

Systems Biology Research Group

MetaScope

User's Guide

Donghyuk Kim
August 2010

Table of content

Figure 1. Four files of MetaScope application.....	8
Figure 2. MetaScope when first launched.	9
Figure 3. The overall look of MetaScope.....	11
Figure 4.Examples of window display styles.....	12
Figure 5. Introduction tab	13
Figure 6. Documentation tab and update tab.....	13
Figure 7. Main Window showing two genomic annotations and two ChIP-chip datasets.....	14
Figure 8. Workspace explorer window	14
Figure 9. Feature property window.....	15
Figure 10. Selected feature window.....	15
Figure 11. Search window	16
Figure 12. Edit history window.....	16
Figure 13. Bookmark window.....	17
Figure 14. Title bar	17
Figure 15. Chromosome tab and its UI components	17
Figure 16. Floating chromosome tab	18
Figure 17. Chromosome tabs attached to the upper and lower sides	18
Figure 18. File > Open menu.....	21
Figure 19. File reading dialog box.....	21
Figure 20. MetaScope with datasets just loaded.	22
Figure 21. "View > Zoom In" and "View > Zoom Out" menu.....	22
Figure 22. Zoom To dialog box.....	23
Figure 23. Horizontal scrollbar when zoomed in	23
Figure 24. "View > Scroll Left" and "View > Scroll Right" menu	24
Figure 25. "View > Position To" menu.....	24
Figure 26. Position To dialog box	25
Figure 27. Bookmark context menu	25
Figure 28. Selecting data tracks.....	26
Figure 29. Selecting one data feature	27
Figure 30. Setting data track editable	27
Figure 31. Selecting multiple data features by dragging.....	28
Figure 32. Selected multiple data features	28
Figure 33. Search dialog box	29
Figure 34. Search result	29
Figure 35. "File > Open Layout" and "File > Save Layout As" menu	30
Figure 36. Set Color context menu	31

Figure 37. Pick Color dialog box	31
Figure 38. Set Height context menu.....	32
Figure 39. Set Height dialog box	32
Figure 40. 4 display styles.....	33
Figure 41. Display context menu.....	33
Figure 42. Move Up and Move Down context menu.....	34
Figure 43. Grouping of two data tracks	34
Figure 44. Group and Ungroup context menu	35
Figure 45. Manual Scale context menu	36
Figure 46. Set Scale dialog box	36
Figure 47. Change Type dialog box.....	36
Figure 48. Hide Track and Close File context menu.....	37
Figure 49. Context menu for feature operation functions.....	38
Figure 50. Uniting multiple data features	39
Figure 51. Merging multiple data features	39
Figure 52. Merge tab in Feature Operation dialog box	40
Figure 53. Filtering multiple data features.....	41
Figure 54. Filter tab in Feature Operation dialog box.....	41
Figure 55. Move and Copy tabs in Feature Operation dialog box.....	41
Figure 56. Edit a Feature dialog box.....	42
Figure 57. Deleting data features.....	42
Figure 58. Context menu for track operation functions.....	44
Figure 59. Averaging data tracks	45
Figure 60. Average tab in Track Operation dialog box	45
Figure 61. Differencing data tracks.....	46
Figure 62. Difference tab in Track Operation dialog box.....	46
Figure 63. Summation tab in Track Operation dialog box.....	47
Figure 64. Filtering data tracks	47
Figure 65. Filter tab in Track Operation dialog box	48
Figure 66. Adjust tab in Track Operation dialog box	48
Figure 67. Assign ID tab in Track Operation dialog box	49
Figure 68. Context menu for RTS integration function.....	50
Figure 69. RTS tab in Integration Operation dialog box	51
Figure 70. RTS integration from RBR and TD	51
Figure 71. Context menu for pORF integration	52
Figure 72. pORF tab in Integration Operation dialog box	52
Figure 73. pORF integration from start and stop codon position and proteomic data.....	52
Figure 74. Context menu for TU integration.....	53
Figure 75. TU tab in Integration Operation dialog box	53

Figure 76. TU integration from TSS, RTS, pORF.....	54
Figure 77. Sample 1. RNAP ChIP-chip	55

Getting Started

Introduction

MetaScope is a powerful software tool which researchers can use in order to:

- Visualize multiple genome-scale datasets along with genomic annotations
- Compare and analyze annotation-anchored experimental data
- Curate and integrate genome-scale datasets
- Build a new annotation with experimental data and canonical annotation such as transcription unit annotation

Overview of MetaScope functionality

General usage of MetaScope includes

- Load genomic annotations
- Load genome-scale datasets including ChIP-chip data, expression profiling data, and TSS (Transcription Start Site) data
- Load processed datasets including predicted binding site from ChIP-chip data, transcription detect signals from expression profiling, and processed TSS data
- Compare and analyze genome-scale experimental datasets with known genomic annotations
- Curate and integrate processed datasets
- Build a new annotation with processed datasets and/or canonical genomic annotations

Unique features of MetaScope

MetaScope supports integrative functions by which multiple genome-scale datasets can be analyzed, compared and integrated. Those integrative functions include:

- Track operation
 - Make average, difference, or sum
 - Merge features from biological replicates
 - Filter features by features on other track.
 - Adjust score values of features, or width of them
 - Assign ID to features in an orderly manner
- Feature operation
 - Filter out features by score, or leave top features with or without sliding window
 - Merge features in feature level
 - Unite multiple features into one
 - Move or copy features into another track
 - Create, edit or delete features
- Integration function
 - Integrate start and stop codon information with proteomic data to generate potential ORF (pORF) annotation
 - Integrate RNA polymerase (RNAP) ChIP-chip binding information with transcript detection signals to generate RNAP-guided transcription segment (RTS)

- Integrate RTS with transcription start site data and pORF data to generate transcription unit annotation.

In addition, MetaScope provides overlapping multiple data tracks, splitting one data window into two, workspace management by which user can store information about data files open, track setting, overlapping information and other configuration for each project.

Documentation

The most recent version of user's guide can be found on the website below:

<http://gcrg.ucsd.edu/Downloads/MetaScope>

If you have any question, please contact Donghyuk Kim (dok023@ucsd.edu).

Document conventions

A expression “A > B” when describing user interface means “click the first item A; this makes the next item B visible; then click B.”

Examples:

- “**File** menu > **Open Workspace**” means “click the **File** menu first, and then choose **Open Workspace**.”
- “**Average** tab > **An existing file**” means “click the **Average** tab and then click the **An existing file** combo box.”

Screen shots in this document were taken at the time this document was written, thus there might be small changes in positions or texts in the user interface, because MetaScope will be updated on a regular basis. MetaScope runs on .NET framework, however there also might be some differences on different versions of Windows operating system.

Installation

System requirements

MetaScope runs on .NET framework 4.0 or higher, thus can run on any computer with .NET framework installed. Currently, .NET framework supports only Microsoft Windows operating system, thus MetaScope cannot run on Linux or Mac machines.

.NET framework is freely available on Microsoft website below:

<http://www.microsoft.com/downloads/details.aspx?FamilyID=9cfb2d51-5ff4-4491-b0e5-b386f32c0992&displaylang=en>

Installation of .NET framework may require rebooting of your system once.

The recommended minimum hardware requirements are Pentium 1 GHz or higher with 512 MB RAM or more. However, if large amount of datasets will be loaded and visualized, then a system with Intel Core processor family or higher and 2 GB RAM or more is recommended.

MetaScope is implemented with WPF (Windows Presentation Foundation) graphical library. WPF renders user interfaces in Windows-based application, and directly utilizes DirectX, rather than relying on the older GDI subsystem. Thus graphic card of the system also affects the overall visualization performance of Metascop.

Typical installation

In order to install MetaScope, visit the MetaScope site:

<http://gcrg.ucsd.edu/Downloads/MetaScope>

Download the “.zip” file containing following 4 files, and unzip those files, and place them whichever folder they need to be placed, as shown in Figure 1.

- MetaScope.exe
- AvalonDoc.dll
- MetaScope.workspace
- MetaScope.Layout.xml

“MetaScope.exe” is the main executable file of MetaScope application. “AvalonDoc.dll” is also a part of MetaScope software, and contains a library for user interface of MetaScope. “MetaScope.workspace” is a default workspace file, which only indicates default user interface layout file “MetaScope.Layout.xml”.

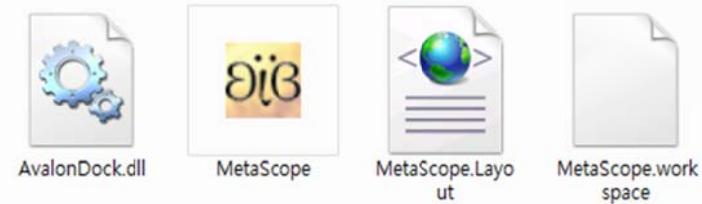


Figure 1. Four files of MetaScope application.

AvalonDock is a WPF controls library which can be used to create a docking layout system like that is present in VisualStudio. It supports fly-out panes, floating windows, multiple docking manager in same window, styles and themes and it can host WinForms controls. This library runs on new BSD license as below:

Copyright (c) 2007-2009, Adolfo Marinucci
All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- * Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- * Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- * Neither the name of Adolfo Marinucci nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Launching MetaScope

The procedure for launching MetaScope is quite straight-forward. Double-clicking “MetaScope.exe” launches MetaScope application. Figure 2 shows MetaScope when it is first launched.

Multiple instances of MetaScope can be launched at the same time, the user should make sure that each instance of MetaScope does not share workspace and data files with other instances.



Figure 2. MetaScope when first launched.

Getting familiar with MetaScope

Terminology

Genomes and chromosomes

Genome and chromosome can slightly different meaning dependent on the context where they are used. In MetaScope a **chromosome** means any single piece of sequence and single continuous region of genomic positions where that sequence is located. Each chromosome should have the unique chromosome ID, and one example of chromosome ID is NCBI reference sequence ID which starts with “NC_”.

A **genome** refers to any group of those chromosomes described above.

For example, the genome of *Escherichia coli* has only one chromosome, the NCBI reference sequence ID of which is NC_000913.

(http://www.ncbi.nlm.nih.gov/nuccore/NC_000913.2)

Another example is the genome of *Klebsiella pneumoniae* MGH 78578 has one main chromosome (NC_009648) and 5 plasmids (NC_009649, NC_009650, NC_009651, NC_009652, NC_009653) which are considered 5 difference chromosomes in MetaScope.

MetaScope recognizes and categorizes all datasets by chromosome ID given in the GFF files.

Annotations

Currently, MetaScope only support data files in GFF format, and does not distinguish experimental raw data, processed data or annotation if they share the same chromosome ID. Anyway, the definition of annotation, or rather genomic annotation, used here is the known or suspected locations of genomic features, such as genes, mRNAs, sRNAs, predicted coding regions, pseudogenes, promoter regions, transcription start sites, RNA polymerase binding regions and others.

GFF (General feature format)

According to the Wikipedia, the general feature format is a file format used for describing genes and other features of DNA, RNA and protein sequences. The filename extension associated with such file is “.GFF”, and currently there are two versions of the GFF file format in general use: GFF version 2 and GFF version 3.

The more information about GFF can be found in Wikipedia website, Sanger institute website and Sequence Ontology Project website below:

http://en.wikipedia.org/wiki/General_feature_format

<http://www.sanger.ac.uk/resources/software/gff/spec.html>

<http://www.sequenceontology.org/gff3.shtml>

The overall look

Figure 3 shows the overall look of MetaScope, displaying the sample dataset 1 “Sample 1. RNAP ChIP-chip” dataset.

The MetaScope application user interface consists of 8 distinct windows. As shown in Figure 3, these windows are:

- Main window
- Workspace explorer window
- Feature window
- Selected feature window
- Search window
- Edit history window
- Bookmark window

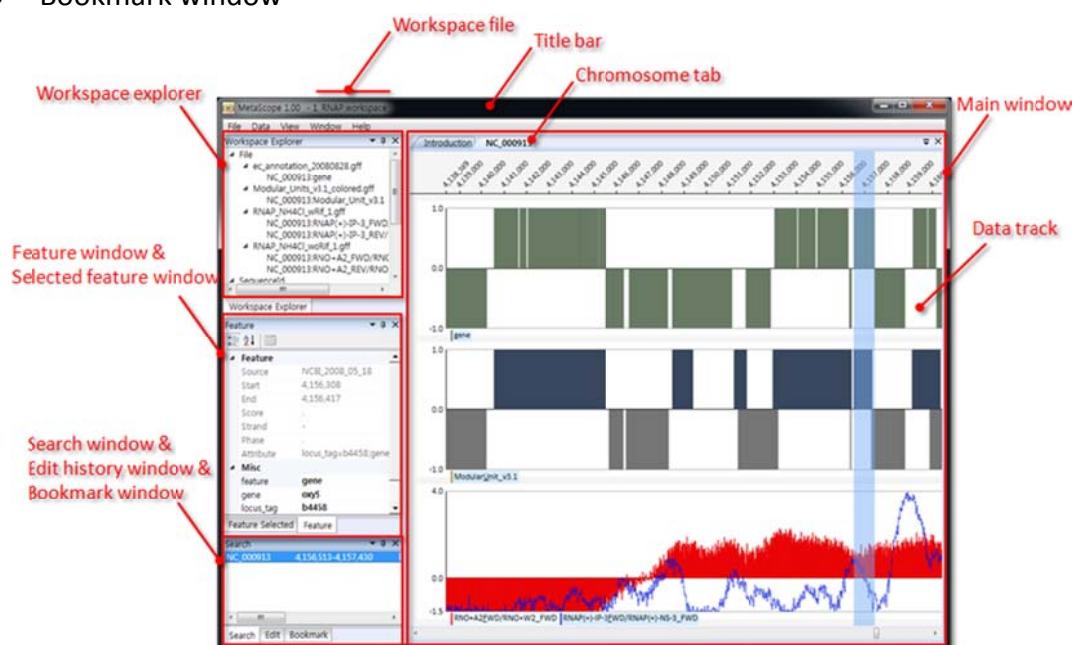


Figure 3. The overall look of MetaScope.

The UI of MetaScope is highly flexible and configurable; thus each window can be selected to be shown or hidden, and when shown, it can be “attached”, “tabbed”, or “floated.” (Figure 4)

Chromosome tabs are detachable as well, thus if research is using more than one monitors, then he/she can move one chromosome tab in one monitor and place the other chromosome tab in another monitor, maximizing MetaScope and monitor usage.

Display style of windows can be saved in the designated application layout file like “MetaScope.Layout.xml”, and can be loaded to restore the previous application layout later on. The application layout file is in XML format, and will be discussed in detail later.

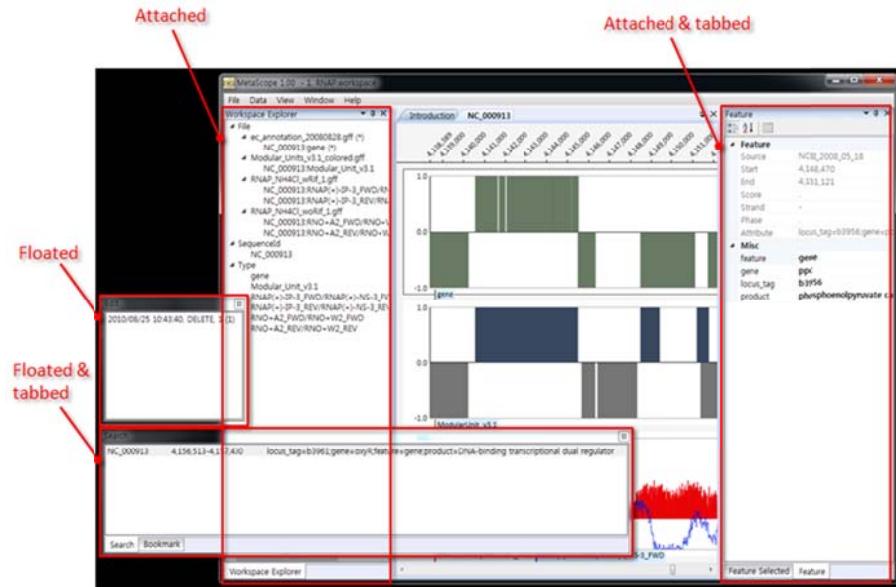


Figure 4.Examples of window display styles

Windows and other UI elements

Main window

Main window contains zero or more tabs, and there are two types of tabs: non-chromosome tabs and chromosome tabs.

Currently there are three types of non-chromosome tabs:

- Introduction tab
- Documentation tab
- Update tab

Introduction tab is the default tab shown when MetaScope is launched, and shows the brief information about MetaScope, as shown in Figure 5. **Documentation tab** and **update tab** are referring the same website, the official MetaScope webpage for now, as shown in Figure 6. The MetaScope user's guide document and the most recent version of MetaScope can be found on that website. Update tab will be updated in the future to refer user's guide document directly.

Chromosome tab displays all genomic annotation and data sharing the same chromosome ID, which is shown on the tab label. In Figure 7, the chromosome ID, NC_000913, is shown on the chromosome tab. Chromosome tab has a genomic position ruler, by which the exact genomic positions of every features visualized can be recognized.



