Chromosome View

The *Chromosome View* in Partek® Genomic Suite™ (Partek GS) (Figure 1) is a visualization tool for genomic data. It is initialized by **View > Chromosome View.** It can be accessed from certain workflows under the *Visualization* tab by selecting **Plot Chromosome View**, or by <right-clicking> on a gene row header and selecting **Browse to Location**. The view can display multiple levels of genomic data simultaneously, including *Heat Maps, Reference Sequence Transcripts, Reference Genomes, Amplification* and *Deletion* sites and more. The chromosome view initially displays chromosome one (1) by default. The panel on the left allows you to customize the way the data is presented in the view.
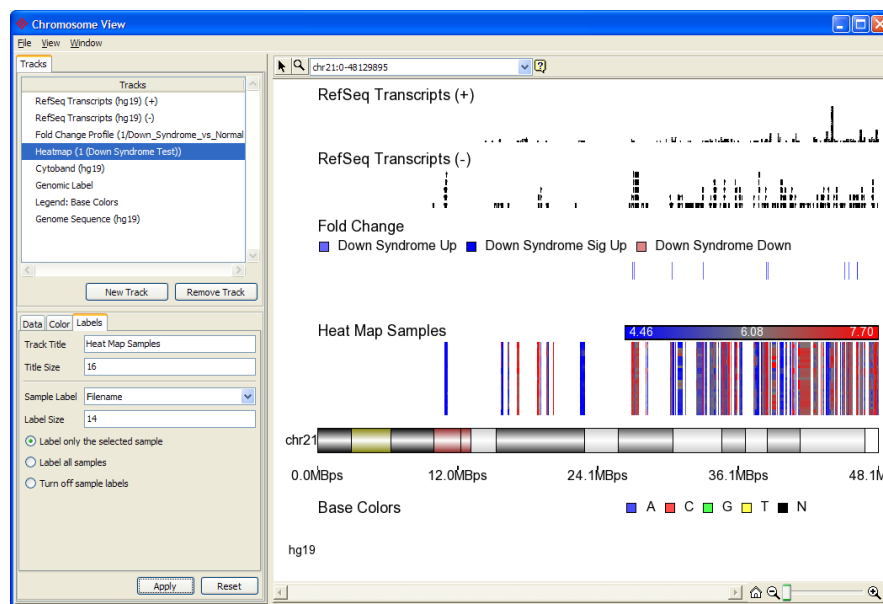

Figure 1: Viewing the Chromosome View

When the *Chromosome View* is first initiated, you will be asked to choose a default annotation if one has not already been chosen though the analytical process (Figure 2). There are three (3) options available for the human genome build 19 in Figure 2, *RefSeq Transcripts, Ensembl Transcripts*, and *Do not download any file at this time*. Partek GS will attempt to automatically download the chosen annotation data. If *Do not download any file at this time* is chosen, only a Cytoband file from UCSC will attempt to be downloaded.
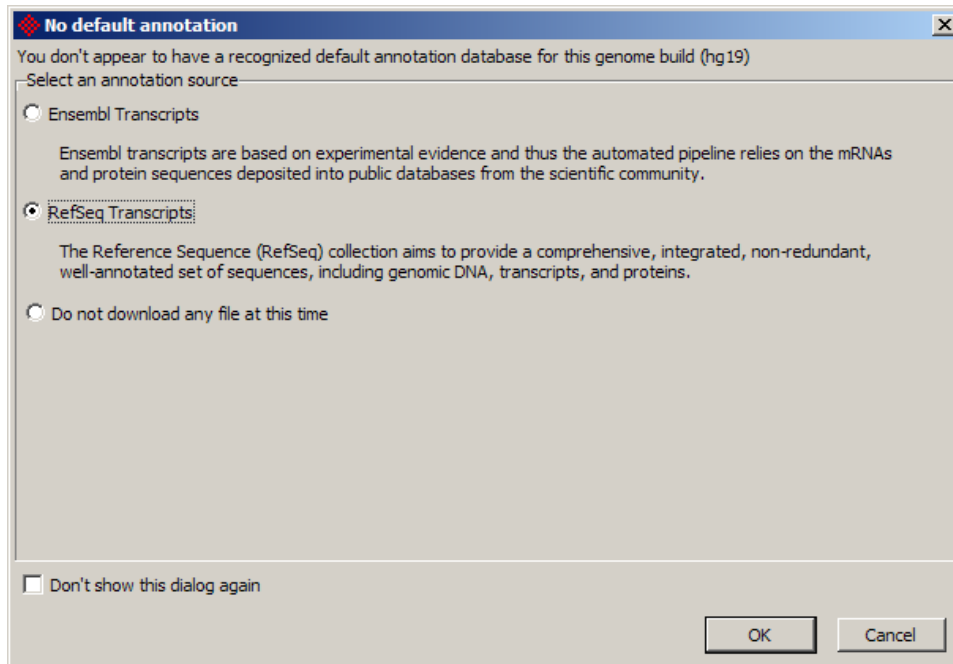
Figure 2: Viewing Annotation options when opening the Chromosome View for the first time

Search Bar

The search bar can be found above the view (Figure 3).

Figure 3: Viewing the Search bar of Chromosome View

Use the search bar to zoom to genomic features that are available in annotation tracks. Type or paste in genomic positions such as chr6:40544957-49169085 or VEGFA (Figure 4).
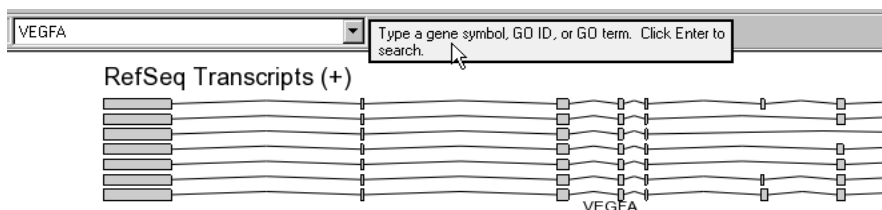

Figure 4: Finding genomic features with the search box

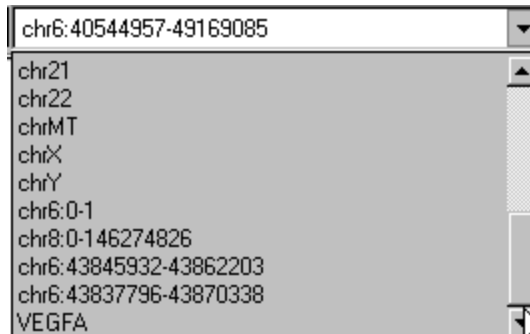The *Search* bar will display a dropdown list of the last ten searches (Figure 5).

Figure 5: Viewing the Search bar history

Modes

Selection Mode

<Left-click> on a track in selection mode (Figure 6) to select the track in the track panel and to edit the properties of the selected track. Individual tracks have unique editable settings. <Right-click> on a track for an option to remove it. <Left-click> on heat maps to select samples. <Left-click> and drag on the Cytoband track to zoom to that region.
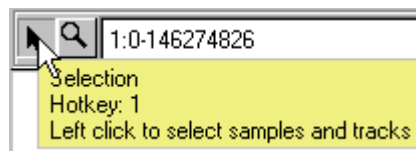

Figure 6: Entering Selection Mode

Navigation Mode

In *Navigation* mode (Figure 7), <left-click> to zoom in on a selected region. <Right-clicking> will re-center the plot.
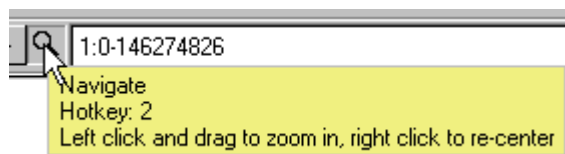

Figure 7: Entering Navigation Mode

The mouse wheel will zoom in and out in both *Selection* and *Navigation* modes.

Mouse Over

Information about features under the mouse cursor (i.e. genes, sample ID, chromosome position) is displayed in the top-right corner (Figure 8).
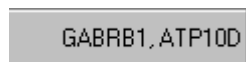

Figure 8: Viewing MouseOver information

Use the slider (Figure 9) or the magnifying glass button in the bottom right corner of the view to zoom in and out. Select the *Home* ⌂ button to reset the view to the full chromosome.

Figure 9: Viewing the Zoom Slider

Tracks

From the *Tracks* tab, you can add or remove tracks and configure the properties of tracks in the view (Figure 10). With multiple tracks selected, changing common properties such as title or font size will be applied to all selected tracks, but unique properties such as track title will only be applied to the bottom selected track. Each track has editable parameters controlled by the tabs below track list. If changes are made to the track, select **Apply** to update the track with the changes. Select **Reset** to change the values back to the default settings.
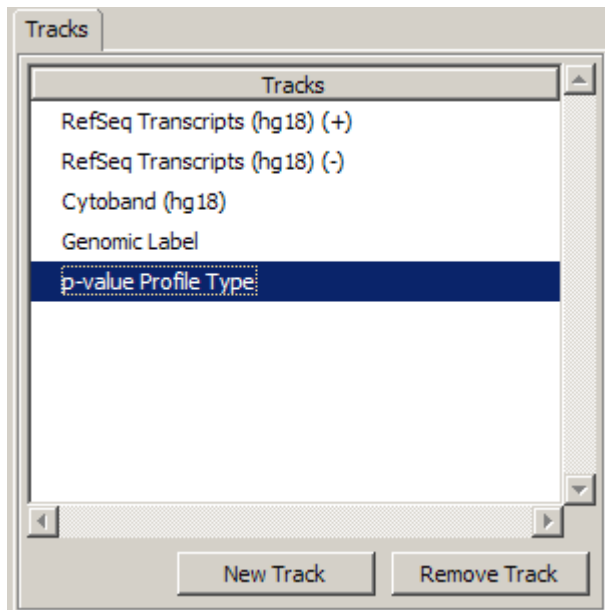

Figure 10: Viewing the Tracks panel, which shows a list of tracks displayed in the view

New Track Options

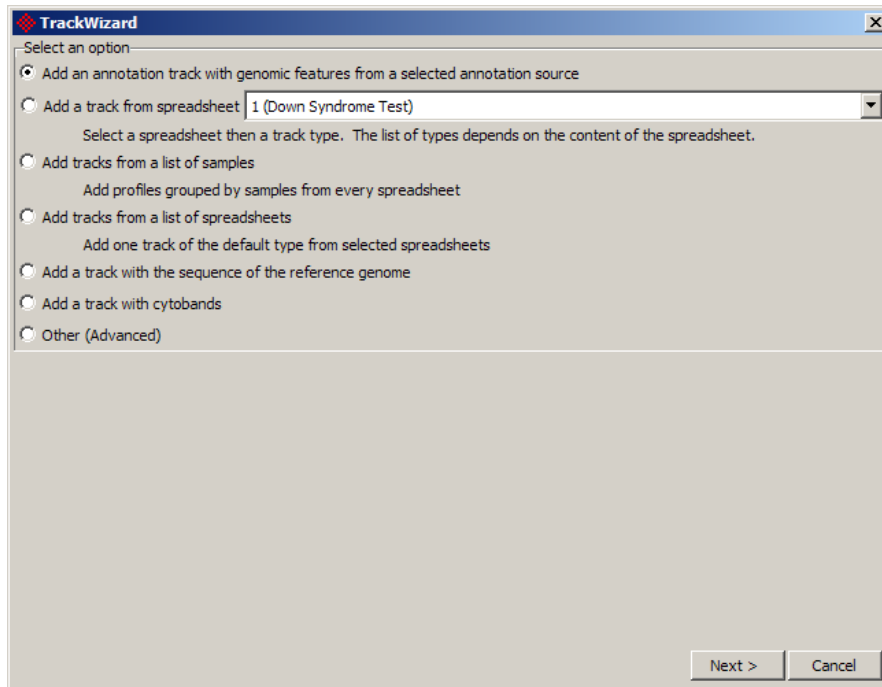Select **New Track** to prompt the *TrackWizard* dialog (Figure 11), which shows the options to add new tracks.

Figure 11: Adding a new track with the TrackWizard

1) *Add an annotation track with genomic features from a selected annotation source* for a list of available annotations. Choose an available annotation and select *Create,* or select *Manage available annotations* to add a new annotation. Partek will attempt to automatically download the annotation file chosen if it is not already available on the local system, as indicated by the *Download Required* message highlighted in red (Figure 12)
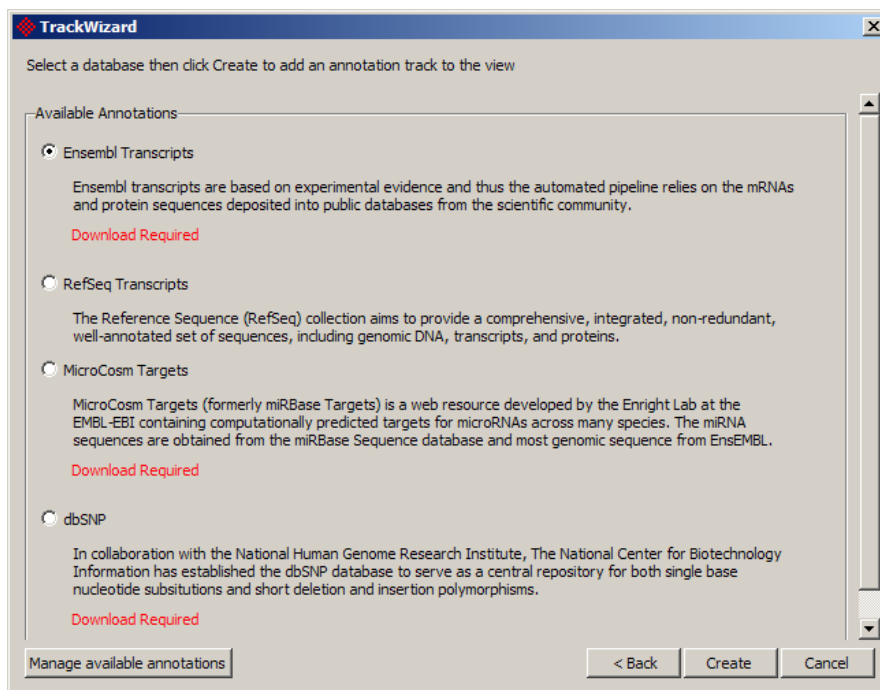
Annotation Track

Figure 12: Adding an annotation track to the Chromosome View

Figure 13 is an example of one of the available annotation tracks Partek GS will attempt to download, *RefSeq*. For RefSeq, two gene annotation tracks are added, one filtered to the positive (+) strand (5' on the left and 3' on the right) and the other track filtered to the negative (-) strand (3' on the left and 5' on the right). An option to display both positive (+) and negative (-) strands on one track is available from the *Strand* drop-down in the *Tracks* panel (Figure 15).
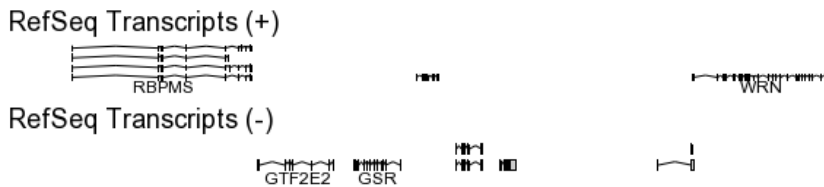

Figure 13: Viewing Gene annotation tracks of positive and negative strands with every gene labeled

By default, each stack of genes is labeled with the *Gene ID* at the bottom of the track (Figure 14). This is a result of the *Label every gene* option selected as the default setting. *Label every isoform* will draw the transcript id on top of each transcript.
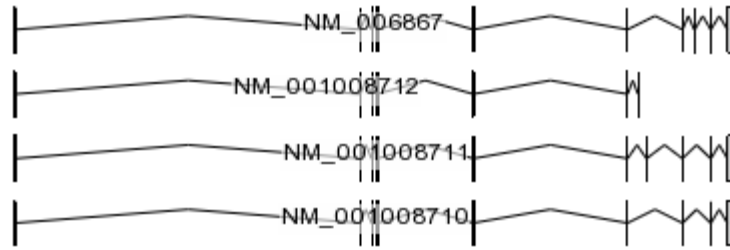
Figure 14: Viewing known isoforms of gene location

The track height slider adjusts the relative height of the track (Figure 15). Select the track from the drop down box of the track to change the height of that track. To increase the track height, move the slider bar to the left and select **Apply**. To decrease the track height, move the slider to the right and select **Apply**.
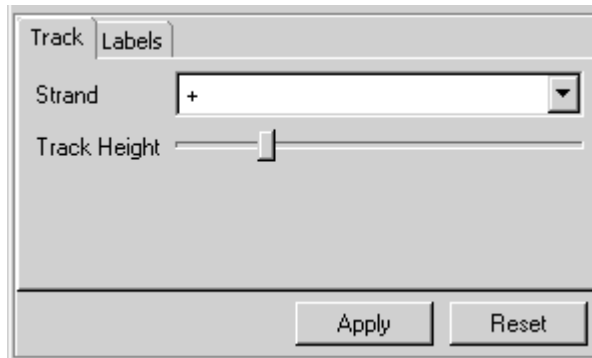

Figure 15: Configuring the track height

The *Labels* tab configures the track title attributes and how to display gene or isoform labels (Figure 16).
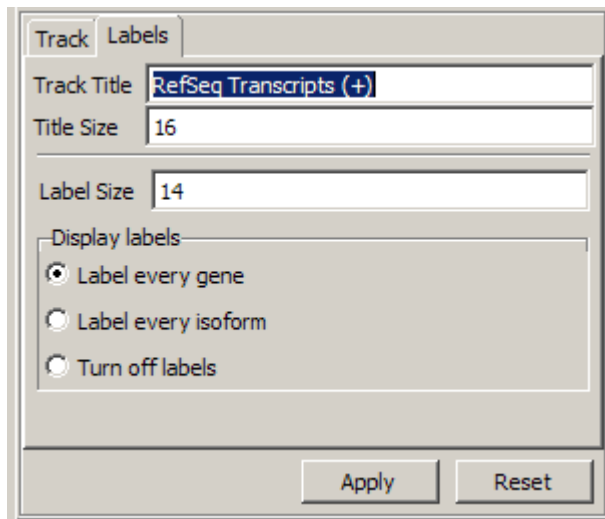

Figure 16: Editing the Annotation Track label properties

The default title size and label sizes can be configured from **Edit > Preferences** from the main window.

