

# ZMap User Manual

Written by and contributions from:

Gemma Guest <[gb10@sanger.ac.uk](mailto:gb10@sanger.ac.uk)>

Charlie Stewart

Laurens Wilming

Ed Griffiths

James Gilbert

Jennifer Harrow

Wellcome Trust Sanger Institute

16 March 2015

## 1 Revision History

Revision	Date	Author
First revision (ZMap v1.5.0)	16/03/2015	Gemma Guest
Updated for ZMap v2.5.1	12/02/2016	Gemma Guest

## Contents

<b>1 Revision History</b>	<b>2</b>
<b>2 Introduction</b>	<b>5</b>
<b>3 Starting ZMap</b>	<b>6</b>
3.1 Running ZMap on a GFF file . . . . .	6
3.2 Choosing sources on startup . . . . .	6
3.3 Running ZMap with a config file . . . . .	7
3.4 Operation within Otter . . . . .	9
<b>4 The ZMap interface</b>	<b>11</b>
4.1 Navigating around the canvas . . . . .	11
4.2 Zooming . . . . .	11
4.3 Toolbar . . . . .	11
4.4 Coordinate display (ruler) . . . . .	11
4.5 Navigator pane . . . . .	17
4.6 The Focus Feature . . . . .	17
4.7 Marking a region . . . . .	17
4.7.1 Mark a feature . . . . .	17
4.7.2 Mark an area . . . . .	19
4.7.3 Manual cropping of the marked borders . . . . .	19
4.7.4 Unmark a feature . . . . .	19
4.8 Splitting the window . . . . .	19
<b>5 Feature display</b>	<b>21</b>
5.1 Bumping features . . . . .	21
5.2 Feature styles and strand . . . . .	21
5.3 Column configuration . . . . .	24
5.3.1 Selecting rows . . . . .	24
5.3.2 Search and filter rows . . . . .	26
5.3.3 Track visibility . . . . .	26
5.3.4 Assigning groups . . . . .	26
5.3.5 Assigning styles . . . . .	26
5.3.6 Reordering tracks . . . . .	26
5.3.7 Saving configuration . . . . .	27
5.4 Style configuration . . . . .	27
5.4.1 Editing styles . . . . .	27
5.4.2 Saving styles . . . . .	27
<b>6 Operations</b>	<b>28</b>
6.1 Show feature details . . . . .	28
6.2 Exporting features for gene objects . . . . .	28
6.3 Searching for a sequence in ZMap . . . . .	28
6.4 Searching for a feature in ZMap . . . . .	28
6.5 Selecting single or multiple features and hiding/showing them . . . . .	28
<b>7 Annotation</b>	<b>35</b>
7.1 Standalone ZMap . . . . .	35
7.1.1 Enable annotation . . . . .	35

7.1.2	Create a temporary feature . . . . .	35
7.1.3	Editing the coordinates . . . . .	35
7.1.4	Setting the CDS . . . . .	37
7.1.5	Editing the transcript's sequence . . . . .	37
7.1.6	Editing the attributes . . . . .	37
7.1.7	Saving the temporary feature . . . . .	38
7.1.8	Other operations . . . . .	38
7.2	Variant construction with Otter . . . . .	38
<b>8</b>	<b>Multiple views</b>	<b>41</b>
8.1	Multiple views from GFF . . . . .	41
8.2	Launching in a ZMap from Otter . . . . .	41
<b>9</b>	<b>Tips for a speedier ZMap</b>	<b>43</b>
<b>10</b>	<b>Keyboard and mouse shortcuts</b>	<b>44</b>
10.1	All windows . . . . .	44
10.2	ZMap Window . . . . .	44
10.3	Alpha-numeric keys . . . . .	44
10.4	ZMap Mouse Usage . . . . .	45

## 2 Introduction

ZMap is a software package that provides a visualisation and editing tool for genomic features. The software is written in C/C++, utilising the gnome toolkit (GTK2) to draw features on a canvas.

ZMap accepts input from multiple sources in multiple formats across multiple genomes and is written in a way so that the addition of further formats is made as trivial as possible. Currently the list of formats includes GFF and DAS (which may reside in a file, an acedb instance, or an http server). Multiple genomes and their associated features can be displayed in a single view as aligned blocks providing support for comparative annotation. ZMap provides facilities for editing and creating features, which can then be saved to a GFF file.

ZMap is part of the Otter annotation suite and contains a protocol for two-way communication with Otter so that it operates seamlessly as an additional display. ZMap can also call tools from the SeqTools suite of sequence visualisation tools to analyse alignments in more detail.

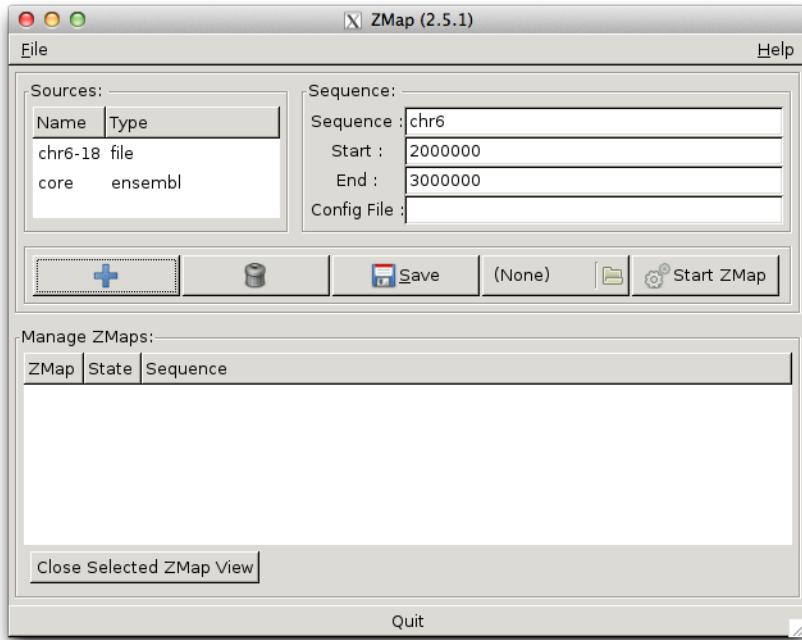


Figure 1: ZMap’s Main Window. Use the ‘+’ button to add new sources, enter the sequence details for the region you want to view, and then click ‘Execute’.

### 3 Starting ZMap

There are several ways to run ZMap depending on the data sources you want to use.

#### 3.1 Running ZMap on a GFF file

To run ZMap standalone on one or more GFF files, simply pass the file name(s) on the command line:

```
zmap file1.gff file2.gff
```

The sequence region will default to the region available in the GFF file. Alternatively, you can specify the sequence/region by passing the `--sequence`, `--start` and `--end` arguments on the command-line.

If all of the files contain the same reference sequence, then their contents will be shown in the same view. If the files contain multiple sequences/regions, then they will be shown in multiple views.

#### 3.2 Choosing sources on startup

If you run ZMap without any arguments then you will be presented with the Main Window where you can choose data sources to load (see Figure 1).

Click on the ‘+’ button to add a new source. This will open the Create Source dialog. From the drop-down menu, select the type of source. Choose ‘File’ for a local/remote file (figure 2) or ‘Ensembl’ for an Ensembl database (figure 3).

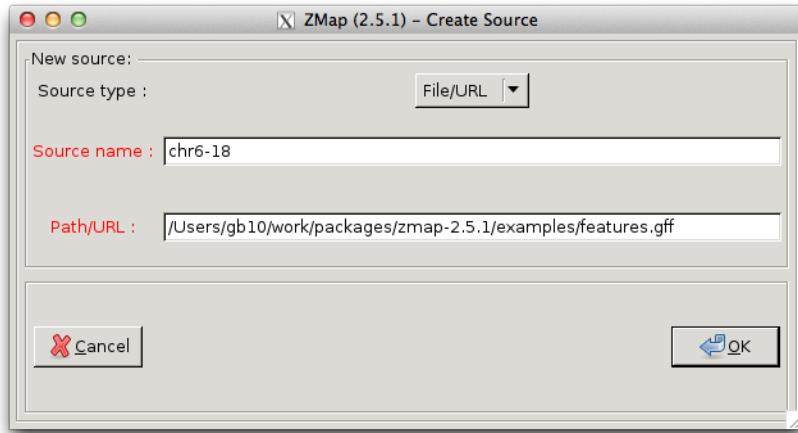


Figure 2: Create a new file source

**Local/remote file** For File sources, you just need to enter the path of the file (or URL if it is a remote file). This file can be in any of the file formats supported by ZMap, e.g. GFF3 or BAM. You also need to enter a name for the source - this can be anything you like but should be a short, descriptive name with no special characters (except underscores or dashes).

**Ensembl database** For Ensembl sources, the public host details are pre-filled for you by default. You may change these if you wish but the public host should be fine for most users (note that a password is not required). You can click the magnifying glass icon next to the 'Database' box to search for available databases in the host (figure 4). You may optionally also enter a list of featuresets or biotypes that you are interested in - in this case, ZMap will only load features of those types; otherwise, ZMap will load all of the features from the database. You can click the magnifying glass icon next to the featuresets/biotypes boxes to search for available featuresets/biotypes. Note that you can select multiple featuresets/biotypes from the search box by holding down Ctrl (for individual selection) or Shift (for range selection) when you click with the mouse button. You can also enter multiple featuresets/biotypes directly into the Create Source dialog as a semi-colon separated list.

After you click OK on the Create Source dialog, the new source will appear in the Main Window. You can enter multiple sources in the same way. If you want to remove a source, select it in the list and click the delete button. You also need to enter the sequence details for the region you want to load. Then click 'Execute' to open ZMap on that region. You can edit the sources/sequence details at any time and click 'Execute' again to start a new ZMap on a different region.

You can save the source details to a config file so that they can easily be re-loaded later. Simply click the Save button and select a file location. To re-use the config, you can either open the Main Window again and select the file using the file chooser button, or you can run ZMap on the file from the command-line as follows:

```
zmap --conf_file=/path/to/file
```

### 3.3 Running ZMap with a config file

If you have several sources, it can be convenient to specify them in a config file rather than typing them on the command-line. You can also then set up many different options and styles to use for

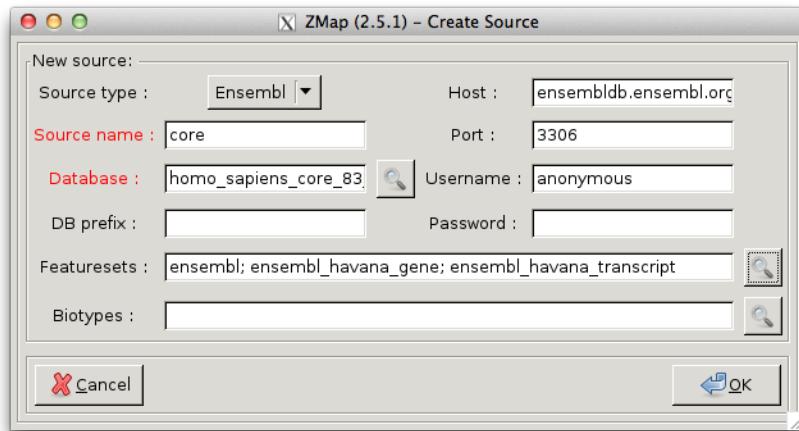


Figure 3: Create a new Ensembl source. Click on the magnifying glass icons to search for databases/featuresets/biotypes to load.

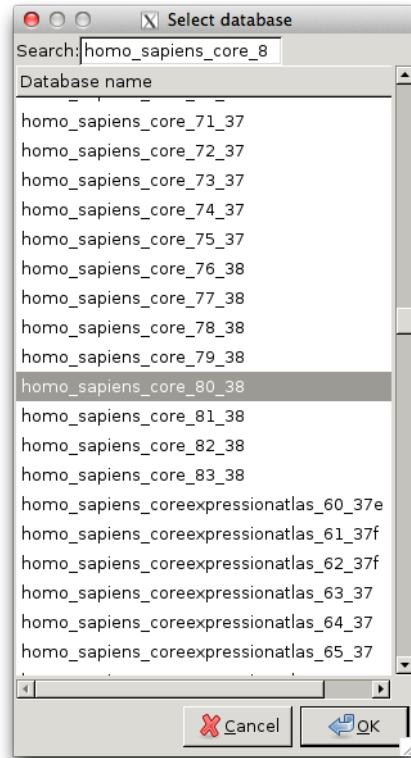


Figure 4: Search a list of available Ensembl databases. Type text in the Search box to jump to items containing that text. You can use the up/down arrows to jump to the next/previous item with the same text.

these sources. To use a local GFF file as a source, a simple config file would look something like this:

```
[ZMap]
# The list of sources
sources=chr6-18

# Sequence name/start/end can be specified here, on the command-line, or
# omitted if they are to be taken from the file
sequence=chr6-18
start=2696324
end=2864370

# Source stanzas (one per source)
[chr6-18]
url=file:///Users/gb10/work/checkout/zmap/examples/features.gff
```

ZMap can then be started by passing the config file on the command line:

```
zmap --conf_file=/path/to/file
```

Alternatively, the config file can be saved as `~/.ZMap/ZMap` and will be picked up by default.

### 3.4 Operation within Otter

ZMap is opened via the Tools menu bar in Otter (figure 5).