Figure 5. Introduction tab

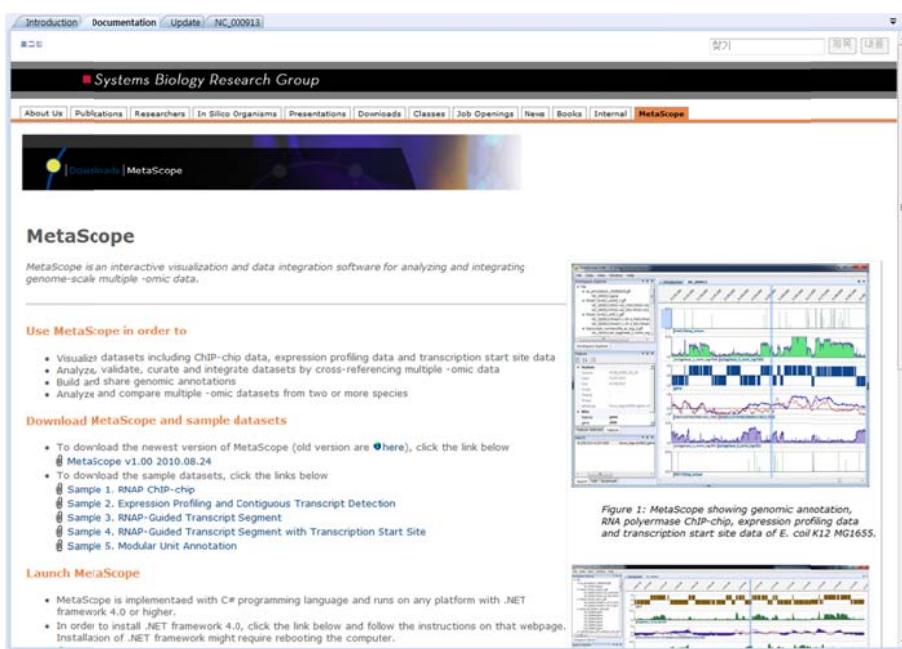


Figure 6. Documentation tab and update tab

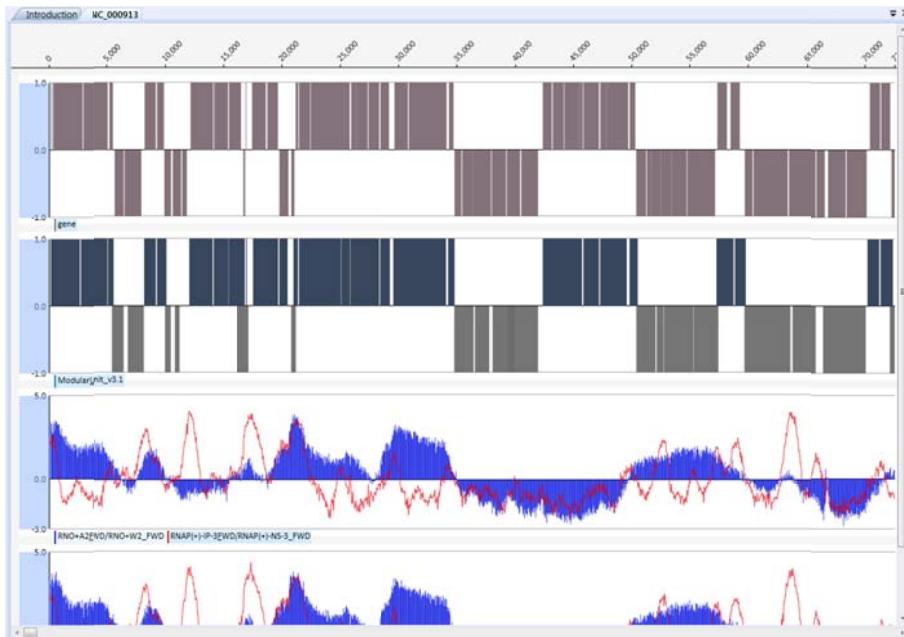


Figure 7. Main Window showing two genomic annotations and two ChIP-chip datasets

Workspace explorer window

Workspace explorer window shows brief information about the current workspace, which includes all files open, and chromosome IDs and types recognized, as shown in Figure 8.

For example, “ec_annotation_20080828.gff” file has one chromosome ID, NC_000913, and one data type, “gene”. Similarly “RNAP_NH4Cl_woRif.gff” data file contains the same chromosome ID, however it has two different data types.

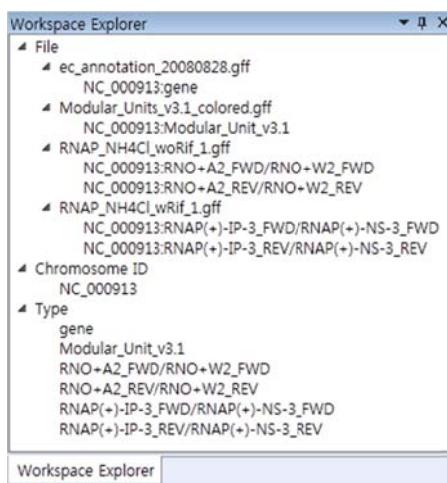


Figure 8. Workspace explorer window

Feature window

Feature window or feature property window shows all the information of one data feature which the mouse is hovering over, as shown in Figure 9. Basically, this shows feature information described in its GFF file: source, starting position, ending position,

score, strand, phase and attribute.

Feature section of **Feature window** shows the information enumerated above, and the Misc section displays parsed list of attribute section of that information. For example, Figure 9 shows 4 parsed attributes: feature, gene, locus_tag and product.

Feature	
Source	NCBI_2008_05_18
Start	138,835
End	141,225
Score	.
Strand	-
Phase	.
Attribute	locus_tag=b0124;gene=gcd;feature
Misc	
feature	gene
gene	gcd
locus_tag	b0124
product	glucose dehydrogenase

Figure 9. Feature property window

Selected feature window

Selected feature window shows the information of one or more selected features, as shown in Figure 10. When only one feature is selected, this window looks like the left panel in Figure 10, displaying the count number, starting position, ending position, score, strand, phase and attribute. In contrast, when two or more features are selected, this windows looks like the right one in Figure 10, showing the count number, starting position and ending position, accordingly.

Feature Selected	
Statistics	
Count	1
Feature	
Start	137,083
End	138,633
Score	1
Strand	+
Phase	.
Attribute	locus_tag=b0123;gene=cueO;fea

Feature Selected	
Feature Group	
Count	10
End	145,017
Start	134,788

Figure 10. Selected feature window

Search window

Search window shows the result of keyword search, such as gene name or gene ID. Given a keyword, MetaScope looks up attribute column of GFF format, in a case sensitive or insensitive way, and shows up every matching case.

For instance, Figure 11 shows two search results with the keyword of "oxyR", which is a well-studied transcription factor, found in both *E. coli* and *K. pneumonia*. The first

column displays the chromosome IDs of features found, the second column represents the starting and ending positions of features, and the last column shows attribute sections of those features.

Searching will be discussed in more details later in this document.

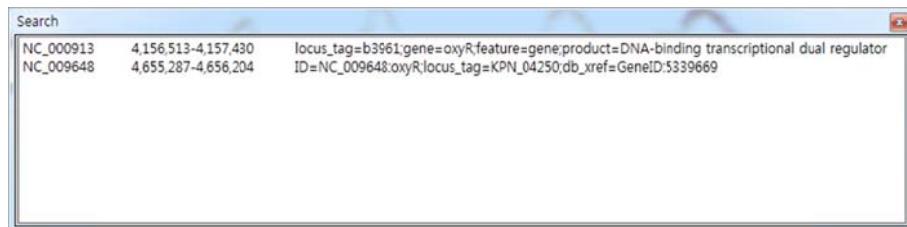


Figure 11. Search window

Edit history window

Edit history window literally shows the history of editing activity, mainly for feature operations. Data operability will be discussed in more details in (???).

The first column shows the time when the data was edited, and the second column means what type of editing was performed. There are three types of editing activity defined in MetaScope: ADD, EDIT and DELETE. The last column shows how many data tracks are involved in the particular editing and how many features are edited.

For example, “2010/08/27 09:21:50, DELETE, 801 (3)” represents deleting 801 features over 3 data tracks is happened at the time of 2010/08/27 09:21:50.

Edit		
2010/08/27 09:21:50	DELETE	12 (1)
2010/08/27 09:21:57	DELETE	801 (3)
2010/08/27 09:22:06	DELETE	705 (3)
2010/08/27 09:22:13	ADD	1 (1)
2010/08/27 09:22:32	EDIT	1->1
2010/08/27 09:22:43	DELETE	1663 (4)
2010/08/27 09:22:47	DELETE	548 (5)

Figure 12. Edit history window

Bookmark window

Bookmark window shows the current list of bookmarks, as shown in Figure 13. This list will be saved in the workspace file, and can be loaded when its workspace file is loaded.

The first column means the starting position of genomic region displayed when the bookmark was made. The second column shows the name of the bookmark, and the last column represents the zooming level, with which the genomic region will be visualized.

Double-clicking the item shown in **bookmark window** commands MetaScope navigate to the given position with the described zooming level.

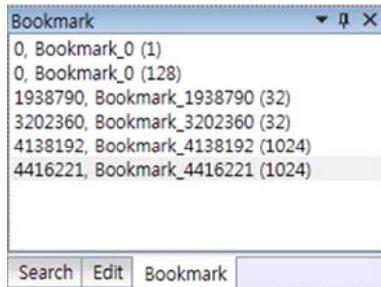


Figure 13. Bookmark window

The title bar

The **title bar** is located at the top area of the application. The **title bar** shows:

- MetaScope icon
- Application name and version information
- Workspace file which is currently open



Figure 14. Title bar

Main window in further details

Main window will be described in further details in terms of data visualization and navigation.

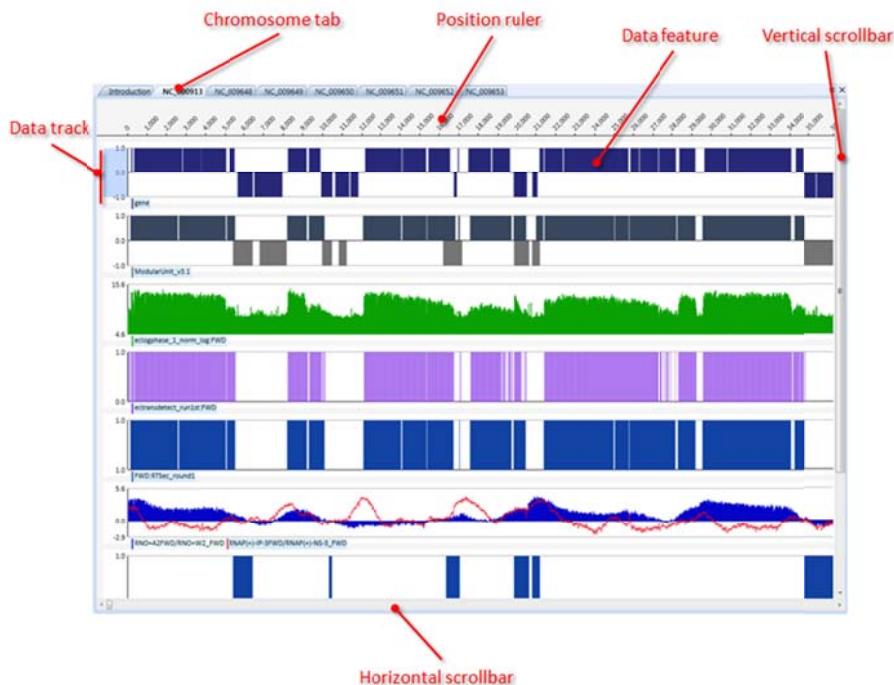


Figure 15. Chromosome tab and its UI components

Chromosome Tab

Chromosome tab is located on top of Main Window, and shows the chromosome ID. As

MetaScope recognizes multiple chromosome IDs, then multiple chromosome tabs will be automatically generated and shown. When the user clicks one of **chromosome tabs**, the **chromosome tab** is activated and focused, showing all datasets and annotations belonging to the corresponding chromosome ID.

Each **chromosome tab** might be attached and tabbed as shown in Figure 15, or floated by drag-and-drop as shown in Figure 16, or rearranged as shown in Figure 17.

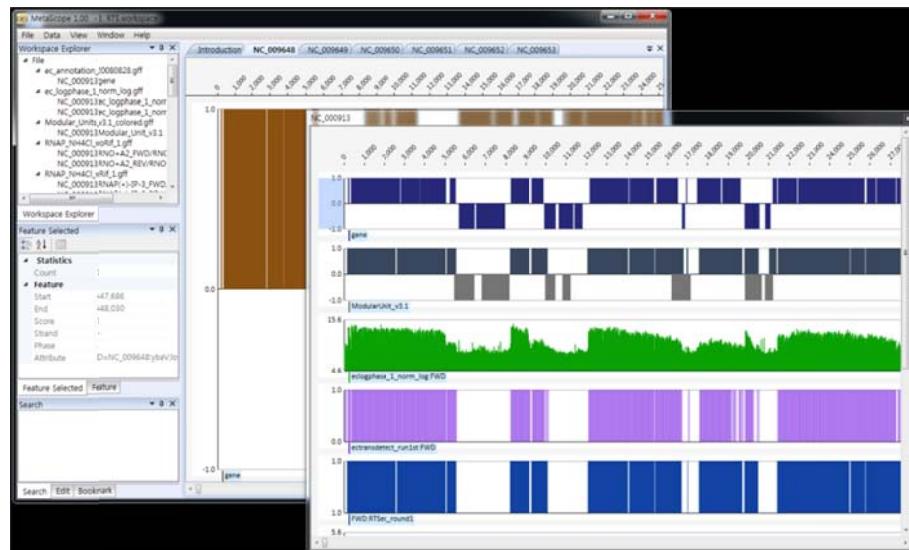


Figure 16. Floating chromosome tab

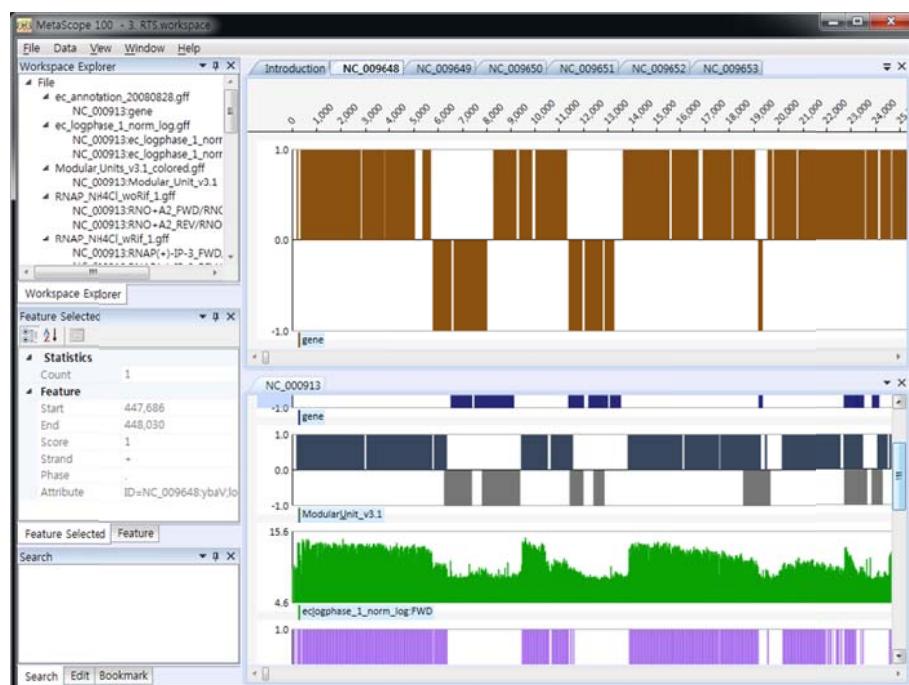


Figure 17. Chromosome tabs attached to the upper and lower sides

Position ruler

Position ruler shows the range of genomic position within which data features are

located. That range starts from 0 and ends at the ending position of the last data feature loaded. As the **chromosome tab** zooms in/out or navigate left/right, the genomic position numbers are automatically recalculated and displayed.

Data track

One chromosome tab can contain multiple **data tracks**, each of which corresponds to a set of data features that share the same data type. The data type represents the third column of GFF file. The more information about this column and the overall explanation of GFF file can be found at the websites listed in terminology section of this document.

Data feature

Data feature is an atomic data structure which are dealt and displayed in MetaScope, and it corresponds to each row in GFF file. Each data feature has 9 distinct fields: sequence ID, source, type, start, end, score, strand, phase, and attribute.

The detailed information and more precise definitions of them can be found at the websites about GFF file, but the brief explanation of them are:

- Sequence ID: This ID means the chromosome ID.
- Source: This field is intended to describe the algorithm or operating procedure that generated this feature.
- Type: This field shows the type of the feature. It can be “gene” or “CDS” for protein coding region or anything else.
- Start: This is the starting position of the data feature in the genome.
- End: This is the ending position of the data feature in the genome.
- Score: This field represents the score of the data feature, which can be used to indicate the signal of the feature, or E-value for sequence similarity feature.
- Strand: This is the strand of the data feature: “+” for positive strand, “-” for minus strand, and “.” for features that are not stranded.
- Phase: This is for “CDS” features, indicating where the feature begins with reference to the reading frame.
- Attribute: This contains a list of feature attributes in the key-value format of “tag=value”.

Horizontal scroll bar

Horizontal scrollbar is located on bottom of **chromosome tab**, and used for navigating across the genomic position.

Vertical scroll bar

Vertical scrollbar is located on the right side of **chromosome tab**, as shown in Figure 15. It is used for scrolling down or up chromosome tab, in order to shift the range of view over data tracks. When there are few data tracks in one **chromosome tab**, vertical scrollbar is hidden. However, as more data tracks are loaded and displayed in the **chromosome tab**, **vertical scrollbar** shows up automatically.

Navigation

Loading datasets

Loading by application menu

One of ways to open and load data files is using application menu “File > Open”, as shown in Figure 18. When clicked, the file open dialog box shows up and allows selecting one or more data file. The workspace file can be opened by this procedure as well.

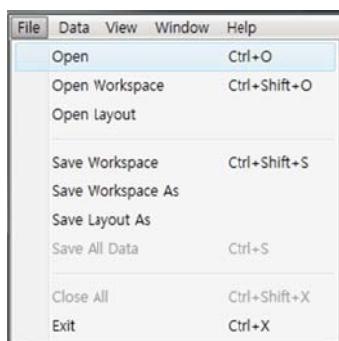


Figure 18. File > Open menu

Loading by drag-and-drop

The easier way is just drag-and-dropping data files or workspace file onto the MetaScope application. They can be dropped onto **workspace explorer**, **introduction tab** or other **chromosome tab**.