2) *Add a track from spreadsheet* to create a track based on samples of the spreadsheet. When adding a track from a spreadsheet, the list of options is determined by the type of spreadsheet. If there is only type of track appropriate for the spreadsheet, then selecting *Next* will add the track. Figure 17, Figure 18, and Figure 19 give examples of track options available from the different spreadsheet types.
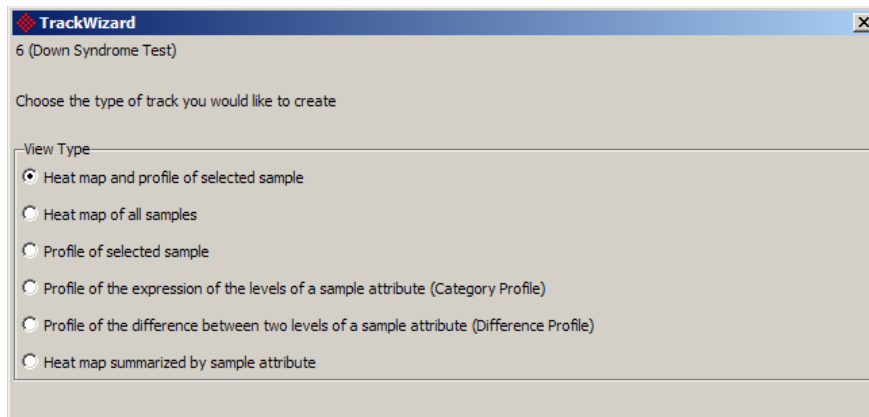

Figure 17: Adding a New Track using the option "Add a track from spreadsheet" with a Sample spreadsheet selected from dropdown menu. Descriptions of these tracks are mentioned below
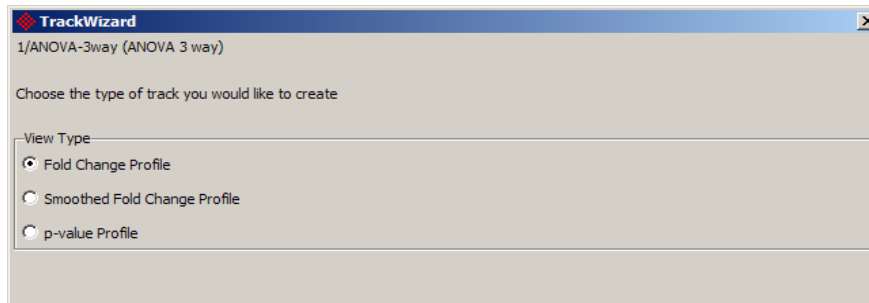

Figure 18: Adding a New Track using the option "Add a track from spreadsheet" with an ANOVA results spreadsheet selected from dropdown menu
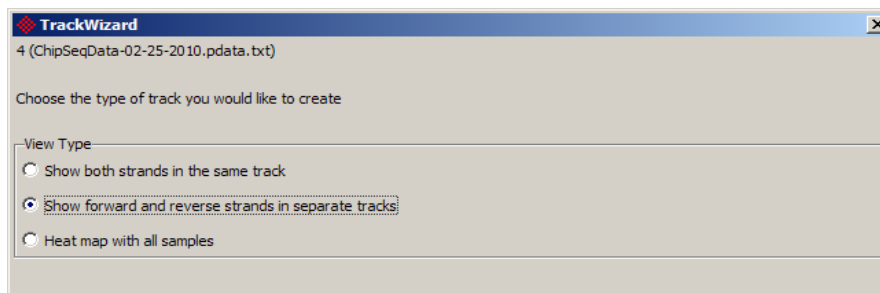

Figure 19: Adding a New Track using the option "Add a track from spreadsheet"

with a ChipSeq reads spreadsheet selected from dropdown menu

3) *Add tracks from a list of samples* to create a profile track by selecting samples (Figure 20). Profiles are grouped by samples from every spreadsheet. Choose the samples by individually checking them and selecting the *Create* button, or choose them by a sample attribute and level attribute and select the *Check* button to select all the samples with the specified attribute. Profile tracks are generally most appropriate for Copy Number visualization, but have a broader purpose. Please see Copy Number tracks section for more information on Profile tracks.
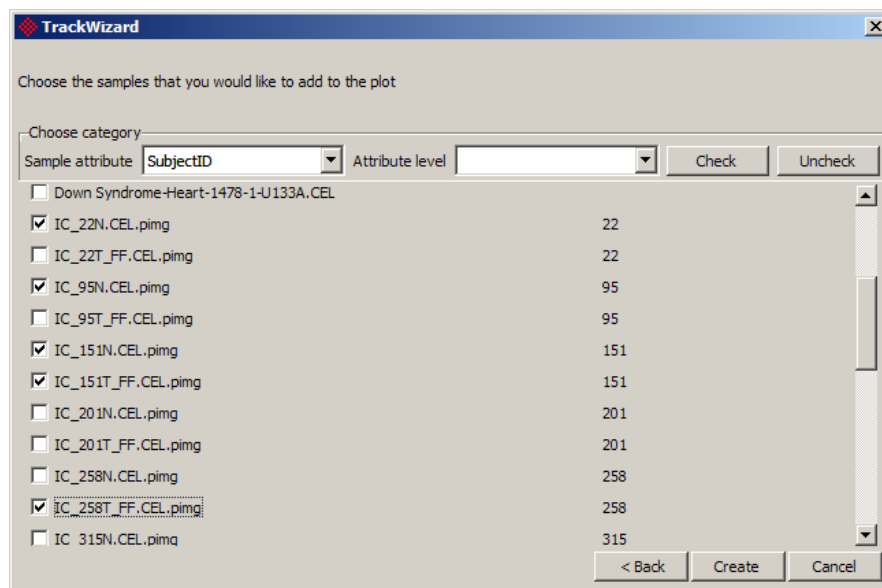

Figure 20: Adding tracks from a list of samples

4) *Add tracks from a list of spreadsheets* to create a track from the spreadsheet list (Figure 21). The option adds one track of the default type from the selected spreadsheets. The spreadsheet list contains all the spreadsheets currently open
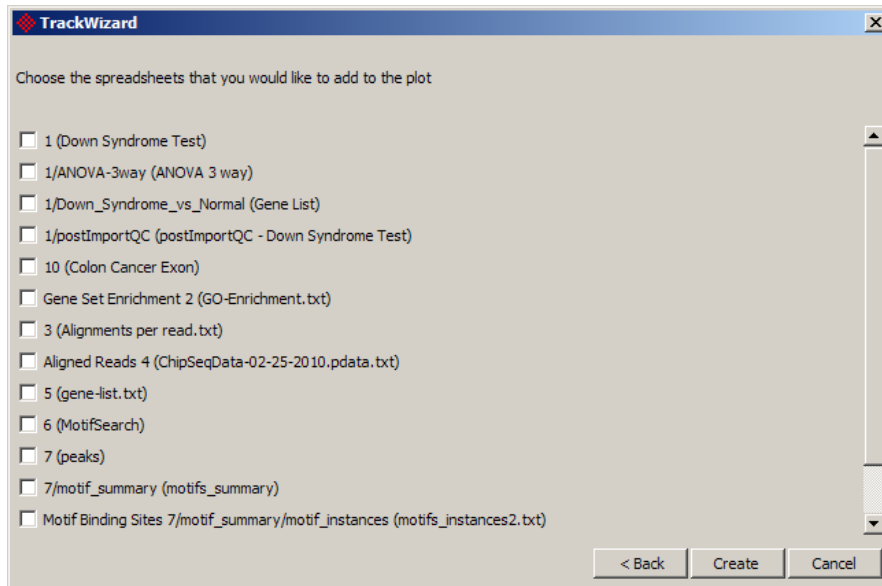
Figure 21: Adding a default track from the list of spreadsheets

> 5) *Add a track with the sequence of the reference genome* to create a track with a known genome such as the human genome 18 (hg18). If one is not available, you will be asked to download or create a .2bit file from fasta files (Figure 22).
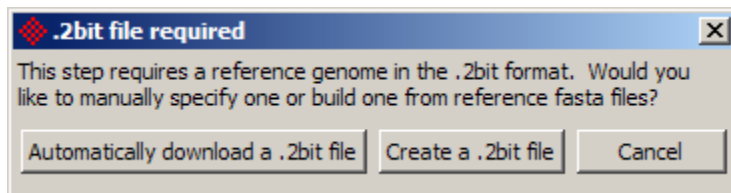

Figure 22: Viewing the .2bit file dialog, Partek Genomics Suite attempts to download a .2bit file for the reference genome track

Reference Genome Track

The *Reference Genome* track displays the individual base pairs of the imported reference genome (Figure 23). To import this track into the view, select *New Track* and then *Add a track with the sequence of the reference genome*. The base pairs will not be visible until the view is zoomed in.

Figure 23: Viewing the base pairs displayed from the hg18 Reference Genome track

The *Track Height* slider can be used to adjust the height of the *Reference Genome* track (Figure 24). Moving the slider to the right will increase the track height, moving the slider to the left will decrease the track height.
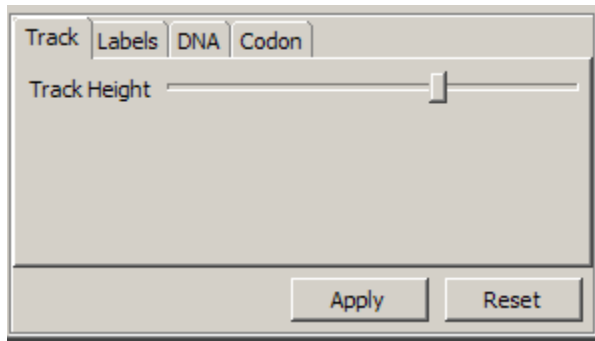
Figure 24: Adjusting the Reference Genome track height

The Reference Genome *Track Label, Label Font Size and Base Font Size* can be adjusted from the *Labels* tab (Figure 25).
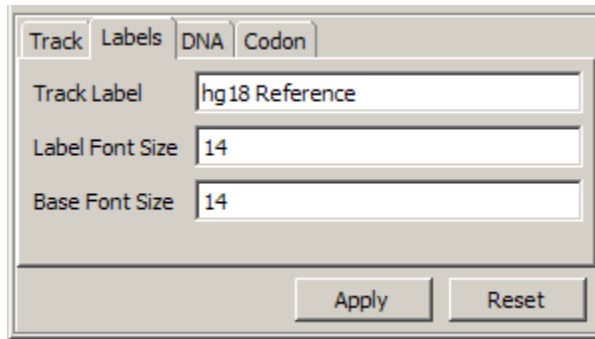

Figure 25: Editing the label properties of the Reference Genome track

Uncheck *Show Bases* from the *DNA* tab to hide the reference sequence. Use this option if you want to only display the codons of the reference sequence (Figure 26). Changing the color here will change base colors on the *Color* tab for the Base Colors Legend. Codons can be displayed to determine if a given mutation results in a change in protein. Select *Configure base colors* to change the colors of the bases from the default colors.
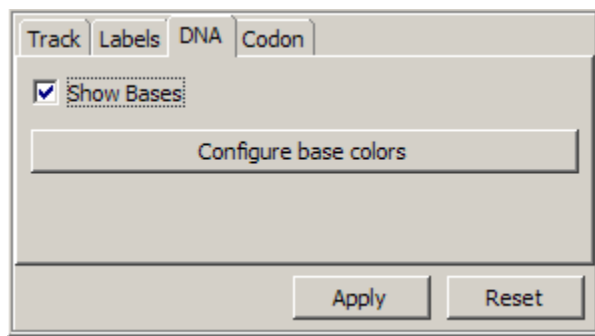

Figure 26: Showing bases & configuring base colors of the Reference Genome track

The *Codon* tab is used to configure how the codons are displayed (Figure 27).
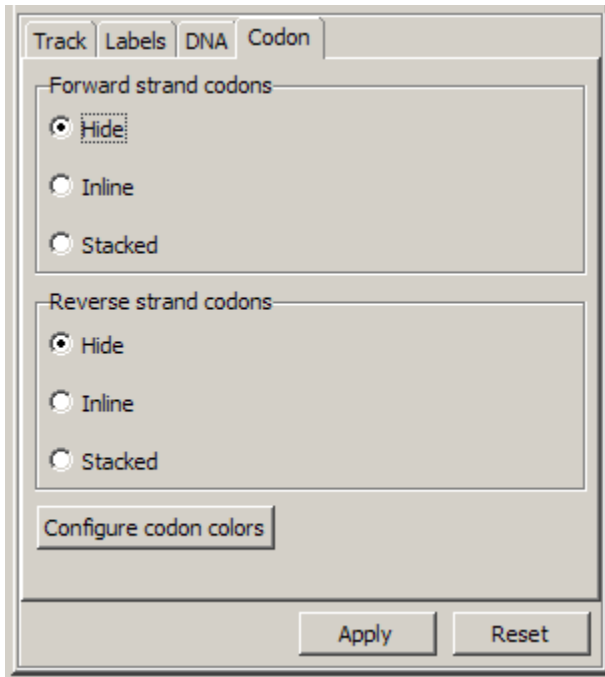
Figure 27: Changing the codon display options of the Reference Genome track

The *Stacked* codon view displays three rows, each row revealing the potential starting base of the codon (Figure 28). Each codon will be drawn over the three (3) bases it covers.
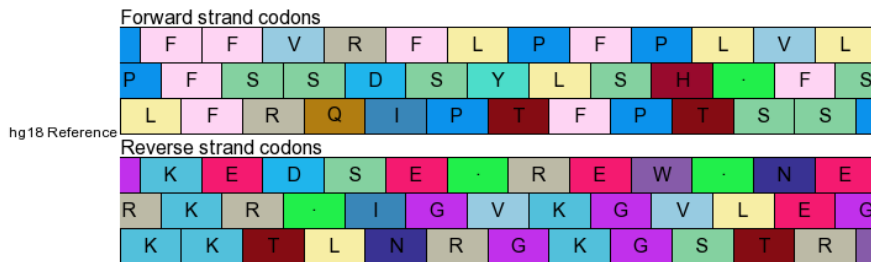

Figure 28: Viewing a Stacked codon view, which is showing forward and reverse strand codons

The *Inline* codon view displays the codon to be transcribed if that base is the first base of three in the codon (Figure 29).
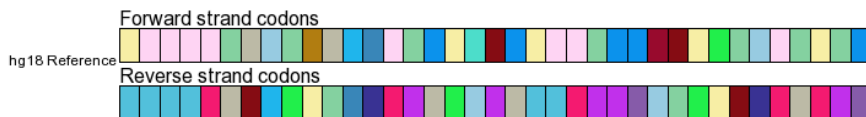

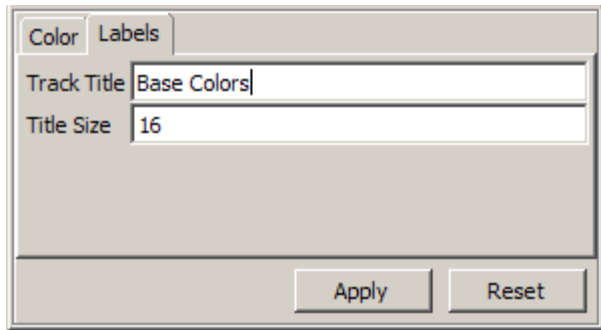*Figure 29: Viewing the Inline codon view, which is showing forward and reverse strand codons*

Selecting the *Configure codon colors* (Figure 29) will open a color palette used to choose a color for each codon (Figure 30).

Figure 30: Viewing the Color palette to change the codon colors

The *Base Color Legend* track displays the color palette being used for the *Reference Genome* track (Figure 31). It can be changed by selecting the *Configure colors* button on the *Color* tab (Figure 32).

Figure 31: Viewing the Legend for base colors of the Reference Genome tracks

Figure 32: Configuring base colors for the Reference Genome and Legend

The *Labels* tab can be used to configure the legend *Track Title* for the base pair colors (Figure 33).

Figure 33: Configuring the label for base colors

6) *Add a track with cytobands* to create a cytoband track. If one is not currently available, Partek GS will attempt to automatically download one.

Cytoband Track

The cytoband displays the chromosomal bands for the current view (Figure 34). The chromosome number is displayed on the left side of the cytoband.

Figure 34: Viewing a Cytoband with a chromosome label

The *Style* tab controls the brightness and labeling of the Cytoband track (Figure 34). Check *Label Cytobands* to display the chromosomal band description below the cytoband (Figure 35).

Figure 35: Viewing a Cytoband with chromosomal band descriptions

The *Center Brightness* slider will adjust the brightness of the Cytoband (Figure 36). Moving the slider to the left will decrease the brightness. Moving the slider to the right will increase the brightness.
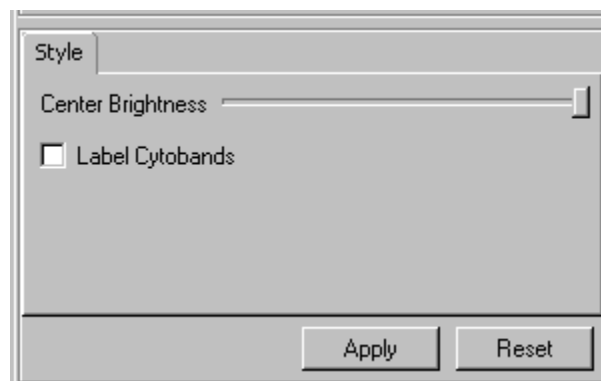

Figure 36: Viewing the Cytoband track properties with Center Brightness to the

right

Genomic Label Track

The *Genomic Label* track displays the relative base scale of the current view (Figure 37). It is typically loaded along with the Cytoband track.

