

Faster results,
better answers,
more biology.



Genome Browser Introduction

Introduction	3
Getting started	5
The Genome Browser window	9
The toolbar	9
The general annotation tracks	12
Annotation tracks	12
The 'Sequence' track	13
The 'Position' track	14
The user data tracks	14
The scale	15
Customizing Genome Browser	17
Adding tracks	17
Removing tracks	17
Changing the height of a track	17
Changing the color of a track	17
Changing the scale of user data tracks	18
Changing the chart type of user data tracks	20
Changing the order of tracks	21
Superimposing tracks	21

The Genomatix Genome Browser

Introduction

The Genomatix Genome Browser is our new, intuitive tool for visualizing regions of the genome and the associated annotation overlaid with multiple instances of your own positional data (e.g. from NGS experiments). You can browse a genome via position or go directly to a gene of interest, get an impression of locus complexity, easily access general annotation and proprietary data from Genomatix and add or combine your own data tracks in various display formats. Whether you want a detailed overview of the various transcripts of a gene, combine ChIP-Seq and RNA-Seq data in genomic context or export a graphic for your next paper, these are all tasks you can achieve with Genome Browser.

This guide will introduce you to the main features of the software for a quick and efficient start with Genome Browser.

We hope it will help you get the most from our software. You can also find information in the online help.

If any questions remain or if you run into any problems you are always welcome to contact us at:

support@genomatix.de (via email)

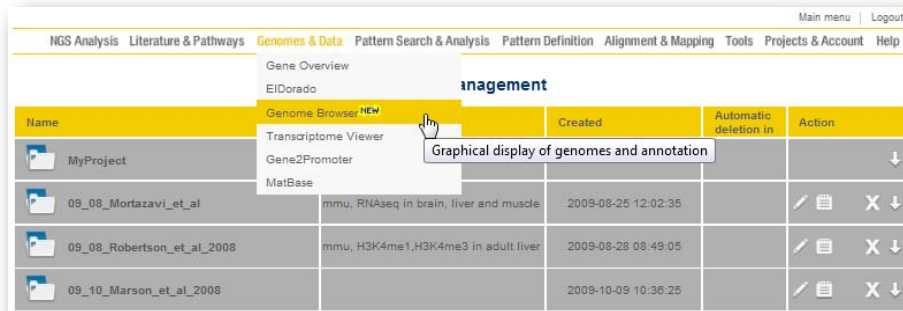
support@genomatix-software.com (via email, US and Canadian users)

+49 89 599766 0 (via phone)

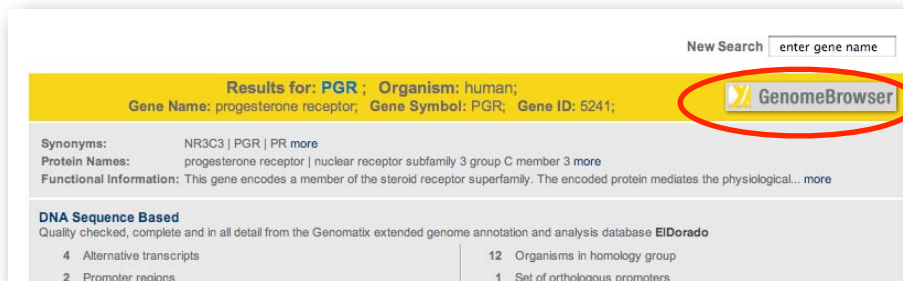
+49 89 599766 55 (via fax)

Getting started

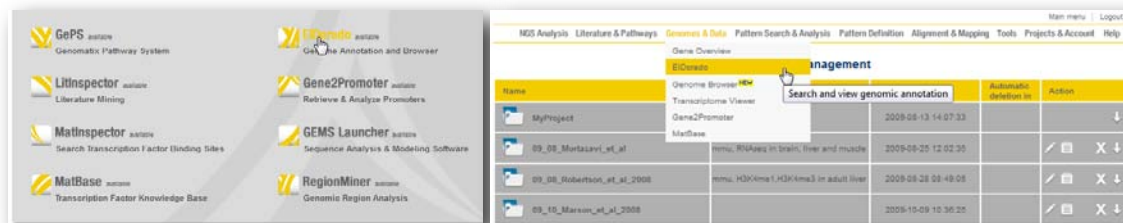
Genome Browser is part of the Genomatix Software Suite and replaces the old graphical representation ('Detailed Graphics') of EIDorado. It can be accessed directly from the navigation bar under 'Genomes & Data':



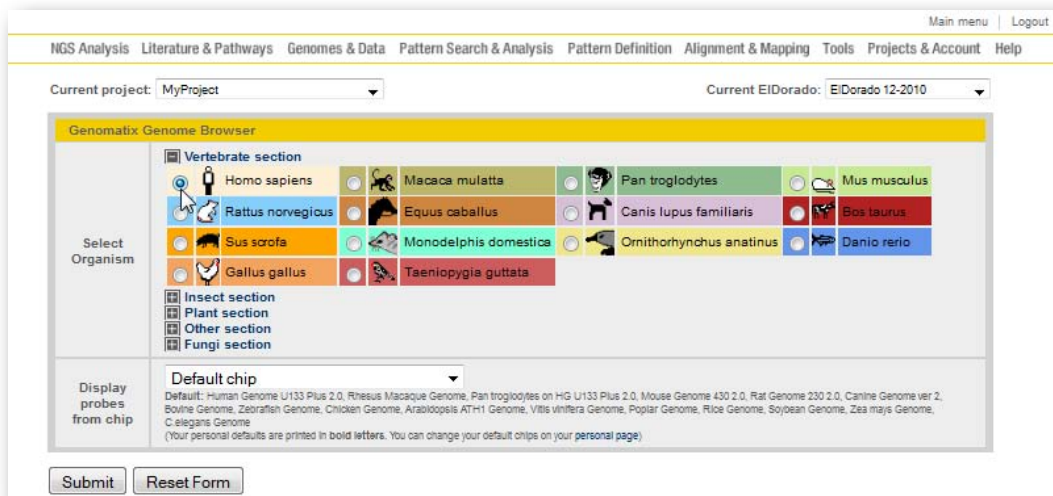
via a direct link (e.g. in a RegionMiner or Gene Overview result):



or via the EIDorado links on the main menu and in the navigation bar:

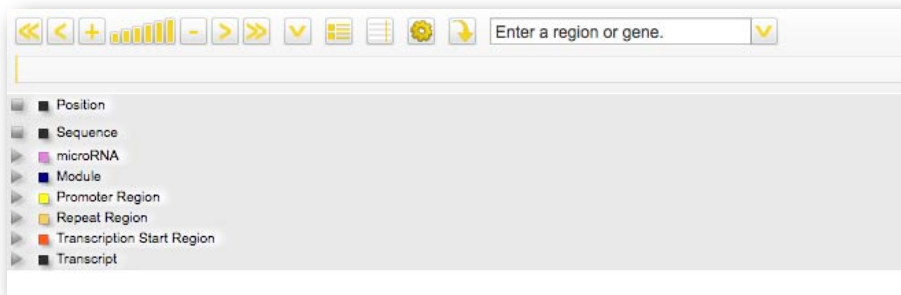


The direct link leads to the Genome Browser overview page, where you have to select the organism you want to browse. In addition, you can select a microarray chip for which probe positions should be displayed:

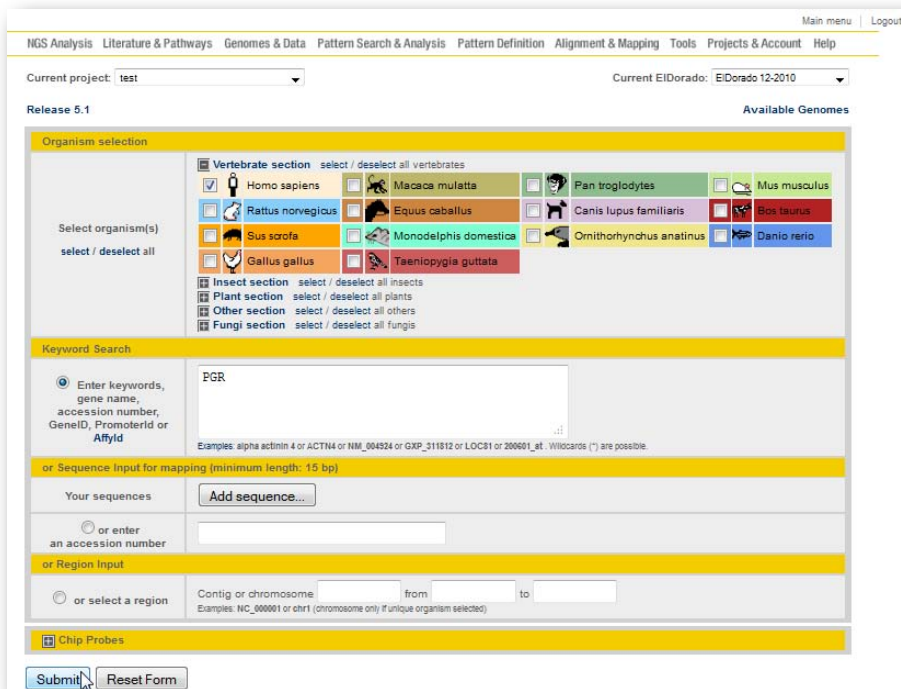


Getting started

Clicking the 'Submit' button will then open the Genome Browser view with no specific region (you can start to view data by entering a region or gene in the toolbar):



If you access Genome Browser through EIDorado you will get to the EIDorado start page where you can select the organism and your gene or region of interest. Alternatively you can also submit the sequence of the region you want to view. If you want to visualize your own data these must be present as BED files in one of your projects. Choose the according project from the 'Current project' drop-down menu on the top left to be able to view them.



After clicking the 'Submit' button you will get to the overview page from where in addition to the visualization you can access detailed information on transcripts, tissue expression profiles, comparative genomics and associated literature (see next page).

Result based on EIDorado 12-2010

PGR found on chromosome 11 of Homo sapiens, NCBI build 37, EIDorado 12-2010
 Extracted region: NC_000011 between 100850700 and 101050699 (200000 bp).
 Probes displayed are from chip Human Genome U133 Plus 2.0.

Genomic Context

Genome Browser
 Annotat
 Detailed transcript info
 Tissue Profiles
 SNP Analysis
 Annotated Sequence

Gene Oriented Analysis

The following genes are annotated in this region:

ARHGAP42 PGR TMEM133

Alternative Transcripts
 Comparative Genomics
 More Gene Info
 Literature Analysis

For comments, questions, or bug reports, please contact support@genomatix.de. © Genomatix Software GmbH 1998-2011 - All rights reserved. License Agreement

To start the Genome Browser, simply click on the 'Genome Browser' button on top of the list. Coming from EIDorado, the region or gene you specified will be shown:

genomatix software suite v2.3.0.10

GenomeBrowser

NGS Analysis Literature & Pathways Genomes & Data Pattern Search & Analysis Pattern Definition Alignment & Mapping Tools Projects & Account Help

chr11: 100,875,183 - 101,026,216

Position 100889 kb 100927 kb 100965 kb 101003 kb

Sequence
 microRNA
 Module
 Promoter Region
 Repeat Region
 Transcription Start Region
 Transcript

ENST00000293463 (PGR) ←
 NM_000926 (PGR) ←
 AK304853 (PGR) ←

Transcript
 PGR / OX1_23209427 (Homo sapiens cDNA FLJ51472 complete cds, highly similar to Progesterone receptor, AK304853)
 (start: 100909411, end: 101000458, length: 91048, strand: -, quality: gold)

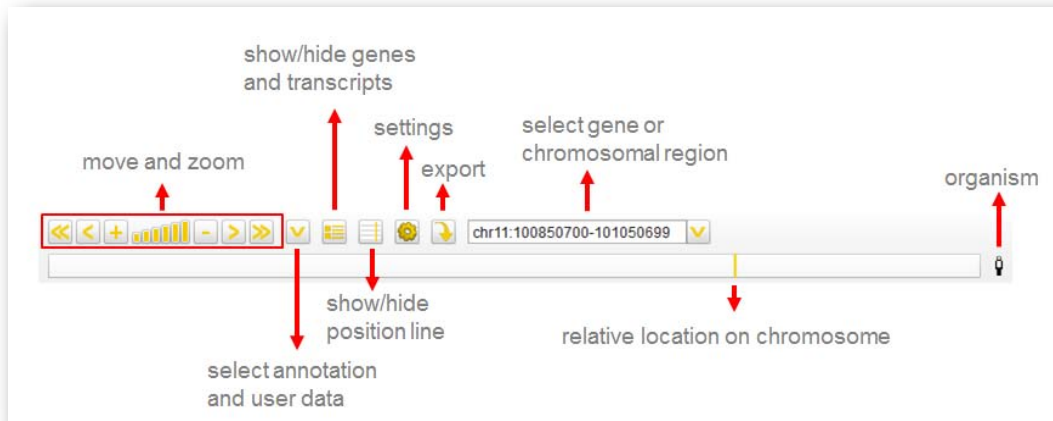
The Genome Browser window

The browser window consists of four parts – the toolbar on top, the general annotation tracks and the user data tracks in the middle and a scale on the bottom. While the user data is displayed below the annotation data per default you can move tracks around between the two. We will introduce these sections in more detail in the next few paragraphs.