File Reading Dialog Box

Once loaded by application menu or drag-and-drop, **file reading dialog box** shows up, and displays the whole list of files to open and the progress of reading data files, as shown in Figure 19.

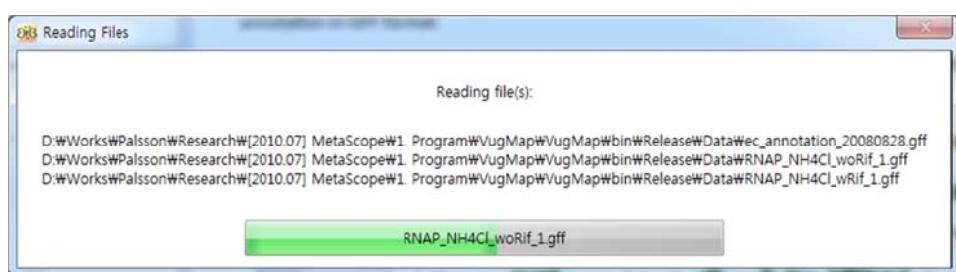


Figure 19. File reading dialog box

Once the loading is done, MetaScope shows the outlook of overall data features, as shown in Figure 20. In this example, 3 data files were loaded as in Figure 19. One is *E. coli* annotation file and two are RNA polymerase ChIP-chip data files. 5 data types are recognized and displayed in 5 separate data tracks.

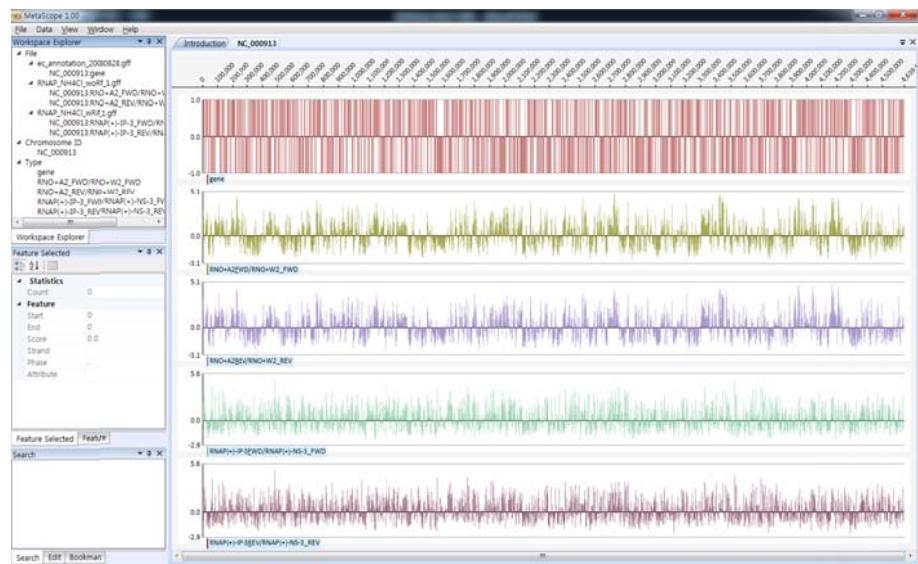


Figure 20. MetaScope with datasets just loaded.

Loading datasets

Once data files are loaded and displayed, **chromosome tab** is zoomed out most, showing every data features at once. The first thing that the user might want to do after this is zooming in or out to narrow down and focus the genomic area of his or her interest.

There are 4 ways to zoom in or out: using application menu, shortcut key, scrolling with control key or zooming to a specific level.

Zooming by application menu

The first of zooming in or out is using application menu, as shown in Figure 21. When it is maximally zoomed out, “View > Zoom Out” menu is inactivated, and similarly when maximally zoomed in, “View > Zoom In” menu is inactivated.

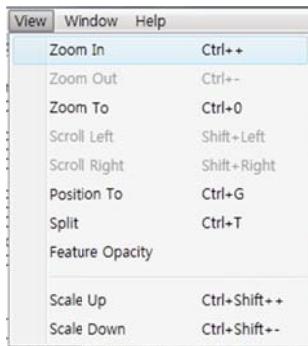


Figure 21. “View > Zoom In” and “View > Zoom Out” menu

Zooming by shortcut

As shown in Figure 21, pressing “+” key with control key makes the view zoomed in, and similarly pressing “-“ key with control key makes it zoomed out.

Zooming by scrolling

The third way, and the easiest way, is scrolling up or down with control key to zoom in or out, accordingly.

Zooming to a specific level

Zooming in or out by those three ways changes the zooming level by 2 fold. The initial zoom level is 1, and it becomes 2 by zooming in once and 4 by zooming in twice. However, there is another way of changing zooming level to a specific level. It is “View > Zoom To” menu, and “Ctrl+0” shortcut does the same thing. When this menu is clicked, Zoom To dialog box shows up, where user can select the zooming level to minimum, maximum or a specific one, as shown in Figure 22.

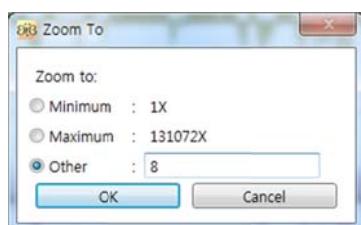


Figure 22. Zoom To dialog box

Scrolling horizontally

Horizontal scrollbar is the basic way of navigating data features and data tracks over genome. When the view is maximally zoomed out, the horizontal scrollbar is expanded out the whole way, as shown in Figure 20. As the view zooms in, the horizontal scrollbar gets shorter and allows user to scroll left or right to change the genomic range within which data feature are displayed, as shown in Figure 23.

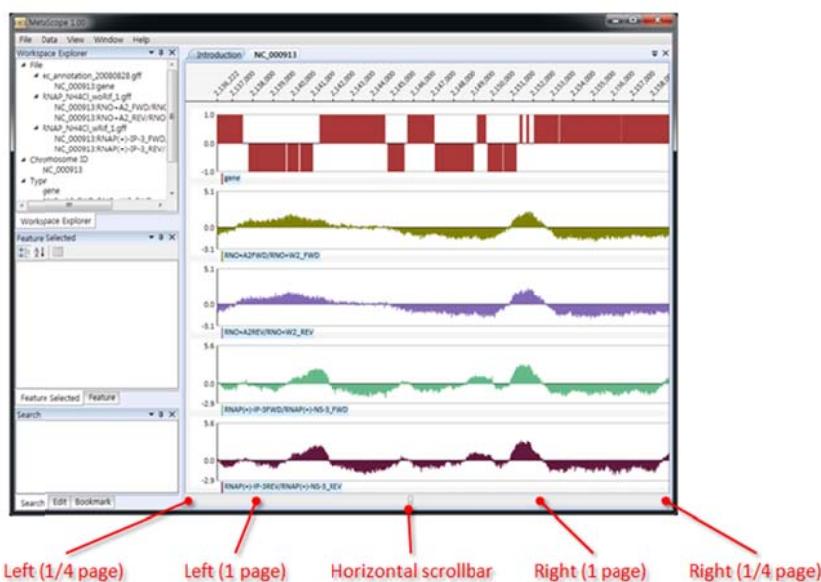


Figure 23. Horizontal scrollbar when zoomed in

User can scroll left or right by dragging the horizontal scrollbar. When the left arrow of

the scrollbar is clicked, the view scrolls left 1/4 of the way, and when the right arrow is clicked, the view scrolls right 1/4 of the way. Similarly, when the middle region between the horizontal scrollbar and the left arrow is clicked, the view scrolls left by one page, and when the right middle region is clicked, the view scrolls right by one page.

Another way of scrolling left or right is using application menu: “View > Scroll Left” and “view > Scroll Right”, as shown in Figure 24.

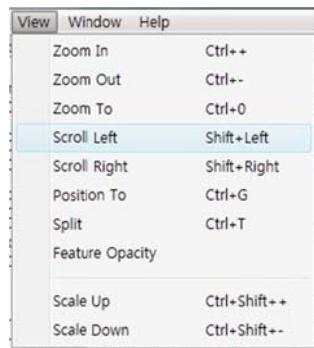


Figure 24. “View > Scroll Left” and “View > Scroll Right” menu

There is another, and the easiest, way of scrolling left or right, which is mouse scrolling with shift key on. When mouse is scrolled up with shift key, the view scrolls left 1/2 of the way, and when mouse is scrolled down with shift key, the view scrolls right 1/2 of the way.

Scrolling vertically

Once many data tracks are loaded and vertical scrollbar is on, just mouse scrolling up and down move the vertical scrollbar.

Going to a specific genomic position

As in the similar way of zooming to a specific level, there is a way of moving to a certain genomic position promptly: “View > Position To” menu. Figure 25 shows this application menu, and “Ctrl+G” shortcut does the same function.

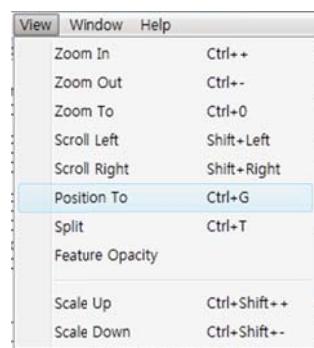


Figure 25. “View > Position To” menu

When “View > Position To” menu is clicked or “Ctrl+G” is pressed, Position To dialog box

shows up, where user can choose the start position, the end position or a specific position to move to, as shown in Figure 26.

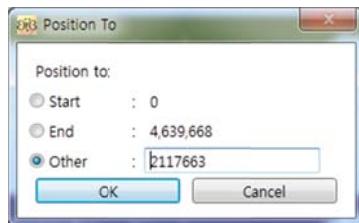


Figure 26. Position To dialog box

Bookmarks

Bookmark can be used for memorizing the genomic position of interest and the zooming level of view. Bookmark can be created by using context menu, as shown in the middle of Figure 27. The list of bookmarks is displayed in Bookmark window, as shown in the left bottom of Figure 27.

Double-clicking the bookmark entity in the list move the genomic position and the zooming level promptly.

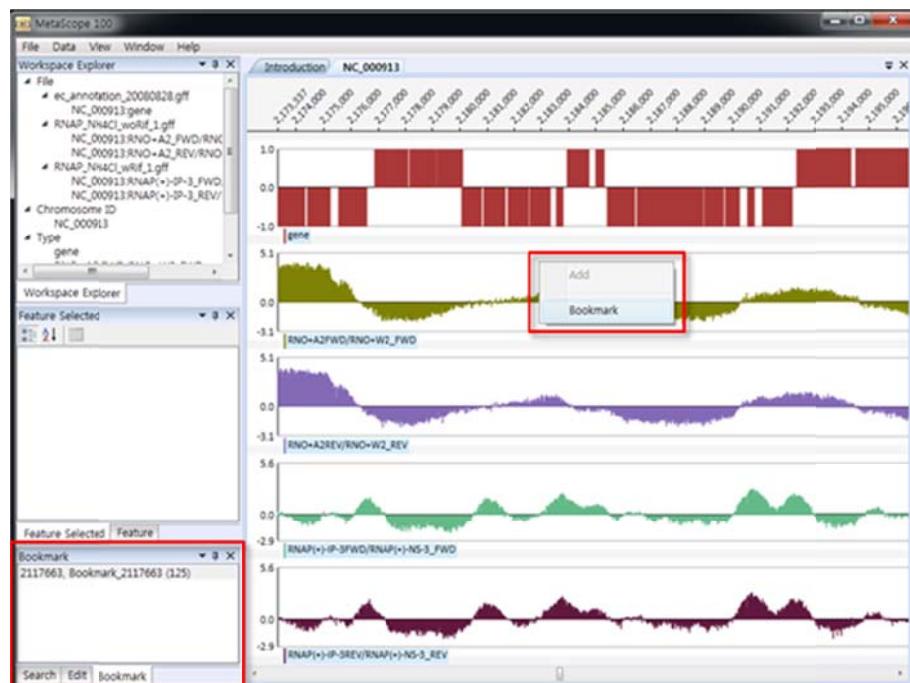


Figure 27. Bookmark context menu

Selecting items

Selecting chromosome tabs

Each chromosome tab can be selected and focused by clicking the label of chromosome tab, showing the chromosome ID.

Selecting data tracks

Clicking the header region, emphasized by red box in Figure 28, selects one data tracks from multiple data tracks displayed in one chromosome tab. Selected data track is distinguished by blue header region of each track.

In order to select more than two data tracks, first select one by clicking the header region, and then clicking with shift key the header regions of other data tracks. Figure 28 shows two data tracks are selected and the other three are not.



Figure 28. Selecting data tracks

Selecting one data feature

One data feature can be selected by simple mouse clicking over the data feature of interest. In Figure 29, one gene in the reverse strand of the *E. coli* genome is selected and indicated by black color.

When one data feature selected, Selected Feature window shows information about the selected feature, as shown in the left side of Figure 29.

Selecting multiple data features

Multiple data features can be selected by mouse dragging, as shown in Figure 31. When user is dragging the mouse, a black rectangle shows up to indicate which region will be selected by dragging.

MetaScope supports selecting multiple data feature of multiple data tracks by one mouse dragging, and requires marking data tracks for selecting multiple features and further data processing. One way of setting data track “editable”, which means this track is susceptible for selecting features and data processing, is double clicking the empty region of the track, or using track header context menu “select to Edit”, as shown in Figure 30. Editable data tracks are indicated by black rectangle surrounding the data tracks.

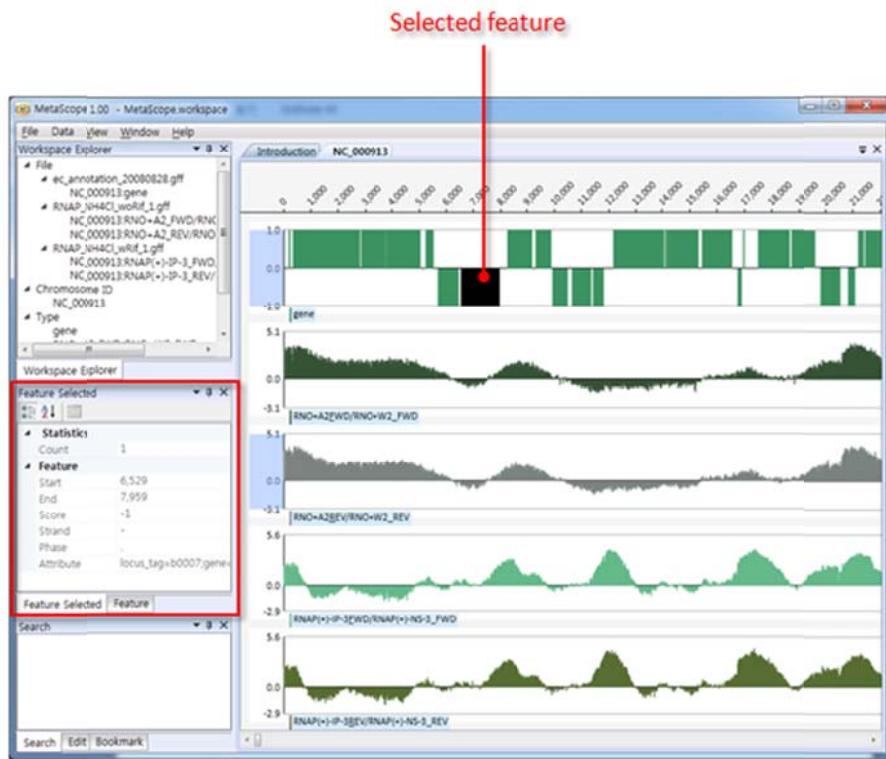


Figure 29. Selecting one data feature

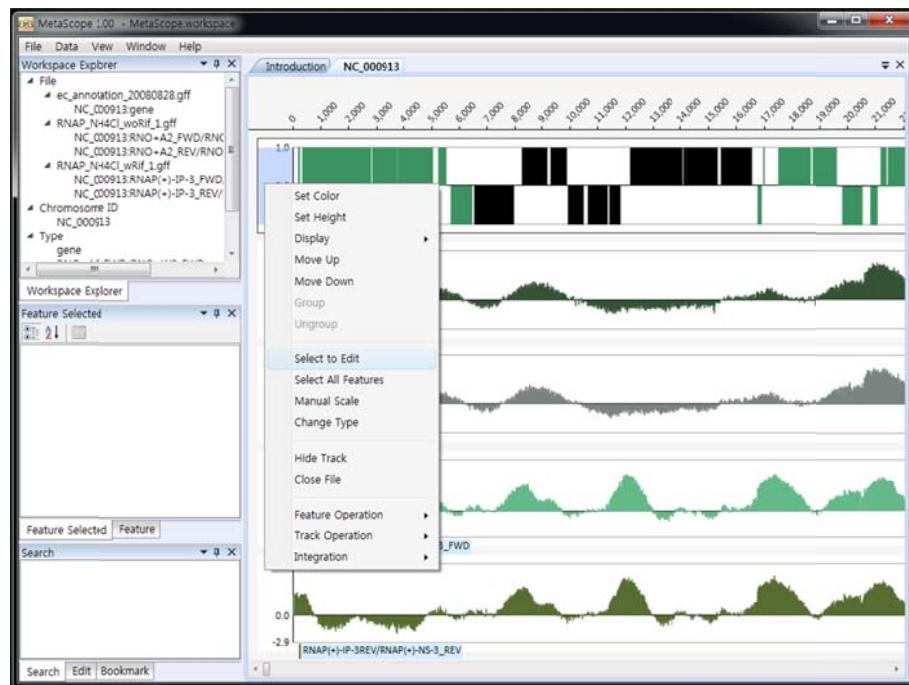


Figure 30. Setting data track editable

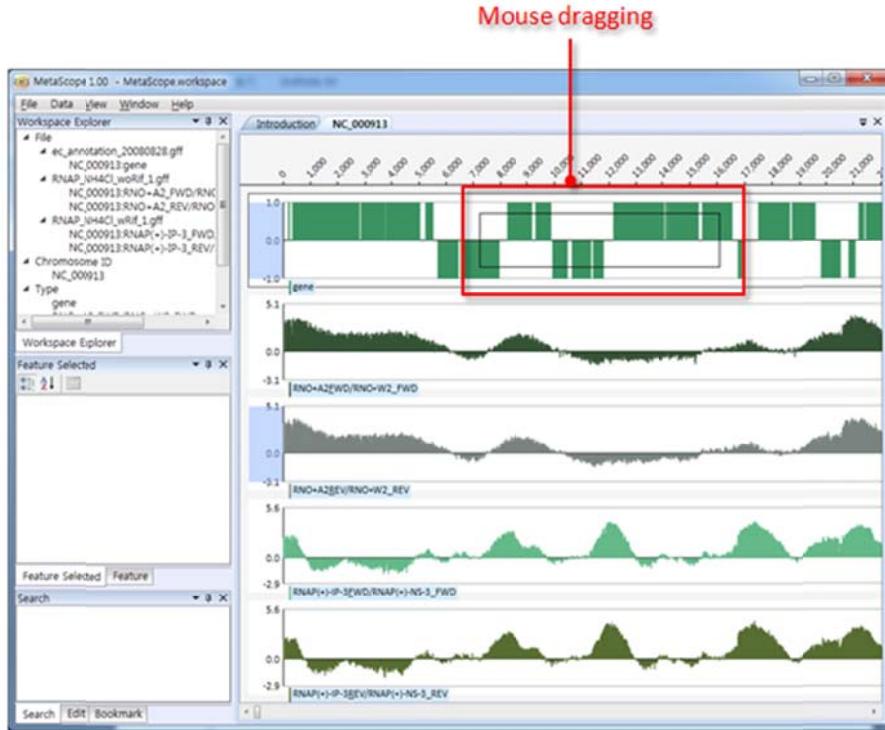


Figure 31. Selecting multiple data features by dragging

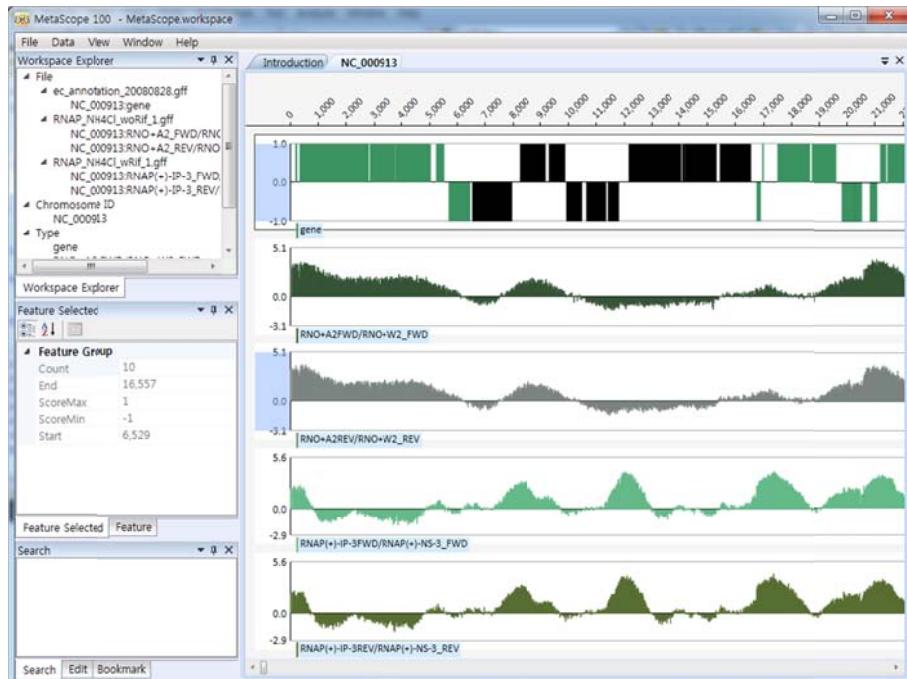


Figure 32. Selected multiple data features

As similar to selecting one data feature, when multiple data features are selected, they are displayed with block color, as shown in Figure 32.

Selecting data tracks and data features is the preliminary step for feature operation, track operation and integration function, which will be discussed in detail later.

Searching

MetaScope supports searching function with which user can search every data features loaded in MetaScope by a keyword like gene name or gene category. Figure 33 shows “Search” dialog box, where user can select a keyword to search with, and whether to use the current chromosome ID or all chromosome IDs.

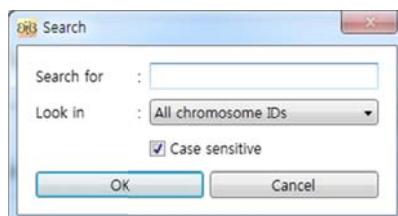


Figure 33. Search dialog box

Figure 34 shows the search result with a keyword of “oxyR”. The Search window shows one matching search result, and double clicking that result moves the genomic position of that data feature and highlights it with transparent blue box.

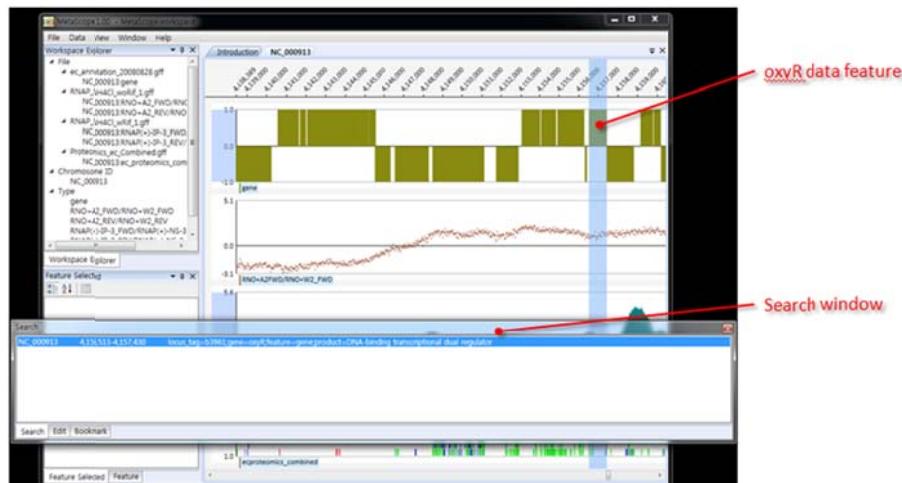


Figure 34. Search result

Customization

Window and tab customization

Every windows and chromosome tabs are rearrangeable as shown in Figure 3, Figure 4, Figure 16, and Figure 17.

The configuration of style and position of each window can be stored in layout file “.xml” in xml format, and can be used to restore the overall layout configuration in the future, by using “File > Open Layout” and “File > Save Layout As” menu, as shown in Figure 35.

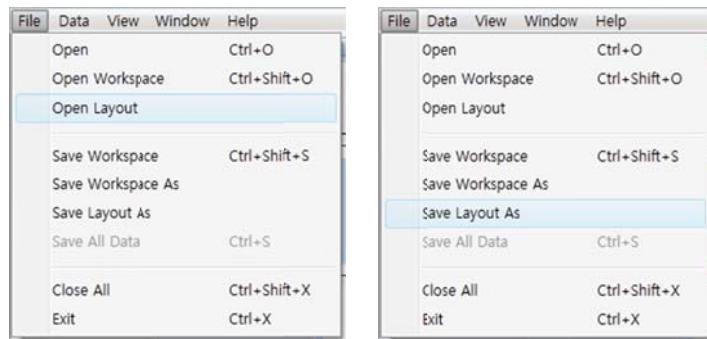


Figure 35. “File > Open Layout” and “File > Save Layout As” menu

Data track customization

Color selection

Data feature can define the color of itself by stating color information in “color=RRGGBB” format in GFF file. If there is no color information about data feature, then MetaScope used randomly chosen color information which is decided when data file is loaded. For instance, 5 separate data tracks are decorated with 5 different colors in Figure 32.

However, user can pick a color for a data track with which data features in that track will be drawn. Figure 36 shows the “Set Color” context menu, by which user can bring up “Pick Color” dialog box, as shown in Figure 37. When user drags the white circle inside the left box, he or she can choose a color to draw data features with. After clicking OK button in Pick Color dialog box, the data track will be automatically updated with a new color.

One thing important to remember is this color information for data track does not override color information described in the attribute section of data feature.

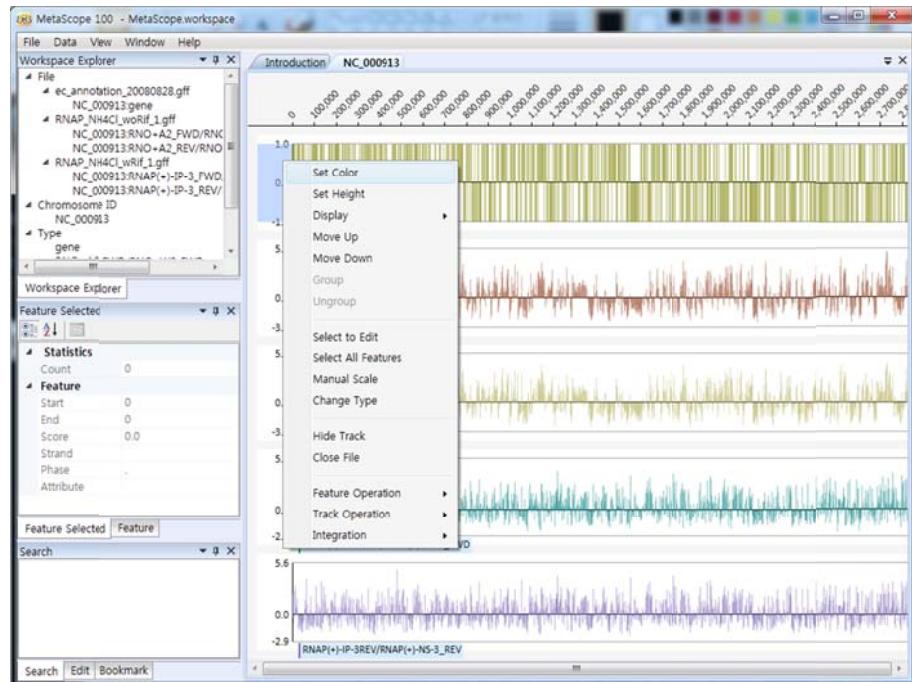


Figure 36. Set Color context menu

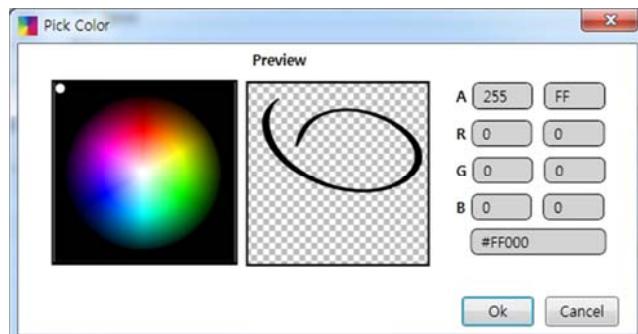


Figure 37. Pick Color dialog box

Setting height

The height of each data track is also configurable, by using “Set Height” context menu, as shown in Figure 38. This menu brings up Set Height dialog box, as shown in Figure 39. In this dialog box, user can choose either automatic setting or manual setting. The minimum height of data track is 50 pixels. Thus any height below this limit cannot be set.

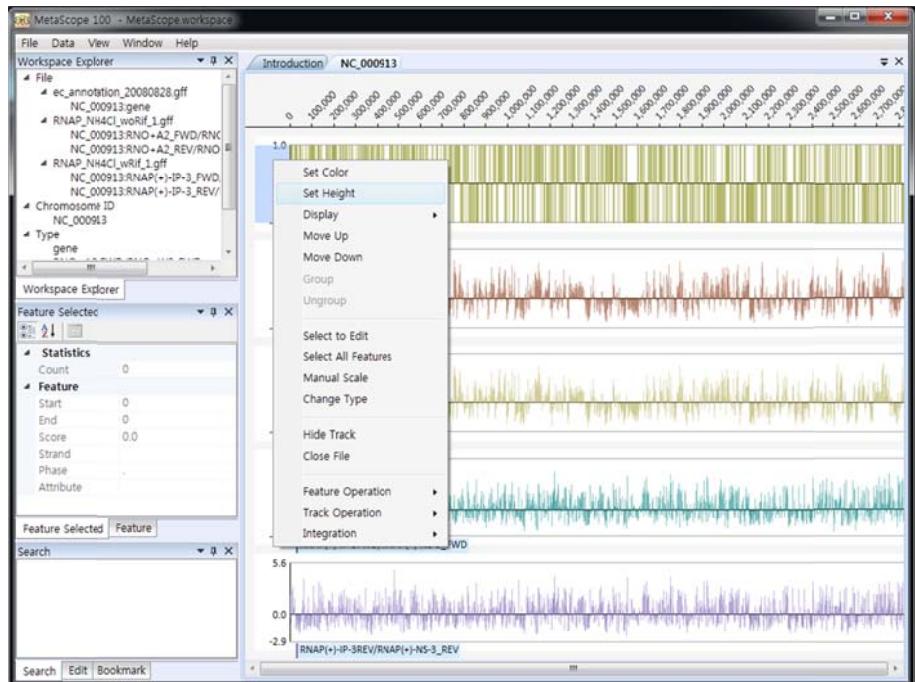


Figure 38. Set Height context menu

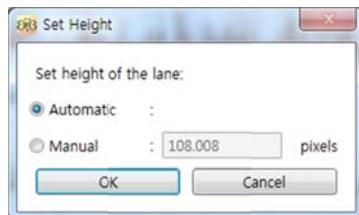


Figure 39. Set Height dialog box

Display style

MetaScope supports 4 display styles: bar, point, line, and stack. Bar style displays every data feature as a box with a height matching score value of that feature, and point style shows feature with a single point corresponding score value of that feature. Line style is similar to point style, but it shows line between every adjacent points. Stack style shows overlapping data features as a stack, so that user can figure out which data feature is overlapping with which one, and this style is good for displaying proteomic data.

Figure 40 shows those 4 display styles. The first track shows *E. coli* genome annotation with bar style. The second and third tracks display RNA polymerase ChIP-chip data by point style and line style, accordingly. The fourth track shows proteomic data with stack style.

These display style for one or more data tracks can be set by using Display context menu, as shown in Figure 41.

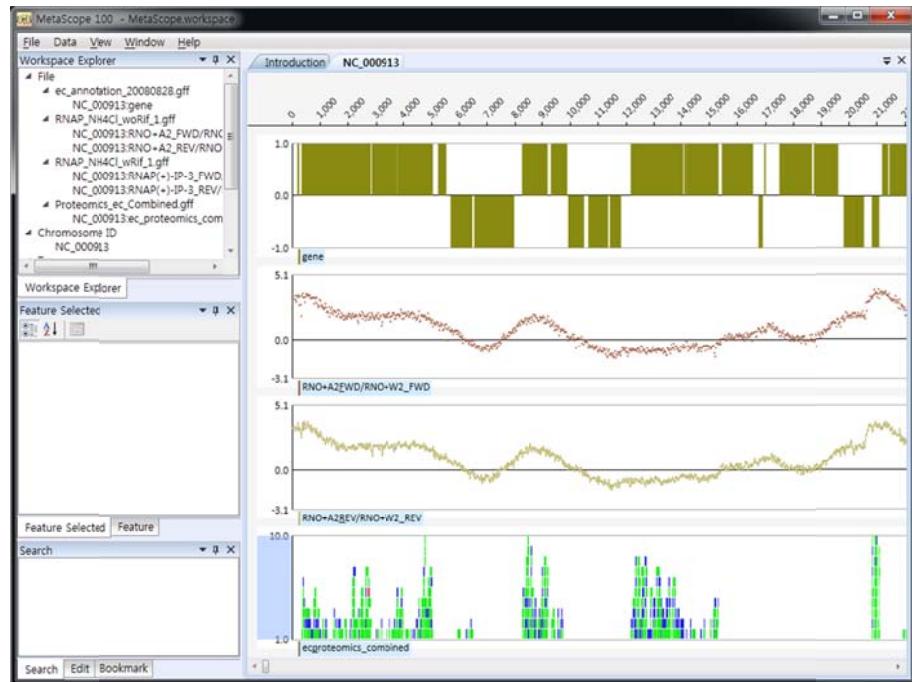


Figure 40. 4 display styles

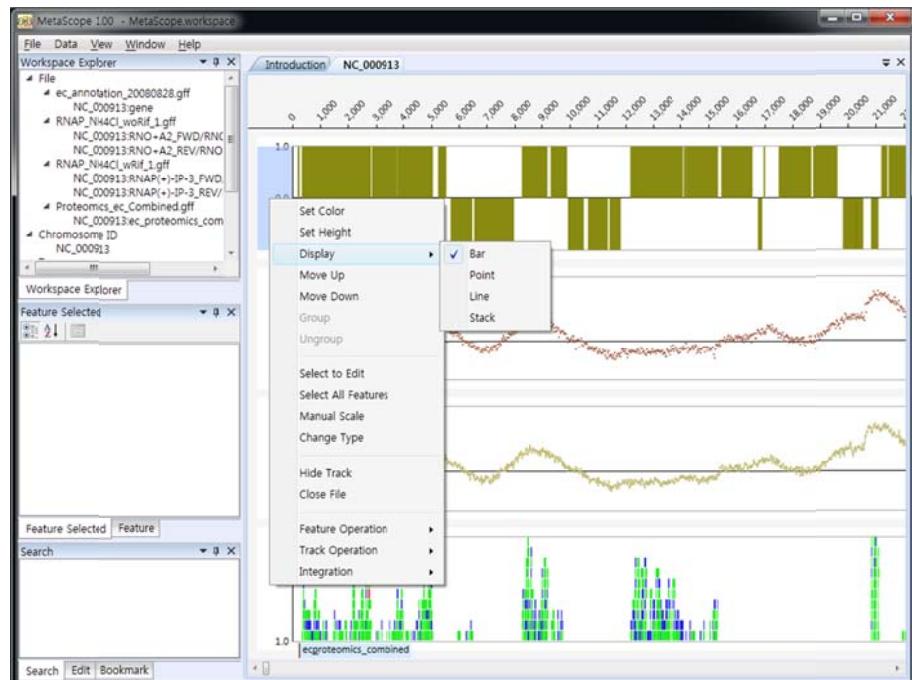


Figure 41. Display context menu

Track rearrangement

The vertical order of data tracks is automatically determined by the order of dataset loading. However, user can rearrange those data track in a new order he or she wants. One way of doing this is using Move Up and Move Down context menu, as shown in Figure 42.