0.0MBps   61.8MBps   123.6MBps   185.4MBps   247.2M

Figure 37: Viewing the Genomic Label Track

The *Text Size* and *Number of tick marks* can be adjusted from the *Axis* tab of the *Genomic Label* track (Figure 38)
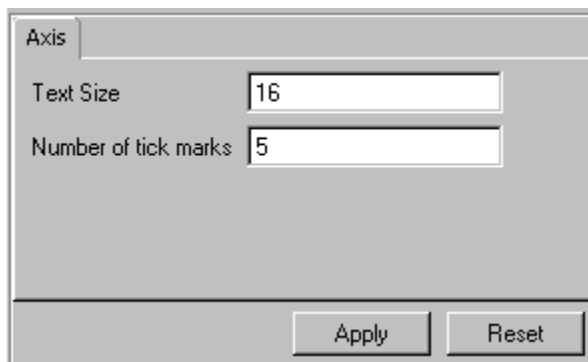

Figure 38: Changing the text size and tick mark frequency of the Genomic Label track

7) *Other(Advanced)* to create a custom track type with specific data. Types of tracks include *Plot Title, Heat Map, Sequence Heat Map, Heat Map summarized by sample attribute, Profile, Annotation, Spreadsheet with genomic regions, Region bar profile, Regions separated by a categorical, Profile split by sample attribute (Category Profile), Color Map,* and *Profile of the difference between two levels of a sample attribute (Difference Profile)* (Figure 39).
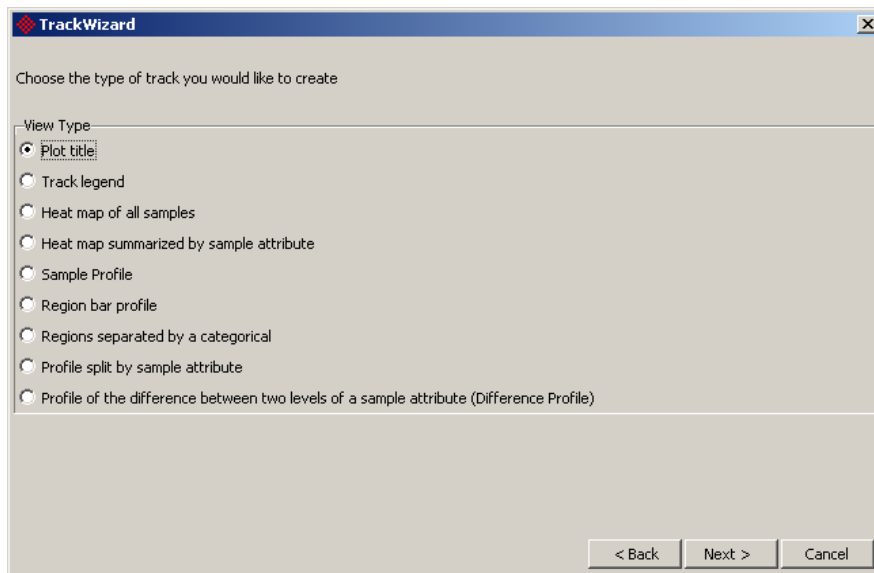
Figure 39: Viewing available track types from the Other(Advanced) option of the New Track Wizard

Plot Title Track

Plot Title option will add a *Title* and *Title Size* to the *Chromosome View* (Figure 40). Select the *Create* button to add the track.
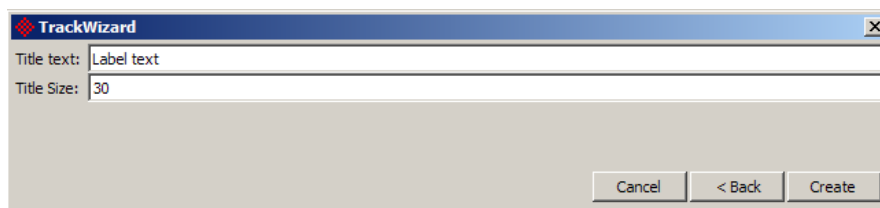

Figure 40: Editing the Title Text and Title Size of the Plot Title track

Even after the *Plot Title* track has been added to the *Chromosome View*, it can still be edited. Figure 41 shows how the labels of the *Title* and *Title Size* can be edited using the *Labels* tab.
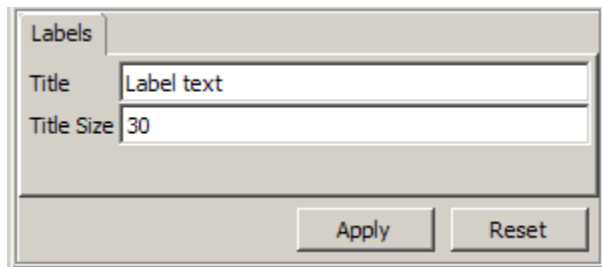

Figure 41: Editing the Title and Title Size properties of the Plot Title track

Heat Map Track

Heat Map option will add a heat map based on the list of available spreadsheet samples. Selecting *Plot Chromosome View* from the workflow or choosing

View>Chromosome View with the sample spreadsheet selected will produce a *Heat Map* (Figure 42).

A *Heat Map* track can be displayed on any spreadsheet that has genomic features (i.e. probesets) on columns. Individual heat maps can be drawn for each sample or for all samples within a selected spreadsheet. The heat map will attempt to display one marker per pixel. If the data is too dense to display and there is more than one marker mapped to a pixel, the color displayed will be the mean of the markers in that pixel.
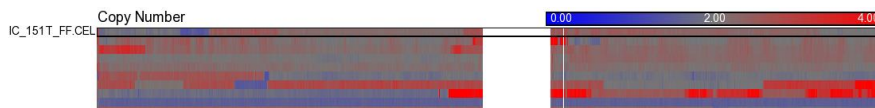

Figure 42: Viewing the Heat Map track; each row corresponds to one sample

Use the *Data* tab to configure the heat map to display all of the samples in the spreadsheet or specify samples to display based on sample attribute (Figure 43). Check the *Smooth Data* checkbox to turn on/off smoothing. Copy Number (and Log2ratio) data is smoothed by default, whereas other data is not. See Chapter 10 of the Partek Documentation for more information on Gaussian smoothing.
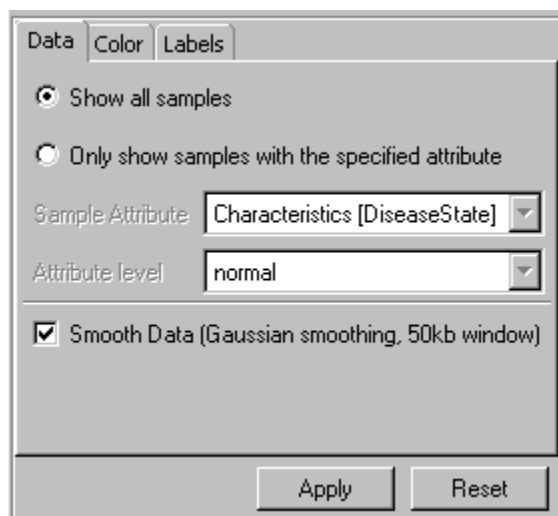

Figure 43: Displaying sample options of the Heat Map track

The *Color* tab can be used to configure the color display of the heat map (Figure 44). The *Min/Max* input options control lets you add the scale for how expression values are displayed corresponding to a given color. The *Min/Max* values are automatically determined by the range of the data.

Figure 44: Configuring the color options of Heat Map track

Use the *Labels* tab to adjust the *Track Title*, *Track Size*, or *Label Size*, and to turn on/off sample labels (Figure 45).
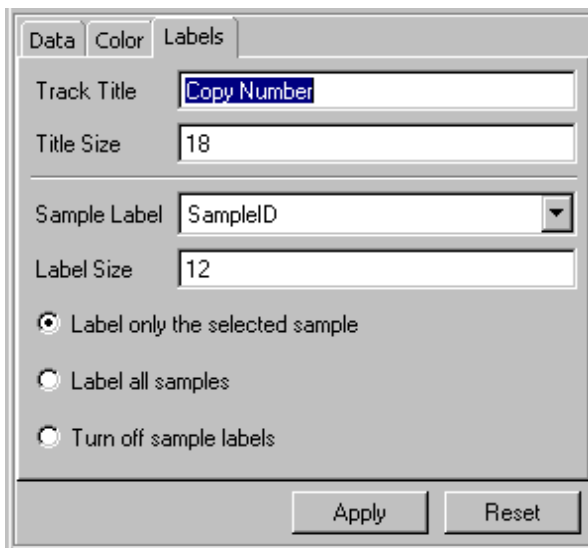

Figure 45: Editing the label properties of heat map

*Add Sequence Heat Map is discussed under **Chip-seq Tracks** under **Workflow Specific Track Description**.

Heat Map Summarized by Sample Attribute Track

The *Heat Map summarized by sample attribute* option displays a heat map track with the expression values summarized across a specified sample attribute (Figure 26).
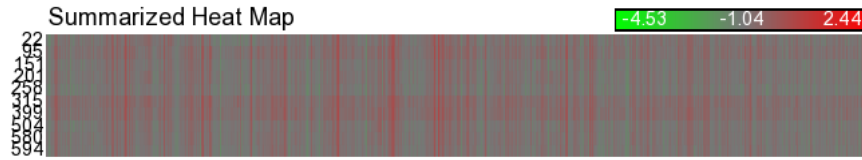
Figure 46: Viewing the Heat Map summarized by sample attribute (Subject ID)

The *Data* tab (Figure 47) lets you control how the *Heat Map* summarized by sample attribute track is to be displayed using the available options of the *Data* tab. The heat map can be changed to display a sample attribute and whether or not smoothing is turn on or off. For more on Gaussian smoothing, please see Chapter 10 of the Partek Documentation.
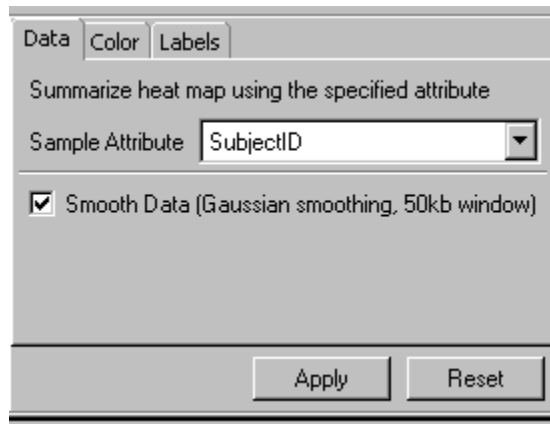

Figure 47: Viewing the Data tab properties of the Heat Map summarized by sample attribute track

The color of the *Heat Map* summarized by the sample attribute can be adjusted using options on the *Color* tab (Figure 48). The *Min/Max* expression intensity scale can be changed as well as the color range.
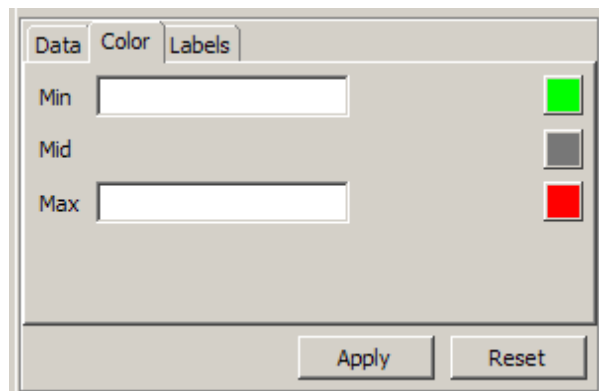

Figure 48: Viewing the Color tab of the Heat Map summarized by sample attribute

The *Labels* tab is used to edit the values of the *Track Title, Title Size,* and *Label Size* of the *Heat Map* summarized by sample attribute (Figure 49).
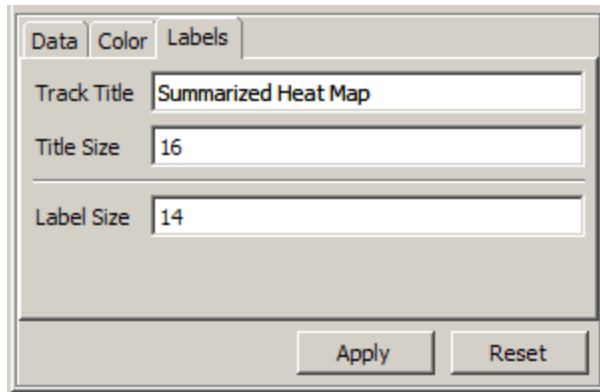
Figure 49: Editing the Track Title, Title Size, and Label Size of the Heat Map summarized by sample attribute

Profile track is discussed under **Copy Number tracks** under **Workflow Specific Track Description.**

Profile Split by sample attribute description can be found under the **Category Profile** > **Gene/Exon Expression tracks** under **Workflow Specific Tracks**.

Legend Track

The *Legend* option allows you to add a descriptive color legend for most of the common tracks (Figure 50). Choose the type of legend you would like to display, add a title to the legend using the *Display name of Legend* input box, and select *Create*.
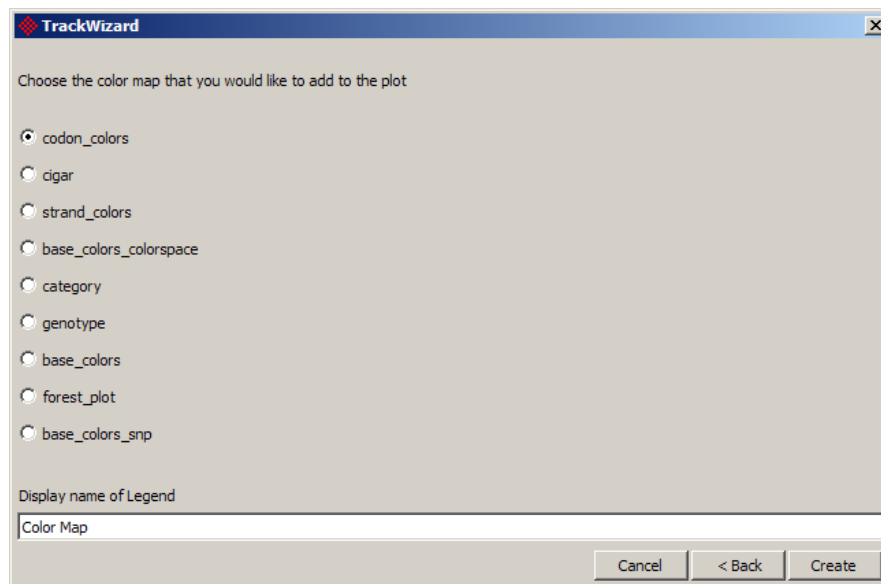

Figure 50: Adding a Color Map track to the Chromosome View

The colors of the *Color Map* track can be changed using the *Color* tab (Figure 51). Select the *Configure colors* button to adjust the way the properties of the *Color Map* are displayed.
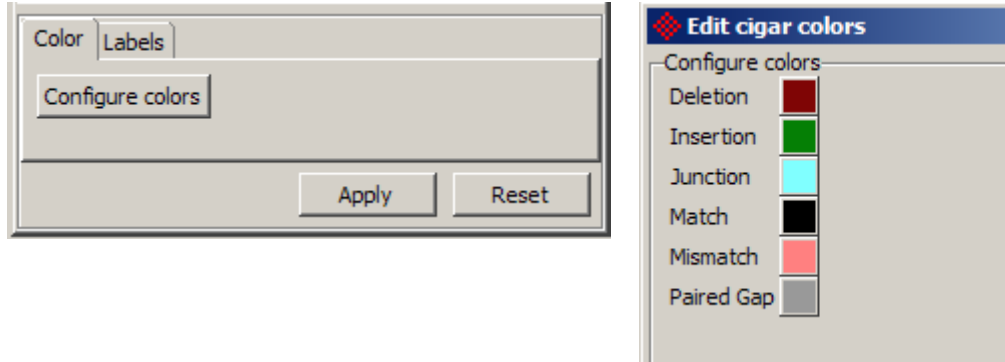
Figure 51: Configuring the colors of the cigar Color Map

The *Labels* tab of the *Color Map* is used to edit the *Track Title* and *Title Size* of the *Color Map* (Figure 52).
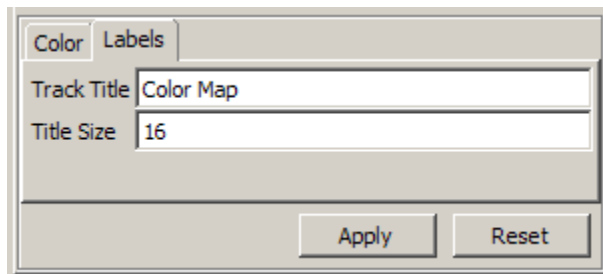

Figure 52: Editing the Track Title and Title Size of the Color Map track

*Profile of the difference between two levels of a sample attribute description can be found under **Difference Profile** > **Gene/Exon Expression Tracks** under **Workflow Specific Track Description.**

Workflow Specific Tracks

The Workflow Specific Tracks section explores tracks that are generally most appropriate for Copy Number, Gene/Exon Expression and Next Generation Sequencing workflows, although not exclusive to those workflows.