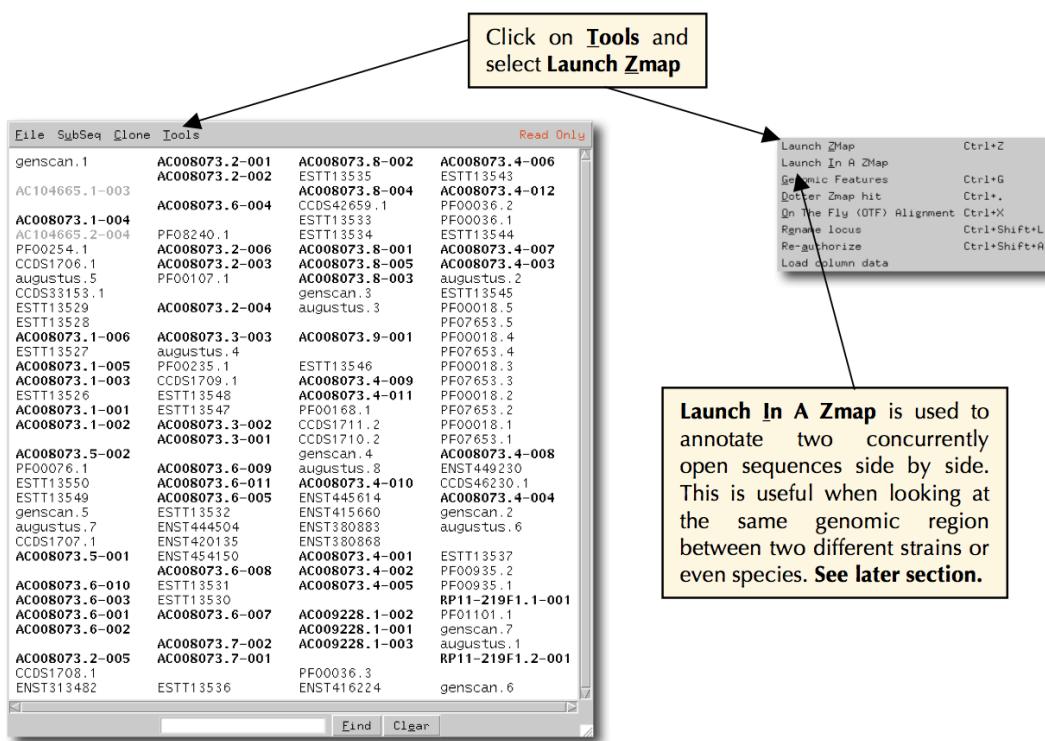


Figure 5: Opening ZMap from Otter

## 4 The ZMap interface

The main ZMap display (figure 6) consists of a canvas, which is a drawing area where any analysis and annotation that is present in your region of interest is shown. There is a toolbar, information bar and menu at the top, and a ruler showing the coordinates at the current position based on the current coordinate system. There is also a navigator pane to the left of the ruler, which is hidden by default.

### 4.1 Navigating around the canvas

Navigate by using the scroll bars or the middle mouse button (figure 7). By clicking the middle mouse anywhere in ZMap you will see a horizontal line. You can move this up and down and the relative position in bp will be displayed along the line.

When the button is released, the window will refresh, centering on the position of the line. You can also click in the window to make it active and use the scroll wheel to navigate up and down or achieve the same result using the scroll bar on the right hand side of the window.

If you release the mouse outside the ZMap window, you can then check the sequence position displayed, without re-centering.

### 4.2 Zooming

See figure 8.

- Zoom in by using the Zoom in/Zoom out buttons on the toolbar
- Draw with the left mouse button to lasso an area to zoom in to
- Press the “z” key on the keyboard to zoom to the current focus feature.
- Use the “Z” key to zoom to a whole transcript if you have one or more exons highlighted, or all HSPs if you have one HSP highlighted (HSPs are the "blocks" that you see in the homology columns, such as ESTs and protein hits).

### 4.3 Toolbar

Many functions are available from the toolbar section in ZMap. Below the toolbar is an information bar, which displays information about the current focus feature. See figures 9 and 10.

### 4.4 Coordinate display (ruler)

The ruler to the left of the canvas shows the coordinates in the current coordinate system. By default, ZMap uses a local coordinate system, which starts at 1 for the first coordinate in the loaded region. You can toggle to the real toplevel coordinates using the menu option `View->Toggle coords`.