The toolbar

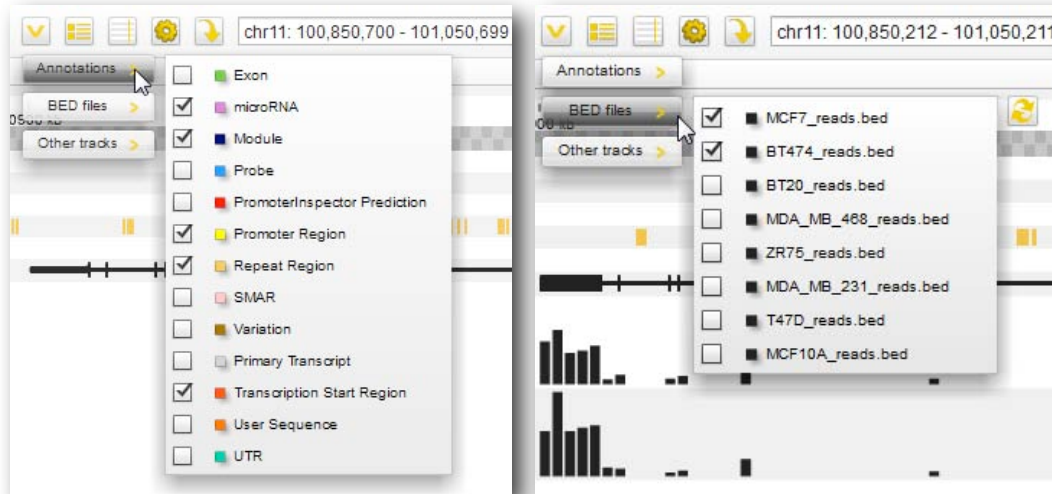
Here is an overview of the toolbar and its elements:



Move and zoom: clicking any of the arrows will shift the display view to the left or right of the current view (single arrows: a half window size, double arrows: a full window size). You can also move the view by clicking in the window, holding and moving the mouse. If there are too many tracks to fit them into your browser window, you can also scroll down using this method.

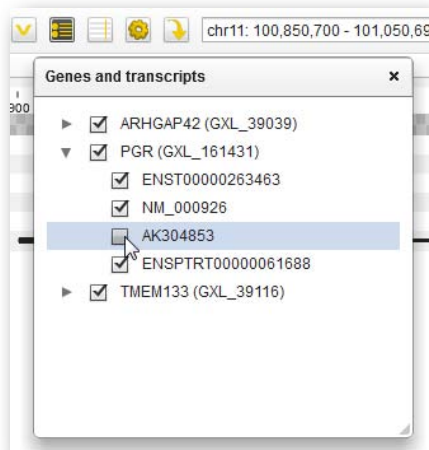
Zooming in or out can be either achieved by clicking on the scaling bars (zoom levels range from 50 bp to 1Mb), by using the + and – buttons or using your mouse wheel.

- Track selection (select annotation and user data): This button allows you to individually select annotation tracks from the EIDorado database, your user data (BED files) or other tracks (like the sequence). If you upload or save BED files in the project management while working with the Genome Browser you can update your BED file list (for the selected organism) by clicking on the reload button (🔄) next to the BED file list.



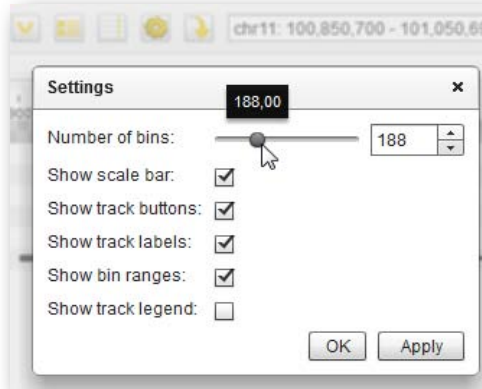
To see the source of the elements annotated in EIDorado just move your pointer over any of the elements or have a look at the overview in the online help or at http://www.genomatix.de/online_help/help_eldorado/eldorado_elements.html. Probes are only shown for the microarray selected on the EIDorado or the Genome Browser start page.

- Show / hide genes and transcripts: using this button will open a small window where you can individually select and deselect genes and transcripts shown in the current window.



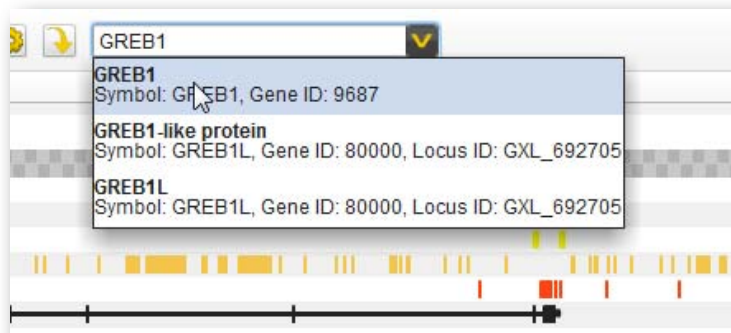
- Show / hide position line: will display a small vertical position line indicating the exact genomic location of its position. It can also be used to visually align different features.