Another way is user can drag and drop one or more data tracks he or she wants to move.

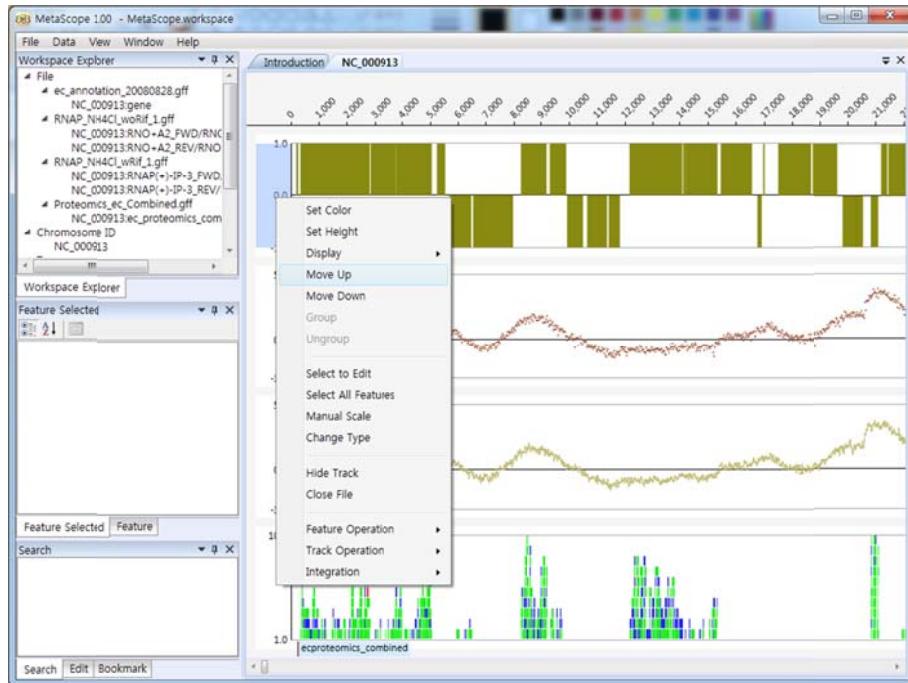


Figure 42. Move Up and Move Down context menu

Grouping and ungrouping

Two or more data tracks can be grouped, so that they can be visualized as in one data track, as shown in the second track of Figure 43. The second track shows two ChIP-chip datasets, one for RNA polymerase static binding (treated with rifampicin) and the other for RNA polymerase dynamic binding (without rifampicin).

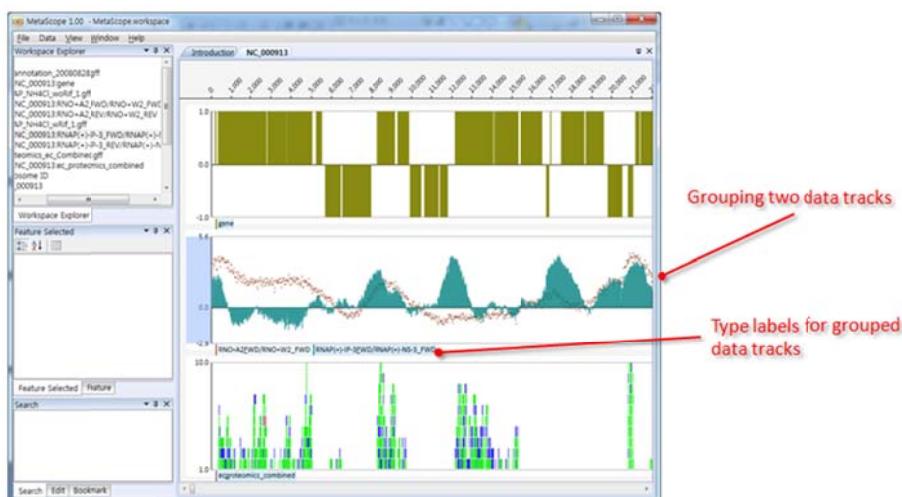


Figure 43. Grouping of two data tracks

As shown in Figure 43, data type labels for grouped data tracks are displayed in order at the bottom of the track. By double-clicking the label, user can choose one data type from multiple data types overlapped here, so that he or she can indicate which data type should be applied by following data processing and integration.

When grouping multiple data track, those tracks must be selected and shown with light blue header, as shown in Figure 44.

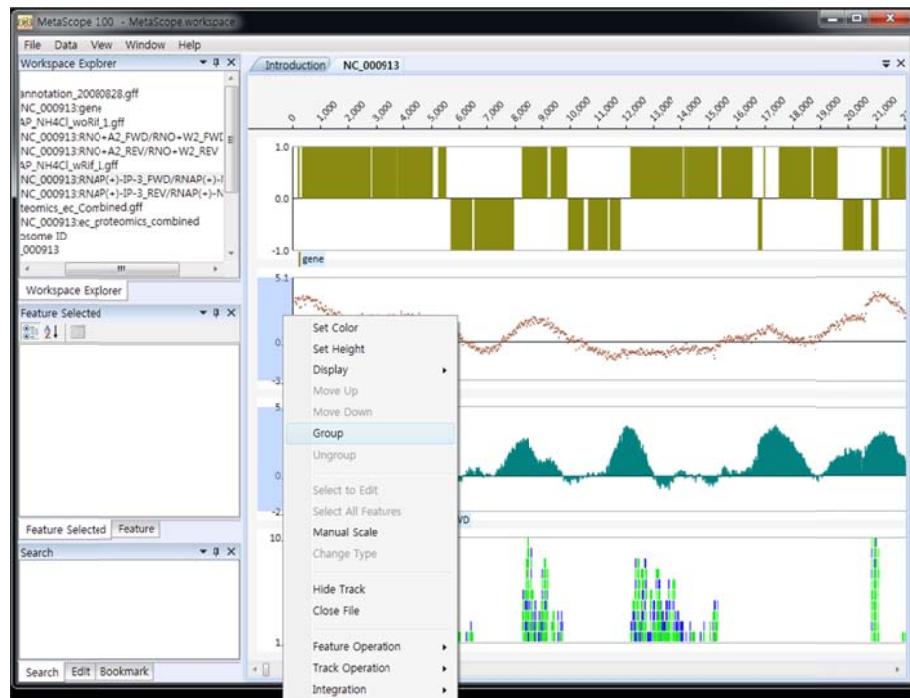


Figure 44. Group and Ungroup context menu

Setting editable

This is explained above in data feature selection. User can use “Select to Edit” context menu or double clicking the empty region of the track in order to set the data track editable.

Scaling manually

The range of scores of data features can be scaled manually so that user can narrow down to the signals range of interest. “Manual Scale” context menu as shown in Figure 45. This menu brings up “Set Scale” dialog box, where user can select no scaling or manual scaling with specific minimum score and maximum score to display, as shown in Figure 46. When the OK button is clicked, MetaScope automatically update the data track with new scaling information.

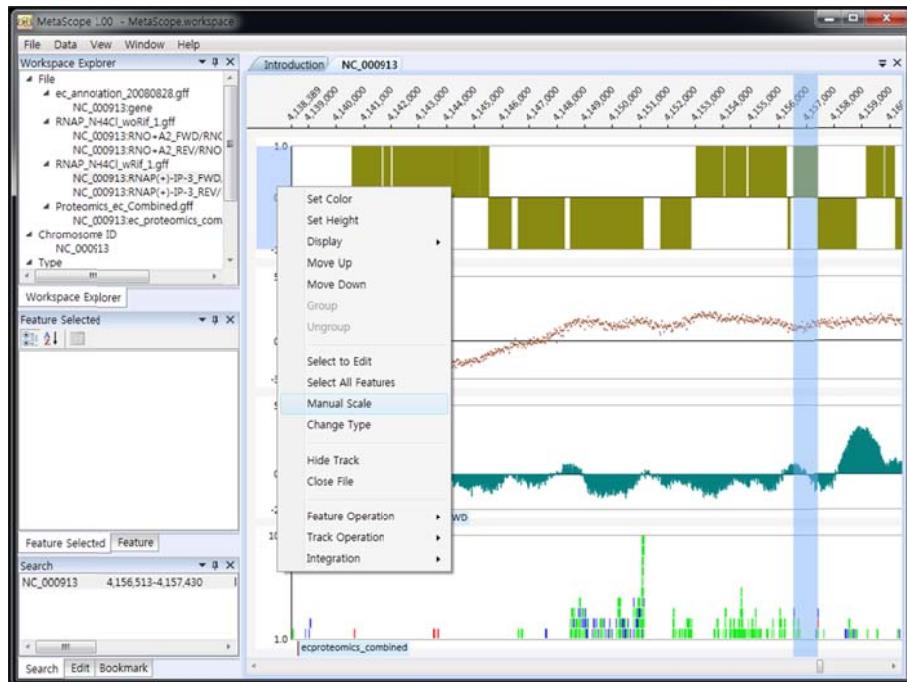


Figure 45. Manual Scale context menu

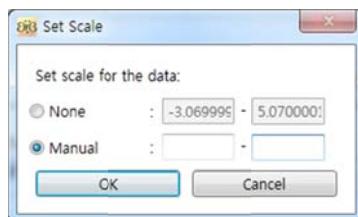


Figure 46. Set Scale dialog box

Data type changing

The data type of data track can be changed, by using “Change Type” context menu, as shown in Figure 45. This menu brings up “Change Type” dialog box, as shown in Figure 47. In this dialog box, user can check the current type, and set a new type.

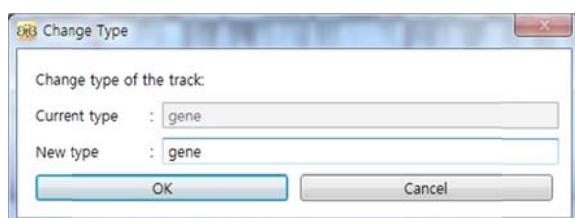


Figure 47. Change Type dialog box

Hiding tracks and closing files

User can choose either to display one or more data tracks by using “Hide Track” context menu, as shown in Figure 48. Workspace explorer can be used to show hidden tracks back, by double clicking data type listed in workspace explorer. Even if the data track is hidden, it does not mean the actual data file containing data feature for that data track is unloaded. Thus user can choose to show the hidden tracks any time he or she wants.

In order to hide the data track and close the data file for good, “Close File” context menu can be used. One thing important to remember about this is if one data track is selected to close the file that containing that data track, it will also hide every data tracks that belong to that particular data file.

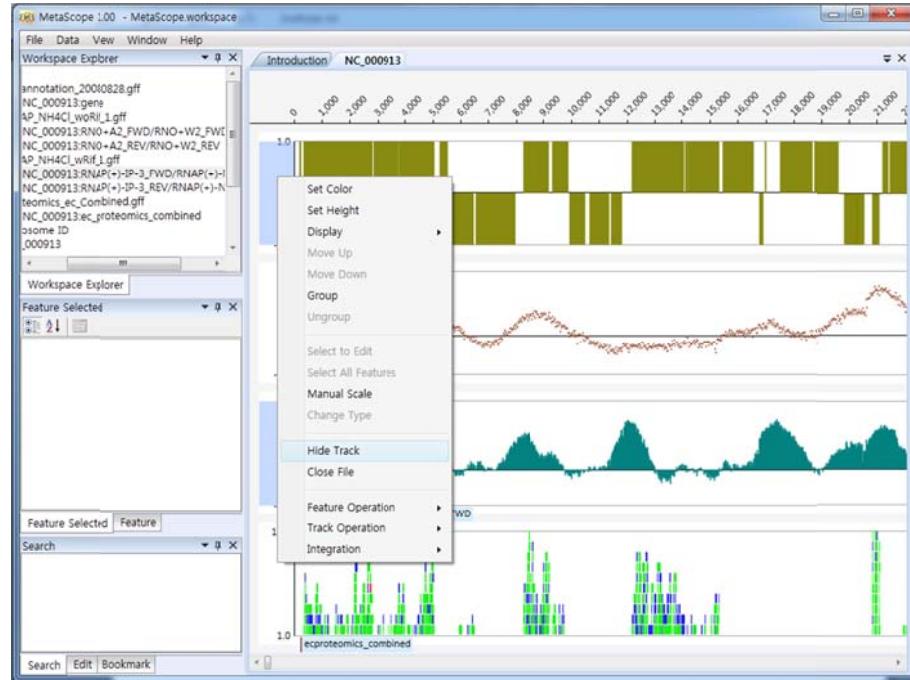


Figure 48. Hide Track and Close File context menu

Workspace file

All configuration explained here can be saved into a “.workspace” file, which is in xml format, and can be used to restore the configuration back. Thus user can use this function usefully, in order to manage multiple projects with different data files and configuration setting for data track and windows.

Feature operation

MetaScope supports 7 functions to handle data in a feature level: uniting, merging, filtering, moving, copying, editing and deleting. These functions can be useful when curating processed datasets like calculated ChIP-chip binding signals, transcript detection signals, transcription start site signals from RNA-seq and building up an annotation from those experimental evidences.

These functions are accessible by context menu on data features, as shown in Figure 49. In order to activate this context menu, one or more data features should be selected by the methods described data feature selection section above, and the data track containing those data features must be set editable. In Figure 49, only one data feature is selected, and all 7 feature level functions are activated.

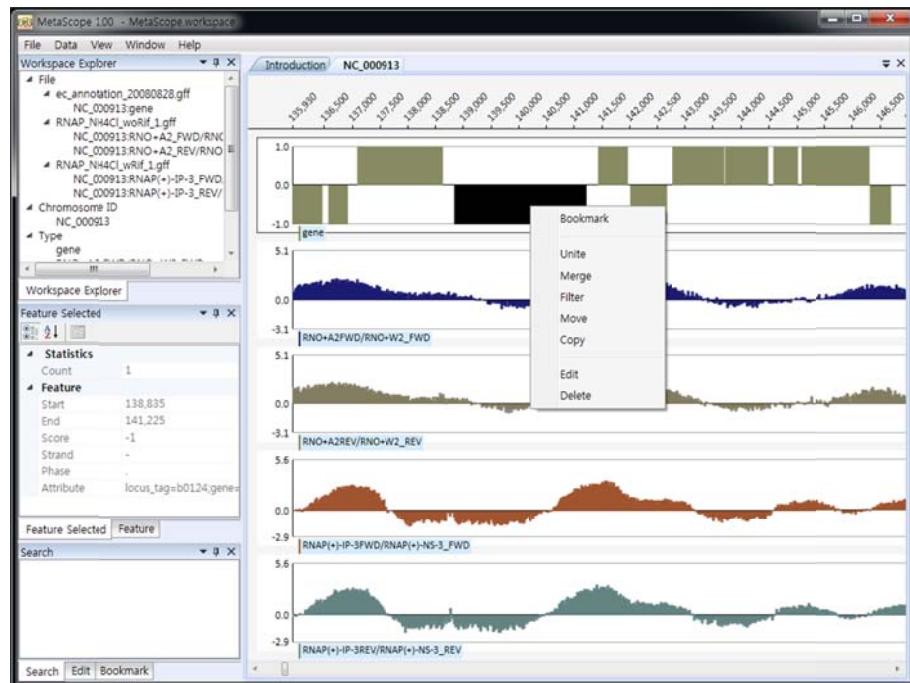


Figure 49. Context menu for feature operation functions

Processing multiple data features

Uniting

Multiple data features in the same strand can be united in to one data feature. In order to unite two or more data features, those data features should be selected first, and click "Unite" context menu.

As shown in Figure 50, this function is useful when processing discrete transcript detection signals into a contiguous transcript segment. One of integration function, building RTS (RNAP-guided transcript segment) from TD (transcript detection) and RBR (RNAP binding region), mainly does this job, however there might be some cases where manual uniting is required, because defining contiguous transcript segment by only

depending RBR information is not complete, especially user is dealing with compact genome, like the genome of *E. coli*.

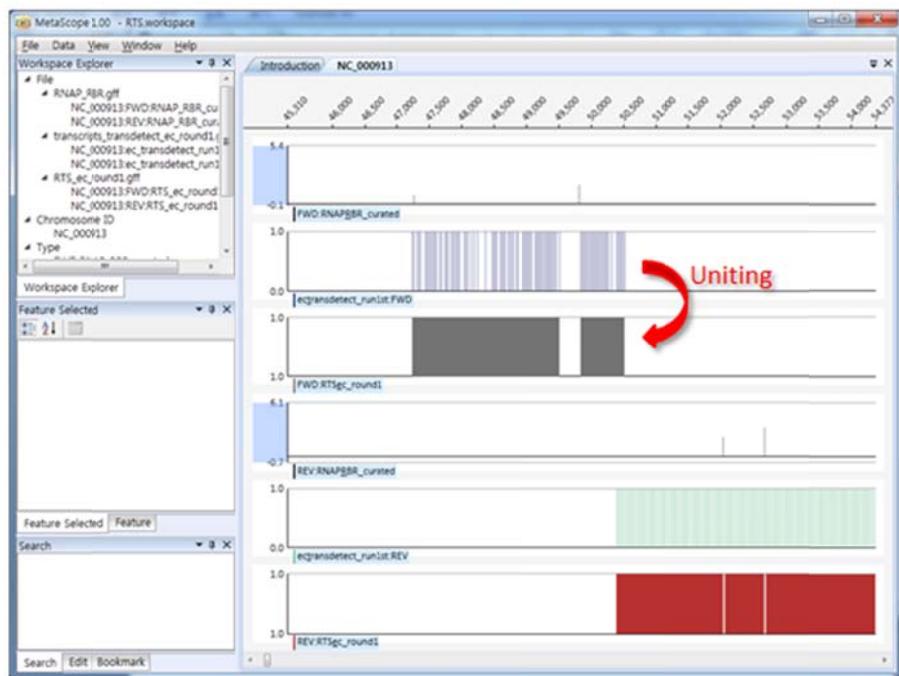


Figure 50. Uniting multiple data features

Merging

Multiple data features overlapped in two or more data tracks can be merged to generate average data features.