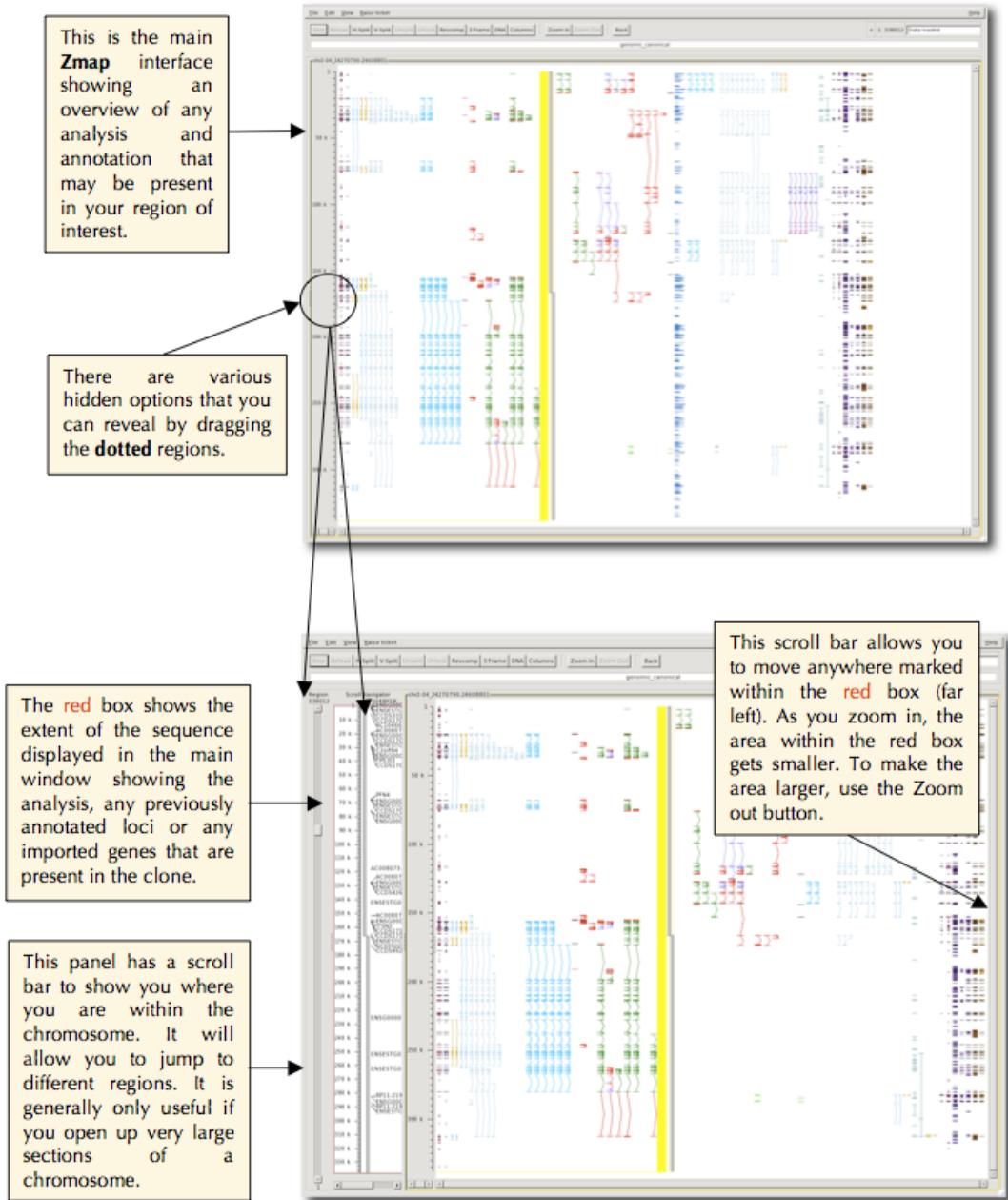


Figure 6: The main ZMap interface

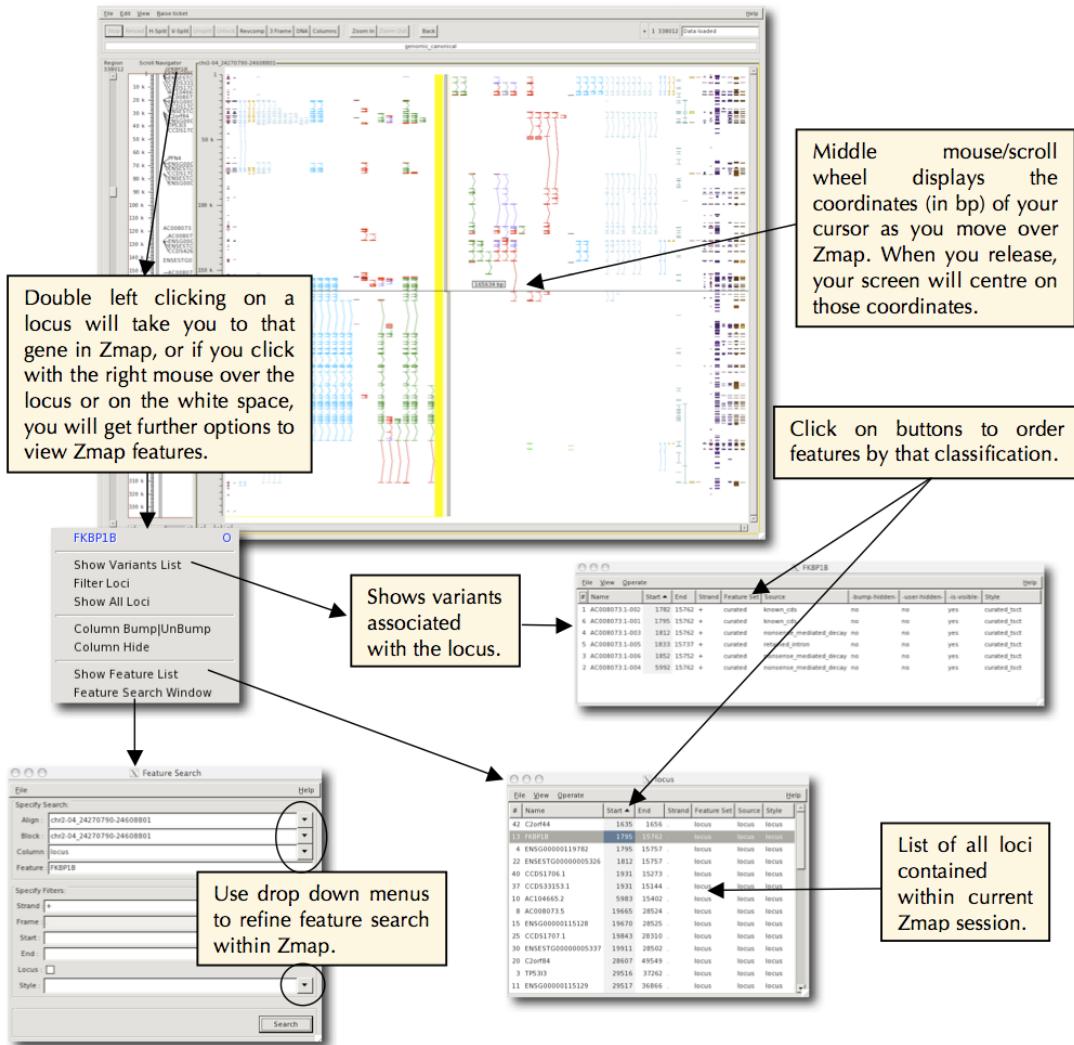


Figure 7: Navigating

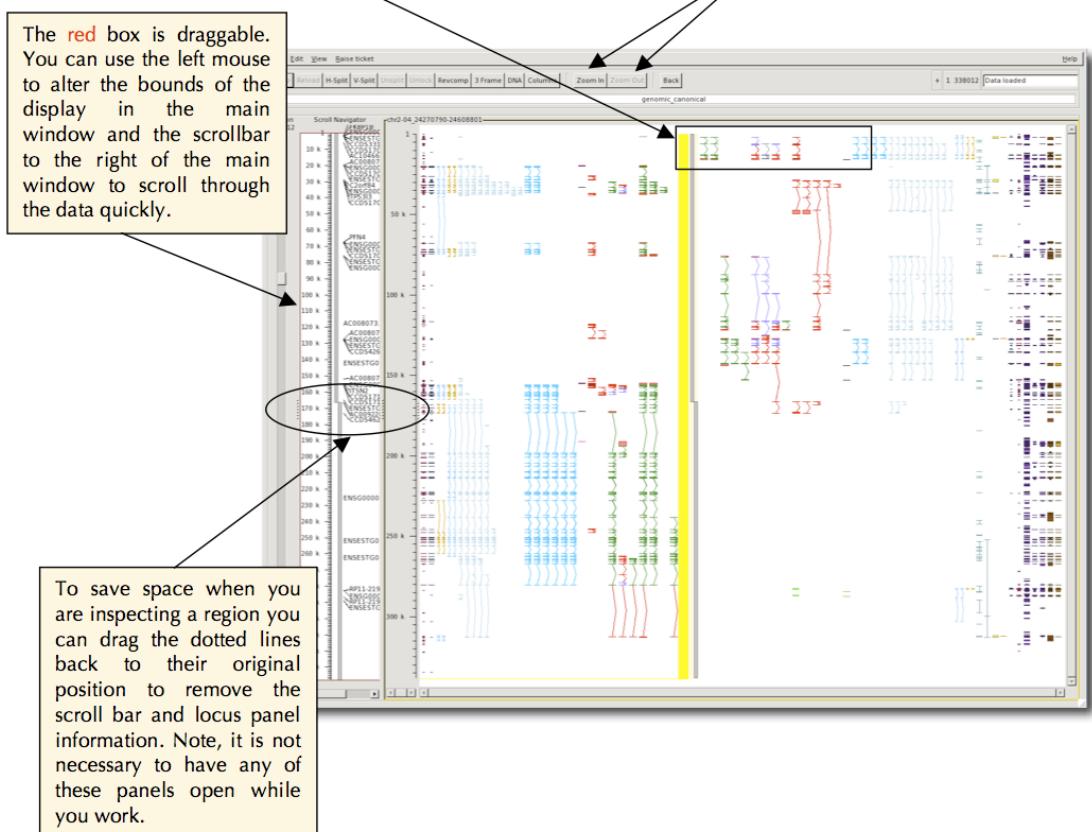


Figure 8: Zooming

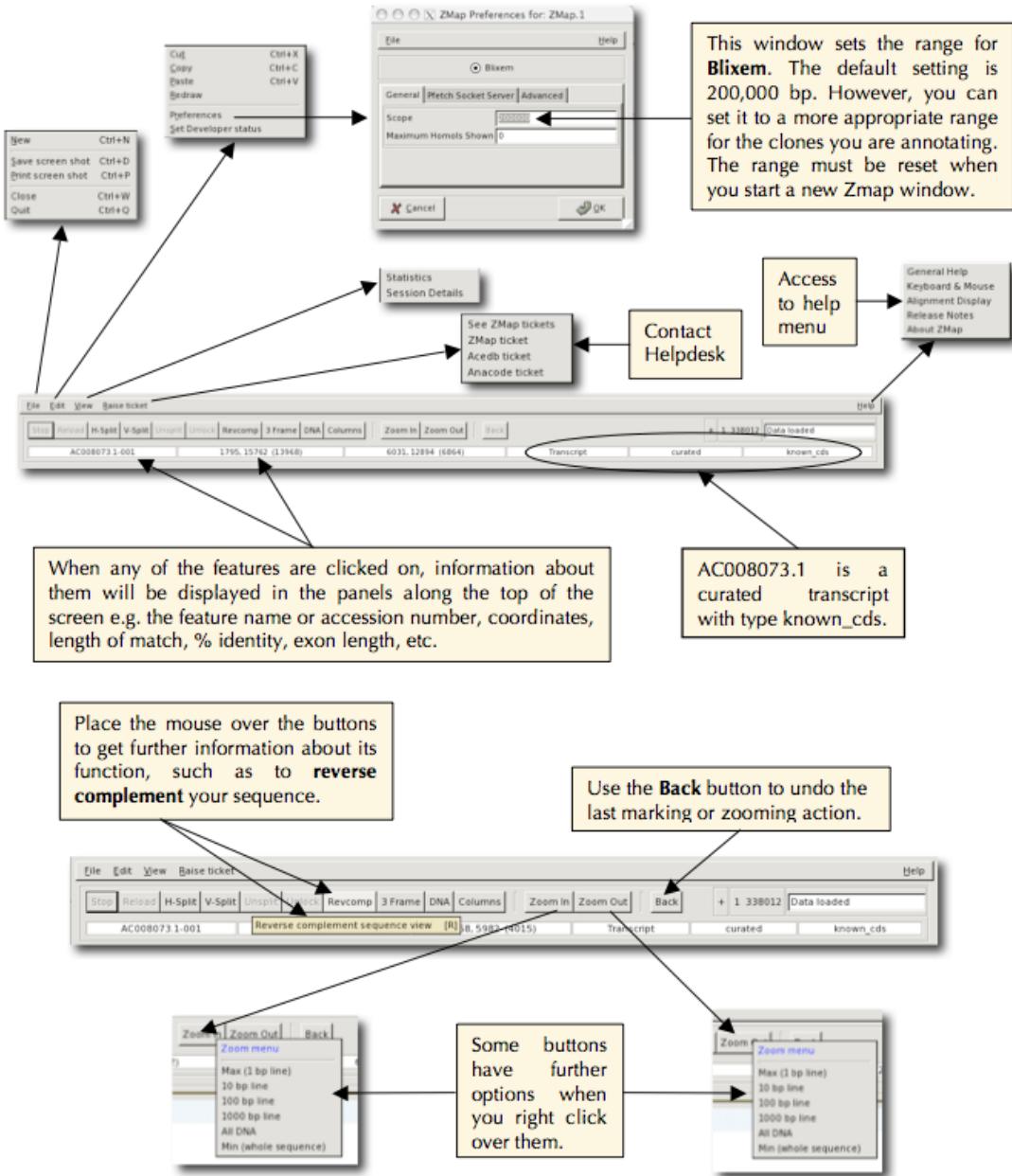


Figure 9: Toolbar feature details

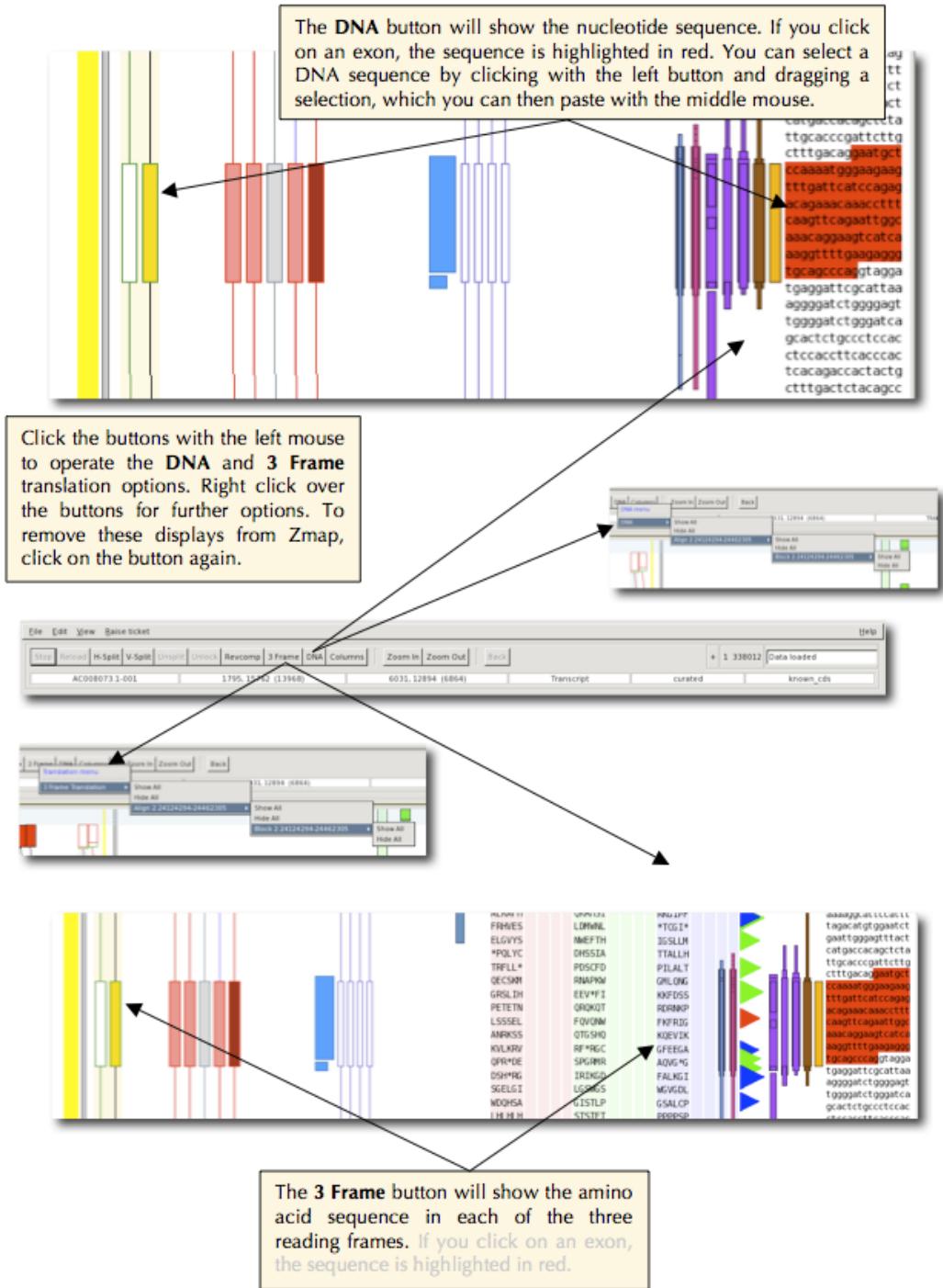


Figure 10: Feature display details

## 4.5 Navigator pane

The navigator pane is on the left of the ZMap display but is hidden by default. It can be exposed by dragging the drag bar to the left of the ruler.

The navigator pane shows the analysis, any previously annotated loci or any imported genes that are present in the clone. Double-clicking on a locus will take you to that gene in ZMap. Right-click to show various other options.

The extent of the current canvas is shown by a red outline in the navigator. This is the region that can be scrolled using the scrollbars in the main canvas area. Zooming in or out on the main canvas changes the size of this box. Alternatively you can drag the red box to scroll the current canvas area, or drag the bounds of the red box to change the zoom level of the main canvas area. See figures [6](#) and [7](#).

## 4.6 The Focus Feature

If you click on a column background then that column becomes the "focus" column and you can do various short cut operations on it such as pressing "b" to bump it. If you click on a feature then that feature becomes the "focus" feature and similarly you can do various short cut operations on it such as zooming in to it. (Note when you select a feature then its column automatically becomes the focus column.)

## 4.7 Marking a region

While the focus-feature facility is useful, the focus changes every time you click on a new feature. Sometimes you want to select a "working" feature or area more permanently. To do this you can "mark" a feature or region and it will stay marked until you unmark it (see section [4.7](#)).

The "marked" area remains clear, while the unmarked area above and below is shaded out with a blue hashed overlay (see figure [11](#)). Working on a marked region is recommended when you have a large amount of data loaded, because operations are faster and the display is less cluttered.

### 4.7.1 Mark a feature

1. Select a feature to make it the focus feature.
2. Press "m" to mark the feature, the feature will be highlighted with a blue overlay.

Feature marking behaves differently according to the type of feature you highlighted prior to marking and according to whether you press "m" or "M" to do the marking:

1. If you press "m", the mark is made around all features you have highlighted, e.g. a whole transcript, a single exon, several HSPs.
2. If you press "M" to do the marking around transcripts the whole transcript becomes the marked feature and the marked area extends from the start to the end of the transcript.
3. If you press "M" to do the marking around alignments all the HSPs for that alignment become the marked feature and the marked area extends from the start to the end of all the HSPs.
4. If you press "M" to do the marking around all other features: the feature becomes the marked feature and the marked area extends from the start to the end of the feature.

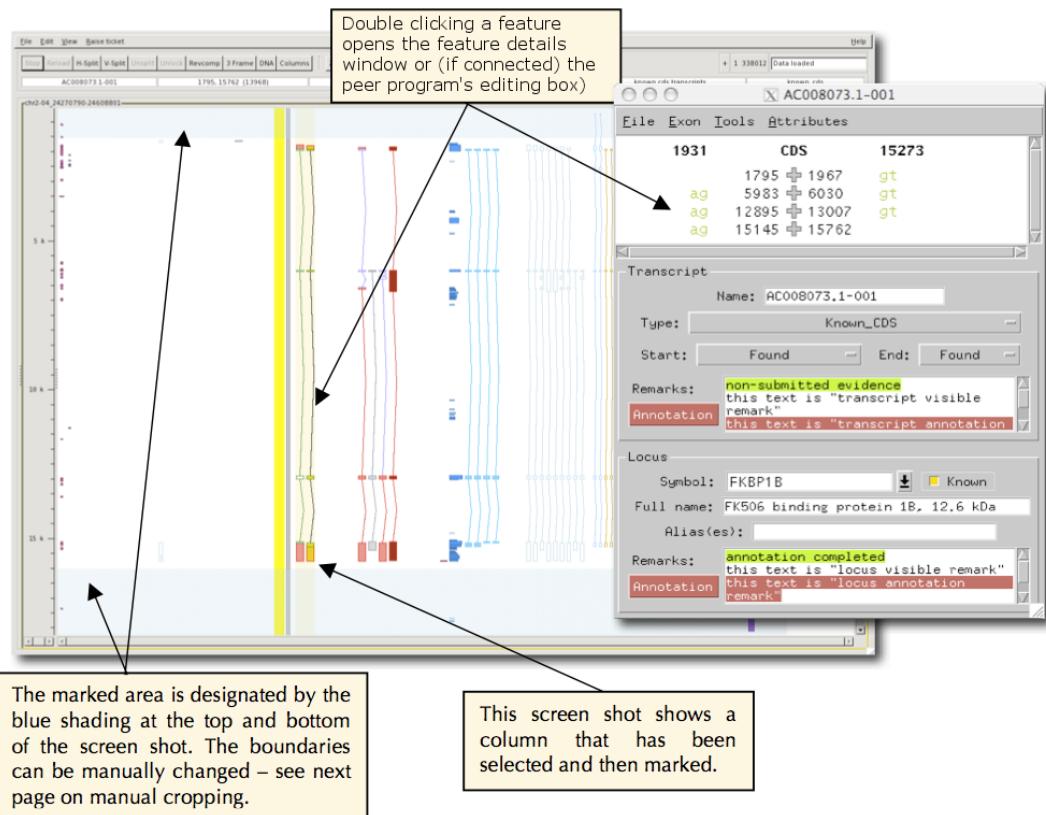


Figure 11: Focus and Mark

5. If no feature is selected but an area was selected using the left button rubberband then that area is marked.
6. If no feature or area is selected then the visible screen area minus a small top/bottom margin is marked.

#### 4.7.2 Mark an area

1. Select an area by holding down the left mouse button and dragging out a box to focus on that area.
2. Press "m" to mark the area.

#### 4.7.3 Manual cropping of the marked borders

You can manually change the borders of the marked area by putting your cursor over this area and using the cropping tool by clicking and holding with the left mouse button and dragging to make the area bigger or smaller.

#### 4.7.4 Unmark a feature

Press "m" or "M" again, i.e. the mark key toggles marking on and off.

### 4.8 Splitting the window

Use the `split` window function to effectively reduce the size of the window when looking at homologies. This is of particular use when you have to deal with very large introns because you can essentially reduce the introns to whatever size you wish, or when there are very many HSPs, because you can keep your gene object in view and static, but still scroll across the evidence. See figure 12.