Copy Number Tracks

The *Chromosome View* can display copy number information in different ways including amplifications, deletions and individual intensity values for region or whole chromosome views, allele specific copy number, and Loss of Heterozygosity.
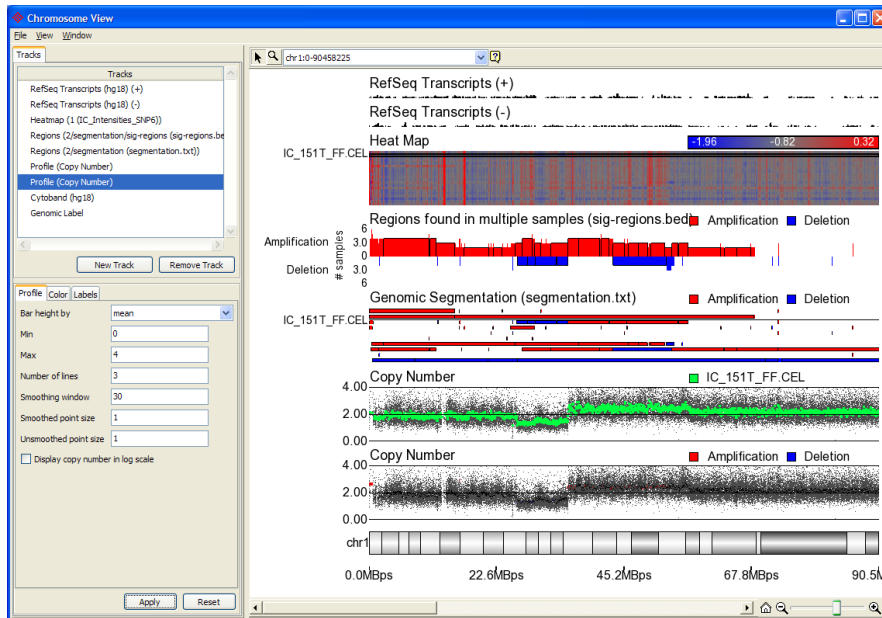
Figure 53: Viewing the Chromosome View featuring Copy Number data

Sig-regions Histogram Track

The result of *Find Regions in Multiple Samples* from the Copy Number workflow will be displayed as a histogram track in the *Chromosome View*. Amplifications extend above the center; deletions extend below the center (Figure 54).
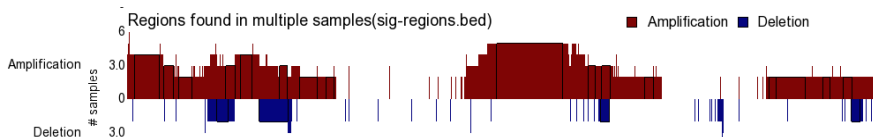

Figure 54: Viewing the Histogram bar height by # Samples (Y-axis)

From the *Profile* tab, check *Separate bars by* to separate the bars by the available attributes in the sig-regions spreadsheet (Figure 55). The histogram bar height is determined by the selected attribute in the *bar height by* drop down menu. The default selection is the *Copy Number* column attribute with the bar height as *# samples*. The *Min* and *Max* values set the Y-axis scale of the track. The *Bars come from* feature sets the baseline from which to extend the histogram height.
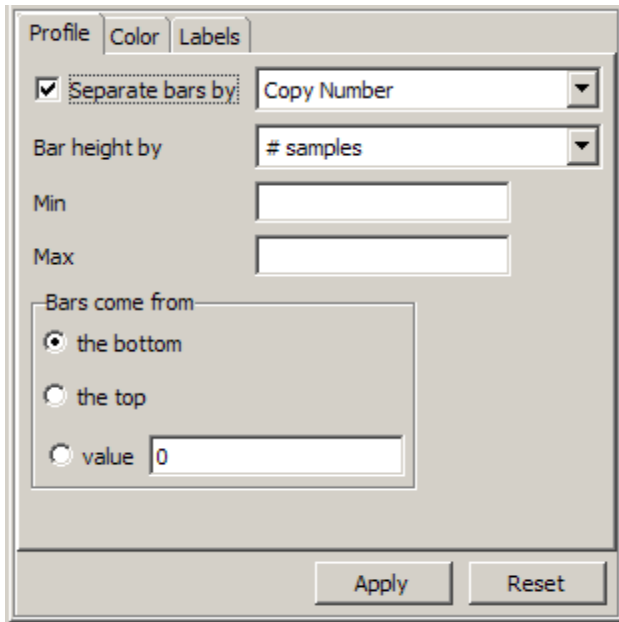
Figure 55: Configuring the sig-regions histogram

Color bars by will determine which attribute is used to color the histogram bars. Select *Configure category colors* to change the colors of the histogram bars by attribute (Figure 56). The default value is Copy Number with amplifications drawn in <u>Red</u> and deletions drawn in <u>Blue</u>. The *Min* and *Max* values set the color range intensity values but will only be noticeable if specific attributes are selected for the *Color bars by* drop down menu.
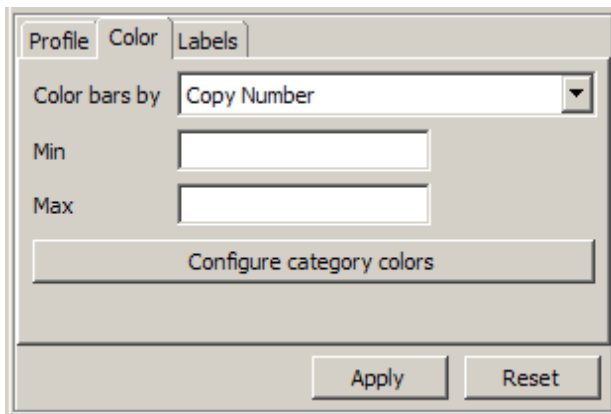
Figure 56: Configuring the Color sig-regions histogram by column attribute

The *Track Title*, *Title Size* and *Label Size* can be changed from the *Labels* tab (Figure 57).
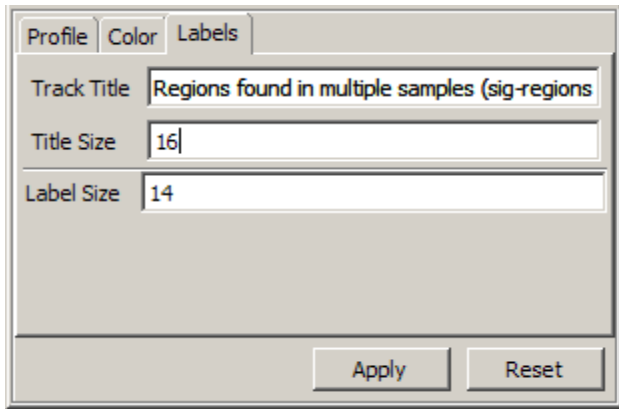
Figure 57: Editing label properties of the sig-regions histogram

Segmentation Track

The *Segmentation* track displays regions of copy number variation. The results are bars plotted with Amplifications and Deletions lengths by sample (Figure 58). By default, one row will be drawn for each sample.
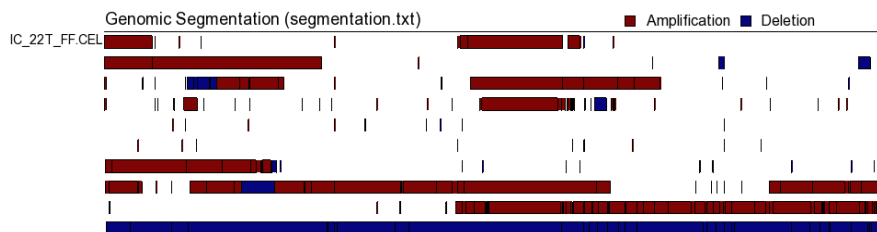

Figure 58: Viewing the Segmentation results; Amplification=red, Deletion=blue

The *Segmentation* track can be displayed by attributes of the spreadsheet. Use the dropdown menu of the *Profile* tab to select bar separation (Figure 59). The *Track Height* can be adjusted using the track height slider. Moving the slider to the right will increase the track height, moving the slider to the left will decrease the track height.

Figure 59: Configuring the Segmentation track display

The *Color* tab gives selections to color the bars. Select **Configure category colors** to change the colors of the bars by attribute (Figure 60). The default value is Copy Number with amplifications drawn in <u>Red</u> and deletions drawn in <u>Blue</u>. The *Min* and *Max* values set the color range intensity values but will only be noticeable if specific attributes are selected for the *Color bars by* drop down menu.
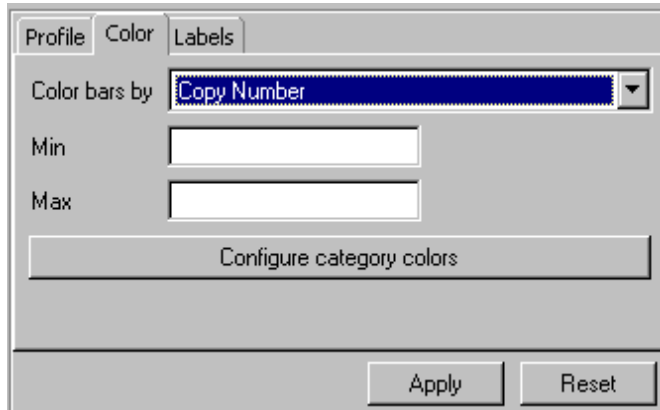

Figure 60: Coloring the histogram by attribute of the Segmentation track

The *Track Title*, *Title Size* and *Font Size* can be changed from the *Labels* tab. By default, the first sample in the spreadsheet is selected (Figure 61). Deselect *Label on the selected sample* to label all the samples.
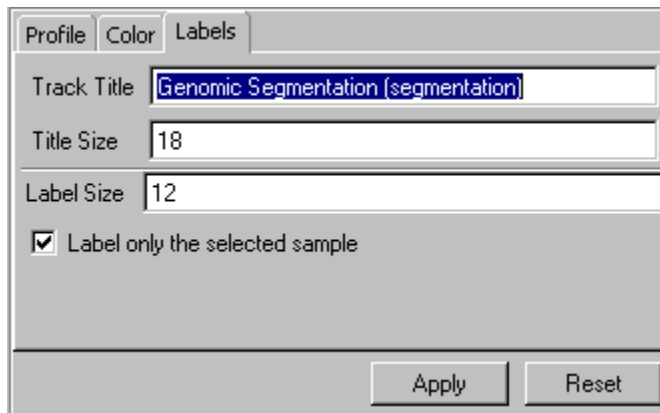

Figure 61: Editing the Track Title and Label Size of the Segmentation track

Profile Track

The *Profile* track displays the expression of individual markers in smoothed and unsmoothed form. The position of the smoothed points is based on the median of the points within a *Smoothing window*. A profile track can be created for each sample, or samples can be displayed overlapping each other. Figure 62 shows a copy number profile for one (1) sample.
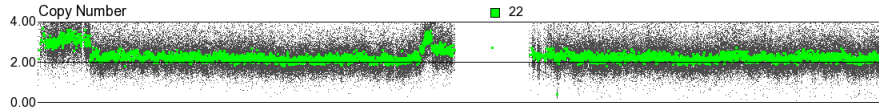
Figure 62: Viewing the Profile track of smoothed copy number sample

Change the *Y-axis* tab configuration options to set the *Min/Max* values of the Y-axis scale, *Number of grid lines*, *Smoothing window* size, *Smoothed point size*, *Unsmoothed point size*, or *Display the copy number profile in log scale* (Figure 63).
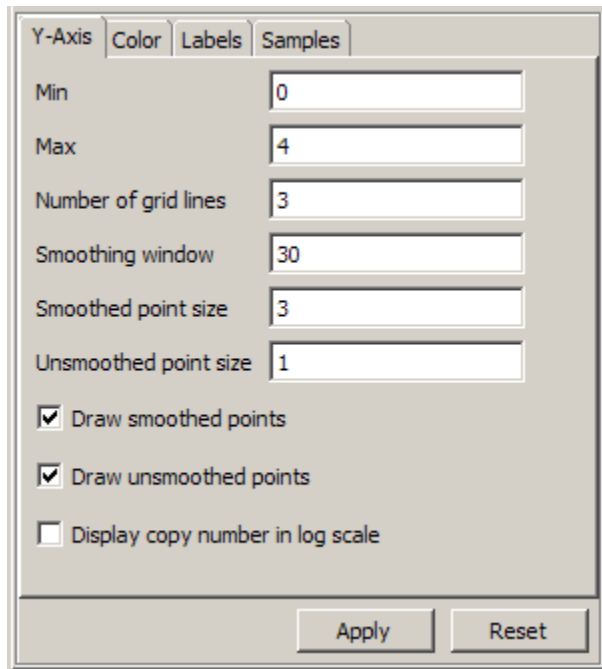

Figure 63: Configuring the Profile track plot options

The *Color* tab can be configured to change the how the colors of the smoothed & unsmoothed points are displayed (Figure 64). Select *Configure category colors* to change how the points are colored.
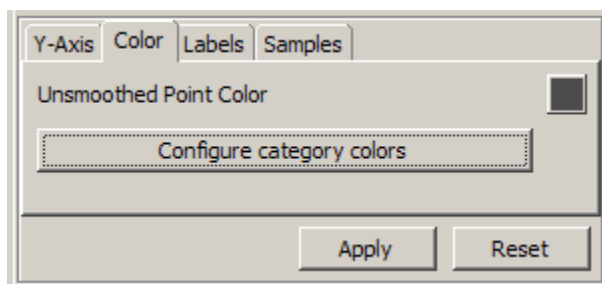

Figure 64: Configuring the color properties of Profile track

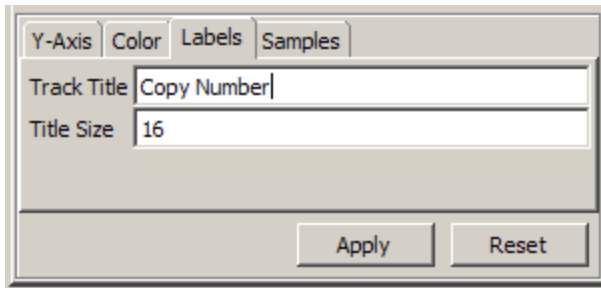The *Labels* tab allows you to edit the *Track Title* and *Title Size* (Figure 65).

Figure 65: Editing Track Title, track Title Size of the Profile track

The *Samples* tab lets you choose which samples are to be displayed on the selected *Profile* track (Figure 66). Samples can be displayed individually or can overlap with each other. Selecting the *Set Samples* button will prompt the same dialog as Figure 20.
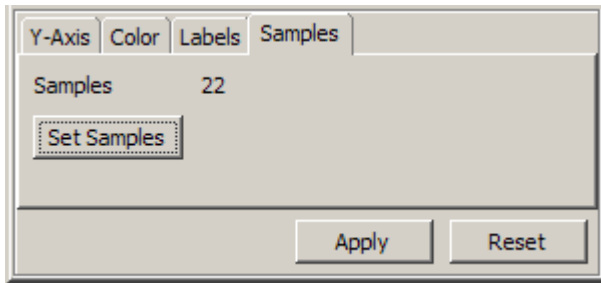

Figure 66: Viewing the Samples tab of the Profile track

The *Profile* track displays the heat map sample selected by the viewer. Figure 67 below shows a highlighted sample on the heat map with a corresponding copy number *Profile* track updated with the sample profile. Other sample profiles can be added (below) but will not be updated with the selection of samples on the heat map.
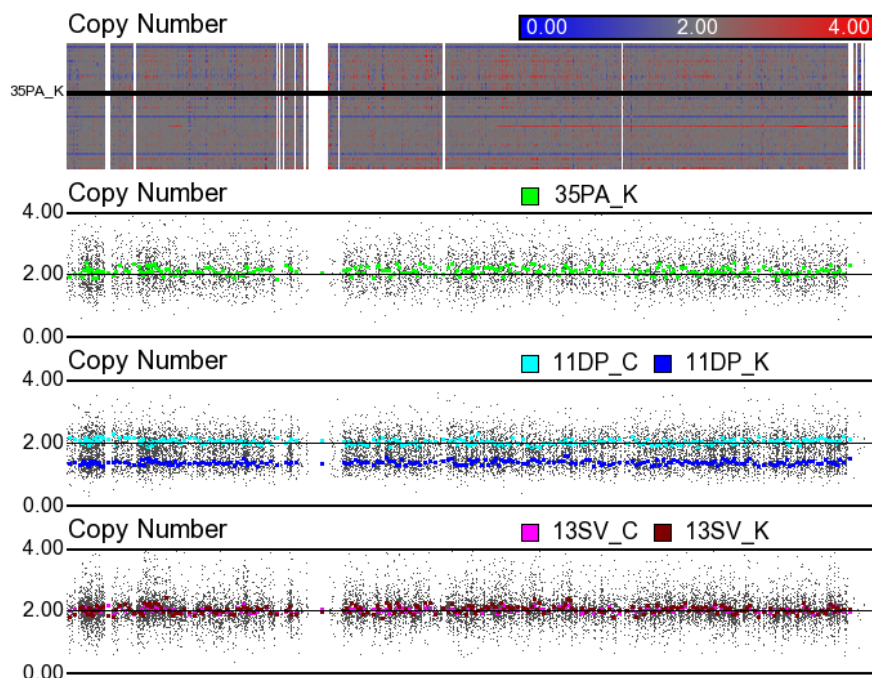
Figure 67: Viewing the heat map that has sample 35PA_K selected. The middle track has manually selected the normal and tumor samples for subject 11DP (the bottom track for subject 13SV). These two tracks will not change as samples are selected in the heat map