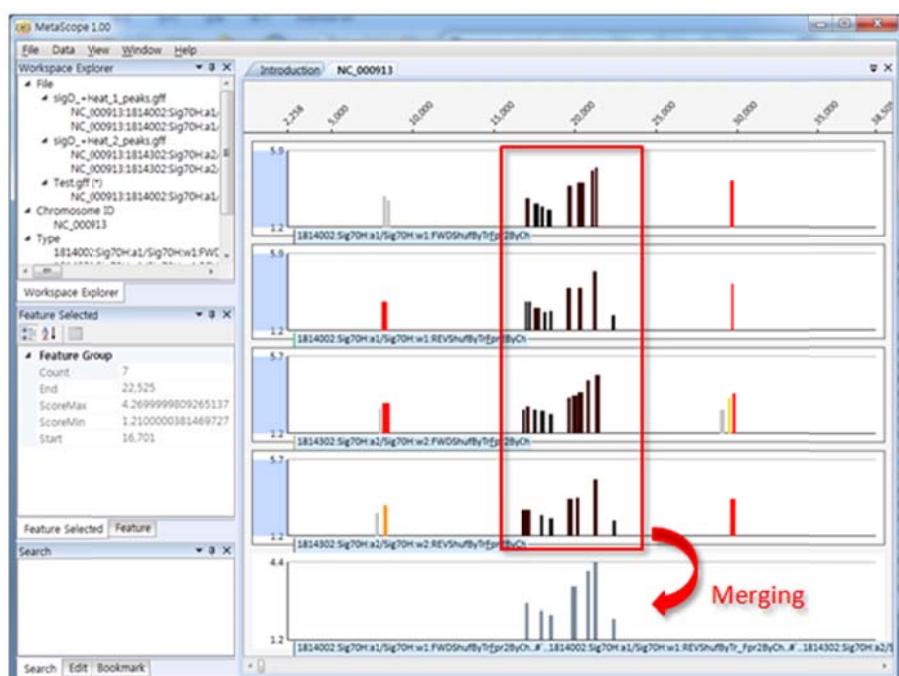


Figure 51. Merging multiple data features

As shown in Figure 51, this function makes average data features from transcription factor binding regions from ChIP-chip experiments for biological replicates. Average data features have average start position, end position and score.

In order to use merge function, multiple data features in several data tracks should be selected first, and clicking “Merge” context menu brings up “Feature Operation” dialog box with “Merge” tab activated, as shown in Figure 52. In this tab, user can choose average or median to calculate those three values from overlapping data features. The newly created average data features can be saved into one of loaded files or into a new one.

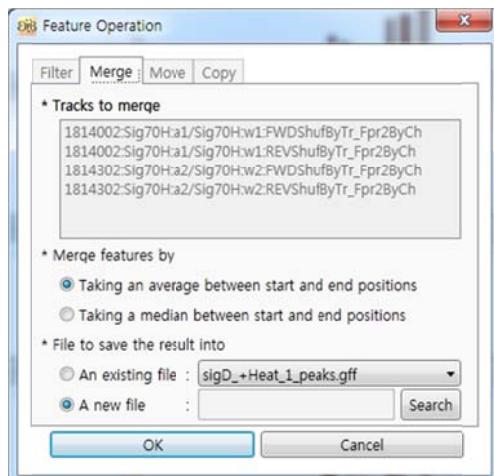


Figure 52. Merge tab in Feature Operation dialog box

Filtering

Multiple data features can be filtered in order to rule out data features that are not significant or less significant in terms of analyzing data.

For example, Figure 53 show raw transcription start site data generated by modified 5' RACE (Rapid Amplification cDNA Ends) followed by deep sequencing. User might be interested in getting major signals, because he or she is interested in finding core promoter motif. In Figure 53, raw TSS data features are filtered out by leaving the only strongest signal in the sliding window with the size of 6.

In order to filter out multiple data features with various criteria, those data features should be selected first, and clicking “Filter” context menu brings up “Feature Operation” dialog box with “Filter” tab on. In this tab, user can see how many data features are selected, and choose one of 6 criteria to filter out with. Selecting N number of top features and selecting N% of top features can be executed with or without sliding window, as shown in Figure 54.



Figure 53. Filtering multiple data features



Figure 54. Filter tab in Feature Operation dialog box

Moving and copying

One or more data features of one data track can be moved or copied into another data track. Figure 55 shows “Move” and “Copy” tabs in “Feature Operation” dialog box.

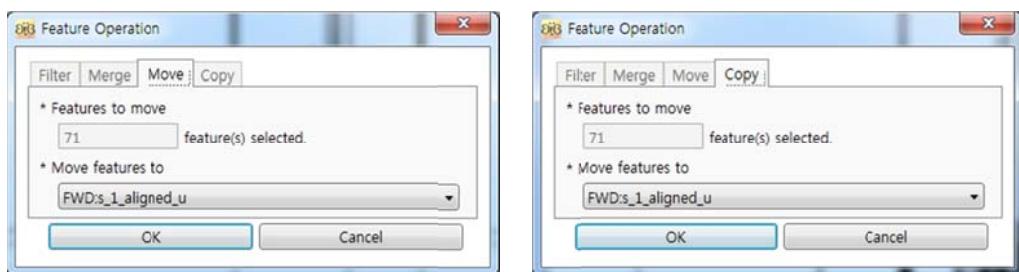


Figure 55. Move and Copy tabs in Feature Operation dialog box

In order to move or copy multiple data features, those data features should be selected first, and clicking “Move” or “Copy” context menu brings “Feature Operation” dialog box

with “Move” or “Copy” tab accordingly.

Processing single data features

Editing

One data feature can be edited, by using “Edit” context menu. Clicking “Edit” context menu brings up “Edit a Feature” dialog box, as shown in Figure 56. In this dialog box, user can edit information about the data feature of interest.

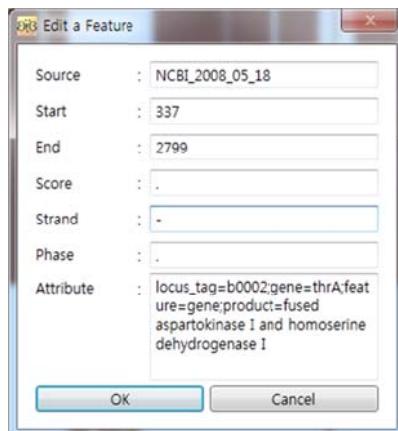


Figure 56. Edit a Feature dialog box

Deleting

One data feature can be deleted by using “Delete” context menu. Multiple data features can be deleted by using “Data > Feature > Delete” application menu or “Ctrl+D” shortcut.

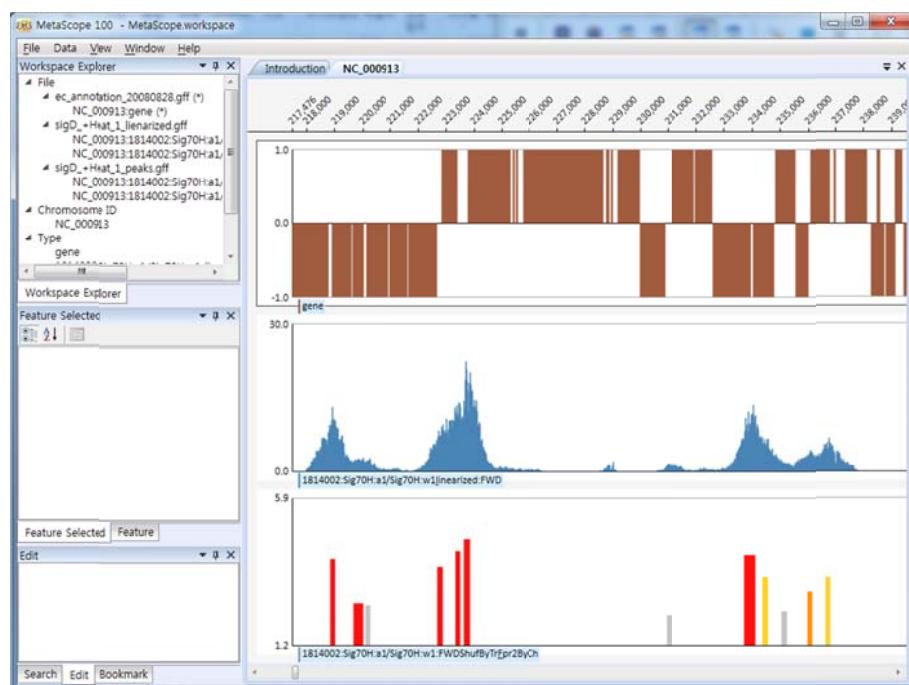


Figure 57. Deleting data features

This function can be used for processing and analyzing ChIP-chip binding signals. The second track in Figure 57 shows unprocessed ChIP-chip data and the third track displays calculated ChIP-chip binding regions. Grey boxes in the third track represent calculated binding regions with lower P-values. Thus user might want to remove those binding signals using this function, in order to more focus on more significant binding signals.

Track operation

MetaScope supports 7 functions to handle data in a track level: averaging, differencing, summing, merging, filtering, adjusting and assigning IDs. These functions can be useful when processing and analyzing unprocessed datasets like expression profiling, ChIP-chip data, transcription start site data by 5' RACE-seq, and proteomics data.

These functions are accessible by context menu on header region of data tracks, as shown in Figure 58. In order to activate this context menu, one or more data tracks should be selected by the methods described data track selection section above. In Figure 58, two data tracks, the second and third ones, are selected, and all 7 track level functions are activated.

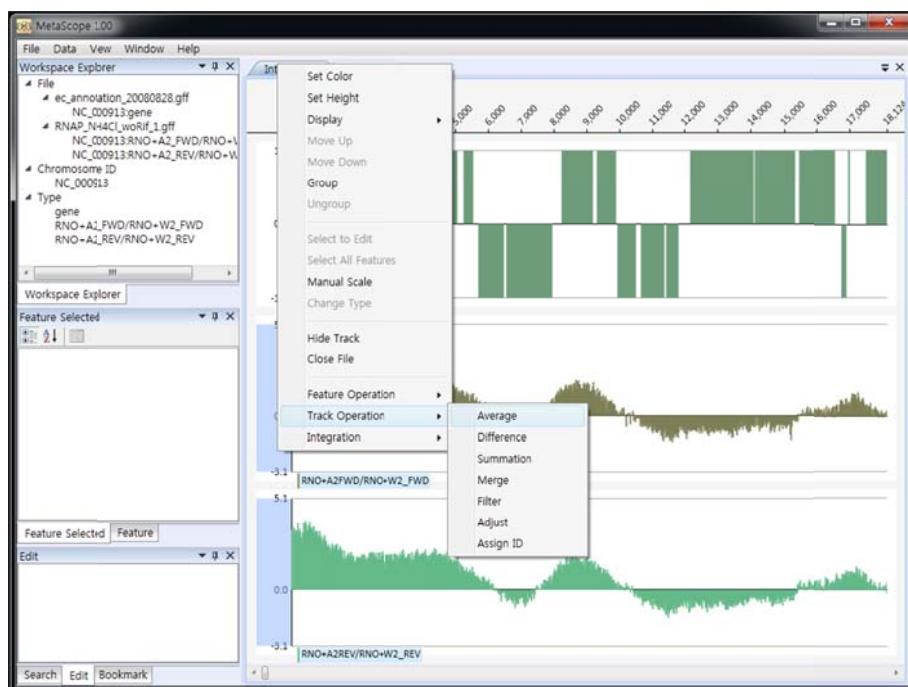


Figure 58. Context menu for track operation functions

Averaging

Unprocessed raw datasets including expression profiling, ChIP-chip data and TSS data are generally generated from biological replicates. Thus at some stage of data analysis, these datasets should be merged in order to build transcript detection signals for expression profiling, calculate binding regions for ChIP-chip data, and produce reproducible signals for TSS data. One of merging those datasets from biological replicates is taking average for matching data features.

Figure 59 shows one example of taking average from two biological replicates of ChIP-chip data. The second and third tracks represent two biological replicates, and the fourth track shows average of those two data tracks.

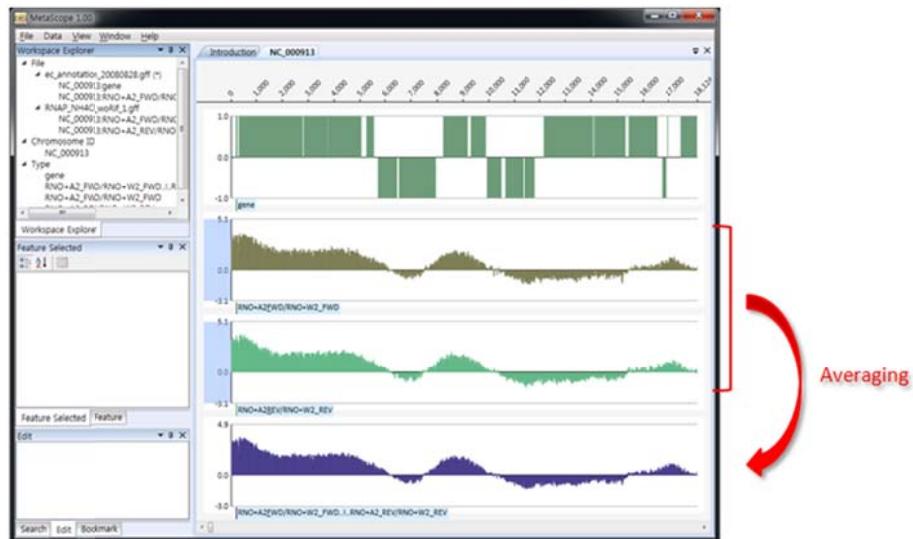


Figure 59. Averaging data tracks

In order to average two or more data tracks, those data tracks should be selected by following the way of selecting data tracks described above, and clicking “Average” context menu. This brings up “Track Operation” dialog box with “Average” tab activated, as shown in Figure 60.

In this tab, user can see multiple data tracks selected to make average, and choose several options for making average. The newly generated data features can be stored in one of existing data files or into a new one.

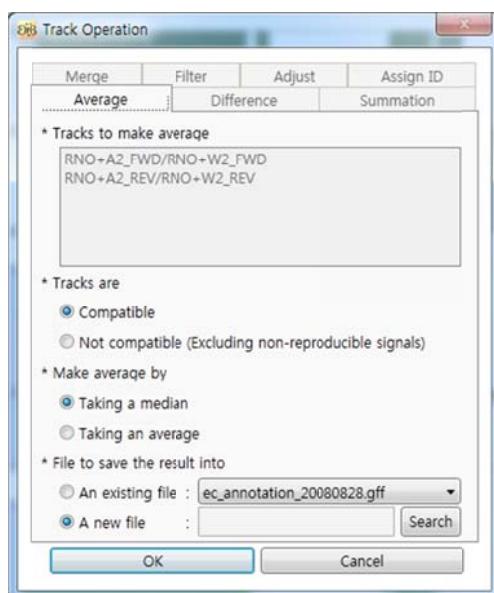


Figure 60. Average tab in Track Operation dialog box

Differencing

How differencing works is basically same as averaging, except for differencing calculates differences between data features from biological replicates. Figure 61 shows

differencing two data tracks from biological replicates. This function can be used for analyzing multiple datasets from biological replicates and feature out and visually see how much differences there are between those replicates.

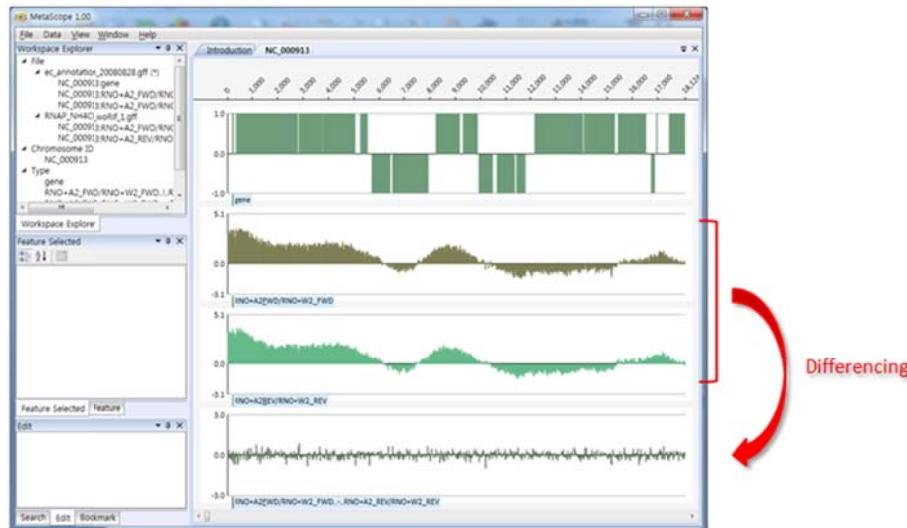


Figure 61. Differencing data tracks

In order to difference two or more data tracks, those data tracks should be selected by following the way of selecting data tracks described above, and clicking “Difference” context menu. This brings up “Track Operation” dialog box with “Difference” tab activated, as shown in Figure 61.