The screen can be split horizontally or vertically (as shown) multiple times. An active window must be selected for **splitting**.

**Unsplit** will remove the last split window.

The windows will be locked together when you first open them. To scroll independently within each window, use the **Unlock** button.

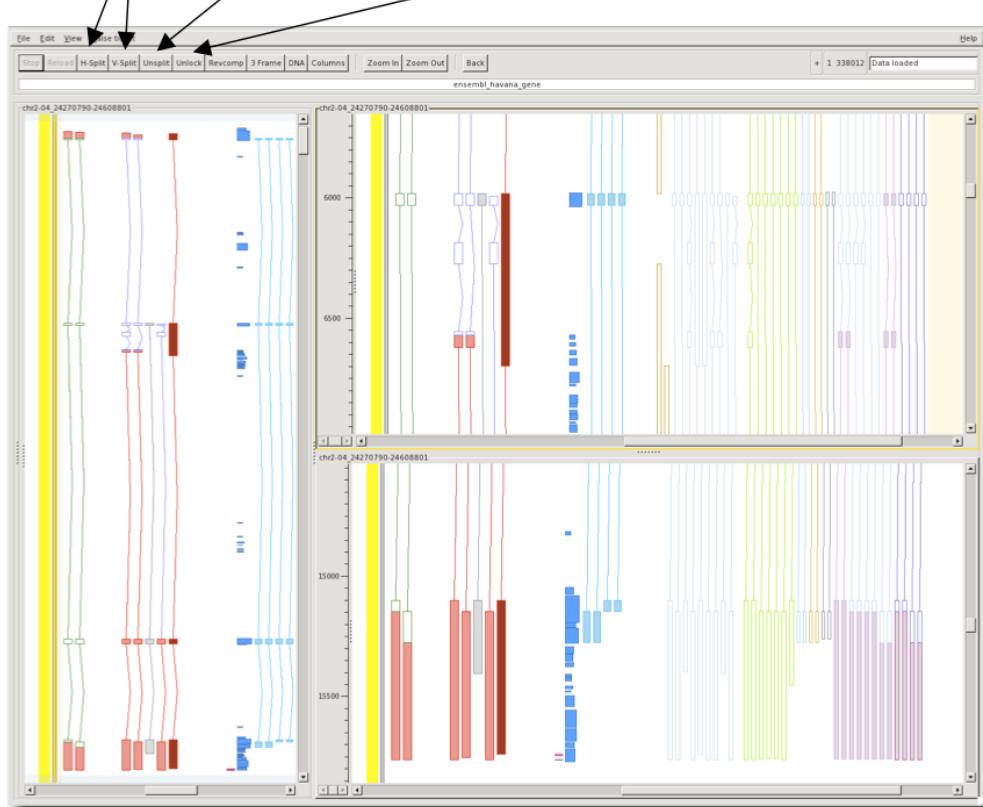


Figure 12: Splitting windows

## 5 Feature display

### 5.1 Bumping features

This section describes how to select a feature, mark it and then zoom in to it and examine evidence that overlaps that feature. The default setting for ZMap is to show HSPs drawn on top of each other. This saves space on the canvas making it easier to see the general features of the region of interest. The bump option allows you to see the HSPs as multiple alignments.

1. Click on the feature you are interested in (perhaps a transcript)
2. Mark it by pressing "m"
3. Zoom in to the feature by pressing either "z" or "Z" (as described previously).

Now when you bump an evidence column to look at matches that overlap the feature you will find that bumping is much faster because only those matches that overlap the feature get bumped and you also have fewer matches to look at. The quickest way to bump a column is:

1. Click on the column to select it.
2. Bump it by pressing "b" (if you press "b" again the column will be unbumped). If you have marked a feature then bumping is restricted to matches that overlap that feature, otherwise bumping is for the whole column.

If you use the default bumping mode (i.e. you pressed "b") then you will find all matches from the same piece of evidence are joined by coloured bars, the colours indicate the level of colinearity between the matches (see figure 13).

1. **Green:** the matches at either end are perfectly contiguous, e.g. 100, 230 —> 231, 351
2. **Orange:** the matches at either end are colinear but not perfect , e.g. 100, 230 —> 297, 351. Matches may also be this color when there are extra bases in the alignment, e.g. around clone boundaries.
3. **Red:** the matches are not colinear, e.g. 100, 230 —> 141, 423

Alignment quality of the HSPs is depicted by the width of every alignment displayed since the width is a measure of that HSP's score. Therefore, the wider it is the closer the score is to 100%. The precise score is displayed in the ZMap details bar by clicking on the alignment. If HSPs are missing either the first or last Blast alignments in the set, they are marked with a red diamond at their start/end respectively. This indicates if they do not start at the first base/amino acid and/or do not end with the last base/amino acid of the alignment sequence. The screen shot below shows what options you get when you right click over a homology - note that you can also select an HSP and type "o". You also get further options such as retrieving the EMBL file for that homology using pfetch and starting **Blixem**, see later section (note, HSPs do not need to be bumped to use Blixem).

### 5.2 Feature styles and strand

ZMap has rich feature display capabilities, which are highly configurable. Different features are displayed in distinct columns, and can be configured with many different styles as shown in figure 14.

1. The thick yellow line represents the genomic sequence; everything to the left represents the negative strand and everything to the right the positive strand. DNA matches (i.e. ESTs,

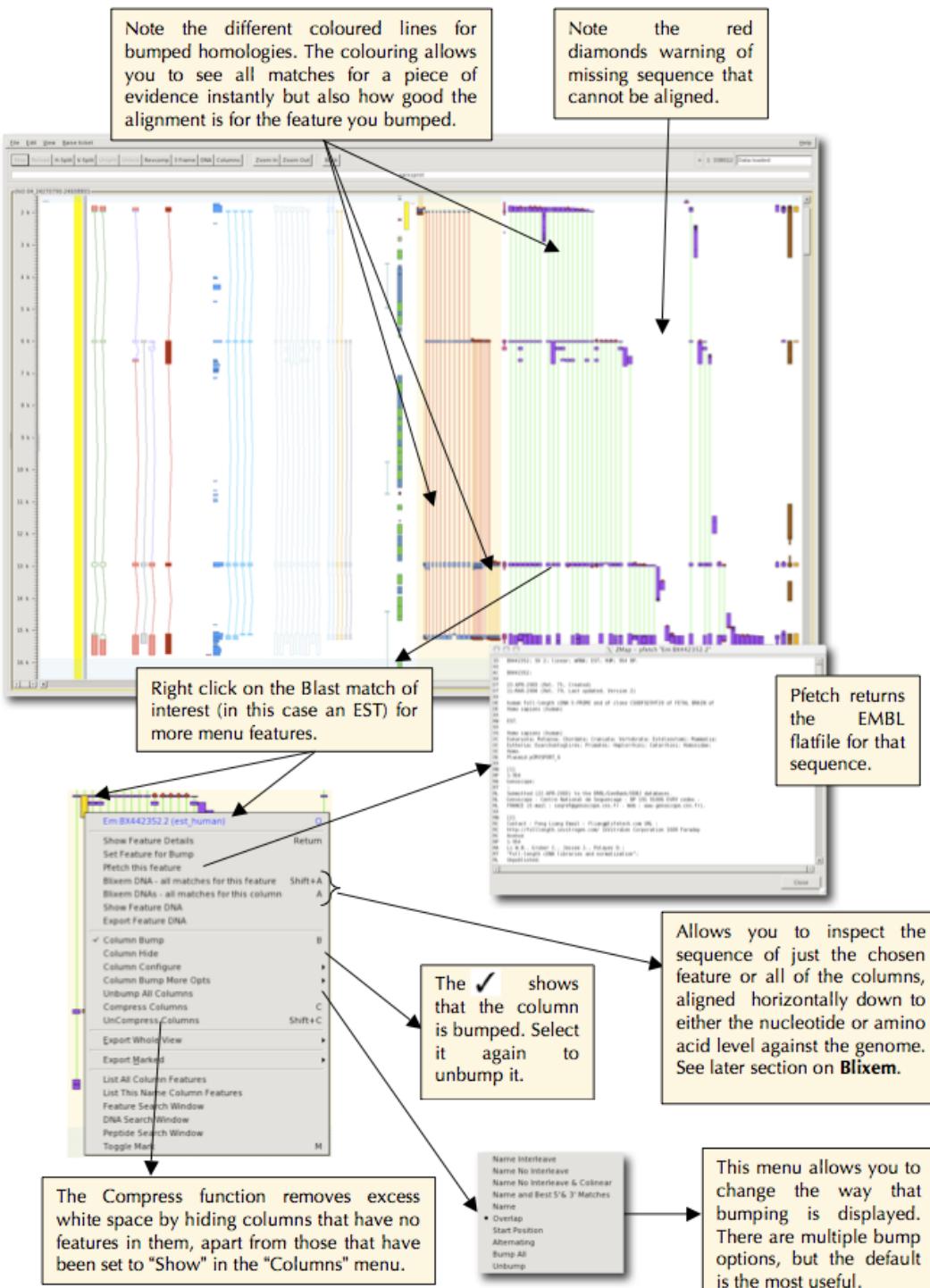


Figure 13: Bumping features

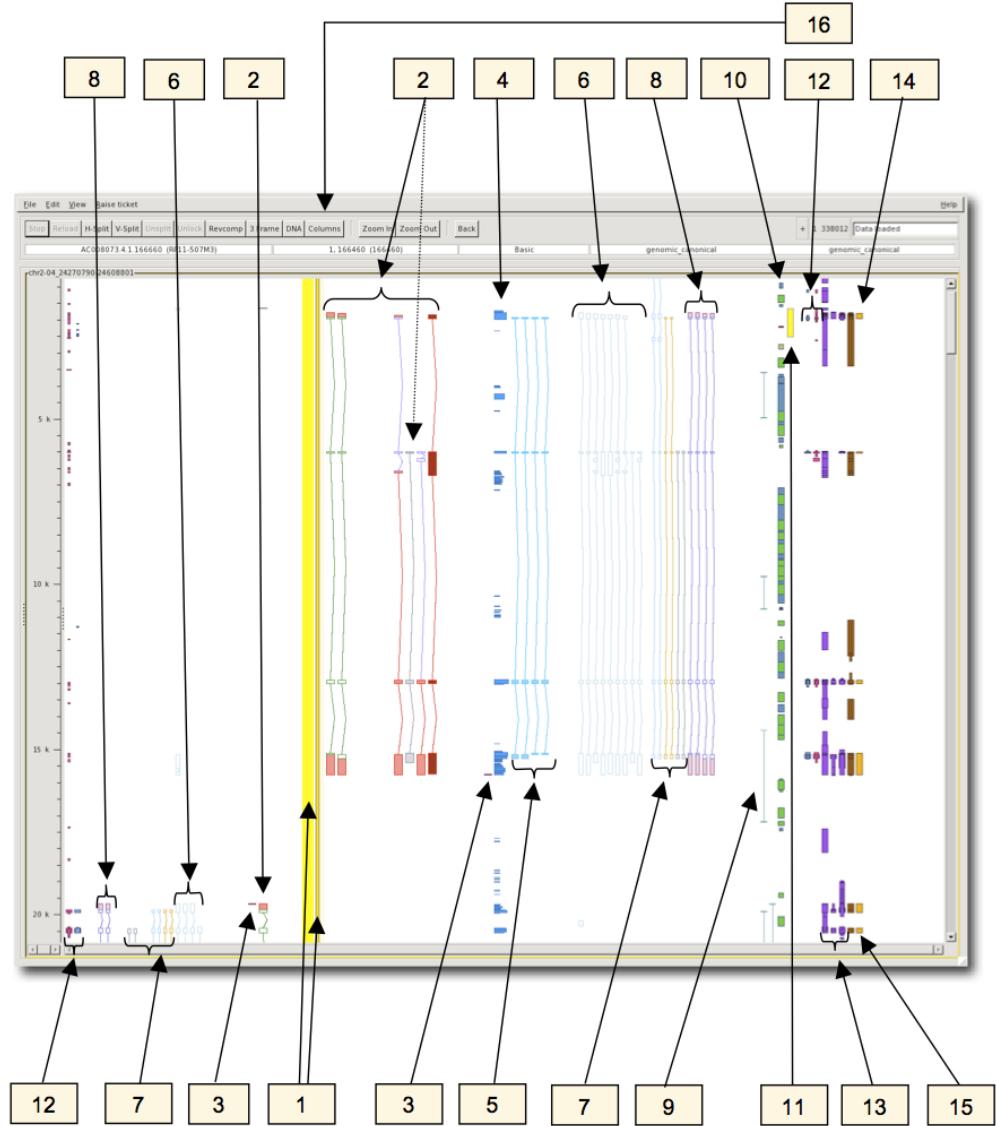


Figure 14: Features

mRNAs and RefSeq) and repeats have been configured so that they are all displayed to the right of the center, although they may align to either strand. Other features are configured to display in their relevant strand. The thin bar to the right is the clone that the genomic sequence is made up from. When connected to the Otter peer, double-clicking on this gives access to the DE editing window.

2. Annotated transcripts; green is coding (CDS), red is non-coding (UTR and transcript variants) and purple shows the "coding" region of NMD variants. Grey transcripts (see dotted line) contain exons outside the sequence slice being viewed and should not be confused with Halfwise hits.
3. Curated features, such as PolyA features are seen as horizontal black lines.
4. Phastcons44 - conserved regions detected using multiple sequence alignments of 44 organisms.
5. Imported annotation from CCDS (human and mouse only).
6. Imported transcripts via DAS source. Here PASA\_ESTs are shown.
7. Predicted transcripts such as Genscan (pale blue), Augustus (gold) and Halfwise predictions of Pfam (grey).
8. Imported annotation from Ensembl.
9. gis\_pet\_ditags and chip\_pet\_ditags are indicators of transcript boundaries.
10. Repeats ( blue=Line , light green=Sine , gold=other ), tandem repeats are red.
11. CpG islands appear as yellow boxes.
12. Protein matches are strand specific - SwissProt are light blue and Trembl pink.
13. EST matches are displayed as purple blocks and are broken down into human ESTs, mouse ESTs, and other ESTs from other organisms. 5' reads are on the left and 3' on the right.
14. mRNA matches contains all species and are displayed as brown blocks,
15. RefSeq matches are the orange blocks.
16. View and configure columns (i.e. the available features and analysis) - see section [5.3](#).

### [5.3 Column configuration](#)

The Column Configuration dialog (figure 15) is used to control which features are displayed, to specify the order of columns, and to edit the style and other properties of the columns. The dialog can be opened from the **Columns** button on the toolbar, or via the right-click menu (`Column Configuration ->Configure this/all columns`).

Each available track is listed and its visibility, group and style are shown.

#### **5.3.1 Selecting rows**

You can select one or more rows in the dialog to perform operations on them. To select multiple rows, you can hold down Ctrl (for one-by-one selection) or Shift (for range selection) while clicking with the mouse, or hold down Shift and press the up/down arrows.

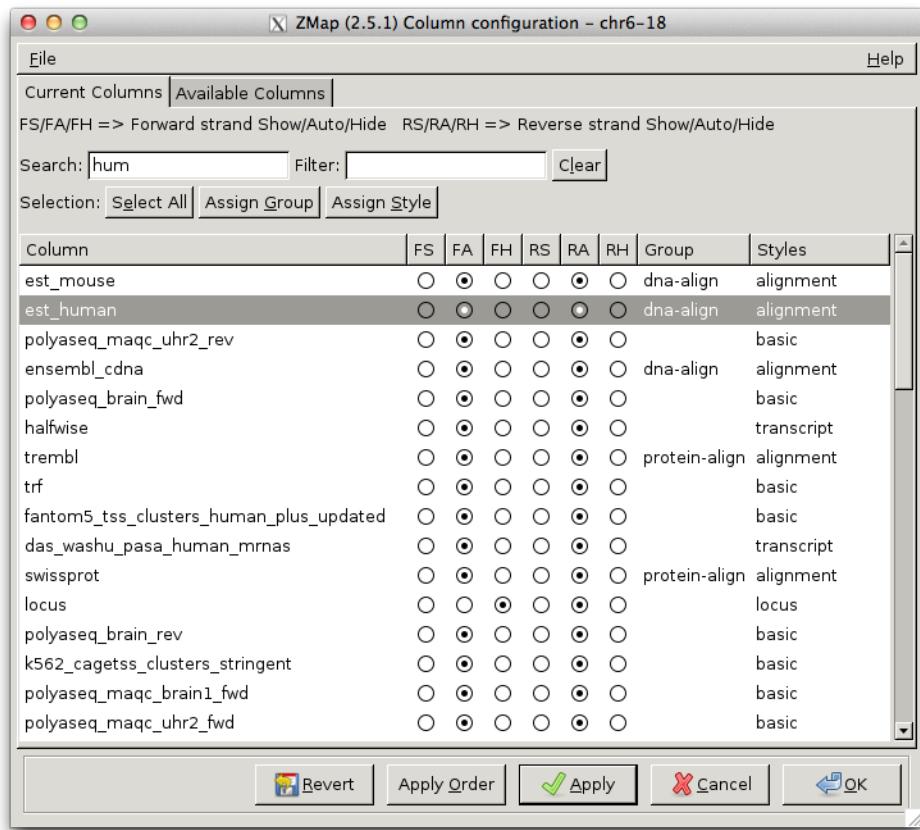


Figure 15: Column configuration dialog

### 5.3.2 Search and filter rows

You can quickly jump to a row containing specific text using the **Search** box. As you type in the box, the row selection will jump to the first row containing that text. You can use the up/down arrows to jump to the next/previous matching row.

You can filter the displayed rows by specific text using the **Filter** box. Type in the required text and press **Enter** and the rows will be filtered. Click the **Clear** button to clear the filter and go back to viewing all rows.

### 5.3.3 Track visibility

The visibility of each track (for both the forward and reverse strand) is controlled by a set of radio buttons with the options “show”, “auto” and “hide”:

- **Show** - the column will always be visible
- **Auto** - column visibility will be controlled automatically depending on the way the column is configured, e.g. the column may only be shown at certain levels of zoom
- **Hide** - the column will always be hidden

To change the visibility, select the track(s) you want to change and then click the relevant radio button in any of the selected tracks. Note that all highlighted tracks will be changed.

### 5.3.4 Assigning groups

You can assign a group to one or more tracks by selecting the tracks and clicking **Assign Group**. You will be prompted to enter a name for the group. The new group name will be populated on the Columns dialog. Remember to click **Apply** or **OK** on the Columns dialog to apply the changes.

Groups are used by various operations in ZMap such as running Blixem - when you run Blixem on a column, you have the option of running it on “associated columns”. This will run Blixem on all columns that are in the same group as the selected column.

### 5.3.5 Assigning styles

You can assign a style to one or more tracks by selecting the tracks and clicking **Assign Style**. This will open the **Choose Style** dialog where you can select the required style, or create a new one (see section [5.4](#)). After choosing a style, the new style name will be populated on the Columns dialog. Remember to click **Apply** or **OK** on the Columns dialog to apply the changes.

### 5.3.6 Reordering tracks

Tracks can be reordered by dragging and dropping rows in the Columns dialog. You can also sort the rows by name, group or style by clicking on the appropriate header. You must click the **Apply Order** button to apply the new track order in ZMap. Note that the main **Apply** and **OK** buttons do NOT apply the column order in ZMap.

### 5.3.7 Saving configuration

Column configuration can be saved in one of two ways. To save the configuration to your general ZMap settings so that will apply in all future sessions, go use the menu option **File->Save** on the **Column configuration** dialog. Alternatively, to save the configuration to a separate config file, so that you can load it selectively in future, use the menu option **File->Export->Configuration** from ZMap's main menu.

## 5.4 Style configuration

You can view and edit the styles of all featuresets in ZMap from the **Styles** dialog. Open the dialog by going to the menu **Edit->Styles**. You can also open the Styles dialog from the **Column configuration** dialog by selecting **Assign Style**.

All featuresets are assigned a default style on startup if they do not already have styles configured (see the `styles.shtml` help file for information about pre-configuring styles). Styles are usually organised into an hierarchy whereby child styles inherit the properties of their parent styles. The Styles dialog shows a tree of the styles hierarchy. The primary default styles, “basic”, “alignment” and “transcript”, are derived from “root”.

Existing styles can be edited by selecting the row in the Styles dialog and clicking the **Edit** button. This will open the **Edit Style** dialog (see section [5.4.1](#)). You can add a new child style to the currently-selected row by clicking the **Add** button, or delete the currently-selected style using the **Delete** button.

### 5.4.1 Editing styles

The **Edit Style** dialog is opened from the **Styles** dialog when you **Edit** or **Add** a style. It can also be opened by right-clicking on a feature and going to the menu options **Column configuration->Edit Style**. When editing an existing style, the new name will default to the same as the existing name. This can be changed if you want to rename the style. The list of featuresets that currently use the style are shown. This is so that you can be sure you want to edit the style for all of those featuresets. If you do not want to change the style for of the featuresets you can create a new child style by clicking the **Create child style** button. This style will be assigned to just the selected featureset if editing the style of a particular feature, or will be assigned to no featuresets otherwise (in which case you can assign it to featuresets later e.g. via the **Column configuration** dialog).

### 5.4.2 Saving styles

Styles can be saved to an `.ini` style configuration file. To do this, go to the menu **File->Save->Styles**. To load the styles in a future ZMap session, set up a configuration file and use the `stylesfile` key (see the `configuration.shtml` help file for details).

## 6 Operations

### 6.1 Show feature details

Right click on a gene object or 'o' key when highlighted to see information on otter IDs and Ensembl IDs. For BLAST hits, double click on the HSP to get the feature interface where you will find details on alignment and on what gene object the HSP has been assigned to as evidence, if any (see figure 16).

### 6.2 Exporting features for gene objects

Figure 17 shows how you can view and export an annotated sequence to your home directory in various different ways, such as dumping features directly.

1. In the main ZMap window, right click on an annotated gene object.
2. From the drop down menu select `Export Feature DNA` and choose sequence required from CDS, transcript, unspliced and with flanking sequence.
3. Alternatively, select `Export Feature peptide` and choose either CDS or transcript.
4. Figure 17 also shows the results of `Show Feature DNA` for annotated gene object AC008073.1-001 in FASTA format; firstly, the section of the transcript that corresponds to the CDS and secondly the whole transcript, including the untranslated region (UTR).

### 6.3 Searching for a sequence in ZMap

DNA and peptide search windows are provided from within ZMap and can be accessed by right clicking on ZMap space and selecting the option at the bottom of the menu. Both search windows are shown in figure 18.

### 6.4 Searching for a feature in ZMap

This option allows you to list all the features contained in a column in one window. There are further options for you to search within these results to find a specific feature. The list of column features can be exported as a GFF file via the File menu. See figures 19 and 20.

### 6.5 Selecting single or multiple features and hiding/showing them

1. If you left click once on a feature in ZMap, you will highlight all of its exons, the coordinates of which are now stored in the paste buffer and can be copied elsewhere, such as into the transcript editing window in Otter.
2. You can select multiple features by holding the Shift key down and left clicking with mouse (same as for multi select on the Mac, Windows etc). This option will highlight a single exon at a time for each feature, but the accession numbers of each feature and the individual exon coordinates are held in the paste buffer. This is a particularly useful way of selecting ZMap hits to use in the OTF alignment tool in Otter, as all selected homologies will be held in the paste buffer and automatically pasted into the OTF accession window. Each of the exon coordinates can also be pasted into the transcript editing window in Otter (see figure 21).