- Settings: The settings window allows you to change some global parameters of the visualization. You can define the resolution for the user data tracks (read more on how this affects the visualization in the '[User data tracks](#)' section below) or switch off the labels and buttons next to the tracks and show the track legend instead. Any settings here affect all tracks shown:

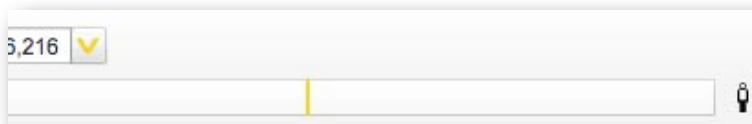


- Export: allows you to export the current view as jpg or png image.

Selecting a gene or chromosomal regions: Using this input you can jump to a different chromosomal region by either entering the chromosomal coordinates (e.g. 'chr1' or 'NC_00009' plus start and end position), a gene symbol or a gene id. If you enter coordinates, just hit the return key. If you enter a gene symbol, a drop down list will pop up, where you can select your gene from:



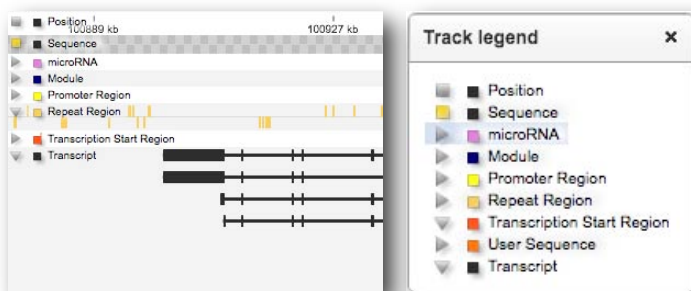
In the lower part of the toolbar you can find the chromosomal location bar and an icon for the currently displayed species:



The yellow indicator shows the relative chromosomal position of the current view. You can also drag it using the mouse to get to another chromosomal position. Moving the mouse pointer over the species icon will show the species name, EIDorado version, project and the currently selected chromosome.

The general annotation tracks

All tracks in the Genome Browser consist of a track button (a small triangle or square to the very left of the track), a track label indicating the type of annotation data or based on the the name of user data and the visualization of the data. If a track can't be shown, e.g. because the selected range is too big to get a meaningful visualization of the data, this will be indicated by a yellow track button. A red-colored track button indicates that data is still being loaded for that track. You can hide the track buttons and labels or choose to use the track legend window instead via the settings menu as described above. Here are screenshots showing the track buttons within the tracks and the track legend window:



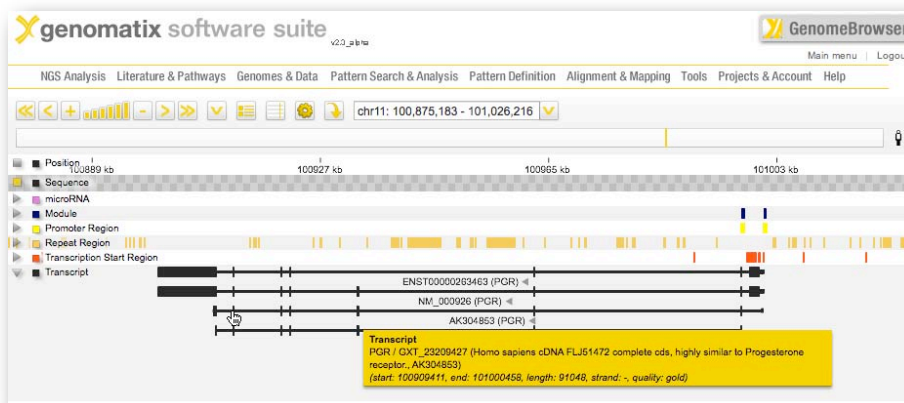
Annotation tracks

All annotation data from Genomatix' EIDorado database can be displayed in this section. The tracks selected by default are microRNAs (from miRBase, pink), modules (known functional transcription factor binding site combinations, dark blue), promoter regions (yellow), repeat regions (light orange), transcription start sites (red) and transcript (black). If you started the Genome Browser via EIDorado, the user sequence track (dark orange) will also be shown, marking the location of the gene you entered or the aligned sequence if you submitted a sequence or accession number.

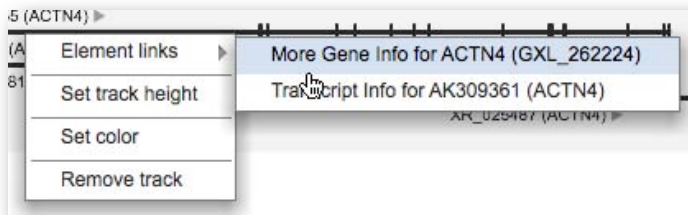
To add or remove tracks use the track selection button (☑) in the toolbar. (*Please note that the 'transcripts' track is located in the 'other tracks' section.*)

Clicking on a track button (the small triangles on the left side of the tracks) will expand the associated track. For most tracks this will result in plus and minus strand being displayed separately, but for the transcript track all transcript variants are shown. Strand orientation for these is indicated by the small grey arrows next to the transcript accession numbers. Tracks with a square instead of a triangle can't be expanded.

If you move your mouse pointer over a transcript, additional information is shown in a popup window:



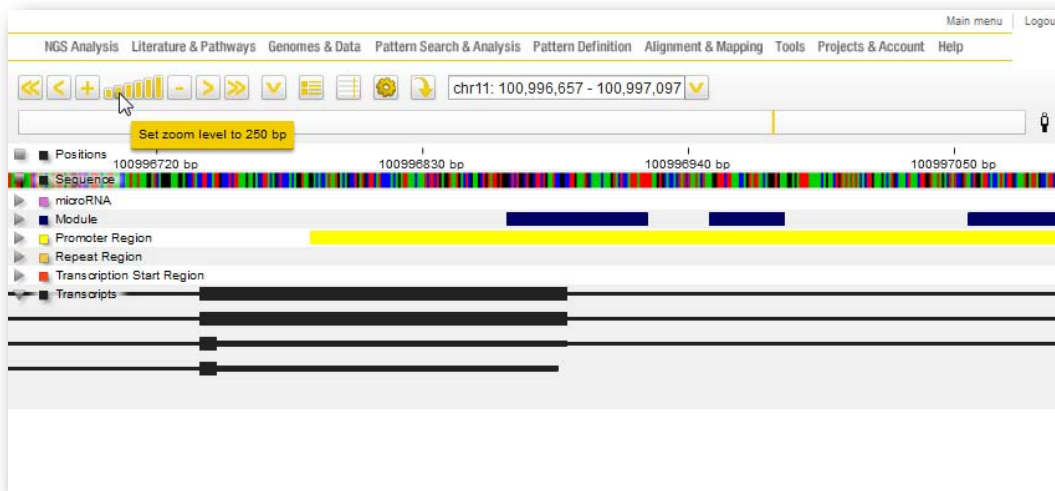
Furthermore, double-clicking the transcript opens a popup menu. Via the 'Element links' option you can go to EIDorado's 'More Gene Info' or 'Detailed Transcript Info' pages for more detailed information on the chosen transcript:



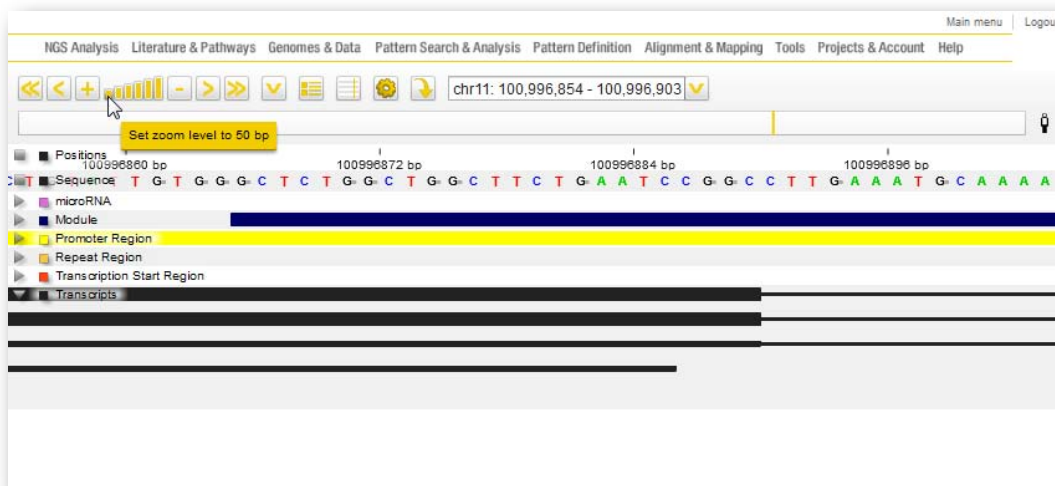
If you want to display a subset of transcripts, use the transcript selection button (☰) as described above to choose them individually.

The 'Sequence' track

The sequence track can be used to display the underlying reference sequence of the region you're looking at. If you zoom in to the 250 bp level (second smallest scaling bar) the sequence will be displayed as color code:



If you zoom in further to the 50 bp zoom level (smallest scaling bar) the nucleotides will be displayed:



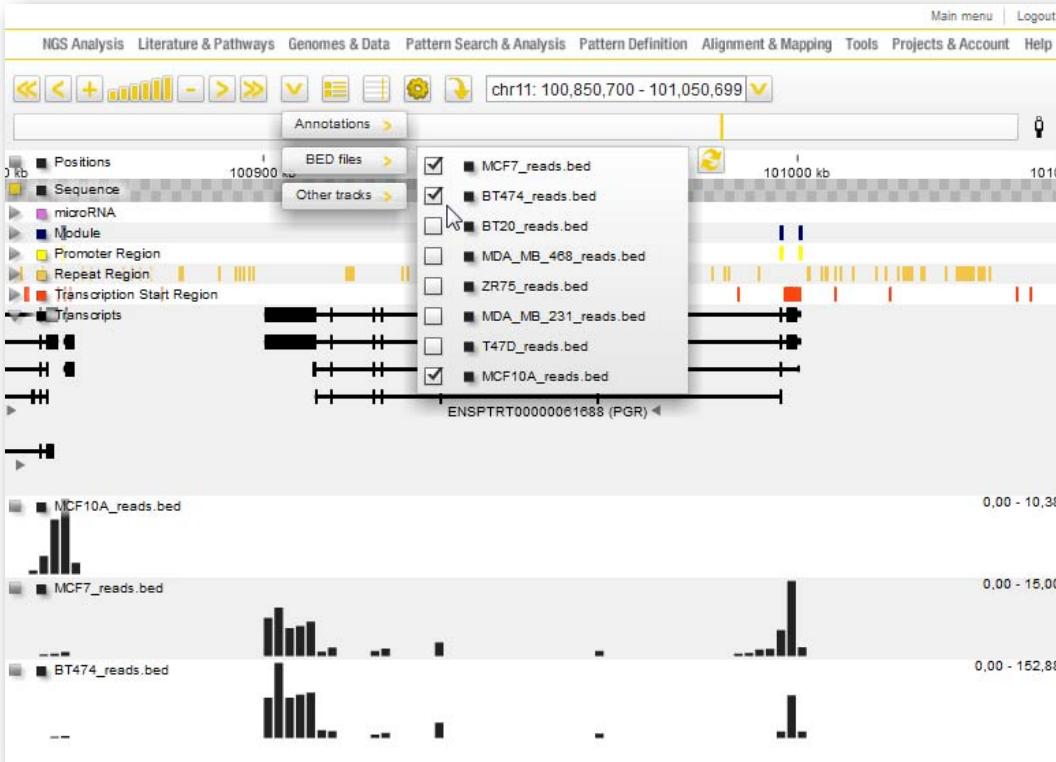
For zoom levels where displaying either the sequence or the color coded nucleotides isn't possible, the sequence track will just show checkered grey squares.

The 'Position' track

The position track shows the exact chromosomal position of the region currently displayed.

The user data tracks

All BED files stored in the project management for the currently selected project can be added to the user data section. To add data, use the track selection button (☐) in the toolbar, then pick the 'BED files' option and tick those files that you want to add as data tracks:



By default the tracks are displayed as bar charts in black and with auto scaling. Each bar represents a distinct region of the underlying sequence, which we call 'bin'. For each bin a coverage value is calculated by dividing the number of 'covered' nucleotides by the total number of nucleotides in the bin. Multiple coverage is taken into account.

Example

```
          GATC   (region 1 - partial)
ACCTCCACA    (region 2)
  CCTCCACTC   (region 3)
AGACCTCCACTCTGATC (bin)
```

Here the bin is 20 nucleotides long. It is covered by 3 regions from the BED file, with a coverage value of $(4 + 10 + 10)/20 = 1.2$

Please note that the bins are recalculated dynamically whenever you change the zoom level or move the view. You can change the number of bins displayed via the 'Settings' menu.