Gene/Exon Expression Tracks

Gene Expression and *Exon* tracks can be added in the *Chromosome View* to visualize up and down regulation of genes, alternative splice events, expression values by categorical attributes, difference profiles between categorical attributes, fold change, p-values, and more (Figure 68).
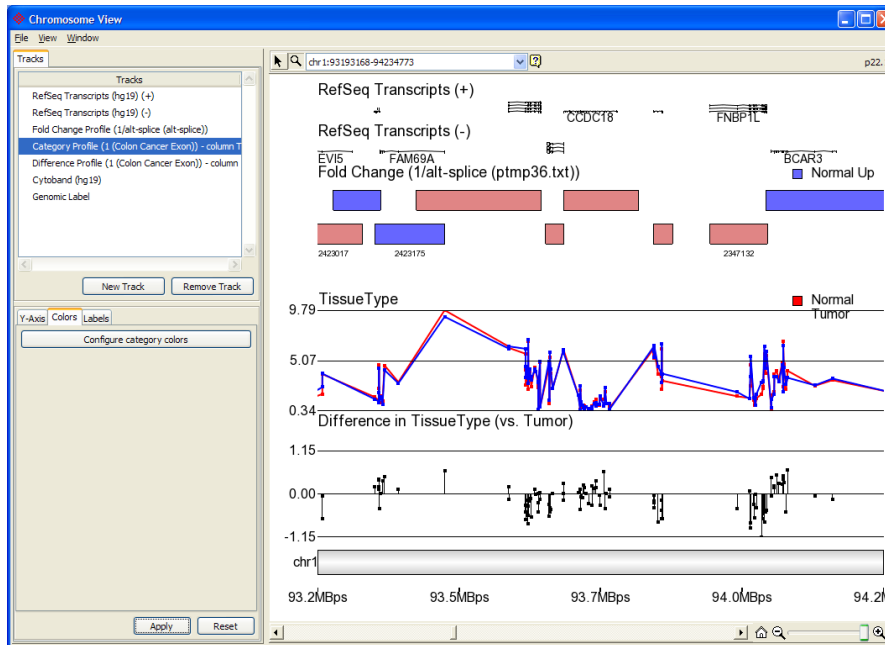
Figure 68: Viewing the Gene/Exon Expression tacks in the Chromosome View

Category Profile Track

A *Category Profile* can be created to display the average expression values across samples at a given probeset to look for possible up/down regulation of genes and alternative splice events. The category profile in Figure 69 shows average expression values of Normal vs. Tumor Tissue.
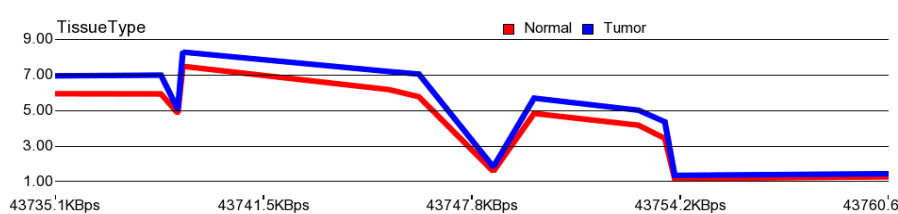

Figure 69: Viewing the Category Profile of Normal vs. Tumor samples

The *Y-Axis* tab is used to configure the track plotting properties (Figure 70). The *Column* drop down menu allows you to choose the column to display the average value of the samples. The *Min & Max* variables set the scale of the track. The position of points is determined by the average value of samples in the same level of a categorical variable. If Min and Max are blank then the y-axis range is automatically set to the range of points within the view. The range of the plot can be manually specified by entering Min and Max values. The grid line increment values will be determined by the difference in the Max & Min divided by the number of grid lines.

The *Smoothing window* option will change the way the probe sets for the selected sample at each location is displayed. This option specifies a window of probe sets to smooth. For every group of probe sets there will be one point drawn based on the

median of the probe sets in the window. There is no overlap between windows. The *Smoothed point size* and *Unsmoothed point size* determine how do display the individual probes. By default for better visualization, the smoothed probes will have a larger point size than the unsmoothed point size. The width of the lines can be increased or decreased using the *Line Width* slider.
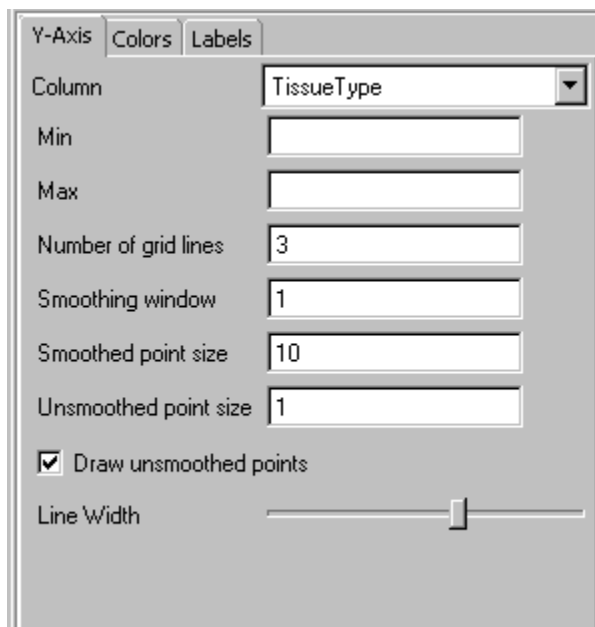

Figure 70: Editing the Y-axis tab of the Category Profile

The *Colors* tab (Figure 71) can be used to configure the color of the *Column* values in Figure 70.
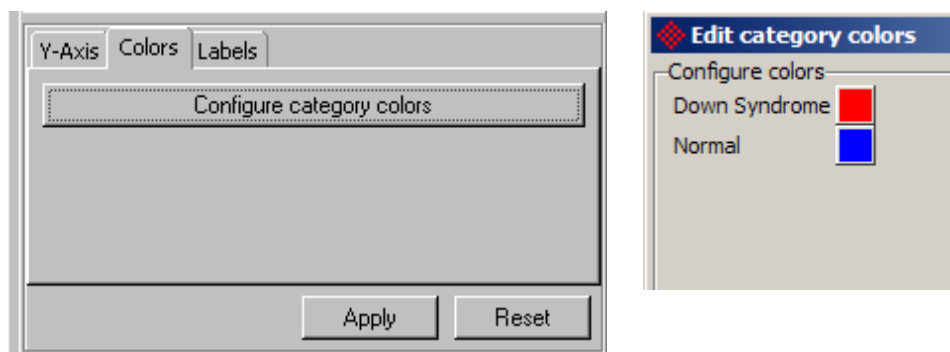
Figure 71: Configuring the category colors from Colors tab

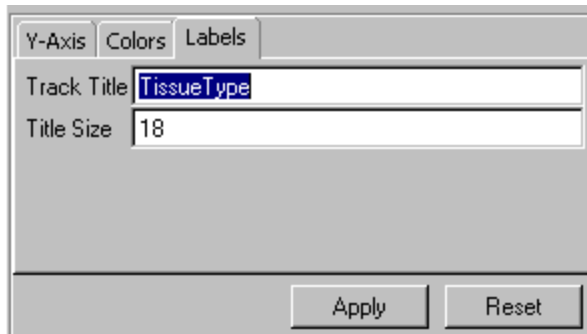The *Labels* tab lets you edit the *Track Title* and *Title Size* (Figure 72).

Figure 72: Editing Track Title and Track Title size of the Category Profile

Difference Profile Track

The *Difference Profile* track displays the difference of the average expression values between a selected attribute category. Figure 73 show the *Difference Profile* between Normal and Tumor average expression values. It can be added by selecting *New Track > Other(Advanced) > Profile of the difference between two levels of a sample attribute(Difference Profile)*.
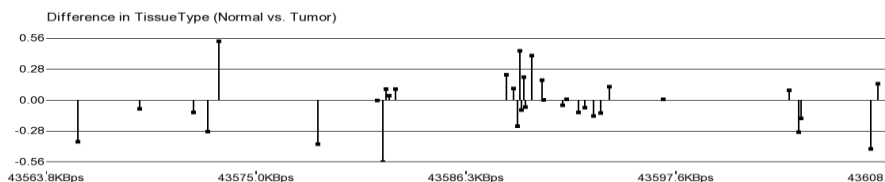

Figure 73: Viewing the Difference Profile between Normal vs. Tumor

Configure the plot properties of the *Difference Profile* using the *Y-Axis* tab (Figure 74). Check the *Split by* checkbox to view differences grouped by levels of a categorical variable. The length is determined by the average value (*Baseline level*) samples subtracted from the average value of the other samples (non-*Baseline level*). The *Compare attribute* drop down menu gives options for which categorical variables to display. The *Baseline level* is the level that comes from the categorical variable specified in *Compare attribute*. The *Min & Max* variables set the scale of the track. If Min and Max are blank then the y-axis range is automatically set to the range of points within the view. The grid line increment values will be determined by:

$$\frac{(Max - Min)}{Number\ of\ grid\ lines}$$

The *Smoothing window* option will change the way the probe sets for the selected sample at each location are displayed. This option specifies a window of probe sets to smooth. For every group of probe sets there will be one point drawn based on the <u>median</u> of the probe sets in the window. There is no overlap between windows.
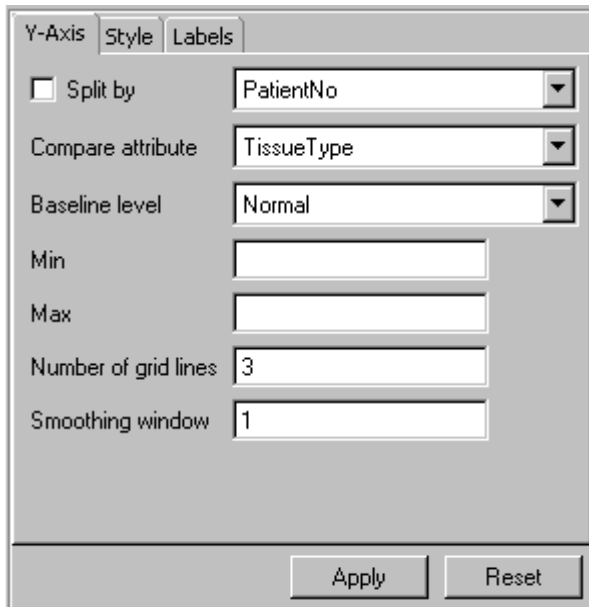
Figure 74: Editing the Y-axis tab of the Difference Profile

Use the *Split by* option to view differences grouped by levels of a categorical variable, determining which *Compare attribute*, and which of the attributes to be chosen as the *Baseline level* (Figure 75).
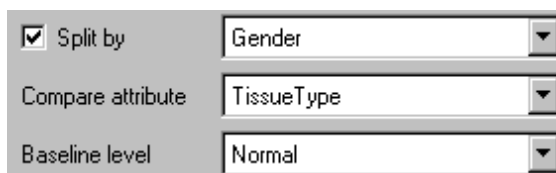

Figure 75: Configuring the Split by option for the Difference Profile track

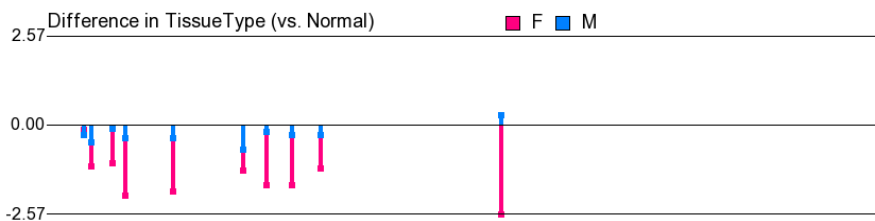Figure 76 shows a *Difference* profile *Split by* Gender, comparing *Normal vs. Tumor TissueType*.

Figure 76: Viewing the Difference profile Split by Gender comparing Normal vs Tumor TissueType. One line for [Avg(Female Tumor) – Avg(Female Normal)]-pink and one line for [Avg(Male Tumor) – Avg(Male Normal)]-blue

Use the *Style* tab to change the way the points are plotted (Figure 77). By default, *Draw line to zero* is checked. Uncheck to view just the points. Use the *Point Size* slider to increase or decrease the size of the points. Check *Connect points* to draw a

line connecting every difference point. Use the *Line Width* slider to increase or decrease the width of the lines.
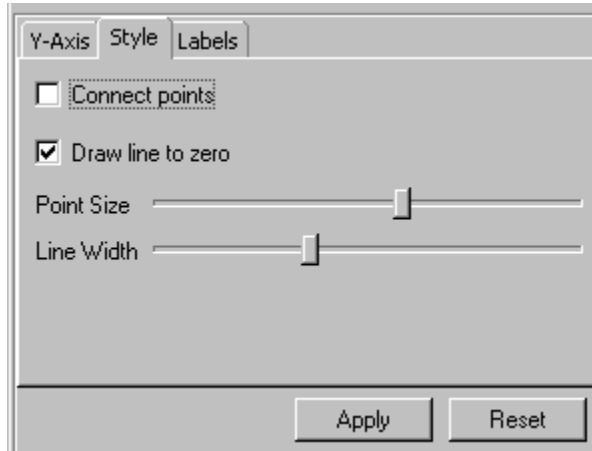

Figure 77: Adjusting the Style tab of the Difference profile

The *Labels* tab lets you edit the *Track Title* and *Title Size* (Figure 78).
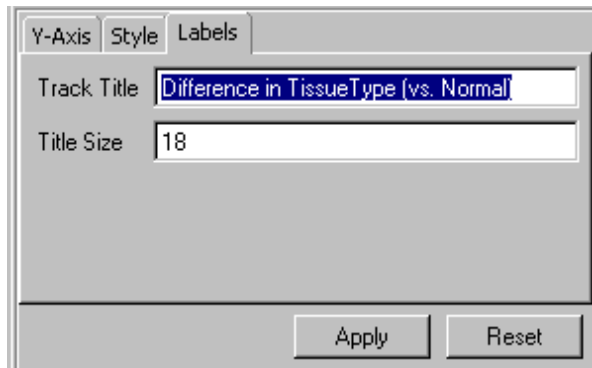

Figure 78: Editing the Track Title and Track Title size of the Difference Profile

Fold Change Profile Track

The *Fold Change Profile* track displays gene or exon regions of up/down & significantly up/down differential expression as displayed by fold change value (Figure 79).

Figure 79: Viewing the Fold change profile track of Normal sample up/down differential expression

Select the *Profile* tab to select the *Factor of interest* to display (Figure 80). All columns with a "FoldChange (Factor)" label format will be listed as available options. Use the *Track Height* slider to increase or decrease the height of the track in the view.
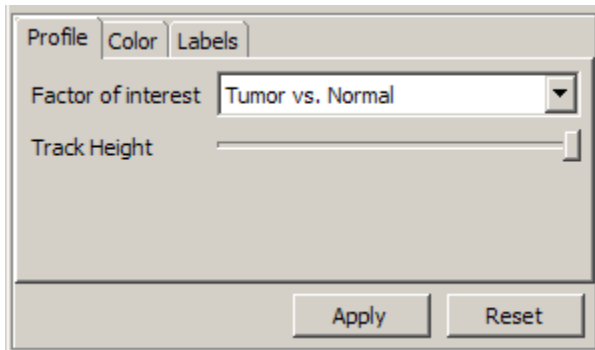
Figure 80: Changing the Fold Change Profile display Factor and Track Height

Select the *Color* tab to choose the color for the positive and negative fold change values. The track can be configured to draw markers a darker color that pass a certain threshold (Figure 81).
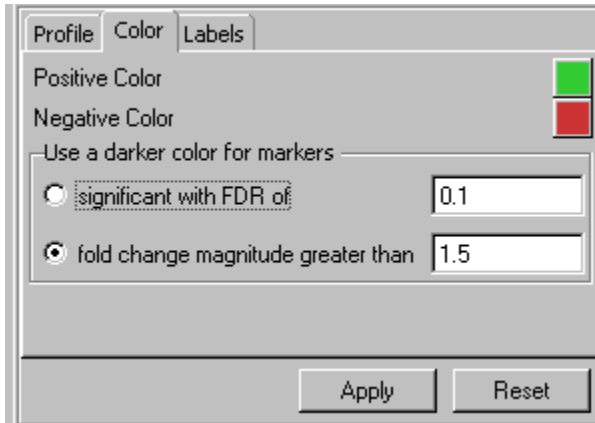

Figure 81: Changing the default colors and by threshold

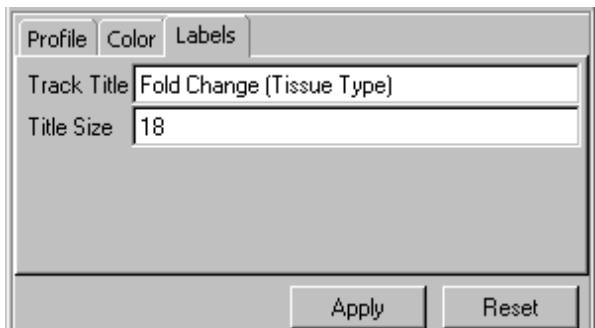Edit the *Track Title* and *Title Size* properties in the *Labels* tab (Figure 82).

Figure 82: Editing the Track Title label and Title Size of the Fold Change Profile

Smoothed Fold Change Track

The *Smoothed Fold Change* track provides visualization for viewing the distribution of nearest markers with positive or negative fold change in each direction (Figure 83).
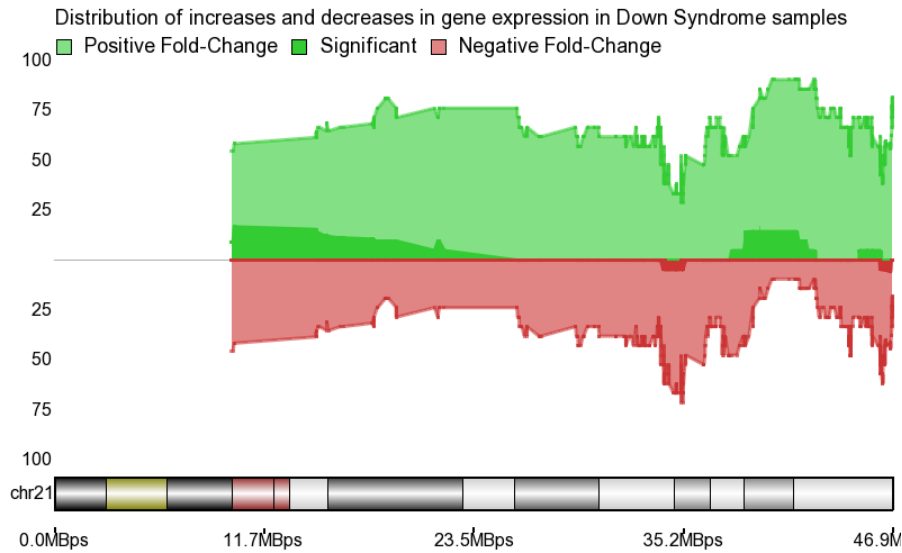

Figure 83: Viewing the Smoothed Fold Change plot of Down Syndrome samples

Use the *Y-Axis* tab to configure the display of the *Smooth Fold Change* track (Figure 84). All columns with a "FoldChange (Factor)" format will be listed as available options. The input box for the *Threshold* will determine the cutoff of significant fold change. The *Min & Max* will set the scale for plot of distribution values. By default, the *Min & Max* values are set to -100 and 100. The *Nearest markers/Base Pairs* values can be chosen to extend or shorten the nearest marker or base pair length. The default setting is the 10 nearest markers in each direction.
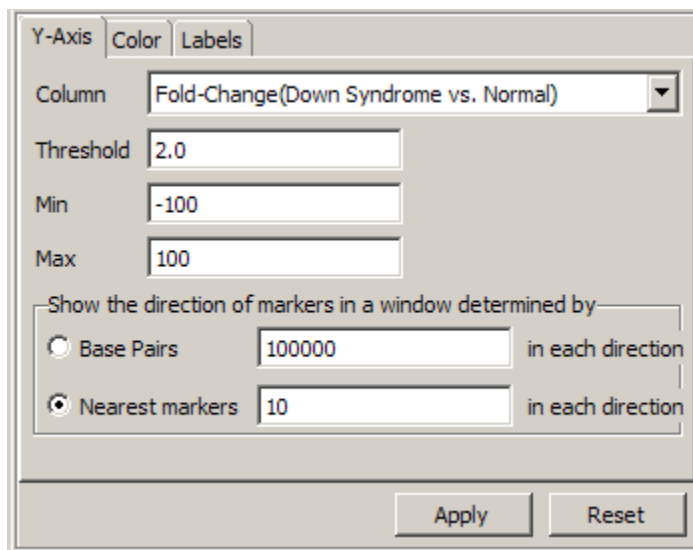

Figure 84: Adjusting plot properties of the Smoothed Fold Change track

Figure 85 shows the *Color* tab for changing the colors used to display the *Positive* and *Negative* fold change distributions.
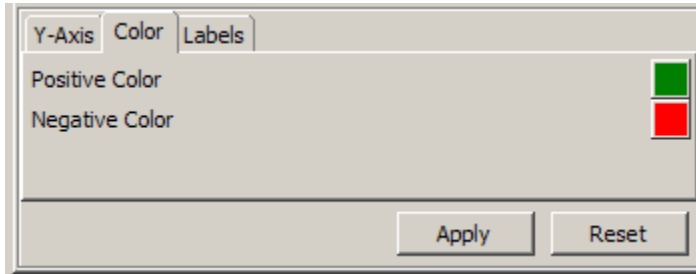

Figure 85: Changing the default colors for Positive/Negative fold change

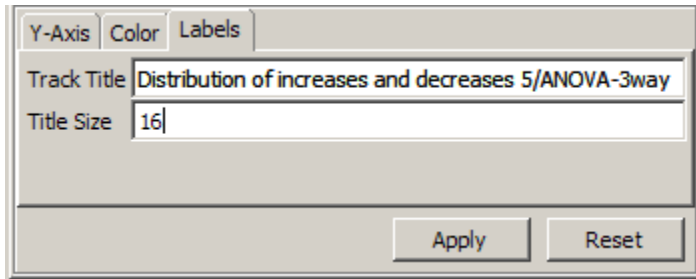Edit the *Track Title* and *Title Size* properties in the *Labels* tab (Figure 86).

Figure 86: Editing the Track Title and Track Title Size of Smoothed Fold Change track

The smooth value at each marker is displayed as a percentage and calculated by:

$$[\frac{g}{2m + 1}] * 100$$

where *g* is the number of nearest markers with a) positive fold change or b) negative fold change; *m* is the number of nearest markers in each direction (default 10); 1 includes the current marker; *100 to turn into percentage.

The significant positive or negative fold change is calculated by dividing the number of significantly a) positive or significantly b) negative markers by the total number of nearest markers (including current marker), as such:

$$[\frac{s}{2m + 1}] * 100$$

where *s* is the number of nearest markers with significantly a) positive fold change or significantly b) negative fold change; *m* is the number of nearest markers in each direction (default 10); 1 includes the current marker; *100 to turn into percentage.

The significant marker frequency is determined by the *Threshold* on the *Y-Axis* tab and is indicated by middle darker section of *Smoothed Fold Change* track.

p-value Profile Track

The *p-value Profile* track displays the p-values at markers for a specified category (Figure 87).

Figure 87: Viewing the p-value Profile track by TissueType

All columns with a "p-value(Factor)" label format will be listed as available options for the *Factor of interest*. Adjust the *Track Height* using the track slider (Figure 88).
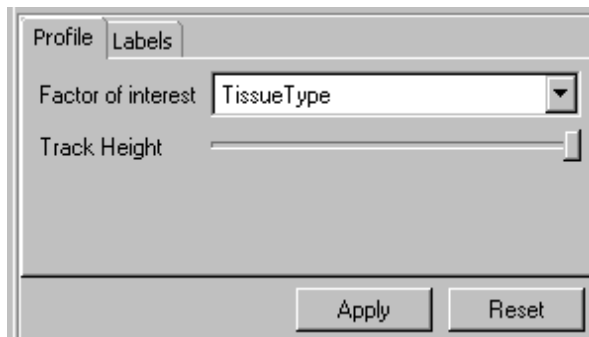

Figure 88: Adjusting the Profile properties of the p-value Profile

Edit the *Track Title* and *Title Size* properties in the *Labels* tab (Figure 89).
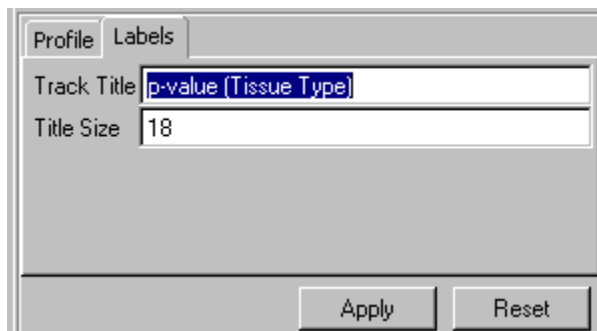
Figure 89: Editing Track Title and Track Title Size of P-value Profile

Correlation Profile Track

The *Correlation Profile* track provides visualization for viewing the distribution of nearest markers with positive or negative correlation in each direction (Figure 90).
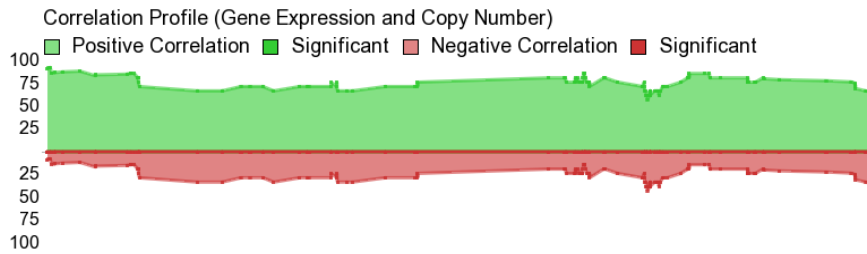
Figure 90: Viewing the Correlation Profile track

The result of the *Correlate Copy Number with Gene Expression* workflow feature plots the *Correlation Profile* track. This plot has similar window parameters as the *Smoothed Fold Change* plot, but instead of fold change, it displays positive and negative correlation. The *Y-Axis* tab is used to change the plotting parameters (Figure 91). The *Column* options permit the display of the linear or rank correlation. The *Threshold* value sets the cutoff for significant positive or negative correlation. The *Min & Max* will set the scale for plot of distribution values. By default, the *Min & Max* values are set to -100 and 100. The *Nearest markers/Base Pairs* values can be chosen to extend or shorten the nearest marker or base pair length. The default setting is the 10 nearest markers in each direction.

See the *Smoothed Fold Change* track section for description of how distribution plot values are calculated.
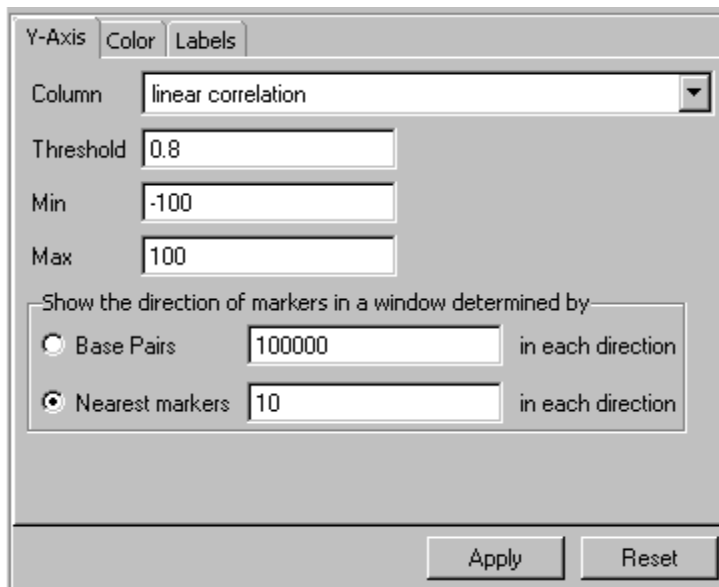

Figure 91: Adjusting plot properties of the Correlation Profile track

The size of the dark section in the middle of the plot is determined by the percentage of correlation values that pass the *Threshold* parameter (Figure 92).
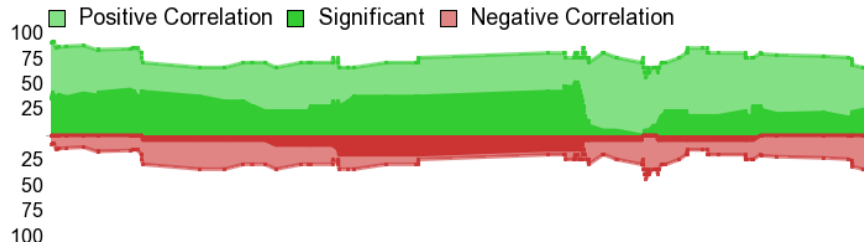
Figure 92: Viewing the Correlation Profile track showing up/down and significant up/down correlation

Figure 93 shows the *Color* tab for changing the colors used to display the *Positive* and *Negative* fold change distributions.

Figure 93: Changing the default colors for Positive/Negative values of the Correlation Profile track

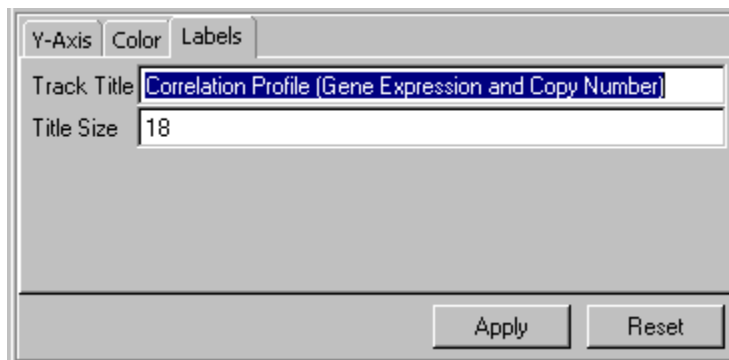Edit the *Track Title* and *Title Size* properties in the *Labels* tab (Figure 94).

Figure 94: Editing the Track Title and track Title Size of Correlation Profile track

Next Generation Sequencing Tracks

RNA-Seq Tracks

RNA-Seq tracks can be added in the *Chromosome View* to visualize mapped read counts along with gene annotation information, cytobands, SNP proportions to find

base pair changes, and isoform proportions locations to look for alternative splicing of genes across the transcriptome (Figure 95).

Figure 95: Viewing the RNA-seq tracks in Chromosome View

Isoform Proportion Track

The *Isoform Proportion* track displays the mapped reads to transcripts and helps visualize differential expression and alternative splicing. The size of each transcript is proportional to the number of reads that map to the transcript. The color indicates the samples for which the reads belong. Figure 96 shows heart and muscle primarily express in NM_005888. Brain and liver primarily express in NM_002635.
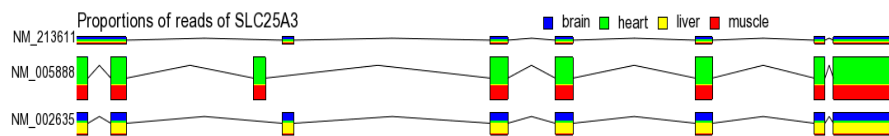
Figure 96: Viewing the Isoform Proportion track showing reads mapped to transcripts

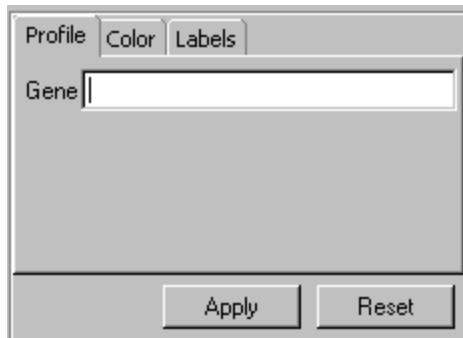The gene symbol can be manually set on the *Profile* tab (Figure 97).

Figure 97: Manually setting the gene symbol on the Profile tab of the Isoform Proportion track

Configure the colors of the *Isoform Proportion* track in the *Color* tab (Figure 98).
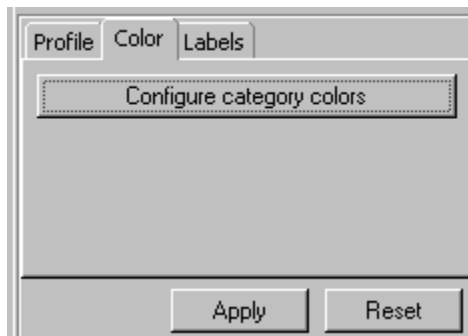

Figure 98: Configuring the category colors from Colors tab

Edit the *Track Title*, *Track Title Size* and *Label Size* of the *Isoform Proportion* track using the *Labels* tab (Figure 99).
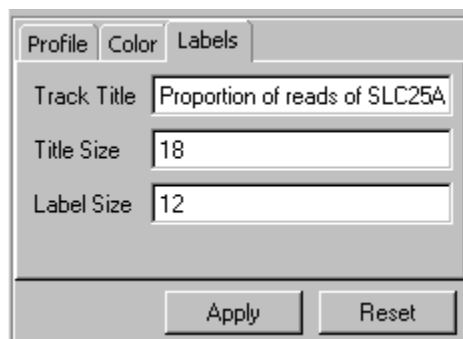

Figure 99: Editing the Track Title, Track Title Size and Label Size of Isoform Proportion track

Alignment Track

The *Alignment* track displays a view of the number of alignments per read (Figure 100). Each alignment track is displayed as an individual sample. By default, the *Histogram* view is displayed with the *alignments colored by sample*.
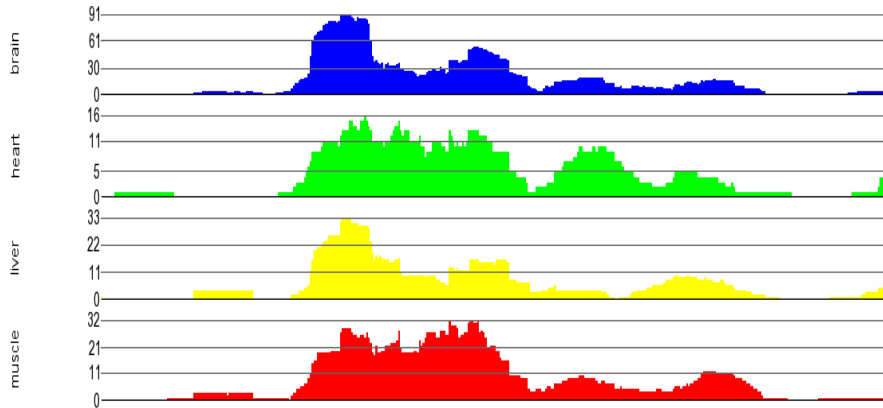
Figure 100: Viewing the Alignment track of RNA-Seq data with Histogram view. The Histogram view is most appropriate for dense data

The *Cigar* track gives a legend of how matches are colored (Figure 101). These include locations where Deletions, *Insertions, Junction, Matches, Mismatches* and *Paired Gaps* occur.

Figure 101: Viewing the Cigar track displayed by default with colors

The *Style* tab of the *Alignment* track controls the way the alignments are displayed (Figure 102). The *Track Height* slider adjusts the height of the track. Moving the slider to the right will increase the track height, moving the slider to the left will decrease the track height. The reads can be displayed as *One per row*, *Fewest number of rows*, and *Histogram*.

The *Histogram* display can be adjusted to have a Maximum Y-Axis scale using the input box for the *Y-Axis Maximum.* Leave this blank to have the maximum automatically determined by the range of the data. From the *Style* tab, the color options for the alignment track can be changed.

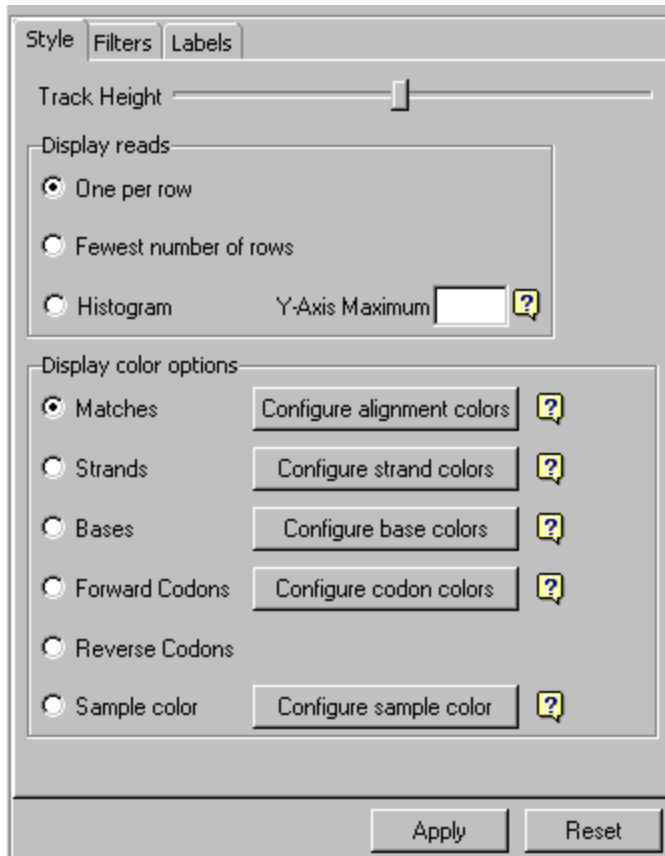The *Histogram* view is useful for viewing regions with the greatest number of reads.

Figure 102: Configuring the Style tab to configure Alignment tracks

Figure 103 shows the *Alignment* track with the *Fewest number of rows* option.
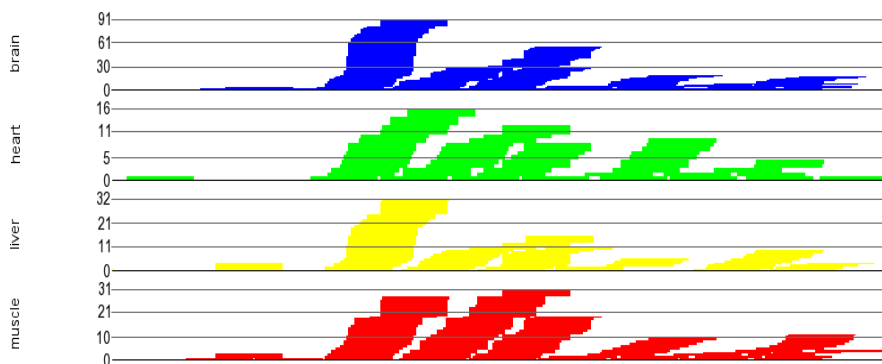

Figure 103: Viewing the Alignment tracks displaying Fewest number of rows option

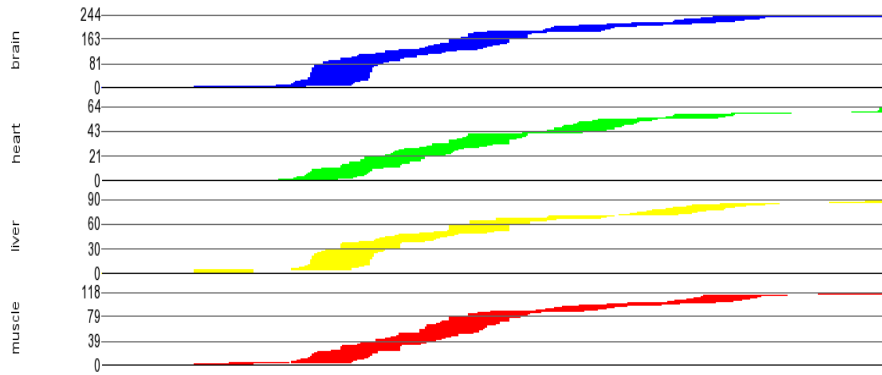Figure 104 shows the *Alignment* track with the reads at *One per row* option.

Figure 104: Viewing the Alignment tracks displaying One per row option

The *Strands* color options will color the results according to the direction of either the forward or reverse read (Figure 105).
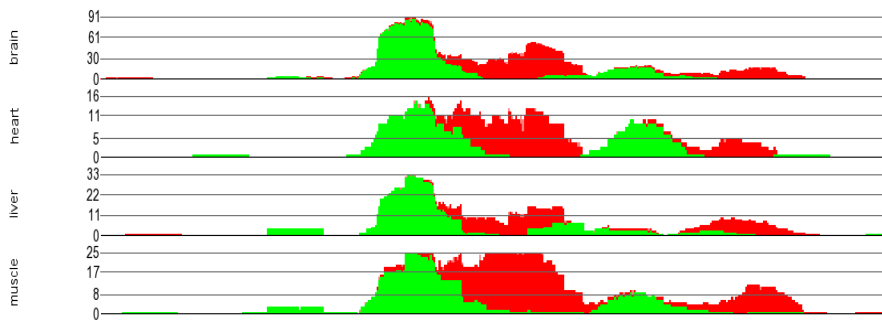
Figure 105: Viewing the Alignment tracks colored by Forward and Reverse strands

The *Bases* color option will color the results according to the base (GATCN) of the read (Figure 106). When the color is set to bases the view must be zoomed in far enough to distinguish base pairs – otherwise the plot will by colored by matches.
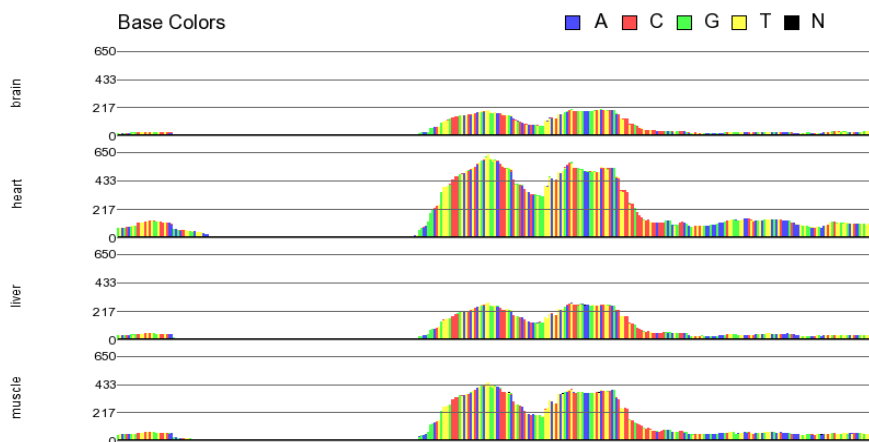

Figure 106: Viewing the Alignment tracks colored by Base Colors

The *Forward Codons* and *Reverse Codons* color options color will be colored according to the codon of the read. It can be used to determine if a mutation causes

a change in amino acid. Figure 107 shows the *Alignment* track color by the *Forward Codons*.
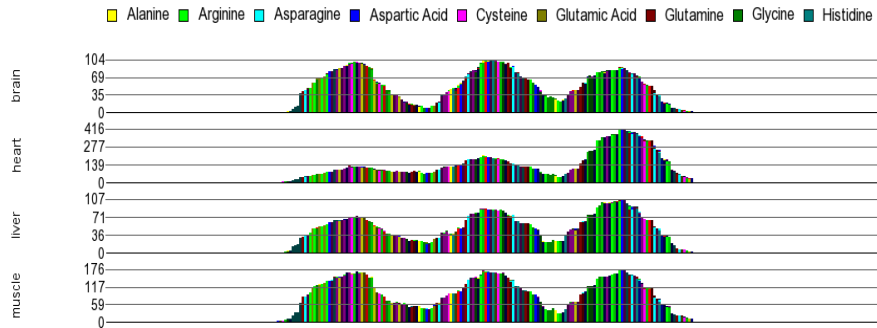

Figure 107: Viewing the Alignment tracks colored by Forward codons

The *Sample color* options will color the results according to the color of the samples in the dataset (Figure 108).
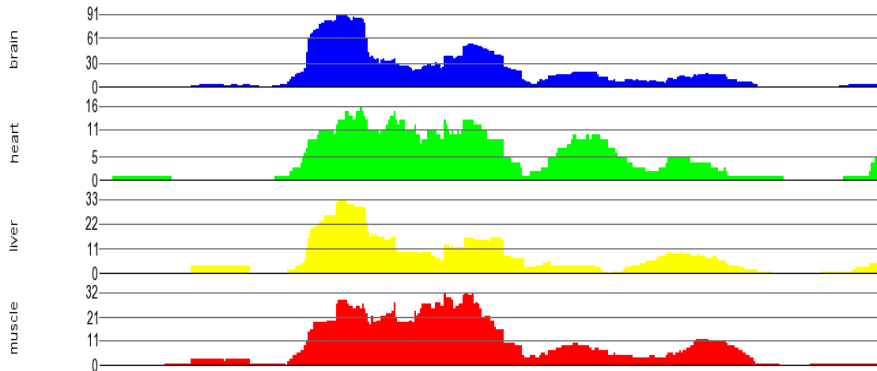

Figure 108: Coloring by sample to match the colors in the isoform proportion view

Classes of reads to display forwards and reverse reads can be configured separately using the *Filter* tab (Figure 109). This includes by *single and forward reads, single end reverse reads, paired end forward-forward reads, paired end forward-reverse reads, paired end reverse-forward reads* and *paired end reverse-reverse reads*.

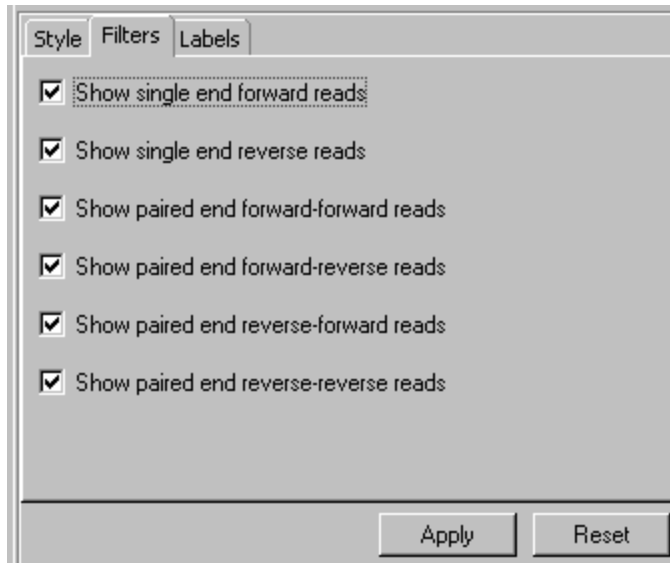*Figure 109: Configuring the Filter tab to display forward and reverse reads*

The *Labels* tab can edit the *Track Title*, *Track Title Size* and *Label Size* of the *Alignment* track (Figure 110).
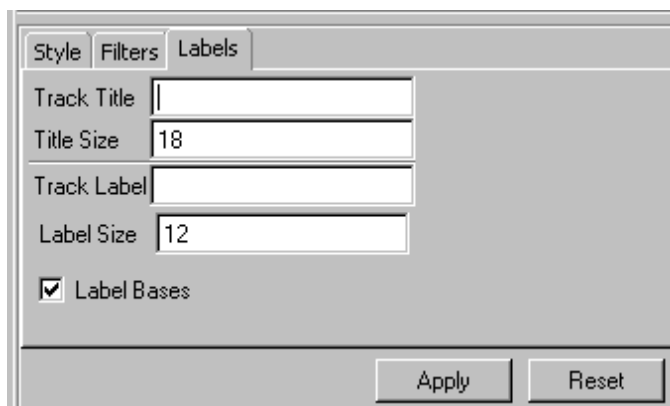

Figure 110: Editing the Label properties of the Alignment track

SNP Proportion Track

The *SNP Proportion* track (Figure 111) gives a graphical representation of the relative SNP abundance for each sample at each genomic location where one is found. The SNP list is created using the *Variations across Samples* option of the RNA-Seq workflow.

Figure 111: Viewing the SNP Proportion track

Configure the colors of the *SNP Proportion* track from the *Color* tab using the *Configure base colors button* (Figure 112).
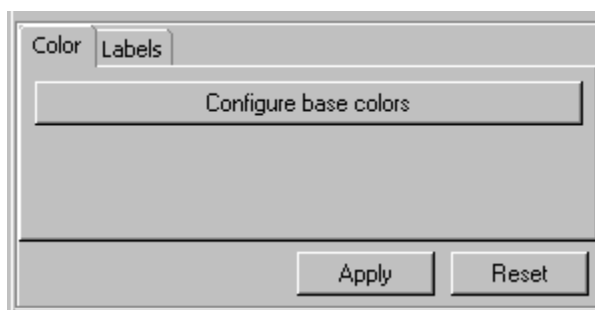

Figure 112: Configuring the base colors of the SNP Proportion track

Edit the *Track Title*, *Title Size* and *Label Size* of the *SNP Proportion* track under the *Labels* tab (Figure 113).
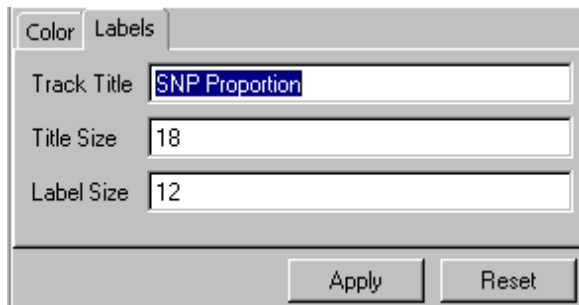

Figure 113: Editing the Track Title, Title Size, and Label Size of the SNP Proportion track

The *SNP Proportion Legend* displays the color configuration of the *SNP Proportion* track (Figure 114). Editing the legend colors will not edit the colors on the *SNP Proportion* track.

Figure 114: Viewing the SNP Proportion Legend of the SNP Proportion track

The *Color* tab allows you to configure the colors of the *SNP Proportion Legend* track using the *Configure colors* button (Figure 115).
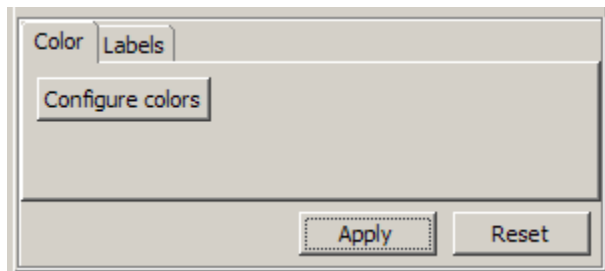

Figure 115: Configuring the base colors of the SNP Proportion Legend track

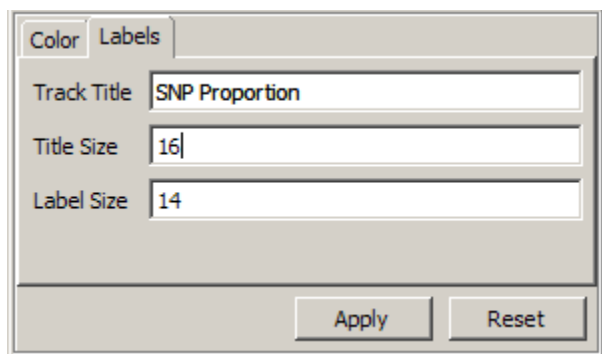Edit the *Track Title*, *Title Size* and *Label Size* of the *Labels* tab (Figure 116).

Figure 116: Editing the Track Title, Title Size, and Label Size of the SNP Proportion Legend track

ChIP-Seq Tracks

The ChIP-Seq tracks are used to identify *in vivo* transcription factor binding sites across the entire genome including motif binding sites and enriched regions (Figure 117).
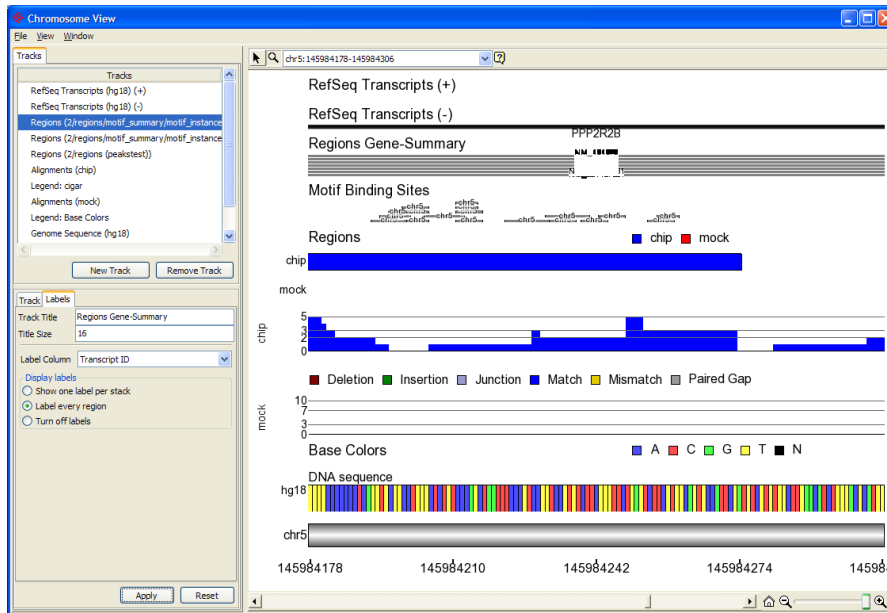
Figure 117: Viewing the ChIP-Seq tracks in the Chromosome View

Motif Binding Sites Track

The *Motif Binding Site* track displays the locations of the instances of detected known or de novo motifs (Figure 118).