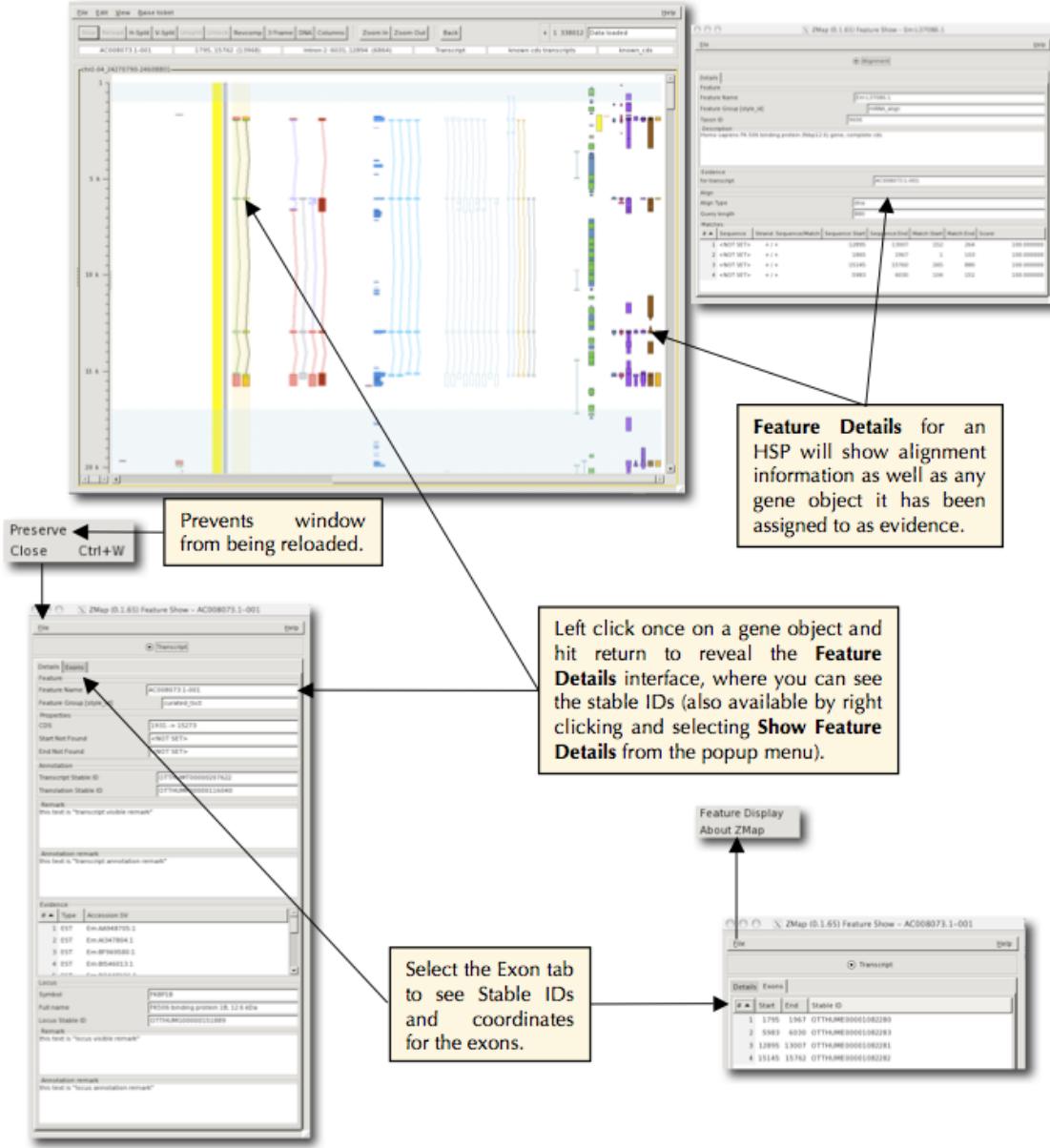


Figure 16: Feature details dialog

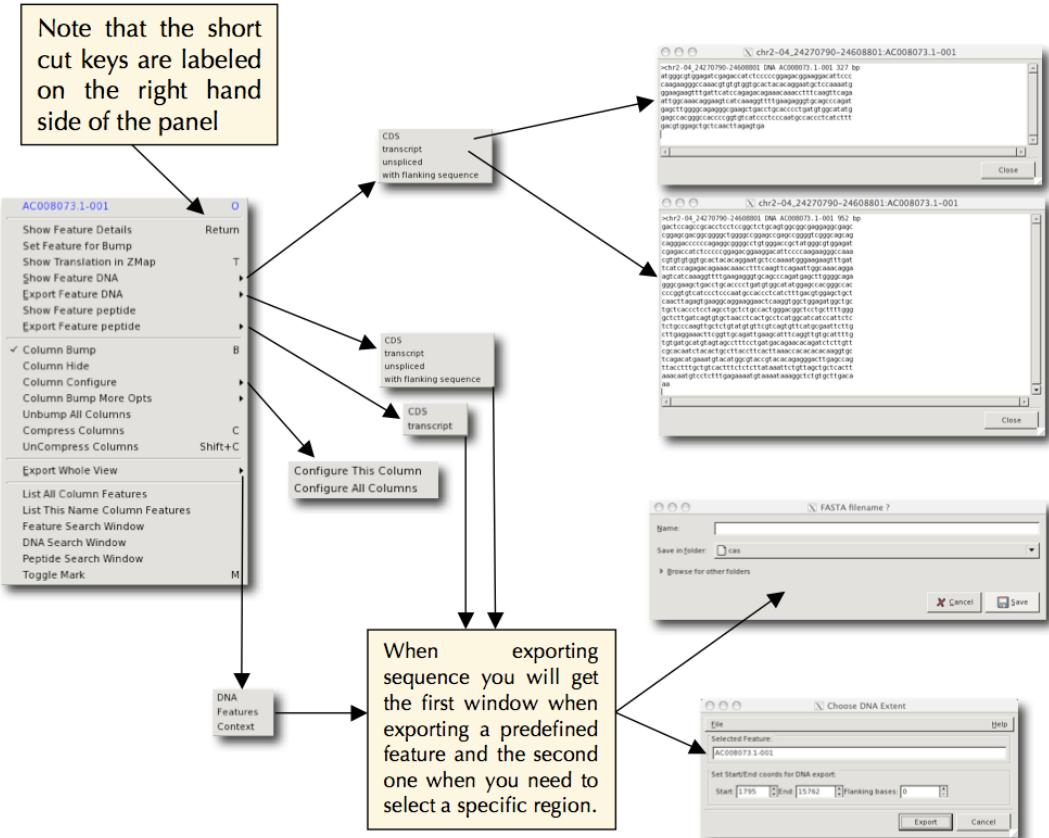


Figure 17: Exporting features

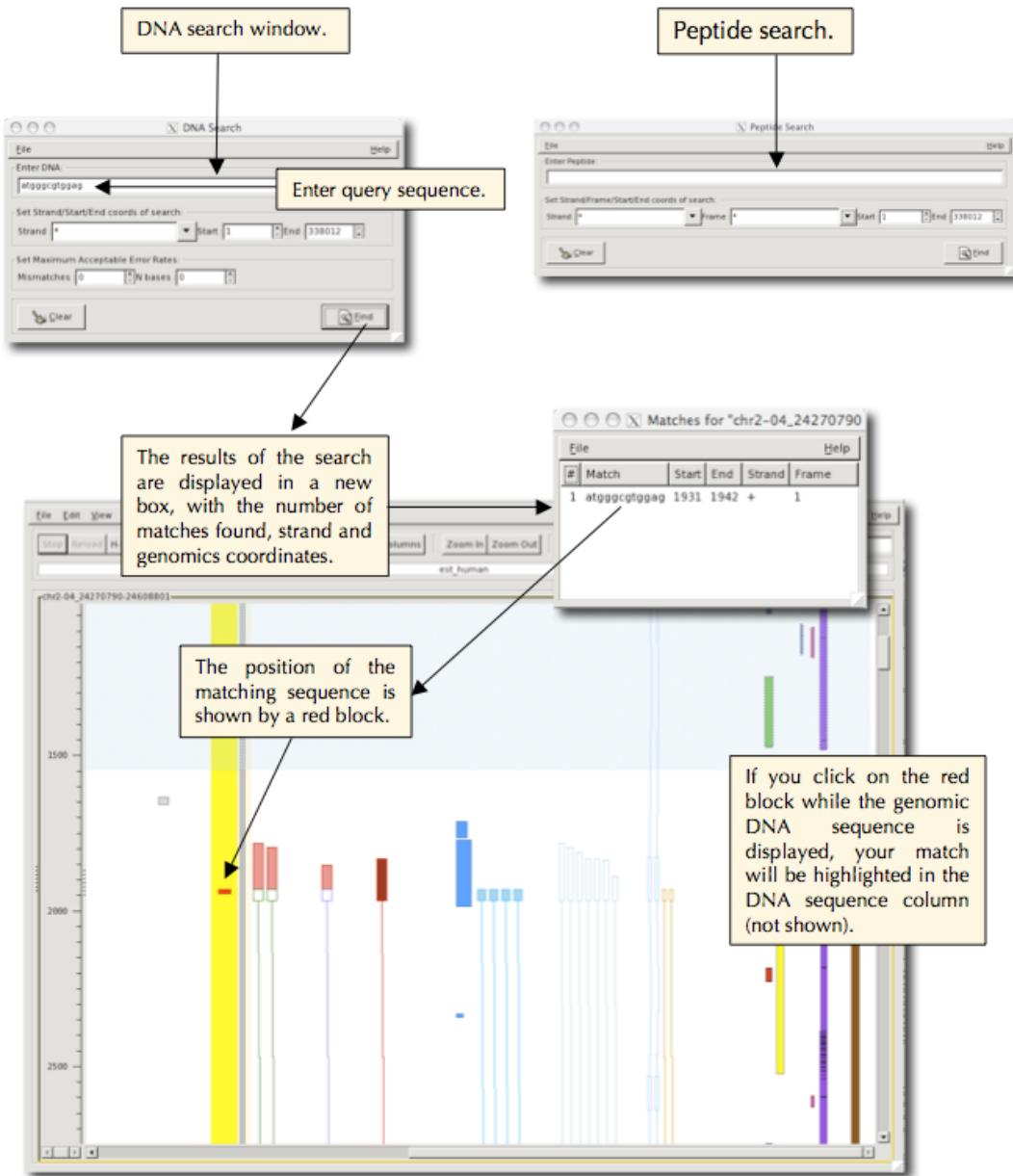


Figure 18: Searching for a sequence

Click over a column with the right mouse to activate this menu. Select **Show feature List**.

Export results as GFF file.

To search for a feature, enter your query here and click on search.

Note, the format needs to be correct for Zmap, so use \* as a wild card. For example accession numbers may have a database prefix and version suffix such as Em:U61167.1, so use the following format \*accession\_number\*, if you are not sure about the database and version.

This lists all the accession numbers and associated information for the column "vertebrate\_mrna". The results can be ordered using the buttons at the top.

The result lists all the exons and associated match information for query accession Em:U61167.1.

#	Name	Start	End	Strand	Query Start	Query End	Query Strand	Score	Feature Set	Source	Mode
9	Em:U61167.1	131013	134280	-	740	4112	+	99.699997	vertebrate_mrna	vertebrate_mrna	mRNA_align
1	Em:U61167.1	138209	198400	-	554	745	+	100.000000	vertebrate_mrna	vertebrate_mrna	mRNA_align
4	Em:U61167.1	388877	198973	-	457	553	+	99.000000	vertebrate_mrna	vertebrate_mrna	mRNA_align
5	Em:U61167.1	288709	200830	-	325	456	+	100.000000	vertebrate_mrna	vertebrate_mrna	mRNA_align
7	Em:U61167.1	200917	200962	-	289	334	+	100.000000	vertebrate_mrna	vertebrate_mrna	mRNA_align
8	Em:U61167.1	204425	204581	-	122	286	+	100.000000	vertebrate_mrna	vertebrate_mrna	mRNA_align
2	Em:U61167.1	208447	208511	-	57	121	+	100.000000	vertebrate_mrna	vertebrate_mrna	mRNA_align
3	Em:U61167.1	399988	210093	-	23	56	+	100.000000	vertebrate_mrna	vertebrate_mrna	mRNA_align

Figure 19: Searching for a feature

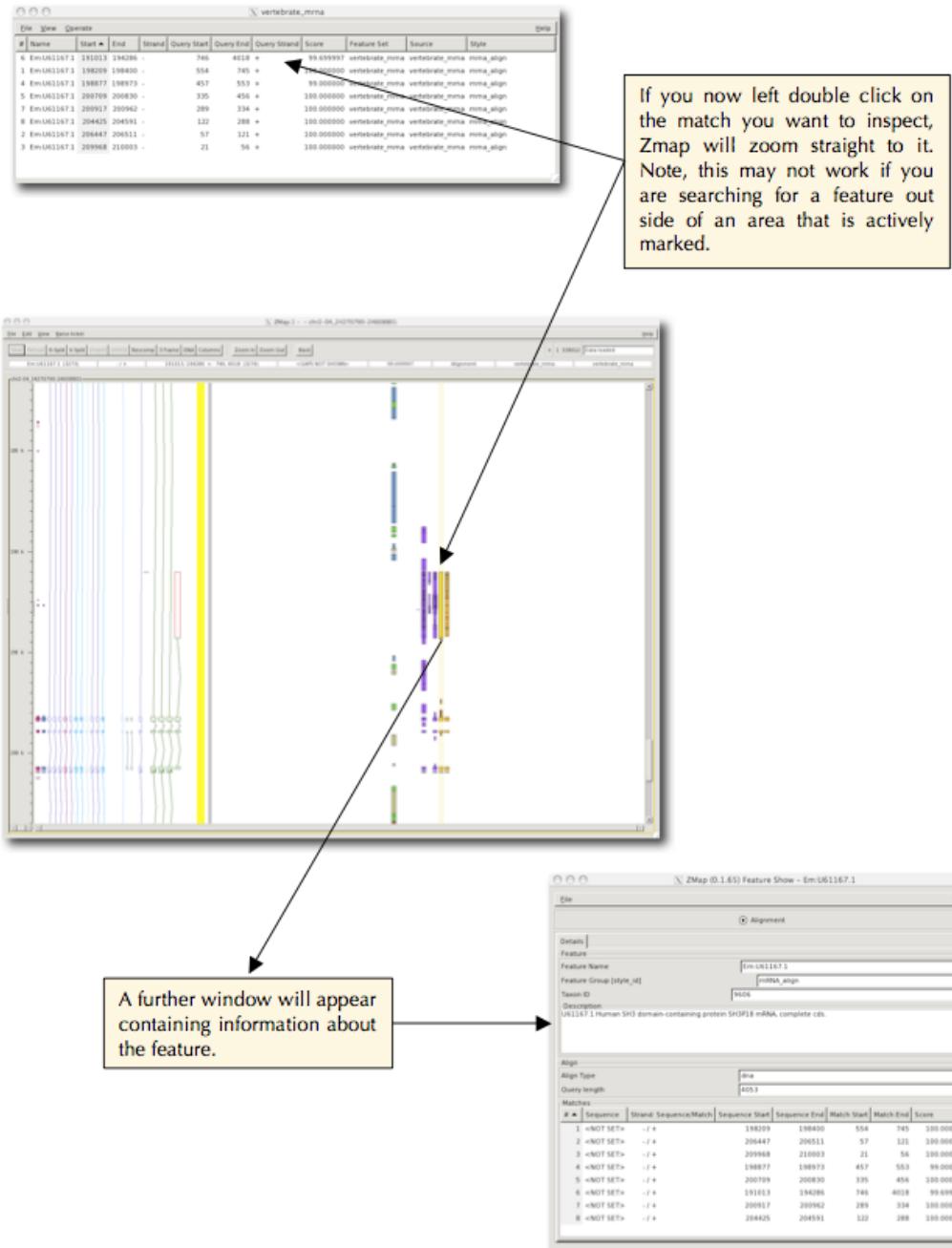


Figure 20: Feature search results

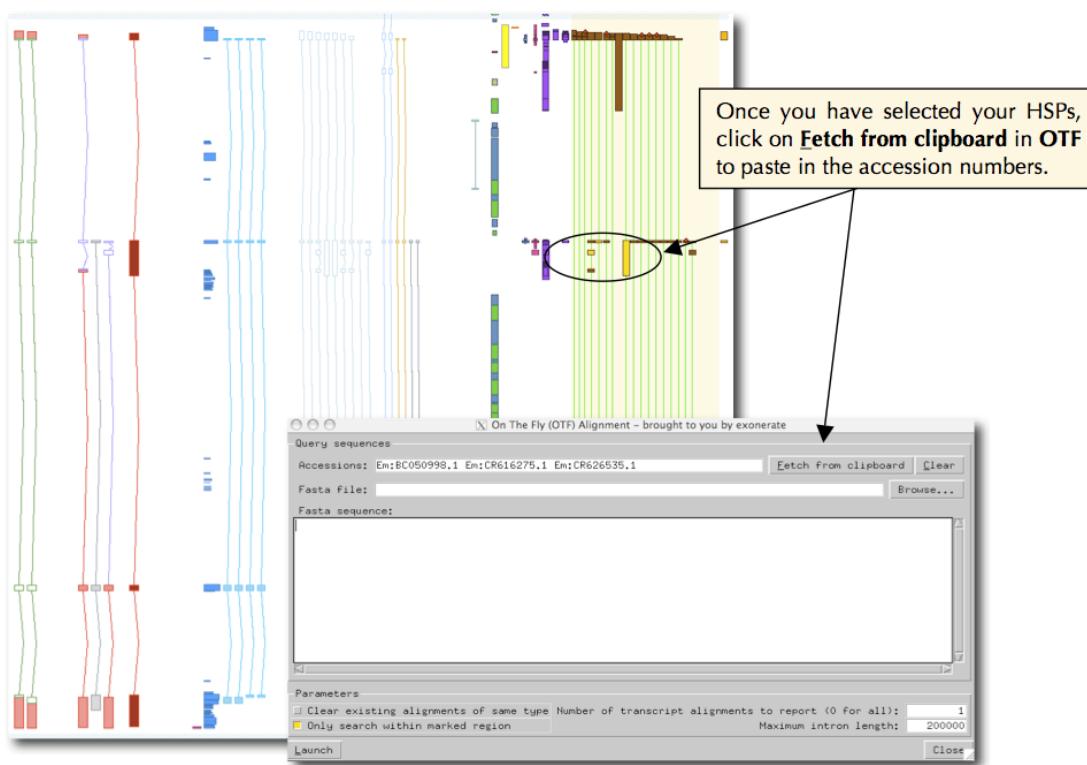


Figure 21: Pasting selected features to Otter's OTF dialog

3. You can remove selected features in ZMap by pressing **Delete** on the keyboard and restore them by pressing **Shift-Delete** (note on the Mac you need to press **Fn-Delete** and **Shift-Fn-Delete**). This is a particularly useful way of removing evidence that you have already assigned to a transcript object.

## 7 Annotation

ZMap can be used to create new features or edit existing features. ZMap can generate variant objects quickly - existing transcript objects can be used as a template for a new object while a ZMap HSP (or any other feature, nucleotide or peptide) can be used to provide the coordinates for the new variant. The annotation process is different depending on whether you are working within Otter or not.

### 7.1 Standalone ZMap

To create/edit features in ZMap, the workflow is roughly as follows:

- Enable Annotation
- Copy a feature with **Ctrl-K** to create a “temporary feature”
- Copy other features/coordinates to edit the temporary feature
- Double-click the temporary feature to save it to a “real” featureset
- Export the new feature to file

#### 7.1.1 Enable annotation

To edit features in ZMap you must first enable annotation. Go to the menu option **Edit -> Preferences** and tick **Enable Annotation**. Alternatively, if you use the keyboard shortcut **Ctrl-K** to copy your first feature to the Annotation column, then annotation will be enabled automatically.

#### 7.1.2 Create a temporary feature

Features are edited by copying them to a **temporary feature** in the **Annotation** column (see figure 22). The temporary feature can be based on any existing feature(s). The process is the same whether editing an existing gene model (in which case you will copy the gene model) or creating a new one (in which case you will copy the evidence features).

To create a new temporary feature, select the transcript (or other feature) you wish to create a variant from and press **Ctrl-K** (or right-click and select **Annotation -> Copy selected transcript(s)**)<sup>1</sup>). Alternatively, you can create a transcript from an alignment (or any other feature) by clicking the alignment and pressing **Ctrl-K** (or right-click and select **Annotation -> Copy selected alignment(s)**).

#### 7.1.3 Editing the coordinates

Once you have created the temporary feature, it can be edited by copying additional features into the Annotation column in the same way. Objects of any type, including splice markers, nucleotide coords, peptide coords etc. can be used to adjust feature/exon coordinates. Select the other features you want to use and use **Ctrl-K** or right-click and use the **Annotation** menu to copy those in to adjust

---

<sup>1</sup>Note that you need to enable the Annotation column from the Preferences dialog for the Annotation menu to appear. **Ctrl-K** will automatically enable the Annotation column if it is not already enabled.

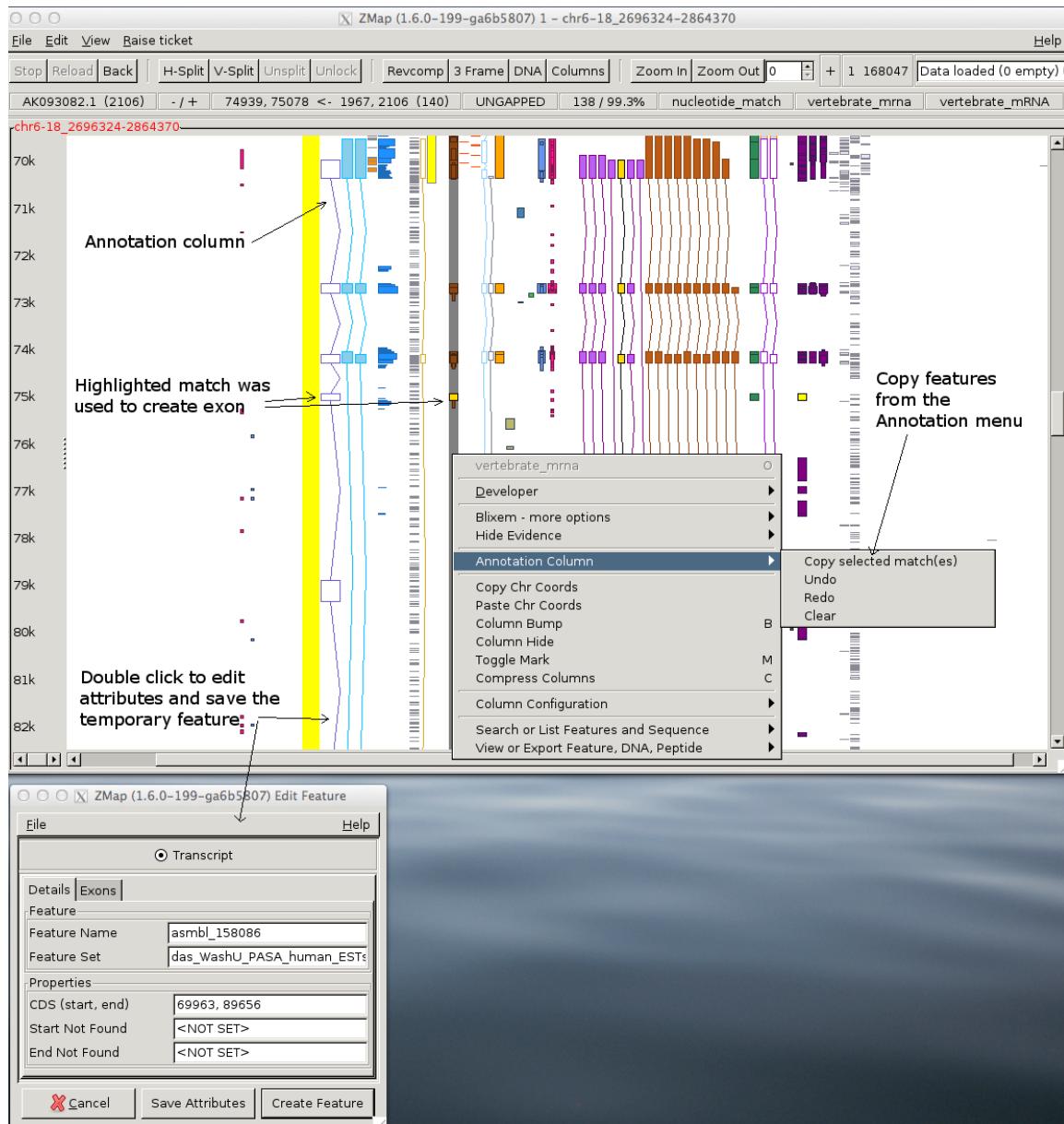


Figure 22: Feature editing in ZMap

the coordinates of the original feature. You can copy the entire feature or just the selected subfeature (e.g. the selected exon from a transcript or the selected hit from an alignment). You can use the keyboard shortcut **Ctrl-K** to copy the entire feature.

ZMap will do its best to make a sensible merge of the new coordinates, e.g. extending the feature extents if a coordinate lies outside the current range. If ZMap cannot automatically merge a coordinate it will ask for more information, e.g. if a coordinate lies within an intron, ZMap will ask whether it should form a 5' or 3' intron boundary.

You can delete exons and introns from the temporary feature by right-clicking on them and using the menu **Annotation -> Delete subfeature**.

You can edit exon coordinates manually in the attributes dialog (see section [7.1.6](#)).

#### 7.1.4 Setting the CDS

When you create the temporary feature, the CDS will be copied from the original feature, if it had one. The CDS can be edited in the attributes dialog (see section [7.1.6](#)). You can also set the CDS based on other features. To do this, right-click on the feature you want to base the CDS on and go to the menu **Annotation -> Set CDS**. From here, you have the option to set just the **Start** or **End** of the CDS from this feature, or to set both (**Range**). There is then a further option to choose whether you want to use the entire feature or just the clicked subfeature to set the CDS from.

#### 7.1.5 Editing the transcript's sequence

You can edit the transcript's sequence to see the effect of variation. This works by copying variation features into the Annotation column so you must have a variation column loaded. You will probably want to show the translation for the temporary feature first so you can see the effect of the changes. Click on the temporary feature and press **t** to show the translation (or right-click on it and select **Show Translation**). You may also want to show the three-frame translation (click the **3 Frame** button on the toolbar to show it).

To apply a variation, right-click on the variation and go to the right-click menu **Annotation -> Copy**. Select **This feature** to copy just the clicked variation, or **Selected feature(s)** to copy all selected features. You can use the keyboard shortcut **ctrl-K** to copy all selected features. The translation will change to show the effect of the variation. It may be useful to undo/redo the operation (see section [7.1.8](#)) to see the translation before and after the change.

#### 7.1.6 Editing the attributes

The coordinates and attributes of the temporary feature can be edited manually. Double-click the temporary feature to open a dialog showing its properties. Click on the **Exons** tab, where you can then edit the exon coordinates. You can also edit other attributes such as the final feature name and featureset on this dialog as well. If you are not ready to save the feature to a featureset yet, you can save these attributes to the temporary feature by clicking **Save Settings**. You can then continue editing the temporary feature, but the attributes will be remembered.

### 7.1.7 Saving the temporary feature

When you have finished editing your feature, you may wish to save it to a “real” featureset in ZMap so that you can clear the Annotation column and work on another transcript. You may then want to export either this single feature, or a group of features, to file. Alternatively, you may just wish to export the temporary feature directly to file. The process is different depending on what kind of save you want to do.

**Export the temporary feature** If you just want to export the temporary feature to a GFF file, right-click on it and use the menu option `Annotation -> Export temp feature`. You will then be able to choose a file to export the feature to. Only the temporary feature will be exported. Note that any attributes you set from the properties dialog.

**Save to a featureset** When you have finished editing your feature, the temporary feature can be saved to a “real” featureset (i.e. to move it out of the Annotation column). Double-click the temporary feature to open a dialog showing its properties. To save it as a new feature, give it a new, unique name. To overwrite an existing feature, give it the same name and featureset as the feature you want to overwrite. ZMap will warn you before it overwrites an existing feature so you don’t have to worry about overwriting one accidentally.

**Export to file** Saving the temporary feature to a featureset moves it into the relevant column in ZMap. However, it is not saved to a file anywhere yet. To save it to file, you must export it to GFF. To export a single feature, right-click on it and use the menu option `View or Export Feature... -> Export -> Clicked feature`. To save all features, go to the main menu and select `File -> Export -> Features`. You will be prompted for a file name to save the feature(s) to.

### 7.1.8 Other operations

**Undo/redo** The right-click Annotation menu offers `Undo` and `Redo` options. These can also be accessed with the keyboard shortcuts `Ctrl-Z` and `Ctrl-Y`.

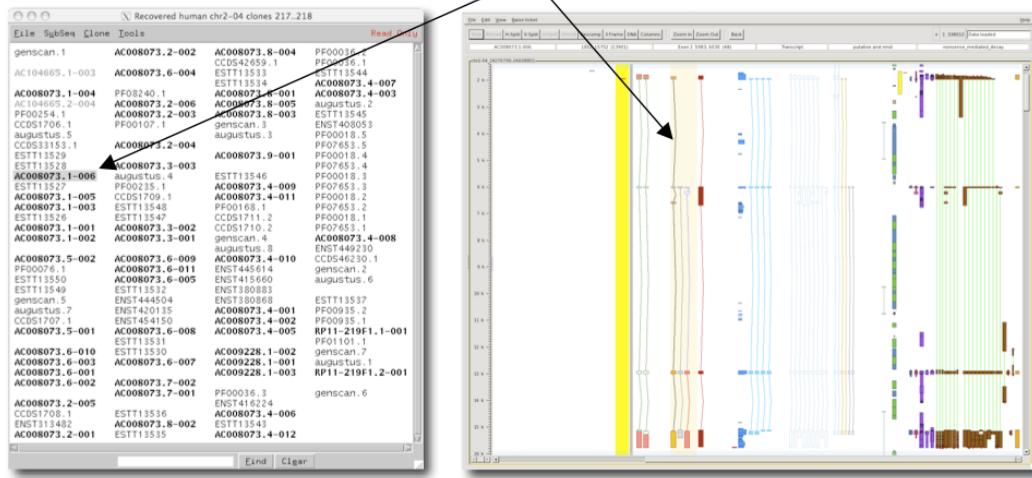
**Clear the Annotation column** To clear the Annotation column ready to start again, right-click anywhere in ZMap and select `Annotation -> Clear`. You may undo this operation if you select it by accident.

**Highlight evidence** This option lets you highlight all of the feature(s) that were used to construct the temporary feature. Right-click on the temporary feature and select `Highlight evidence`. The evidence features will be highlighted in yellow.

## 7.2 Variant construction with Otter

When running ZMap under Otter, variants are constructed in a different manner. See figures 23 and 24.

1) Select the object that will form the foundation to the new variant, either by highlighting the object in Otterlace or clicking on the object in Zmap.



2) Click on the HSP that will give its coordinates to the new variant object.