Moving your mouse pointer over a bar will give you the name of the BED file associated with the track and coverage for that bin:



You can customize each user data track by setting different colors or choosing another chart type. Combining tracks is also possible. To learn more, please refer to the next section 'Customizing Genome Browser'.

The scale

The scale below the tracks helps you to quickly verify the genomic distances you're looking at in the current view. It can be hidden via the settings menu.

Customizing Genome Browser

Adding tracks

To add tracks to either the annotation or user data area use the track selection button (☑) as described in the sections above.

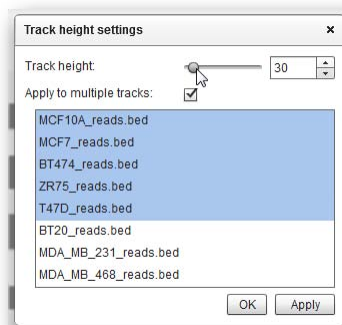
Please note that if you add more tracks than can be shown in your browser window, you will have to scroll down to view those tracks.

Removing tracks

To remove a track, either click on the track selection button (☑) and deselect the track in the popup-menu or double click the track and select the 'Remove track' option from the pop-up menu.

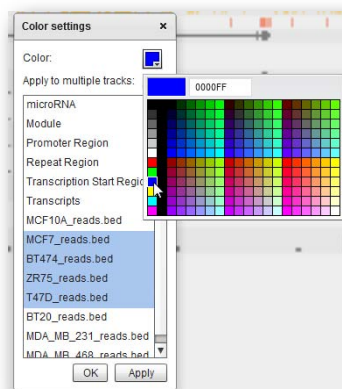
Changing the height of a track

You can change the height of any track in both the annotation and user data area. Double click on the track and select 'Set track height' from the menu. In the upcoming window you can set the track height in pixels. To change the height for several tracks at once, check the box 'Apply to multiple track' and select the corresponding tracks:



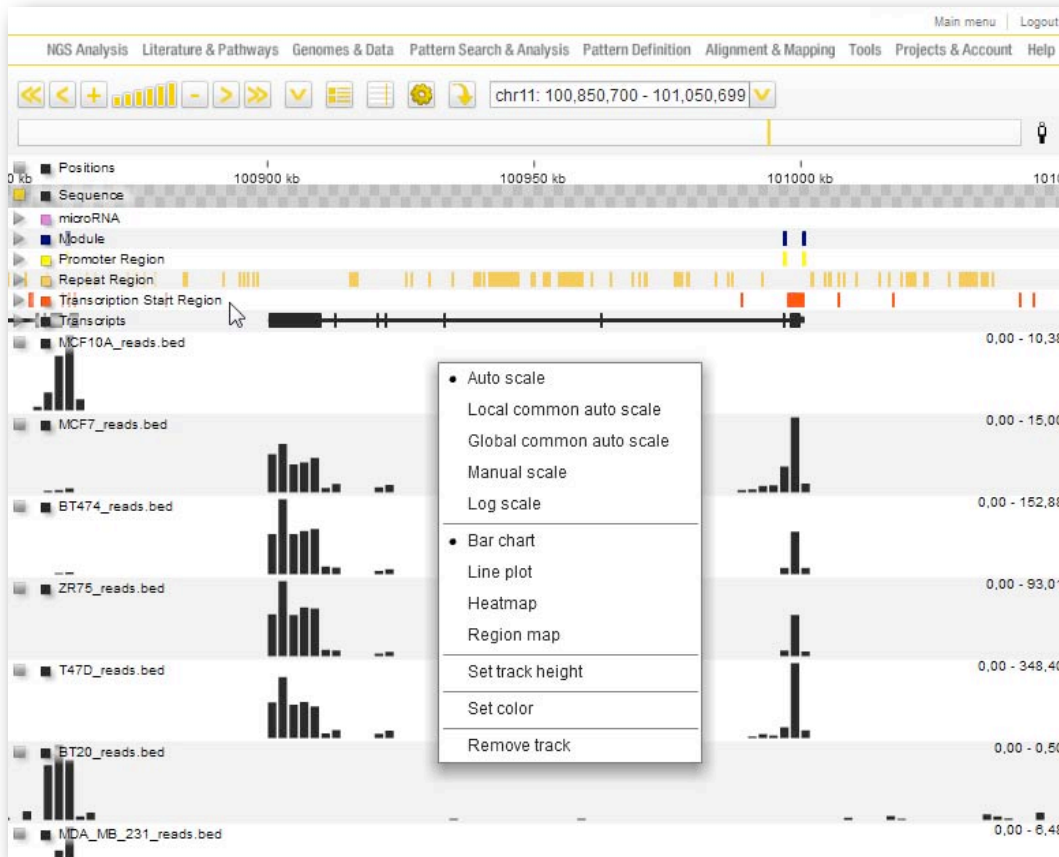
Changing the color of a track

You can also change the color of a track. Double click it, select 'Set color' and then choose a color using the color picker in the upcoming window. To change the color for several tracks at once, check the box 'Apply to multiple tracks' and select the corresponding tracks:



Changing the scale of user data tracks

Double-clicking a user data track will bring up an extended popup menu, where you can choose from different ways to scale the y-axis of a track. The default is 'Auto scale':



With 'Auto scale', the data are automatically scaled, which means that each track is scaled based on its maximum value in the currently displayed region. As a consequence, the maximum values displayed differ between tracks and will also differ if you move to a different location of the chromosome. The displayed data range for each track is shown on the right hand side of the track. In this example the range shows that PGR is much higher expressed in T47D (third from the bottom, range from 0 to 348.40) than in MCF10A (top, range from 0 to 10.38).

If you select 'Local common auto scale', all tracks for which this scaling method was selected will be scaled based on their common maximum in the displayed region. This can be helpful to assess the differences between data in the region you're looking at.

If you select 'Global common auto scale', each lane is in addition normalized based on the total length of regions and thus taking the numbers and lengths of all bed files into account where this scaling method was selected. This will give you more of a global view of your data.