In this tab, user can see multiple data tracks selected to make difference. The newly generated data features can be stored in one of existing data files or into a new one.

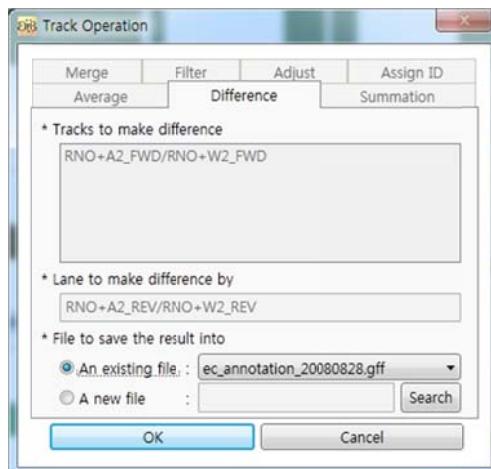


Figure 62. Difference tab in Track Operation dialog box

Summing

Summing works in a similar way of averaging and differencing. This function can be used for adding up all signals from biological replicates.

In order to sum two or more data tracks, those data tracks should be selected by following the way of selecting data tracks described above, and clicking “Summation” context menu. This brings up “Track Operation” dialog box with “Summation” tab activated, as shown in Figure 63.

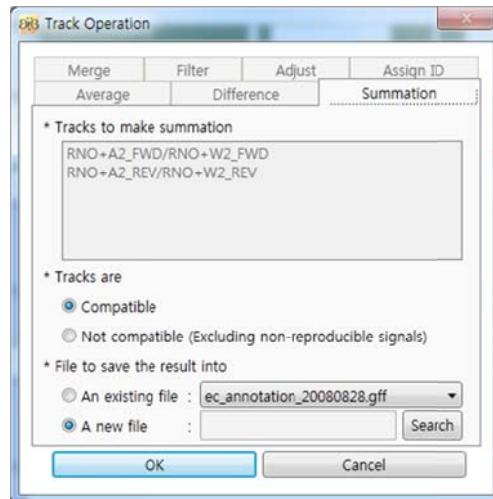


Figure 63. Summation tab in Track Operation dialog box

Filtering

Data track filtering is different from data feature filtering described above. They are same in terms of taking out data features which do not meet a given criteria, but data feature filtering applies to selected data features, and data track filtering deals with whole data features in data tracks. One more difference is one track is chosen to be a filter data track, and the other tracks are subject to be filtered by that filter track.

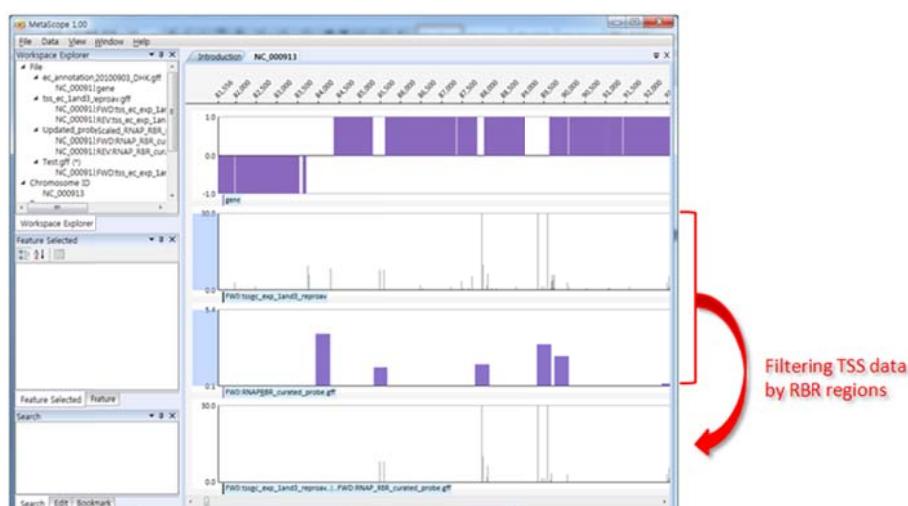


Figure 64. Filtering data tracks

Figure 64 shows one example of data track filtering. The second track in Figure 64 is TSS data and the third track is curated RBR (RNAP binding region). If user wants to leave TSS signals lying inside those RBR regions, then he or she can filter TSS data track by RBR

data track, as shown in Figure 64. The fourth track of Figure 64 represents filtered TSS signals, meaning TSS data that lie in those RBR regions.

In order to perform data track filtering, first select data tracks to filter out and then select the filter data track. After this, clicking “Filter” context menu brings up “Track Operation” dialog box with “Filter” tab activated, as shown in Figure 65. User can choose whether filter out or leave data features inside filter regions.

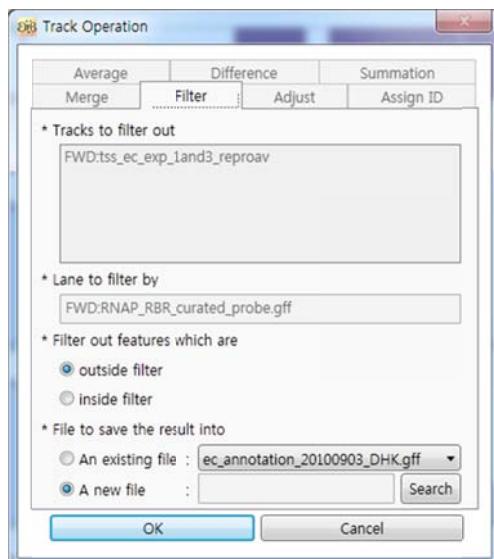


Figure 65. Filter tab in Track Operation dialog box

Adjusting

Score and width of data features in one or more data tracks can be adjusted, by using “Adjust” context menu. Clicking “Adjust” context menu brings up “Track Operation” dialog box with “Adjust” tab activated, as shown in Figure 66.

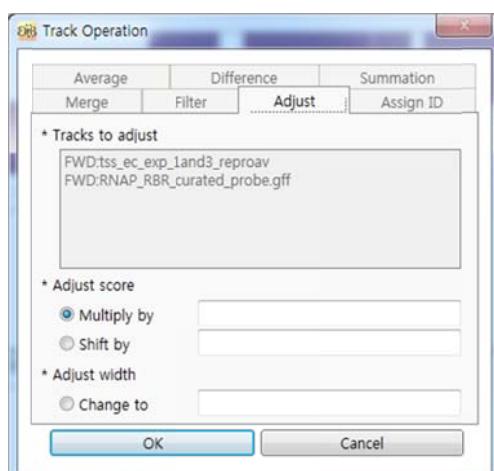


Figure 66. Adjust tab in Track Operation dialog box

Assigning ID

During the course of analyzing, processing and integration of multiple datasets, it might

be required to have a set of IDs for data features. “Assign ID” context menu addresses this problem, and clicking this menu brings up “Track Operation” dialog box with “Assign ID” tab activated, as shown in Figure 67.

This ID will be stored in the attribute section of data feature and GFF file, in the format of “ID=value”.

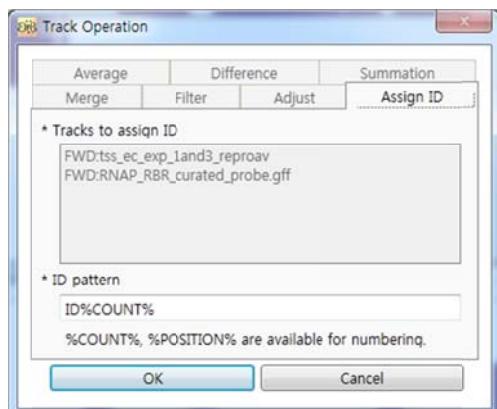


Figure 67. Assign ID tab in Track Operation dialog box

Integration function

MetaScope supports three integration functions: building RTS, pORF and TU. Building RTS (RNAP-guided transcript segment) requires RBR (RNA polymerase binding region) data and TD (transcript detection) data from expression profiling. Building pORF (potential ORF) requires start and stop codon positions and proteomic data. Building TU (transcription unit) requires RTS, TSS (transcription start site) data and pORF information.

Once these datasets are ready, user can build TU annotation using these datasets through step-by-step process, which is described below

Integrating RBR and TD to build RTS

In order to build RTS, datasets for process RBR and TD data should be loaded in MetaScope, as shown in Figure 68. “Integration > RTS” context menu can be used to perform RTS integration, as shown in Figure 68.

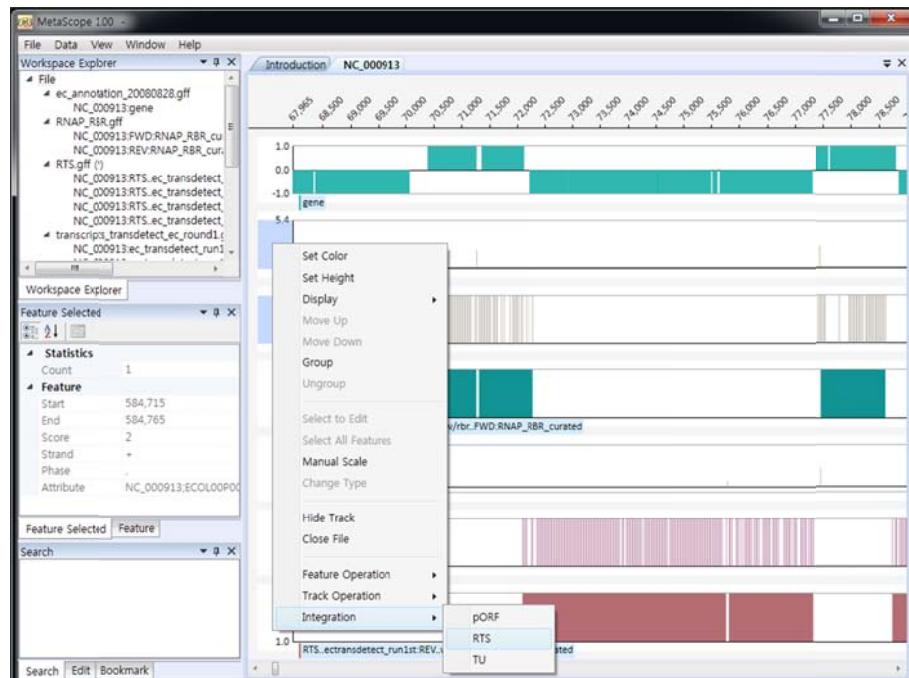


Figure 68. Context menu for RTS integration function

After selecting data tracks for RBS and TD, clicking “Integration > RTS” context menu brings up “Integration Operation” dialog box with “RTS” tab activated, as shown in Figure 69. In this dialog box, user can select tracks for RBR and TD. The resulting RTS data features can be saved in one of existing data files or into a new one.

Figure 70 shows one example of RTS integration from RBR and TD. The second track shows RBR data, and the third track represents TD data. The fourth track is RTS data for the forward strand of the genome, generated from RBR for the forward strand and TD by RTS integration function of MetaScope.

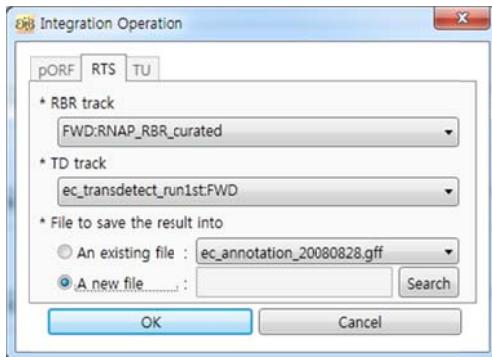


Figure 69. RTS tab in Integration Operation dialog box

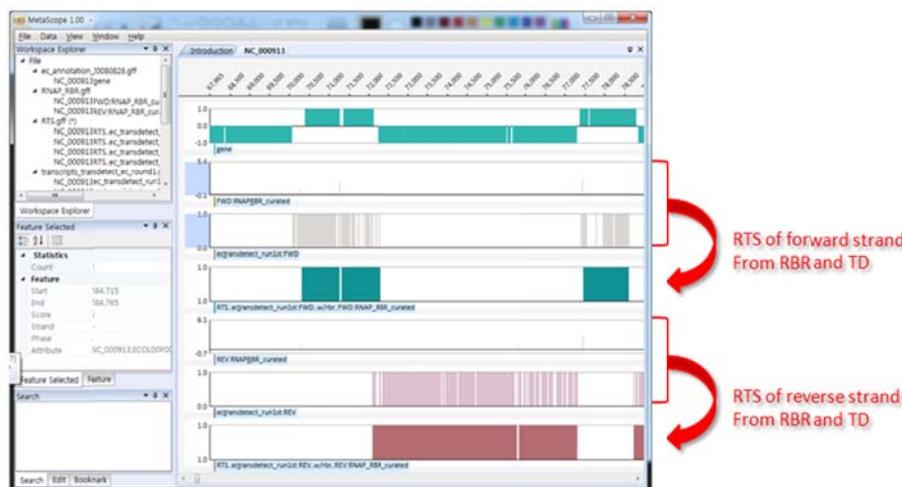


Figure 70. RTS integration from RBR and TD

Integrating start and stop codon position and proteomic data to build pORF

In order to build pORF, datasets for start and stop codon positions of a genome of interest and proteomic dataset should be loaded in MetaScope, as shown in Figure 71. “Integration > pORF” context menu can be used to perform pORF integration, as shown in Figure 71.

After selecting data tracks for start codon positions, stop codon position and proteomic data, clicking “Integration > pORF” context menu brings up “Integration Operation” dialog box with “pORF” tab activated, as shown in Figure 72. In this dialog box, user can select tracks for start codon, stop codon and proteomic data. The resulting pORF data features can be saved in one of existing data files or into a new one.

Figure 73 shows one example of pORF integration. The second track shows start codon position, and the third track represents stop codon position. The fourth track is proteomic data displayed in stack style. The fifth track is pORF data, integrated from those three datasets.

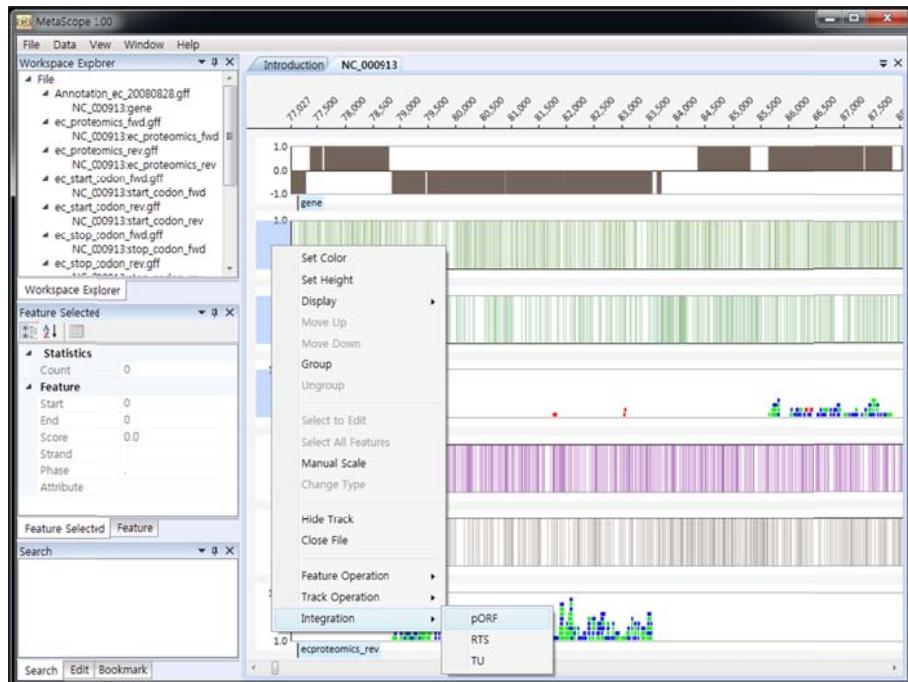


Figure 71. Context menu for pORF integration

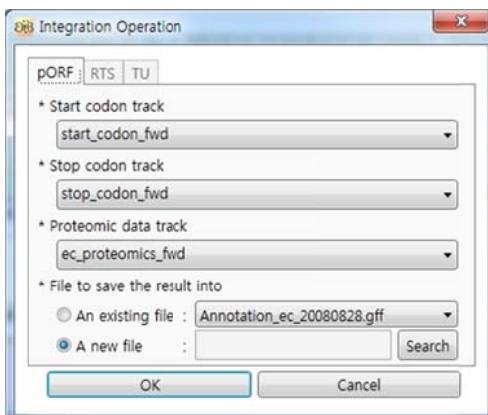


Figure 72. pORF tab in Integration Operation dialog box



Figure 73. pORF integration from start and stop codon position and proteomic data

Integrating RTS, TSS and pORF to build TU annotation

In order to build TU annotation, datasets for processed RTS, TSS and pORF dataset should be loaded in MetaScope, as shown in Figure 74. “Integration > TU” context menu can be used to perform TU integration, as shown in Figure 74.

After selecting data tracks for RTS, TSS and pORF, clicking “Integration > TU” context menu brings up “Integration Operation” dialog box with “TU” tab activated, as shown in Figure 75. In this dialog box, user can select tracks for RTS, TSS and pORF. The resulting TU annotation can be saved in one of existing data files or into a new one.

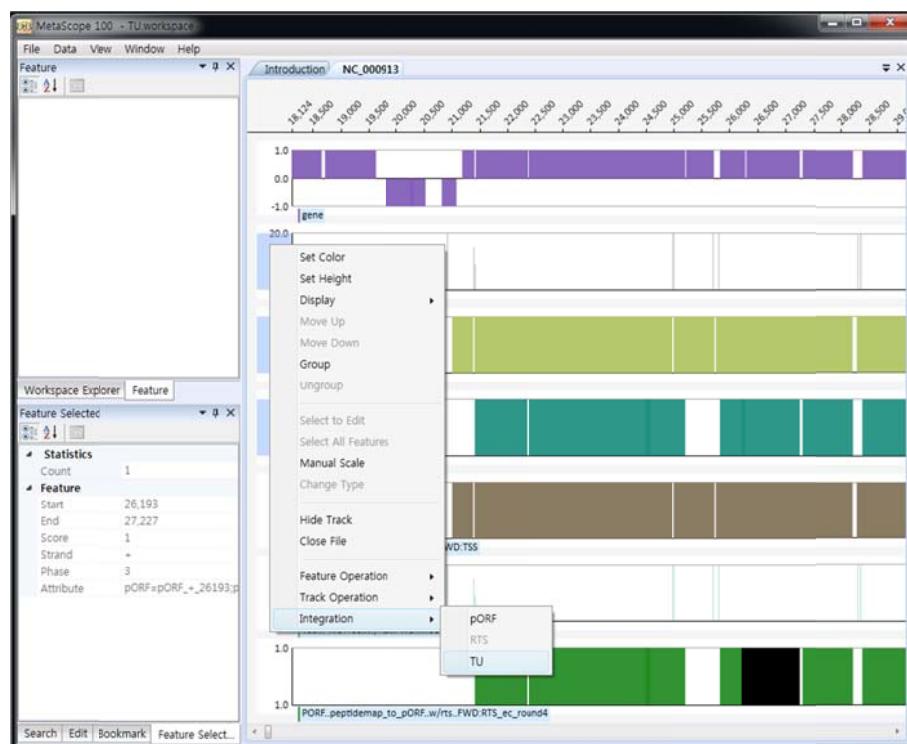


Figure 74. Context menu for TU integration

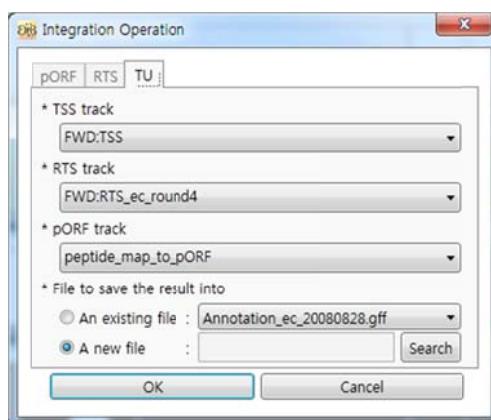


Figure 75. TU tab in Integration Operation dialog box

Figure 76 shows one example of TU integration. The second track shows TSS data, and the third track represents RTS data. The fourth track is processed pORF data. The fifth ,

sixth and seventh tracks are TU annotation, integrated from those three datasets.

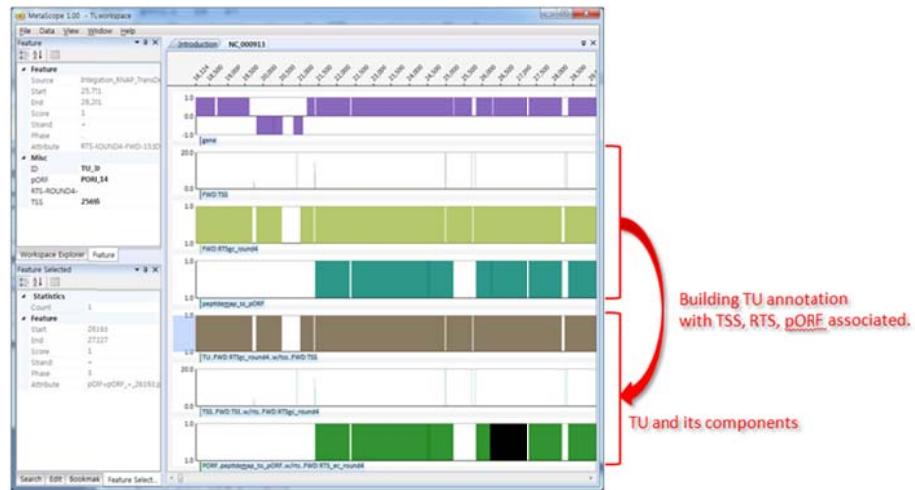


Figure 76. TU integration from TSS, RTS, pORF

Sample datasets

Sample 1. RNAP ChIP-chip

The first sample dataset contains one set of dynamic RNA polymerase ChIP-chip (without rifampicin) data, one set of static RNA polymerase ChIP-chip (with rifampicin) and calculated RNAP binding regions processed by Nimblegen's NimbleScan program.

“RNAP.workspace” file contains all information for loading datasets of this sample.

The second and third tracks are dynamic RNAP data, the fourth and fifth tracks are static RNAP data and the sixth and seventh tracks are calculated RNAP binding regions from static RNAP data.

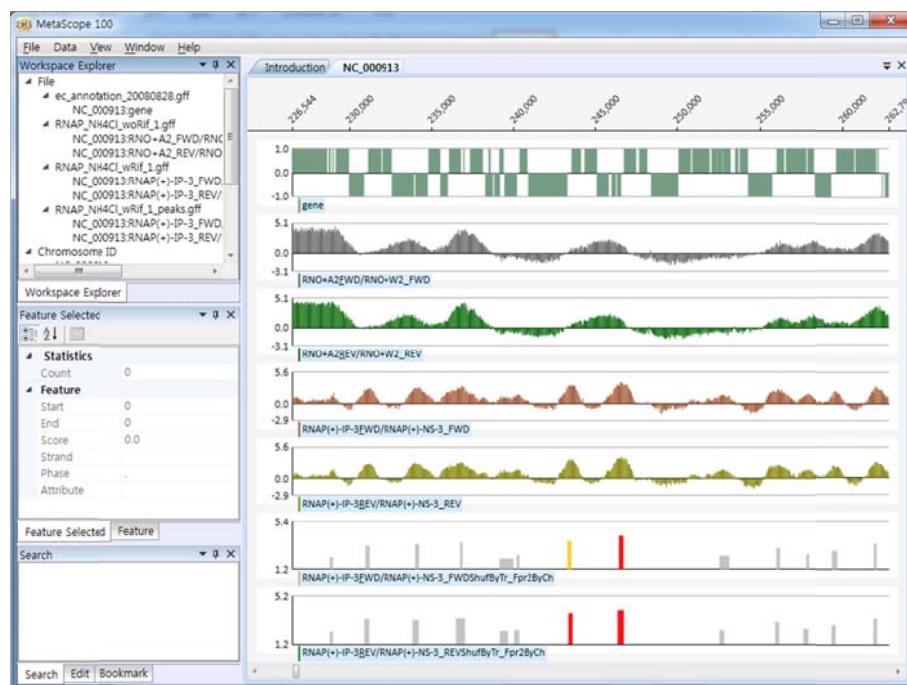


Figure 77. Sample 1. RNAP ChIP-chip

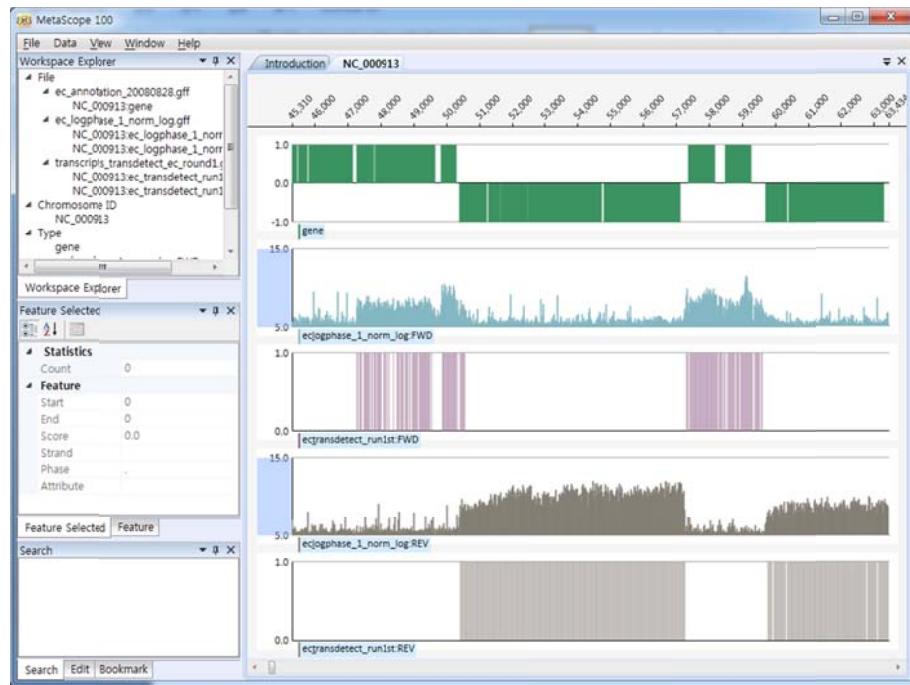
With this dataset, user can visualize genome-wide landscape of RNAP occupancy, and analyze these data with calculated RNAP binding regions. Since peak finding algorithms have some false positives and false negatives, it might be required to curate those RNAP binding peak data. In other cases, user might want to filter out weak binding peak with low P-values, so that he or she can more focus on strong binding regions.

Sample 2. Expression Profiling and Transcript Detection

The second sample dataset contains one set of expression profiling data and one set of calculated transcript detection signal data.

“TD.workspace” file contains all information for loading datasets of this sample.

The second and fourth tracks are expression profiling data for forward and reverse strands respectively. The third and fifth tracks are calculated transcript detection signals, processed by third party algorithm.



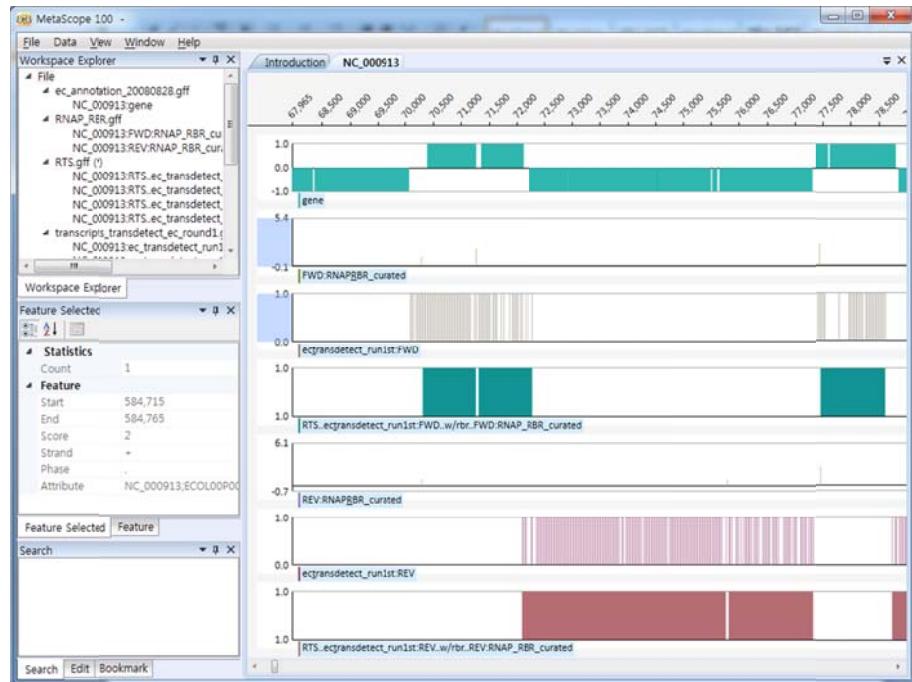
With this dataset, user can visualize genome-wide landscape of gene expression profiling, and analyze these data with calculated transcript detection signals. Since transcript detection algorithms generate slightly different calculated signals depending on specific options. Thus user might want to try various options for transcript detection algorithms, and visualize them along with the original expression profiling dataset, so that he or she can decide which option works best for further analysis.

Sample 3. RNAP-Guided Transcript Segment

The third sample dataset contains one set of processed RBR data and one set of calculated TD data.

“RTS.workspace” file contains all information for loading datasets of this sample.

The second and fifth tracks are RBR data for forward and reverse strands respectively. The third and sixth tracks are calculated transcript detection signals, processed by third party algorithm. The fourth and seventh tracks are integrated RTS from RBR and TD.



With this dataset, user can visualize genome-wide distribution of contiguous transcript segments integrated with RNAP binding information.

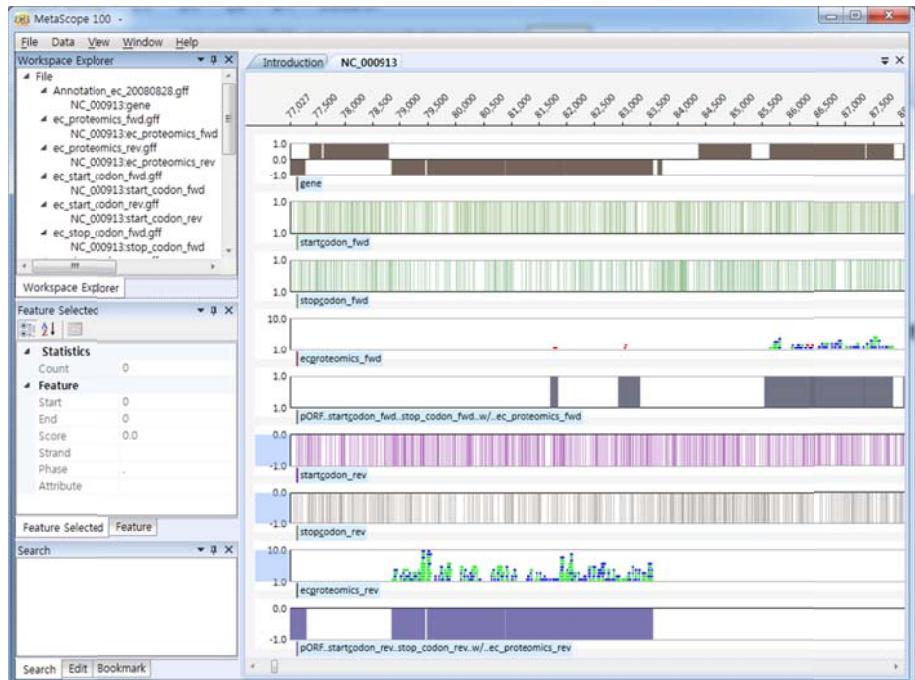
Sample 4. pORF (potential open reading frame)

The fourth sample dataset contains one set of start and stop codon position data and one set of proteomic data.

“pORF.workspace” file contains all information for loading datasets of this sample.

The second and third tracks are start and stop codon position data for forward strand, and sixth and seventh tracks are for reverse strands respectively. The fourth and eighth tracks are proteomic data displayed in stack style. The fifth and ninth tracks are integrated pORF annotation from those datasets.

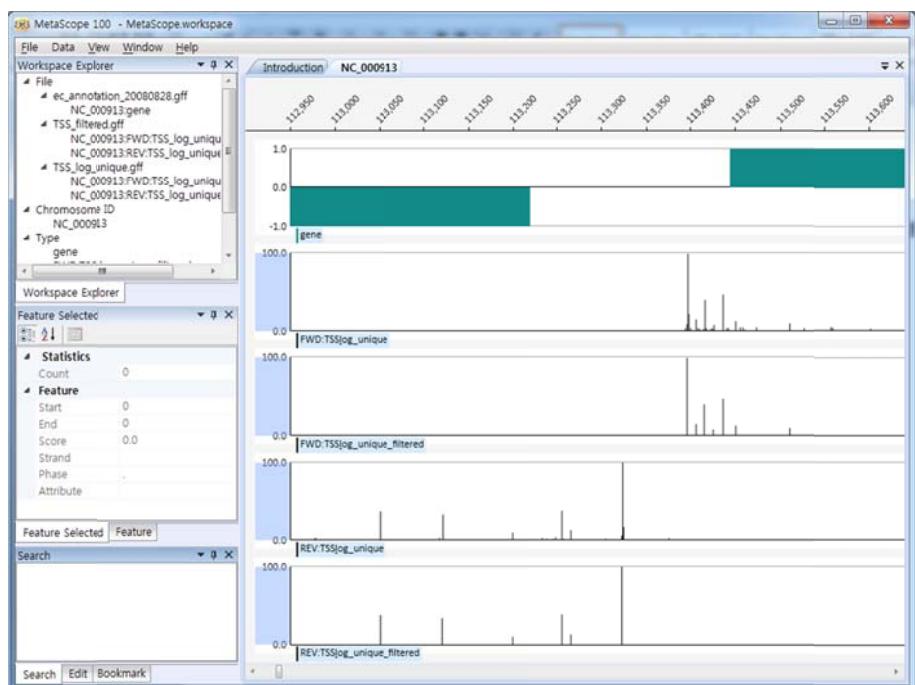
With this dataset, user can try to build pORF based on start and stop codon positions and proteomic data, not dependent on canonical genomic annotation. Further, user can compare the pORF with the known ORF information.



Sample 5. TSS (Transcription Start Site)

The fifth sample dataset contains one set of TSS data generated from modified 5' RACE followed by deep sequencing and one set of processed TSS data.

“TSS.workspace” file contains all information for loading datasets of this sample.



The second and fourth tracks are TSS data for forward and reverse strands respectively. The third and fifth tracks are processed TSS data for forward and reverse strands.

With this dataset, user can try to use feature operation functions in order to filter out noisy TSS signals. If those TSS signals are filtered using known genomic annotation, then it is possible to figure out how many TSS signals are located in intergenic regions and how many are in genic regions.

Sample 6. Transcription Unit Annotation

The sixth, last, sample dataset contains one set of processed TSS data, one set of processed RTS data and one set of processed pORF data.

“TU.workspace” file contains all information for loading datasets of this sample.

The second, third and fourth tracks are TSS data, RTS data and pORF data respectively. The fifth, sixth and seventh tracks are processed TU annotation and its component. Each TU, RTS and pORF get assigned with an unique ID, and those IDs are referred in attribute section of data features for TU, RTS and pORF.

