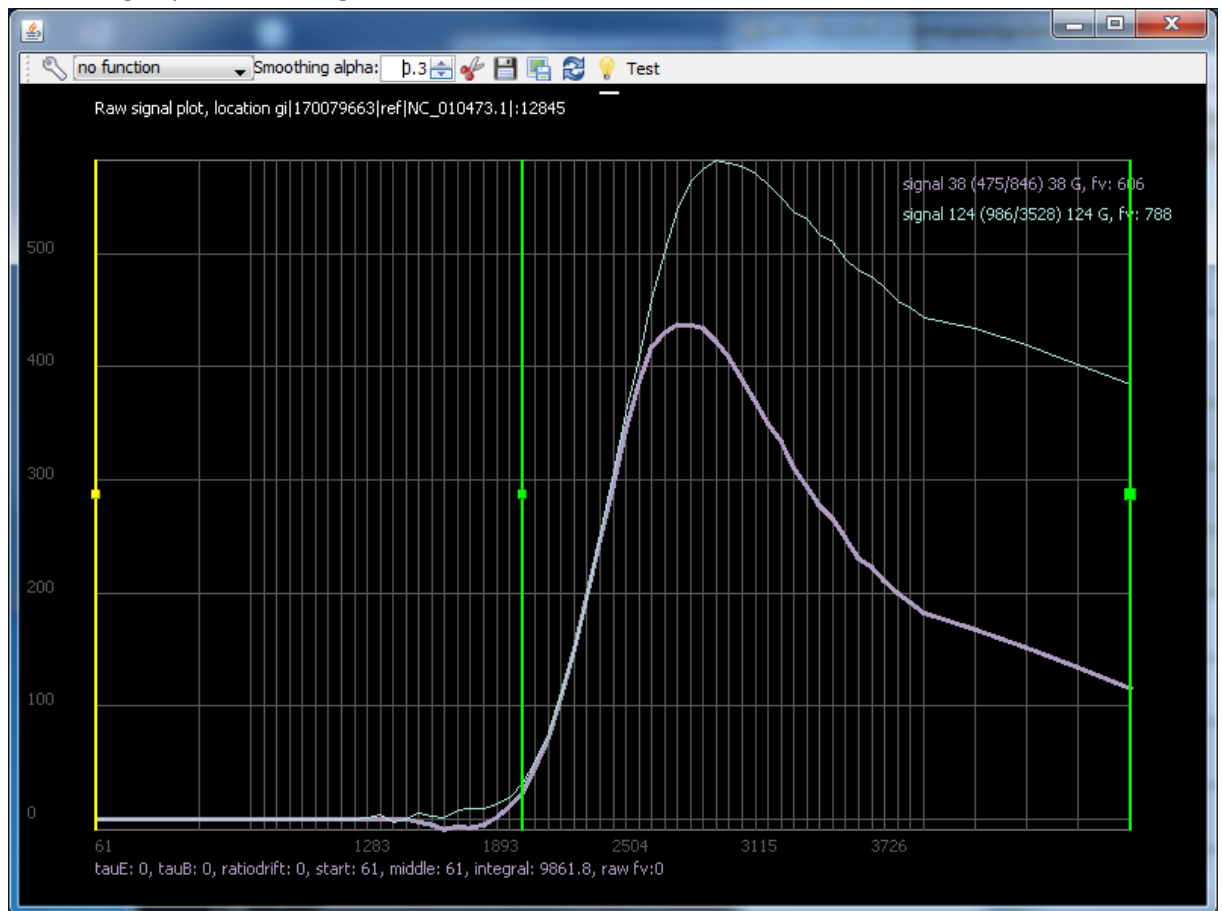
- The export buttons let you save the data in Excel (.csv) format, and also export the image

- If you like, you can overlay a function on top of this, such as the peak function: select the function in the drop down box:
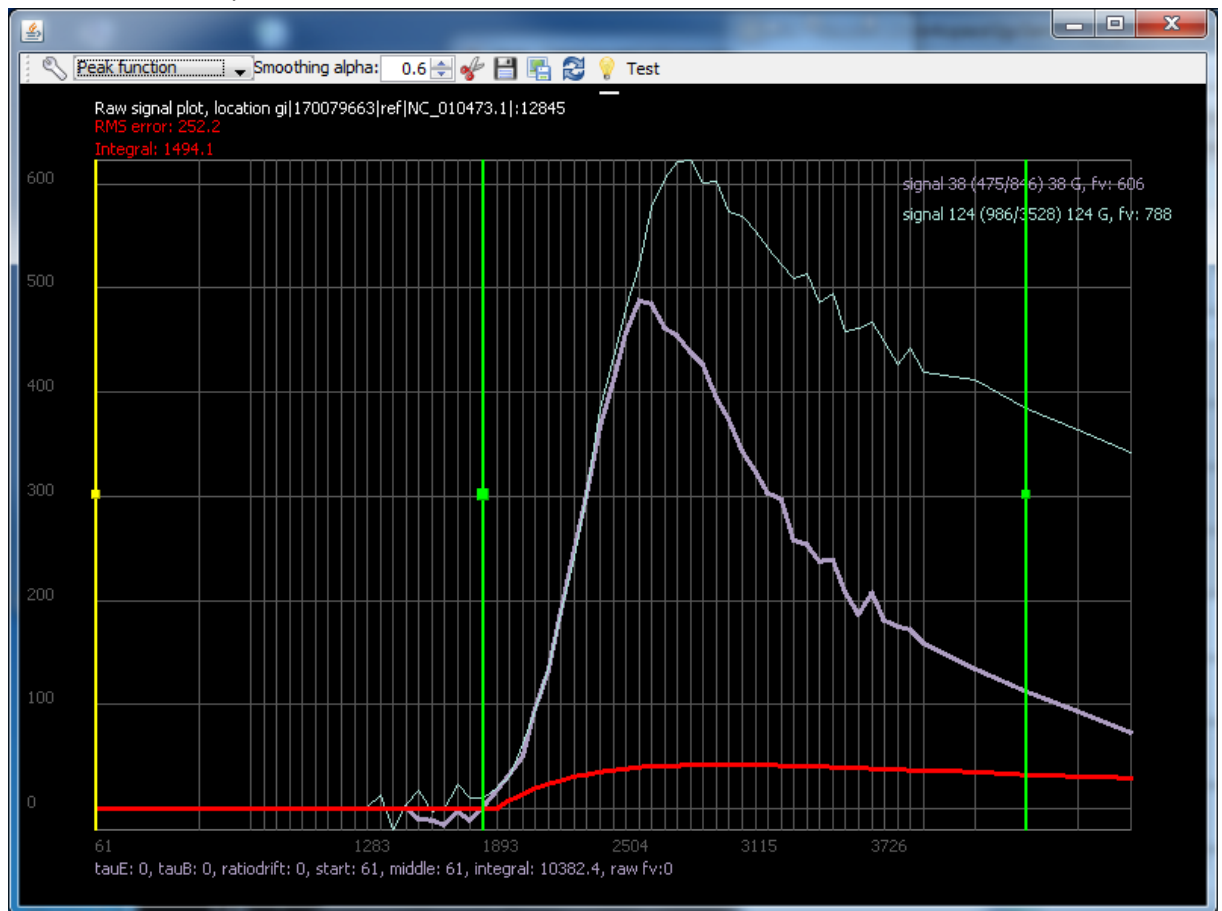
- Open the settings so you can adjust the function to match the selected signal

- Drag the left green bar to move the curve around, and slide the bars with m, k , and y scale to change the shape of the curve
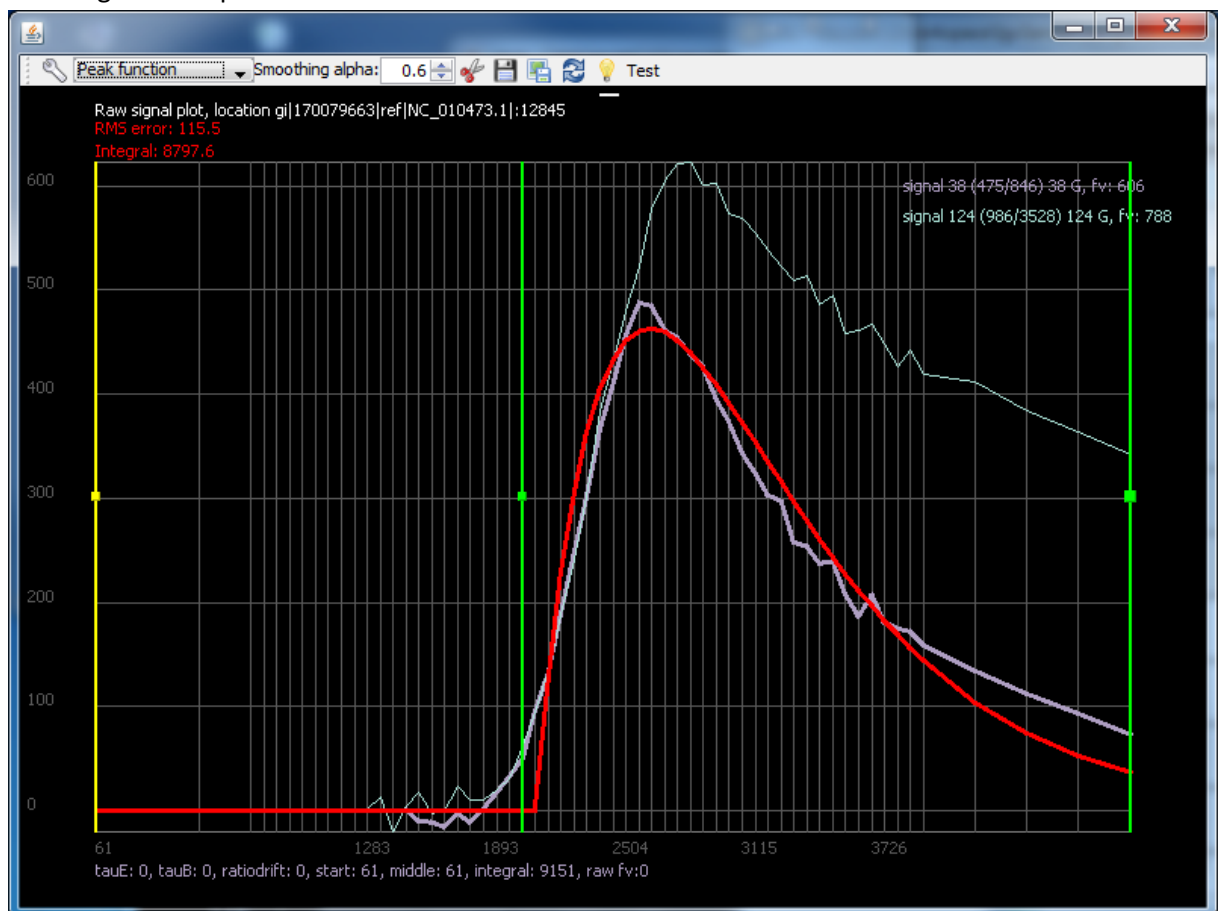
- It computes the integral and the RMS of the curve between the green bars

Raw Data in Excel

This explains how to use IonRetriever to view raw data in Excel.

1. For a run of interest, note the run name (or the primary key, which is the last numeric part of the run name), and note the coordinates and flows of the wells of interest
2. Also note from which torrent server the raw data is (such as ioneast.ite)
3. Open a browser and type a url of the form:
   http://myserver.com:8080/IonRetriever/GetRaw?resulttypes=signal&pk=28599&readflows=600.400.19
   The first part is always the server, port 8080 and GetRaw followed by a question mark:
   http://myserver.com:8080/IonRetriever/GetRaw?
4. The remainder are parameters in the format key=value, and they are separated by a &
5. The most important part is the name of the run or primary key, and of course the list of wells and flows.
6. If you use run name, use runname=nameofrun such as **runname= Auto_user_BEL-734--R160333-318_25_5_5_contra_fl-ATT_15449**
   You can also use just the beginning of the run name, as long as it is unique ☺. So this would

also be ok: `runname=Auto_user_BEL-734--R160333-318_25_5_5_contra`

If you use the primary key use pk=15449
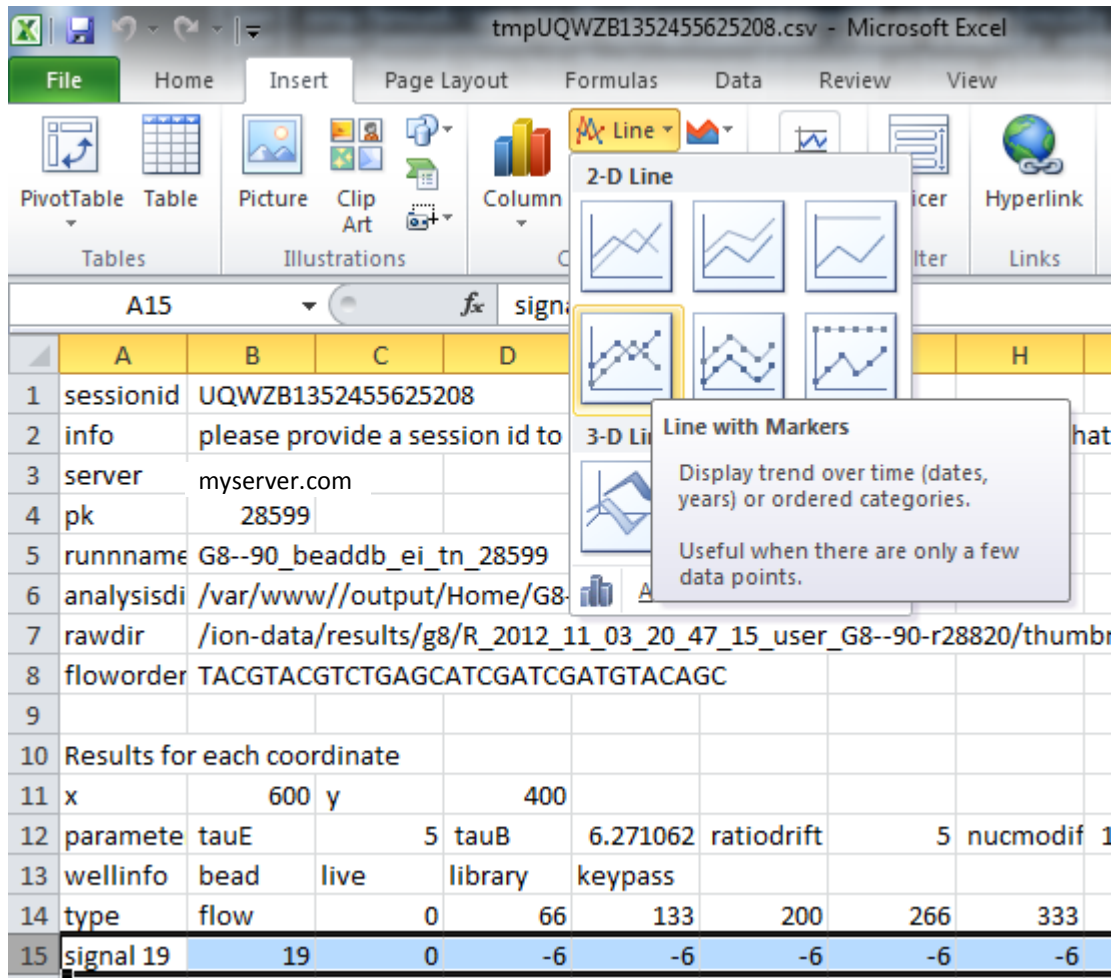
7. The list of wells and flows have the format: readflows=colum.row.flow_column.row.flow_...
Example: 100.200.1 would be the well 100/200 and flow 1.

8. After you hit enter, you will get a .csv file in a few seconds. Depending on your operating system and settings, Excel might open right away, or it might just show you that it downloaded a .csv file, or it may open the result as text:
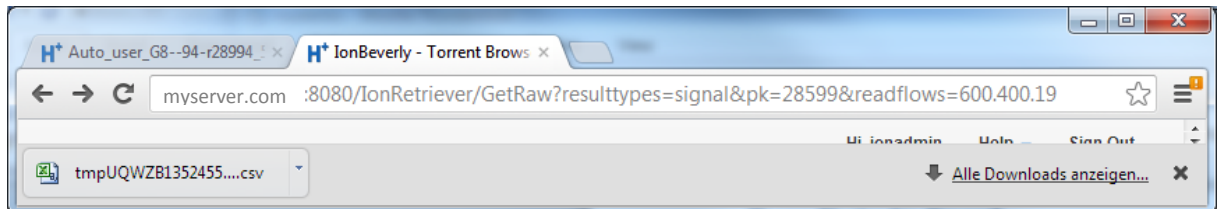
9. In my case, it downloads a .csv file and when I double click, it opens Excel:
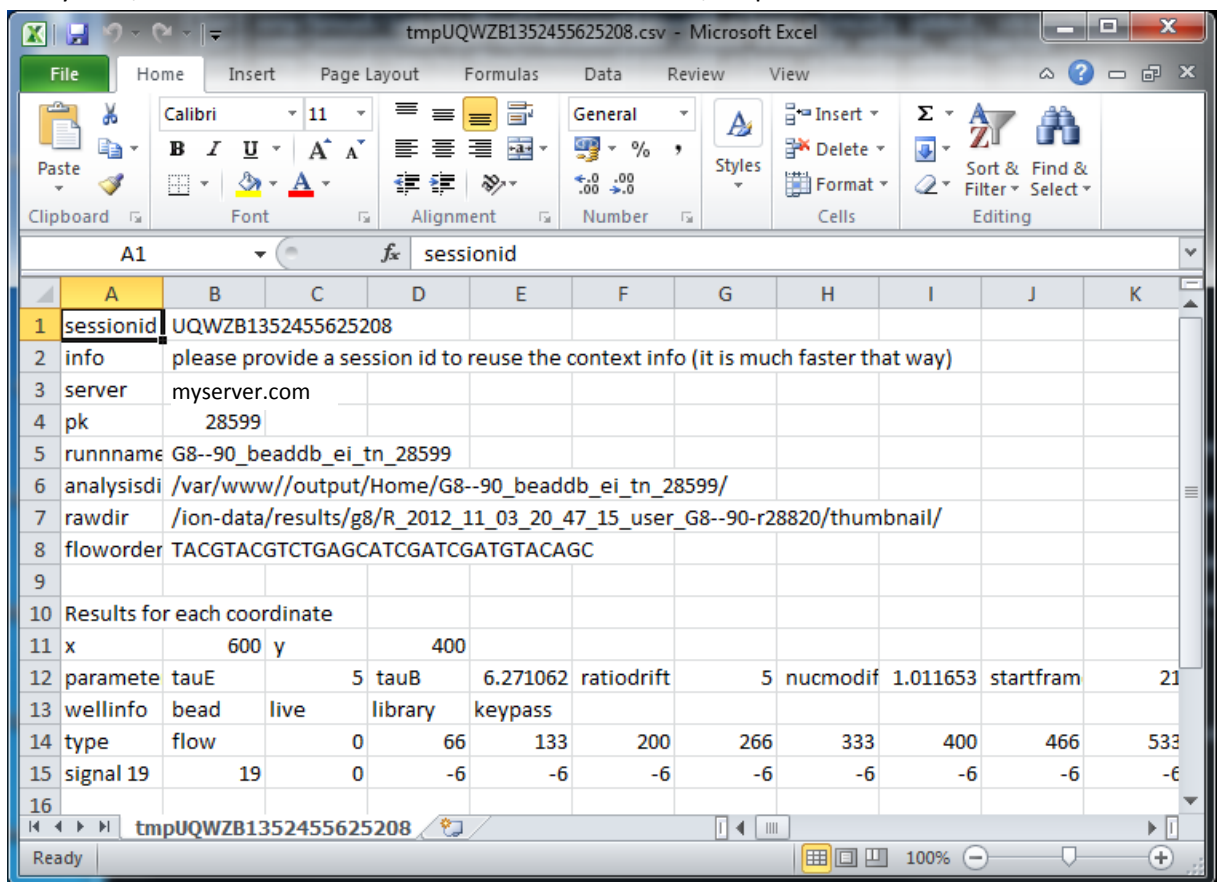
10. To view the data in a chart, select the entire row (such as row 15)

(you can later remove column B since this is the flow number)
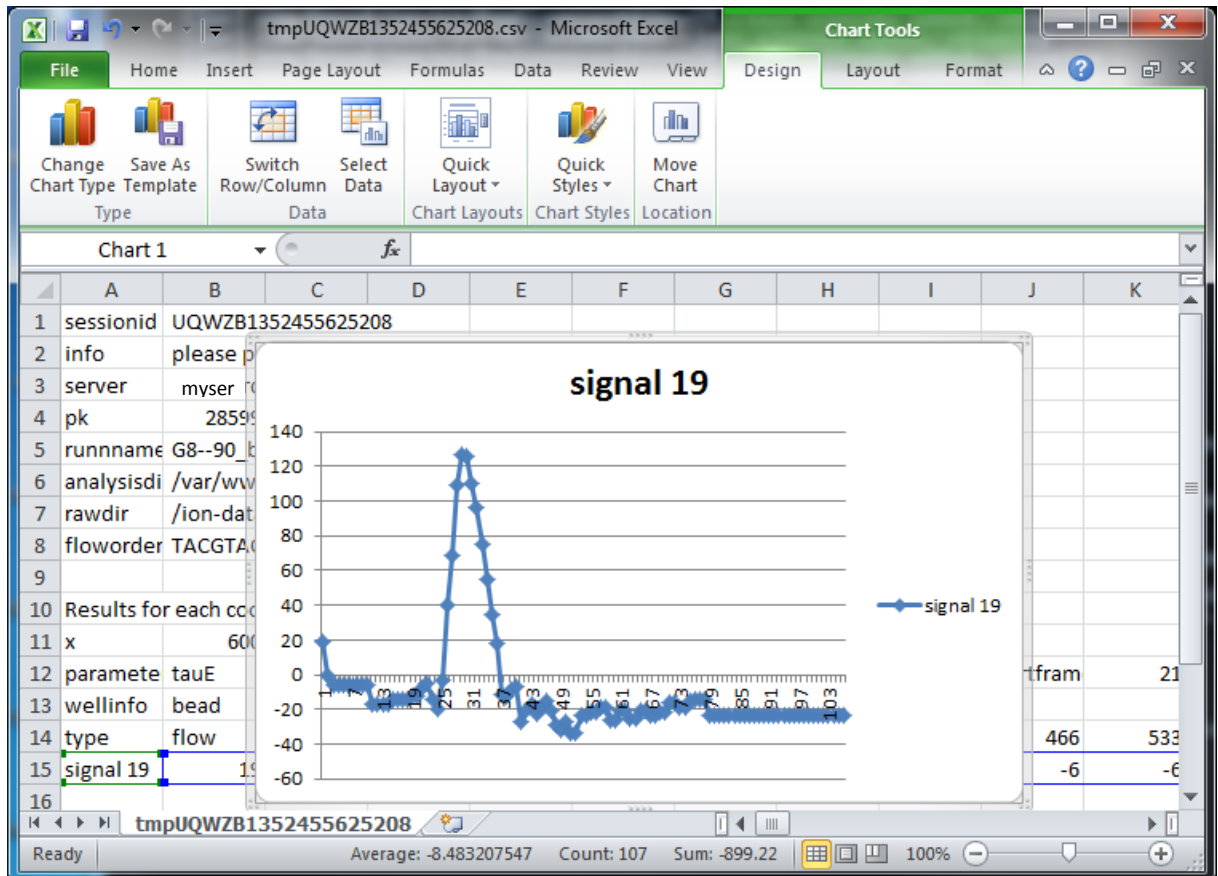
11. You can also ask for the zeromer and the empty signal by using resulttypes:

    resulttypes=signal_empty_zeromer

12. In this case, you would get three rows with data:

IonRetriever with other Clients

If you are a developer and would like to get raw data for a different client (any language), you can also get the data as .json in case that is easier for you to process.

1. To get the result as json format, use format=json
   http://myserver.com:8080/IonRetriever/GetRaw?**format=json**&resulttypes=signal_empty_ze romer&pk=28599&readflows=600.400.19

2. In that case, will look like this:

```
1    "result" {
2        "sessionid" : "YOCYO1352456229882",
3        "info" : "please provide a session id to reuse the context info (it is much faster that way)",
4        "server" :   "myserver.com"
5        "pk" : "28599",
6        "runnname" : "G8--90_beaddb_ei_tn_28599",
7        "analysisdir" : "/var/www//output/Home/G8--90_beaddb_ei_tn_28599/",
8        "rawdir" : "/ion-data/results/g8/R_2012_11_03_20_47_15_user_G8--90-r28820/thumbnail/",
9        "floworder" : "TACGTACGTCTGAGCATCGATCGATGTACAGC"
10       "data" : {
11   "timestamps":
     [0,66,133,200,266,333,400,466,533,600,666,733,800,866,933,1000,1066,1133,1200,1266,1333,1400,1466,
     00,2266,2333,2400,2466,2533,2600,2666,2733,2800,2866,2933,3000,3066,3133,3200,3266,3333,3400,3466,
     00,4266,4333,4400,4466,4533,4600,4666,4733,4800,4866,4933,5000,5066,5133,5200,5266,5333,5400,5466,
     00,6266,6333,6400,6466,6533,6600,6666,6733,6800,6866,6933],
12           "600_400"_19 : {
13               "wellinfo" : "bead,live,library,keypass"
14               "signal 19":
                 [0,-6,-6,-6,-6,-6,-6,-6,-6,-17,-17,-17,-17,-14,-14,-14,-14,-12,-12,-8,-5.11,-13.83,-20.
                 7,75.45,54.48,34.35,18.38,-11.85,-12.62,-8.57,-7.17,-27.05,-20.37,-16.74,-22.43,-19.17,
                 97,-22.99,-21.87,-21.88,-18.42,-18.43,-26.2,-26.23,-20.56,-20.53,-25.69,-25.67,-20.37,-
                 1,-18.69,-14.7,-14.67,-14.66,-14.66,-23.85,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-2
                 4,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-
15               "Median empty 19":
                 [0,-2.15,-2.15,-2.15,-2.15,-2.15,-2.15,-2.15,-2.15,-3.51,-3.51,-3.51,-3.51,-5.12,-5.12,
                 6,3.81,15.6,30.83,49.86,70.67,89.95,107.64,124.88,141.84,157.96,173.92,188.13,200.86,20
                 4,216.39,216.57,215.28,214.79,215.59,215.59,216.99,217,217.88,217.88,218.43,218.43,220.
                 217.09,217.09,217.61,217.61,218.07,218.07,216.69,216.68,214.66,214.66,214.66,214.66,213
                 ,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,21
                 4],
16               "zeromer":