Figure 118: Viewing the Motif Binding Site track showing instances motifs

From the *Track* tab you can adjust the height of the track using the *Track Height* slider (Figure 119). Moving the slider to the right will increase the height, moving the slider to the left will decrease the height.
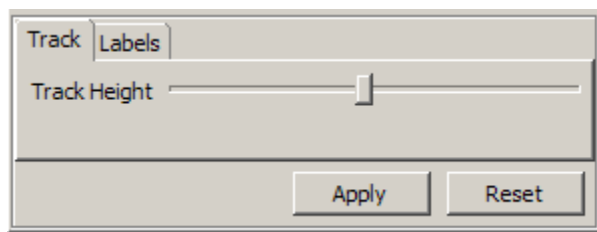

Figure 119: Adjusting the height of the Motif Binding Site track

The *Labels* tab allows you to edit the *Track Title* and *Title Size*, or turn on/off labels per stack or per region (Figure 120). The *Label Column* to display from the spreadsheet attributes can be specified. These will only be visible if you are zoomed in far enough. The labels can be changed to *Show one label per stack*, *Label every region*, or *Turn off labels*.
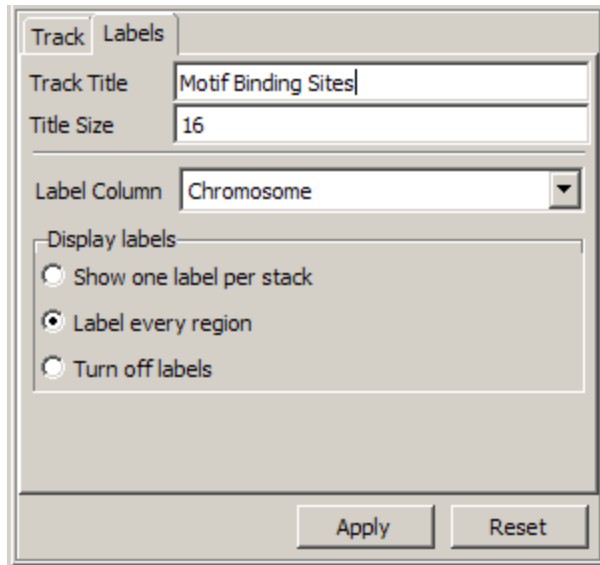
Figure 120: Editing the Track Title, Title Size, and label properties of the motif region

Region Track

The *Region* track in Figure 121 gives the enriched regions detected using the *Create List of Enriched Regions* step from the ChIP-Seq workflow. Regions that contain a binding site for the transcription factor of interest will have many sequence reads mapped to it. Lists of regions are created by looking at the Peaks and identifying regions in or not in a sample and in or not in a control sample.
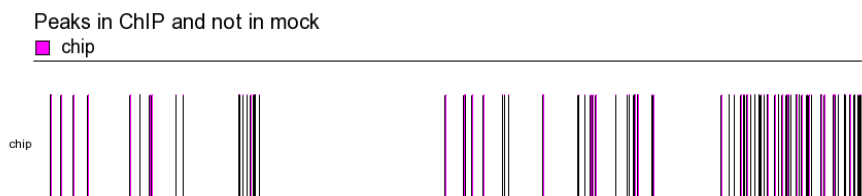

Figure 121: Viewing the Region track showing regions in chip sample but not in mock sample

The *Profile* tab of the *Region Track* allows you to separate the peaks by options in the drop down menu (Figure 122). If a spreadsheet has genomic features on rows and a sample ID column then, by default, there will be one row per sample. The *Track Height* slider is used to adjust the height of the track in the view.
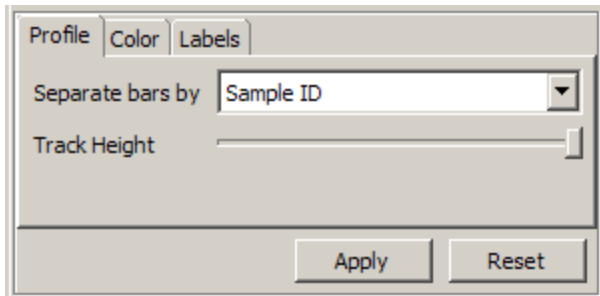
Figure 122: Setting properties of the Profile tab of the Region Track

The *Color* tab is used to adjust how the bars are to be colored (Figure 123). The *Min* and *Max* inputs control the color scale of how the *Color bars by* attribute are drawn. Select the *Configure category colors* button to configure the colors of the attribute selected in the *Color bars by* dropdown menu.
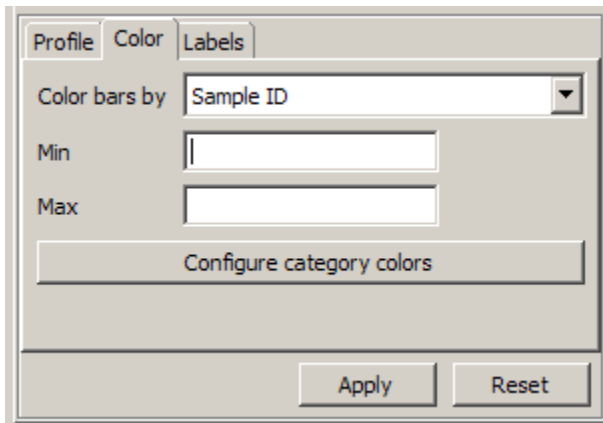

Figure 123: Configuring the colors of the Color tab of the Region Track

The *Labels* tab is used to edit how the labels of the track are displayed, and whether or not to display all or only the selected sample (Figure 124).
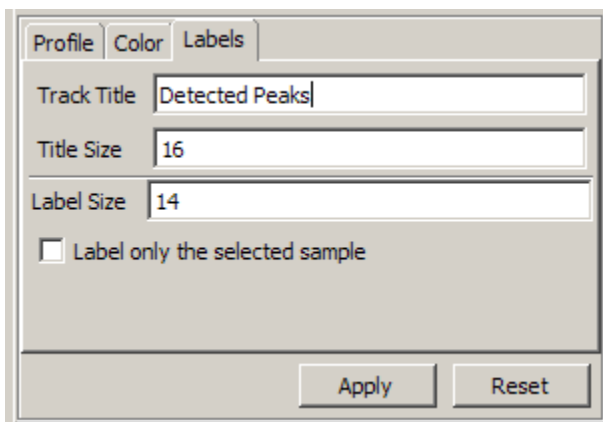

Figure 124: Editing the Track Title, Title Size and Label only the selected sample of the Region track

Region Bar Profile Track

The *Region bar profile* by *New Track > Other (Advanced) > Region bar profile* adds a track which displays the mapped reads coverage and allows you to configure the ways the coverage is displayed (Figure 125).
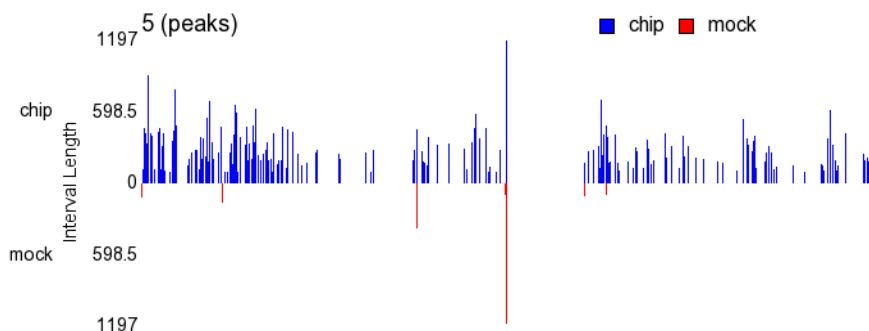

Figure 125: Viewing the Region bar profile track showing chip vs. mock samples with height as interval length separated by Sample ID

From the *Profile* tab, check *Separate bars by* to separate the bars by the available attributes in the spreadsheet (Figure 126). The histogram bar height is determined by the selected attribute in the *bar height by* drop down menu. The *Min* and *Max* values set the Y-axis scale of the track. The *Bars come from* feature sets the baseline from which to extend the histogram height.
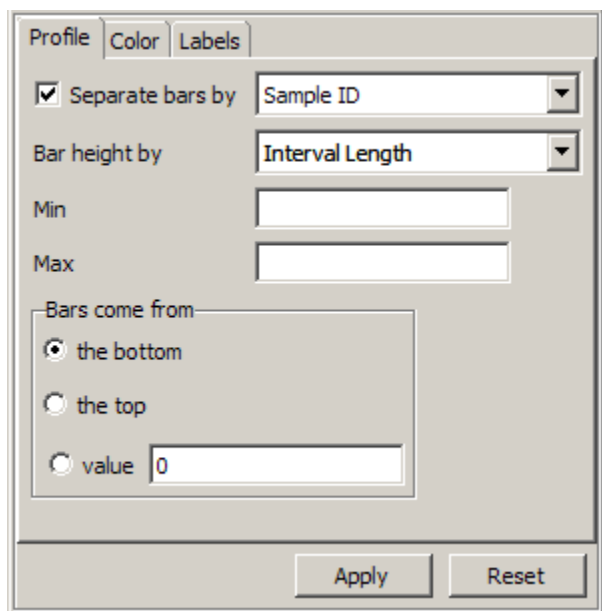

Figure 126: Configuring the Profile tab of the Region bar profile tab

Color bars by will determine which attribute is used to color the histogram bars. Select *Configure category colors* to change the colors of the histogram bars by attribute (Figure 127). The *Min* and *Max* values set the color range intensity values

but will only be noticeable if specific attributes are selected for the *Color bars by* drop down menu.
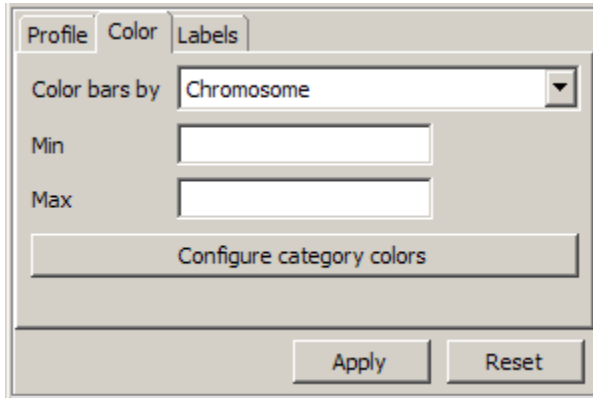

Figure 127: Configuring the Color tab of the Region bar profile tab

The *Track Title*, *Title Size* and *Label Size* can be changed from the *Labels* tab (Figure 128).
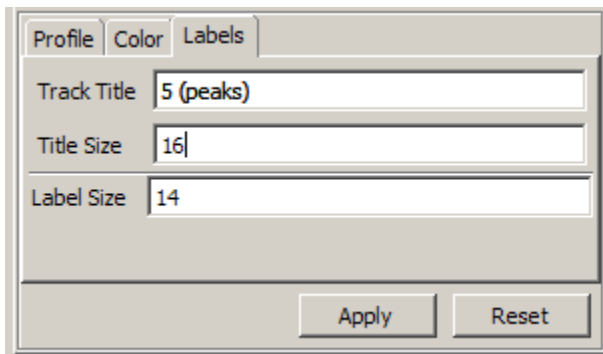

Figure 128: Editing the Track Title, Title Size and Label only the selected sample of the Region track
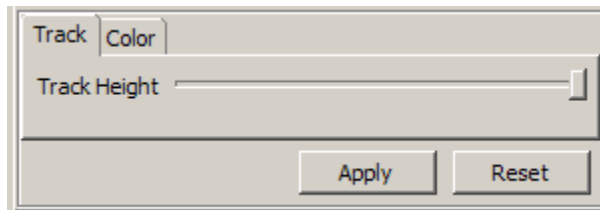
Sequence Heat Map Track

Sequence Heat Map will add a heat map track & alignment track (Figure 129). This option is specifically designed to display a heat map of sequence data. The color of cells is based on reads per kb per million reads (RPKM). Initially the alignment track is empty. Select a sample in the *Sequence Heat Map* to populate the alignment track.
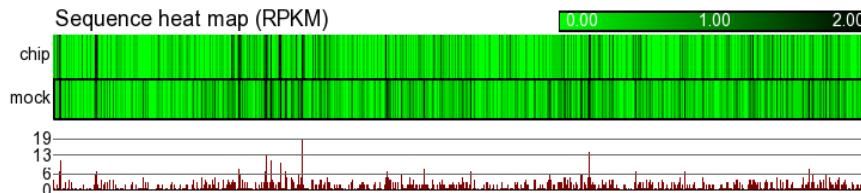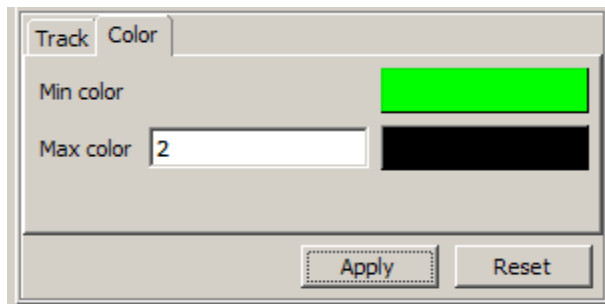
Figure 129: Viewing the Sequence Heat Map with mock sample selected. Each row corresponds to one sample. Alignment track below is populated with alignments from sample as samples are selected in Sequence Heat map

Figure 130 shows how to adjust the height of the *Sequence Heat Map* track using the *Track Height* slider. Moving the slider to the right will increase the height of the track. Moving the slider to the left will decrease the height of the track.

Figure 130: Adjusting the height of Sequence Heat Map track

The color of the *Sequence Heat Map* can be changed using the *Color* tab (Figure 131). The *Min* and *Max* colors can be set, and the heat map *Max color* intensity can be changed.

Figure 131: Adjusting the Color tab of the Sequence Heat Map track

DNA-Seq Tracks

DNA-Seq is useful for looking at Mendelian inconsistencies, SNP, and inheritance information.
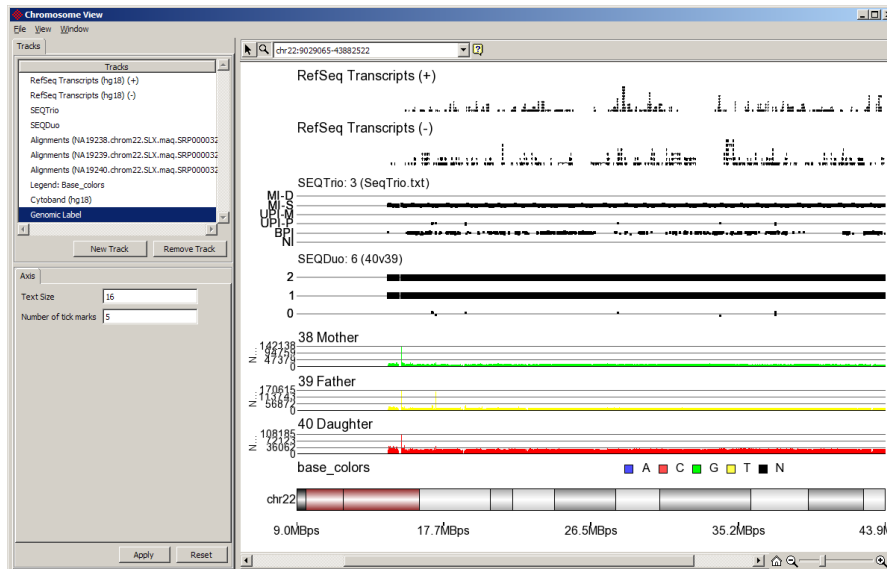
Figure 132: Viewing the DNA-Seq tracks in the Chromosome View

SeqDuo Track

The *SeqDuo* track displays the number of concordant alleles between two samples' genotypes for each SNP. Figure 133 shows a *SeqDuo* track containing an *Idenity by State of 1*.
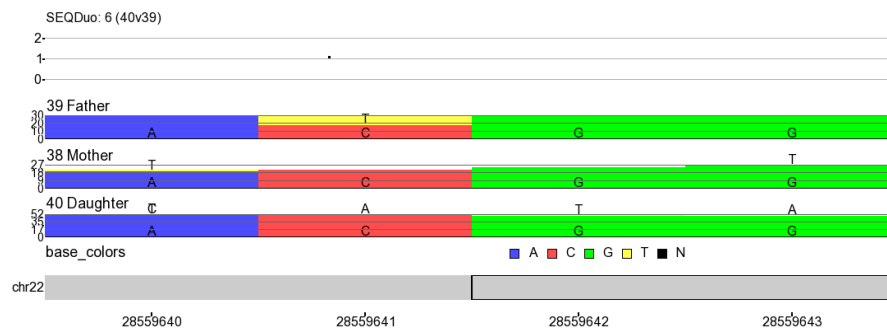

Figure 133: Viewing the SeqDuo track with Identity by state

SNPTrio Track

The *SeqTrio* track displays information about Mendelian allele consistency and inheritance. Figure 134 shows an example of a *SeqTrio* track with *Uniparental Inheritance* from the father's side.
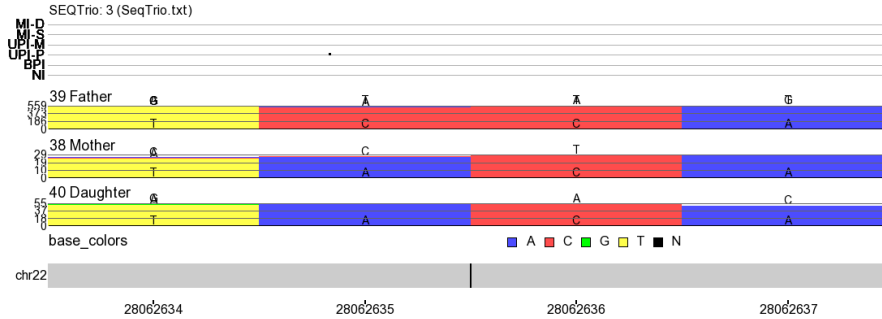
Figure 134: Viewing the SeqTrio track with Uniparental Inheritance-Paternal

The *List 1* gives meanings for the *SeqTrio Scale* parameters.

MI-D	Mendelian Inconsistency – Double Alleles
MI-S	Mendelian Inconsistency – Single Allele
UPI-M	Uniparental Inheritance – Maternal
UPI-P	Uniparental Inheritance – Paternal
BPI	Biparental Inheritance
NI	Not Informative

List 1: Viewing the SeqTrio track scale parameters

More information regarding SNP Trio is available from
http://pevsnerlab.kennedykrieger.org/SNPtrio04.htm.