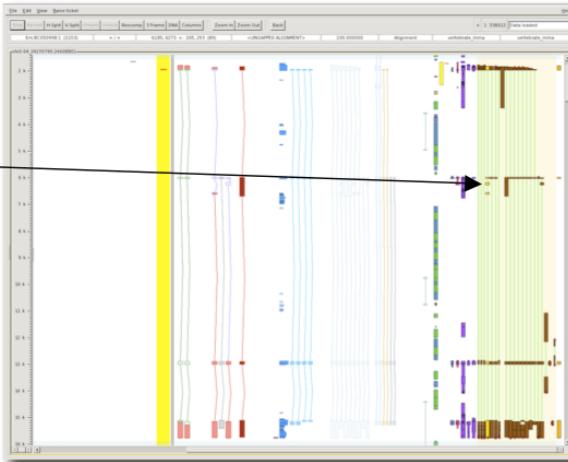


Figure 23: Variant construction in Otter (1)

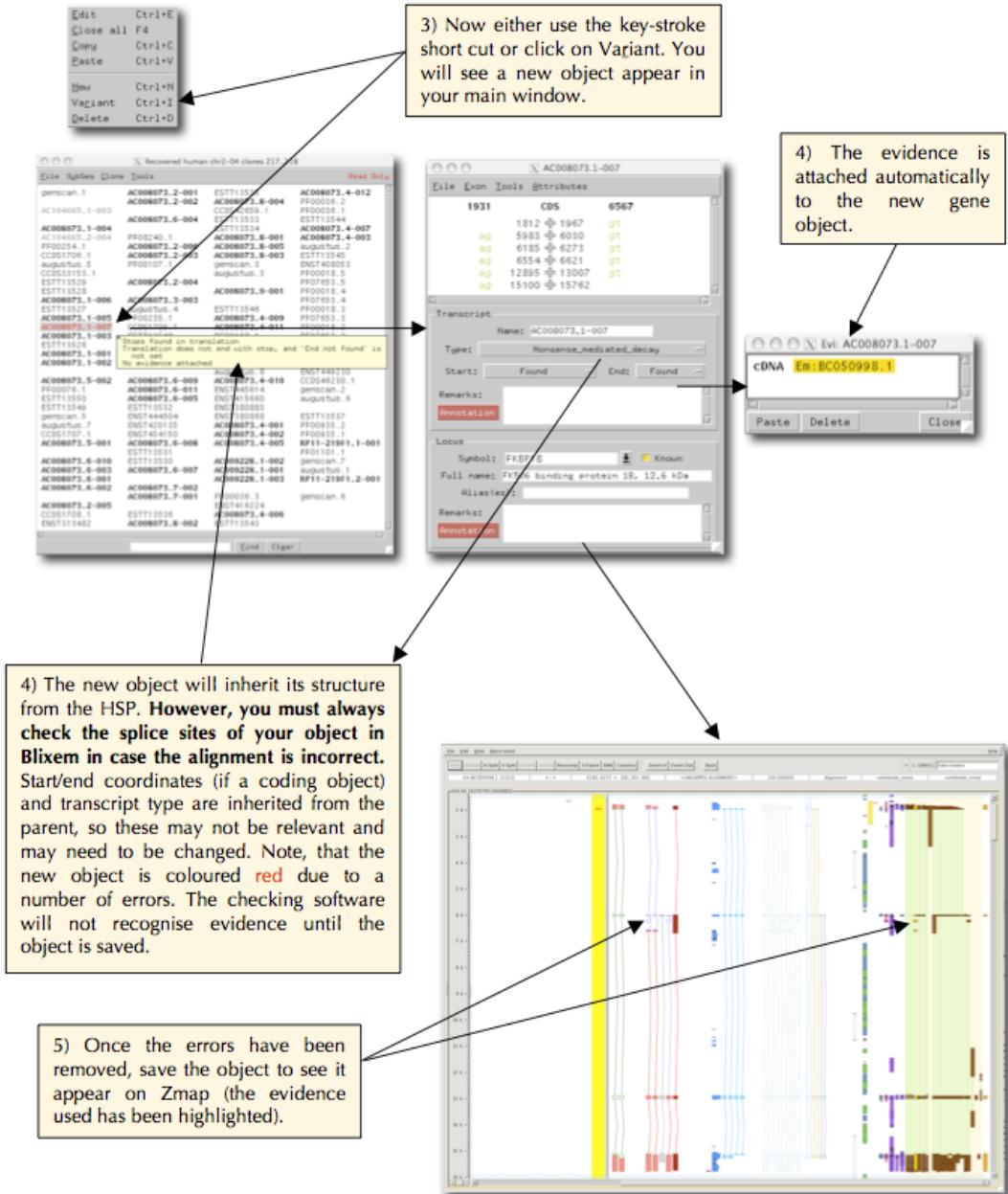


Figure 24: Variant construction in Otter (2)

## 8 Multiple views

This function allows you to open two or more sequences alongside each other (such as a human region and the syntenic region in mouse, or two haplotypes), so that simultaneous investigation can be carried out.

### 8.1 Multiple views from GFF

If you want have two different regions in GFF files, simply pass both GFF files to ZMap on the command line to view them side-by-side in ZMap:

```
zmap file1.gff file2.gff
```

### 8.2 Launching in a ZMap from Otter

To open multiple sequences from Otter, you will need to open both sets of clones in the same Otter session. To open both ZMap windows in one window as shown below, you need to select "Launch In A ZMap" option in one clone set. These clones will open to the left of the already open Otter session. This screen shot shows human gene SF3B14 and the syntenic region in mouse. The gene copy and paste function (referred to in the Otter section) is of much use here, saving time when building gene objects. See figure 25.

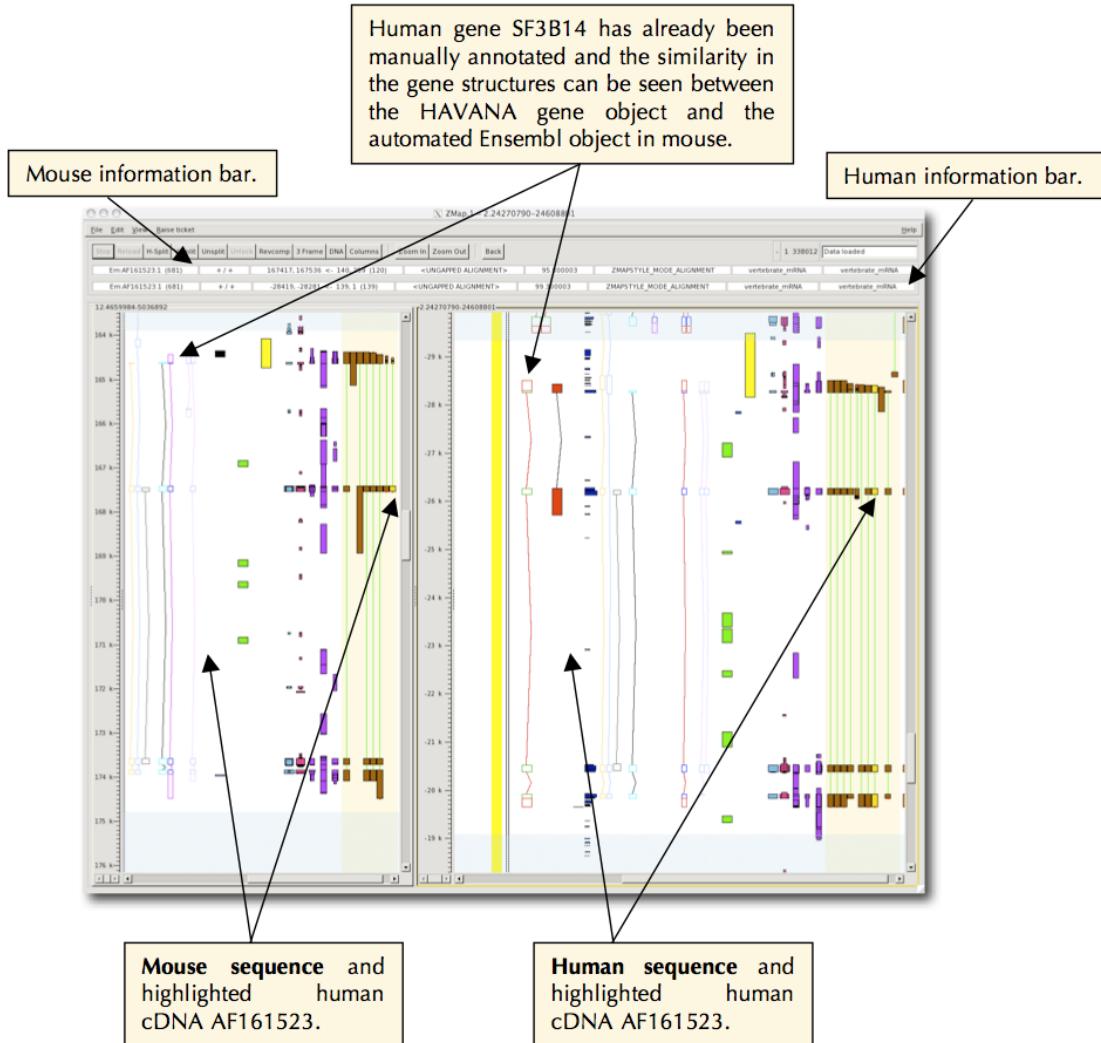


Figure 25: Launch in an existing ZMap

## 9 Tips for a speedier ZMap

1. Specifically: zoom and mark within ZMap early on after launching. Either select a gene object and press 'z' to zoom OR select a rectangle to zoom in by dragging the left mouse button around it. Reverse complement now if necessary, then press 'm' to mark the region.
2. The quickest way to zoom out of ZMap again is to right mouse click on the 'zoom out' buttons at the top of zmap and choose one of the options (this is definitely much quicker than doing individual 'zoom outs' with the left mouse button). Likewise for 'zooming in' again (or use keyboard equivalents).
3. Bump within a marked region only. Bumping without marking is slow and removes the lines connecting Blast matches.
4. When you have finished working within a marked region, unbump the evidence you have been working on (e.g. ESTs) and unmark that region before you go on to select the next region to mark and bump - or you could miss visualising the evidence in the new region.
5. If you want to get rid of some white space try the compress 'c' function or alternatively toggle off some of the columns. **Warning - this may hide features as well.** If a column (e.g ESTs) is bumped and you want to lose it temporarily, it is quicker to turn the column off (when you turn it on again it will still be bumped when it reappears) than unbump then rebump again later.
6. Jumping to genes/objects: If you expand the left hand 'scroll navigator' overview' you can jump directly to genes and objects by double-clicking on them.

## 10 Keyboard and mouse shortcuts

In general ZMap will be faster for zooming, bumping etc if you make good use of the built in short cuts. These can often avoid the need for ZMap to redraw large amounts of data that you may not even be interested in. For example, click once (highlight) on a feature and a carriage return will bring up evidence. Another example is to press T for translation.

### 10.1 All windows

Short Cut	Action
Cntl-W	close this window
Cntl-Q	quit ZMap

### 10.2 ZMap Window

Short Cut	Action
Control keys	
+ (or =), -	zoom in/out by 10%
Cntl + (or =), Cntl -	zoom in/out by 50%
up-arrow, down-arrow	scroll up/down slowly bit
Cntl up-arrow, Cntl down-arrow	scroll up/down more quickly
left-arrow, right-arrow	scroll left/right slowly
Cntl left-arrow, Cntl right-arrow	scroll left/right more quickly
page-up, page-down (Mac users should use fn and up/down arrow)	up/down by half a "page"
Cntl page-up, Cntl page-down	up/down by a whole "page"
Home, End (Mac users should use fn and left/rights arrows)	Go to far left or right
Cntl Home, Cntl End (Mac users will have to configure their keyboards for this)	Go to top or bottom
Delete, Shift Delete	Hide/Show selected features.
Enter	Show feature details for highlighted feature.
Shift up-arrow, Shift down-arrow	Jump from feature to feature within a column.
Shift left-arrow, Shift right-arrow	Jump from column to column.
Cntl-K	Copy selected feature(s) to the Annotation column for editing. Also enables Annotation if not already enabled.

### 10.3 Alpha-numeric keys

Short Cut	Action

a	Blixem all sequences in column
A	Blixem only highlighted sequence in column
b	Bump/unbump current column within limits of mark if set, otherwise bump the whole column.
B	Bump/unBump current column within limits of the visible feature range.
c	compress/uncompress columns: hides columns that have no features in them either within the marked region or if there is no marked region within the range displayed on screen. Note that columns set to "Show" will not be hidden.
C	Compress/unCompress columns: hides all columns that have no features in them within the range displayed on screen regardless of any column, zoom, mark etc. settings.
h	Toggles highlighting (good for screen shots).
m	mark/unmark a range which spans whichever features or subparts of features are currently selected for zooming/smart bumping
M	Mark/unMark the whole feature corresponding to the currently selected subpart (e.g. the whole transcript of an exon or all HSPs of the same sequence as the highlighted one) for zooming/smart bumping
o or O	show menu Options for highlighted feature or column, use cursor keys to move through menu, press ESC to cancel menu.
r	reverse complement current view, complement is done for all windows of current view.
t or T	translate highlighted item, T hides Translation.
w or W	zoom out to show whole sequence
z	zoom to the extent of any selected features (e.g. exon/introns, HSPs etc) or any rubberbanded area if there was one.
Z	Zoom to whole transcript or all HSPs of a selected feature.

#### 10.4 ZMap Mouse Usage

Left	Middle	Right
<i>Single mouse button click</i>		
highlight a feature or column Plus drag: draw a rectangle around an object for zoom	horizontal ruler with sequence position displayed, on button release centre on mouse position. Release mouse outside ZMap window to prevent re-centering.	show feature or column menu - for options such as pfetch, show feature DNA, show peptide, export peptide
<i>Double mouse button click</i>		
display details of selected feature. Double click on object to get edit window	same as single click	same as single click
<i>Shift + mouse button click</i>		
highlight a subpart of a feature (e.g. a single exon or alignment match) OR multiple highlight	same as single click	same as single click