                 [0,2.15,2.15,2.15,2.15,2.15,2.15,2.15,2.15,3.51,3.51,3.51,3.51,5.12,5.12,5.12,5.12,4.55
                 35,-0.4,-0.37,-0.34,-0.33,-0.32,-0.31,-0.31,-0.27,-0.24,-0.16,-0.06,-0,-0.02,-0.01,-0.0
                 2,-0,-0.01,-0,-0.03,-0,0.03,0,0.02,0,0.02,0,-0.01,-0,-0.01,-0,-0.01,-0,0.02,0,0.04,0,0,
                 ,0,0,0,0]
17           }
```
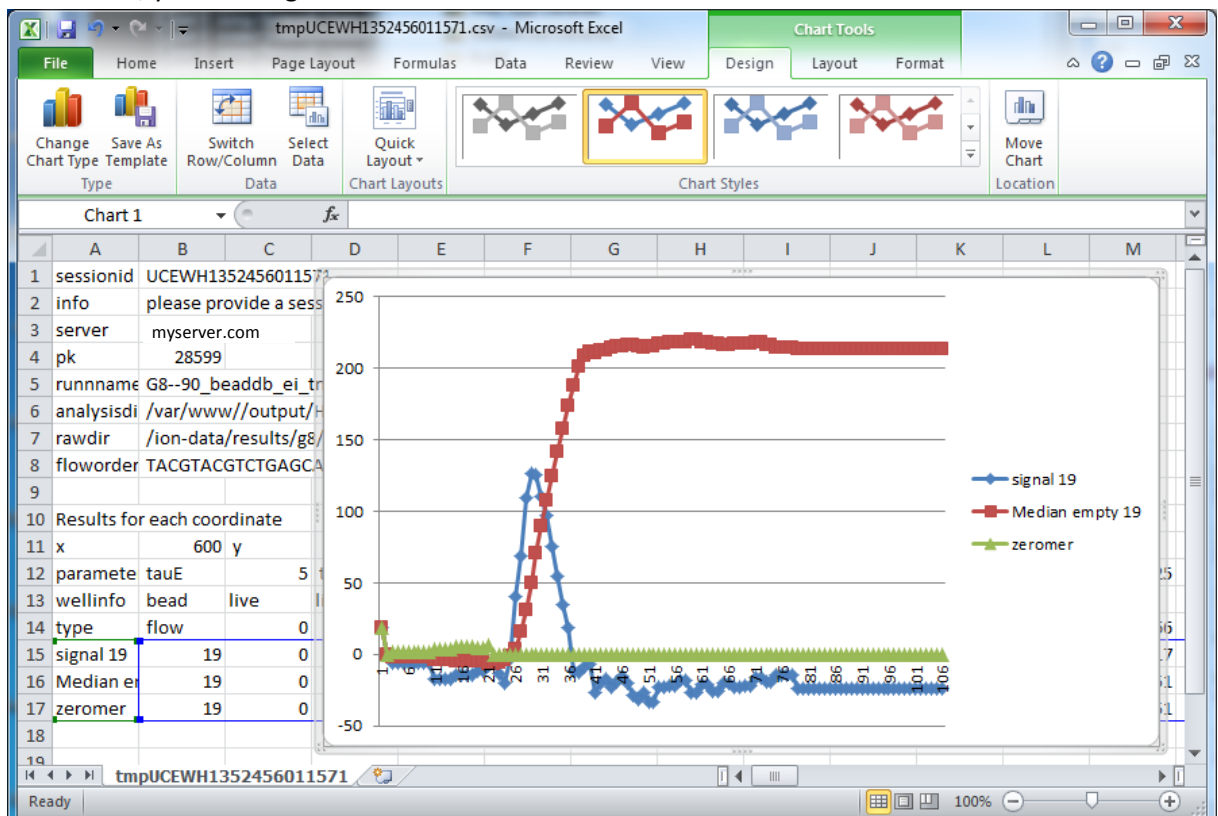
3. If you prefer less verbose output, you can use verbose=false:
   http://myserver.com:8080/IonRetriever/GetRaw?**verbose=false**&format=json&resulttypes=s ignal_empty_zeromer&pk=28599&readflows=600.400.19

```
1   "result" {
2       "sessionid" : "SAKHU1352456340741",
3       "info" : "please provide a session id to reuse the context info (it is much faster that way)",
4       "data" : {
5   "timestamps":
    [0,66,133,200,266,333,400,466,533,600,666,733,800,866,933,1000,1066,1133,1200,1266,1333,1400,1466,
    2266,2333,2400,2466,2533,2600,2666,2733,2800,2866,2933,3000,3066,3133,3200,3266,3333,3400,3466,353
    6,4333,4400,4466,4533,4600,4666,4733,4800,4866,4933,5000,5066,5133,5200,5266,5333,5400,5466,5533,5
    333,6400,6466,6533,6600,6666,6733,6800,6866,6933],
6           "600_400"_19 : {
7               "wellinfo" : "bead,live,library,keypass"
8               "signal 19":
                [0,-6,-6,-6,-6,-6,-6,-6,-6,-17,-17,-17,-17,-14,-14,-14,-14,-12,-12,-8,-5.11,-13.83,-20.
                5.45,54.48,34.35,18.38,-11.85,-12.62,-8.57,-7.17,-27.05,-20.37,-16.74,-22.43,-19.17,-15
                .99,-21.87,-21.88,-18.42,-18.43,-26.2,-26.23,-20.56,-20.53,-25.69,-25.67,-20.37,-20.35,
                -14.7,-14.67,-14.66,-14.66,-23.85,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.
                .84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84],
9               "Median empty 19":
                [0,-2.15,-2.15,-2.15,-2.15,-2.15,-2.15,-2.15,-2.15,-3.51,-3.51,-3.51,-3.51,-5.12,-5.12,
                .81,15.6,30.83,49.86,70.67,89.95,107.64,124.88,141.84,157.96,173.92,188.13,200.86,208.7
                39,216.57,215.28,214.79,215.59,215.59,216.99,217,217.88,217.88,218.43,218.43,220.23,220
                7.09,217.61,217.61,218.07,218.07,216.69,216.68,214.66,214.66,214.66,214.66,213.84,213.8
                84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,
10              "zeromer":
                [0,2.15,2.15,2.15,2.15,2.15,2.15,2.15,2.15,3.51,3.51,3.51,3.51,5.12,5.12,5.12,5.12,4.55
                -0.4,-0.37,-0.34,-0.33,-0.32,-0.31,-0.31,-0.27,-0.24,-0.16,-0.06,-0,-0.02,-0.01,-0.03,-
                0.01,-0,-0.03,-0,0.03,0,0.02,0,0.02,0,-0.01,-0,-0.01,-0,-0.01,-0,0.02,0,0.04,0,0,0,0.01
11          }
12
13  }
14  }
```

4. Note how it keeps saying that you should provide a session ID. If you write a client that will have to ask for raw data multiple times for the same run, add sessionid=id to the URL so that it can reuse cached data (for instance it won't have to query for the experiment every time you call IonRetriever). To make it easier, it already gives you the session id that it created (combination of time stamp and random values), but you can use your own, as long as it is unique for all users that currently use the same service: http://myserver.com:8080/IonRetriever/GetRaw?**sessionid=SAKHU1352456340741**&verbose=false&format=json&resulttypes=signal_empty_zeromer&pk=28599&readflows=600.400.19

5. Now you got rid of that message too:

```
"result" {
    "sessionid" : "SAKHU1352456340741",
    "data" :  {
"timestamps":
[0,66,133,200,266,333,400,466,533,600,666,733,800,866,933,1000,1066,1133,
2266,2333,2400,2466,2533,2600,2666,2733,2800,2866,2933,3000,3066,3133,320
6,4333,4400,4466,4533,4600,4666,4733,4800,4866,4933,5000,5066,5133,5200,5
333,6400,6466,6533,6600,6666,6733,6800,6866,6933],
        "600_400"_19 :  {
            "wellinfo" : "bead,live,library,keypass"
            "signal 19":
            [0,-6,-6,-6,-6,-6,-6,-6,-6,-17,-17,-17,-17,-14,-14,-14,-14,-12
            5.45,54.48,34.35,18.38,-11.85,-12.62,-8.57,-7.17,-27.05,-20.37
            .99,-21.87,-21.88,-18.42,-18.43,-26.2,-26.23,-20.56,-20.53,-25
            -14.7,-14.67,-14.66,-14.66,-23.85,-23.84,-23.84,-23.84,-23.84,
            .84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-23.84,-2
            "Median empty 19":
            [0,-2.15,-2.15,-2.15,-2.15,-2.15,-2.15,-2.15,-2.15,-3.51,-3.51
            .81,15.6,30.83,49.86,70.67,89.95,107.64,124.88,141.84,157.96,1
            39,216.57,215.28,214.79,215.59,215.59,216.99,217,217.88,217.88
            7.09,217.61,217.61,218.07,218.07,216.69,216.68,214.66,214.66,2
            84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213.84,213
            "zeromer":
            [0,2.15,2.15,2.15,2.15,2.15,2.15,2.15,2.15,3.51,3.51,3.51,3.51
            -0.4,-0.37,-0.34,-0.33,-0.32,-0.31,-0.31,-0.27,-0.24,-0.16,-0.
            0.01,-0,-0.03,-0,0.03,0,0.02,0,0.02,0,-0.01,-0,-0.01,-0,-0.01,
        }
```