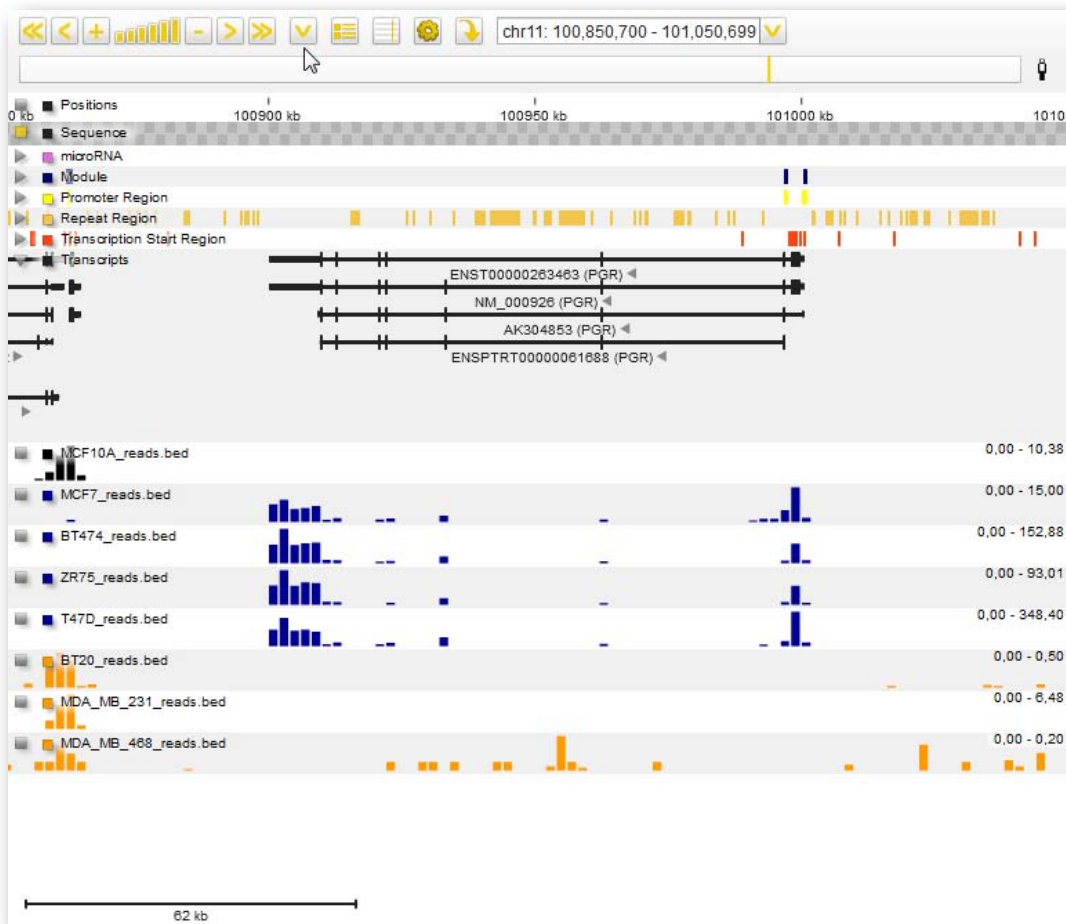
Please note that if you select one of the common scales, data in your current view might be scaled down to being invisible if the maximal values in other tracks are high in comparison. The best way to check for the presence of any data at the current position is to use 'Auto scale', for comparing tracks either the 'local' or 'global common scales' are recommended.

'Manual scale' let's you enter a minimum and maximum value for the y-axis range. *Depending on the values, some data might become invisible due to the scaling!*

Selecting 'Log scale' will change the scale from linear to log 2 based. This can be applied to any of the 4 scaling choices.

Example

For the view below, tracks for estrogen receptor (ER) positive breast cancer cell lines were colored in blue, while ER negative breast cancer cell line tracks were colored in orange. The control cell line (MCF10A) remained black. The track height was set to 30 for all:



Changing the chart type of user data tracks

Besides changing the height or color of a track you can also choose from four different chart types for user data tracks:



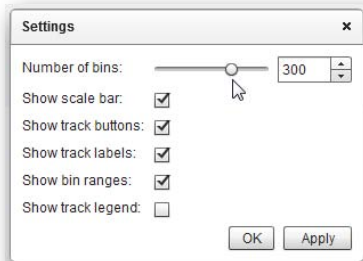
'Bar chart' is the default, where each bar stands for a bin containing regions from the BED file. As in a histogram, the more regions from the BED file fall into a bin the higher the bar will be plotted.

'Line plot' uses the same binning approach as 'bar chart' but will display a line instead of bars. You could think of it as the outline of a bar chart.

'Heatmap' will display the values of the bin in a heatmap style, i.e. the value of each bin is indicated by the saturation of the drawn bar. Higher values will have 'darker' bars, lower values are 'lighter'. All bars are drawn in the same height.

'Region map' will just indicate that some data has been found for this position in the BED file without assigning any value to it.

The resolution of the user tracks, i.e. the number of bars or line segments drawn can be changed via the 'Settings' button (⚙️) in the toolbar:



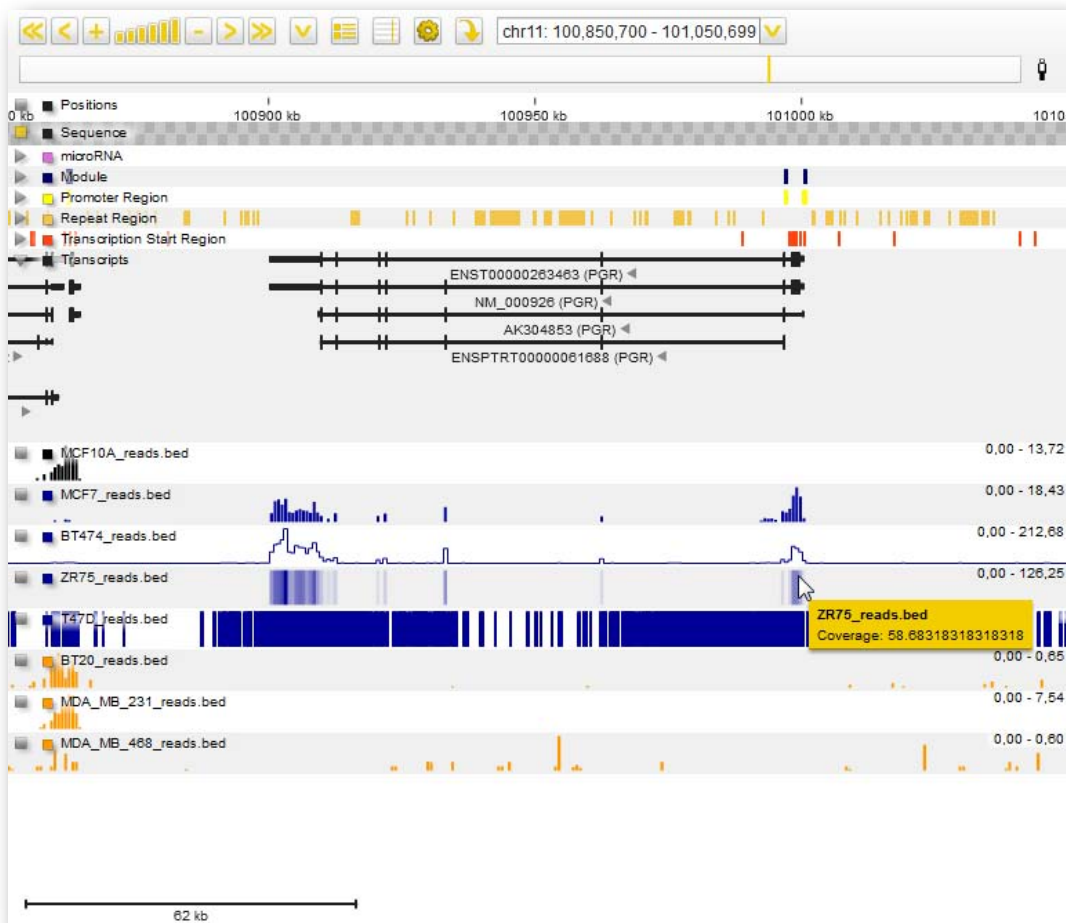
Please note that choosing a high number of bins might slow down the responsiveness of the Genome Browser when displaying many user tracks, as more calculations have to be performed!

Example

On top of the next page you can see an example combining several settings in one view: It shows the four different chart types:

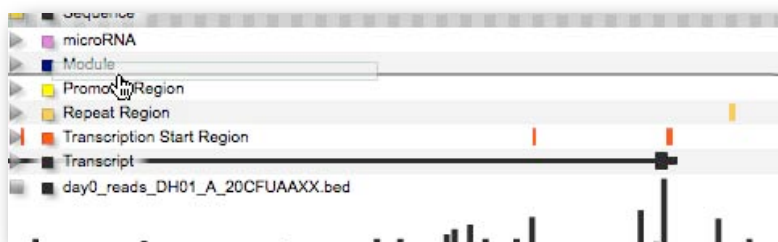
- bar chart (MCF7, first blue track),
- line plot (BT474, second blue track),
- heatmap (ZR75, third blue track), and
- region map (T47D, fourth blue track),

all of them set to a resolution of 300 bins:



Changing the order of tracks

If you need a different order of tracks, e.g. to better correlate a user track with an annotation track, you can simply move tracks around, by clicking on their name and then moving them to the new position and then releasing the mouse button. A line will appear in between other tracks to indicate the position where the track will be inserted:

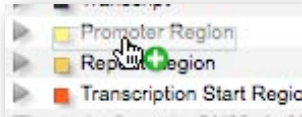


You can move any track to any position in the display,

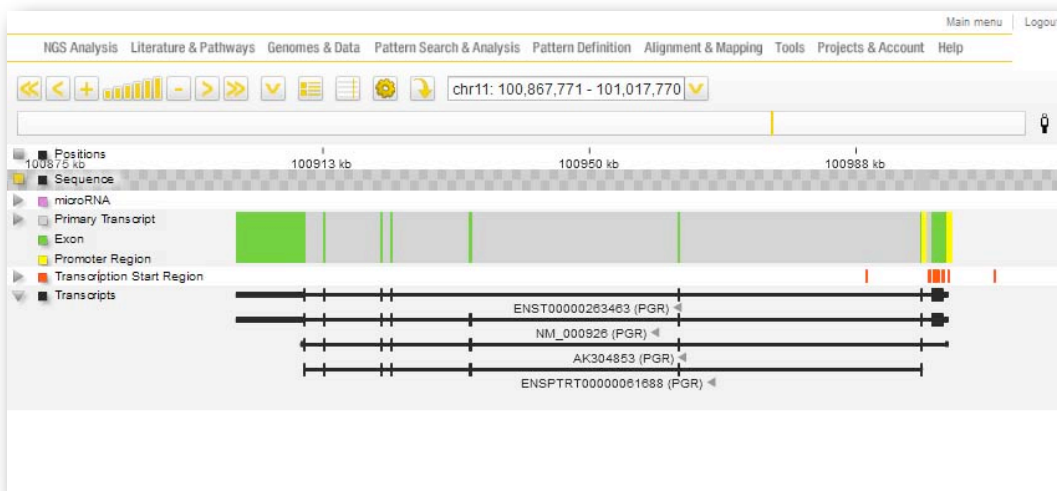
Superimposing tracks

A powerful new feature of Genome Browser is the capacity to superimpose tracks. This enables you to convey multiple levels of information in a single track. You can combine

several annotation tracks or user data tracks. If you prefer a superimposed display for the annotation track (as in the old 'Detailed graphics' in EIDorado), e.g. for primary transcripts, exons and promoters, you can overlay these tracks using simple drag and drop (select a track by clicking on the track label, hold and move the track over the other track and release the mouse button when the small green plus sign appears):



The view below was generated by removing the 'Repeat Region' and 'Module' tracks and adding the 'Primary Transcript' and 'Exon tracks' and then combining those (using drag and drop) with the 'Promoter' track:



To separate any track out of a combined track, simply click on its label and drag it to a different order in the display.

Please note that the 'Sequence', 'Position' and 'Transcript' tracks can't be superimposed on any other tracks.

You can also superimpose user data tracks using the same drag & drop method as for the annotation tracks. In combination with different chart types and/or scalings this can be used to create highly informative combined tracks.

Example

For the view shown on the next page four BED files containing raw reads and clusters from two cell lines were added. The read tracks were then set to 'Common local auto scale'. The cluster track displays were set to 'Region Map' and their color was changed. The resolution was increased and subsequently the read files were superimposed on top of the corresponding cluster files. You can now see the cluster positions, the read distribution within the cluster and the relative coverage all in a single track for each experiment:

