

MacVector 14.5

for Mac OS X

Creating Vector Maps

MacVector, Inc.
Software for Scientists

Copyright statement

Copyright **MacVector, Inc**, 2016. All rights reserved.

This document contains proprietary information of **MacVector, Inc** and its licensors. It is their exclusive property. It may not be reproduced or transmitted, in whole or in part, without written agreement from **MacVector, Inc**.

The software described in this document is furnished under a license agreement, a copy of which is packaged with the software. The software may not be used or copied except as provided in the license agreement.

MacVector, Inc reserves the right to make changes, without notice, both to this publication and to the product it describes. Information concerning products not manufactured or distributed by **MacVector, Inc** is provided without warranty or representation of any kind, and **MacVector, Inc** will not be liable for any damages.

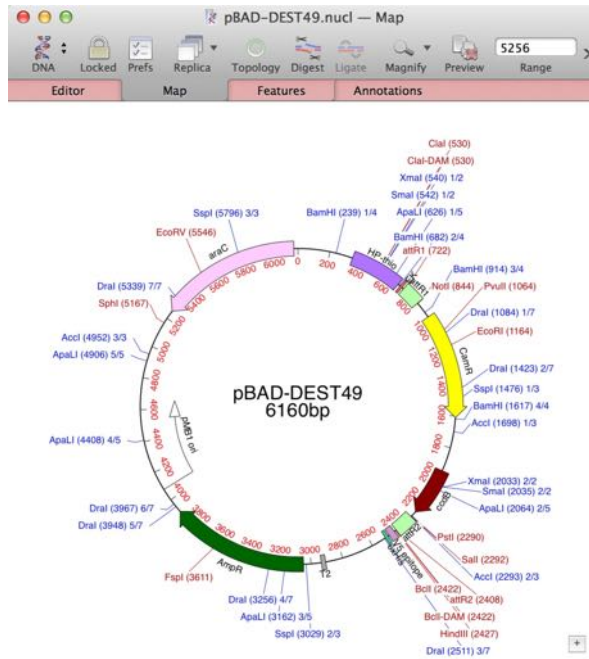
This version of the Creating Vector Maps tutorial was published in February 2016.

Contents

CONTENTS	3
INTRODUCTION	4
SAMPLE FILES	4
CONCEPTS	4
TUTORIAL	5
Creating Features	5
Changing The Appearance Of A Feature	7
Complementary Strand Features	8
Restriction Enzymes	10
Using The Graphics Palette	13
Getting Graphics Out Of MacVector	16
Map View Preferences	18
Enable Shadowing	18
Simplify layout for large sequences	18
Enable Label Hiding	18
Other Settings	19
Understanding How Labels Are Displayed	19
GenBank Feature Format Primer	20
Setting Symbol Defaults	21

Introduction

MacVector is capable of producing stunning high resolution graphical maps of sequences that can be exported as vector graphics to other applications. You have incredible control of many aspects of the way the maps are laid out and the shapes, colors, patterns, fonts and labels associated with the features of the sequence. This tutorial shows you how to get started with MacVector so that you can generate multi-colored maps like this;



Sample Files

For this tutorial, we will work with an unannotated copy of the classic cloning vector pBR322;

```
/Applications/MacVector/Tutorial Files/AutoAnnotation/  
pBR322Ascii.txt
```

Concepts

The most important thing to understand about the way MacVector displays vector maps is that every graphical object, except for the restriction enzyme cut sites, keys off a **Feature**. So, to make something appear on the **Map**, you must first create a **Feature**, defining the start and stop location of where it is located on the sequence. Once you have a **Feature** defined, you can then change its appearance by double clicking on the graphical representation of the **Feature** and changing its **Symbol** settings. **Features** do not have to be displayed on the **Map** – you make the

choice of which **Features** are displayed, where they are displayed and what they look like.

Tutorial

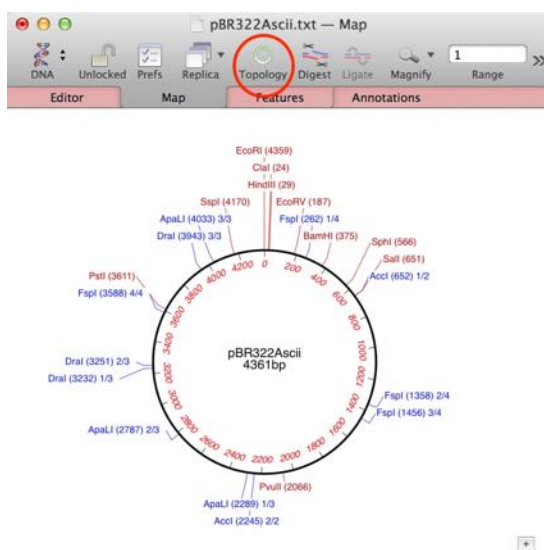
Creating Features

We will start with a “plain” unannotated sequence, create some features and then make them look nice.

Select **File | Open** and navigate to the /Tutorial Files/AutoAnnotation/ folder. Select the pBR322ASCII.txt file and click the **Open** button.

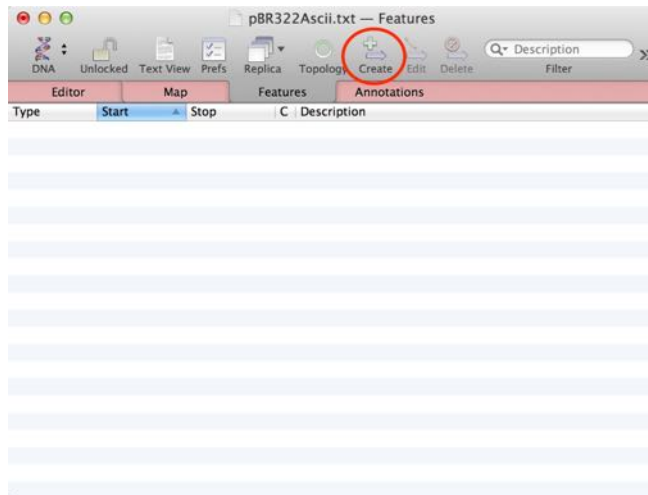
While the file is a simple text file, MacVector automatically interprets the contents and works out that it must be a DNA sequence. By default, MacVector assumes that new sequences have a linear topology. Because pBR322 is a circular molecule, the first thing we must do is tell MacVector that it is actually a circular molecule.

Switch to the **Map** tab. Then click on the **Topology** button so that the imagery changes to indicate the molecule is circular.



At this stage, we see a circular molecule, but it is just a plain circle because there are no Features present.

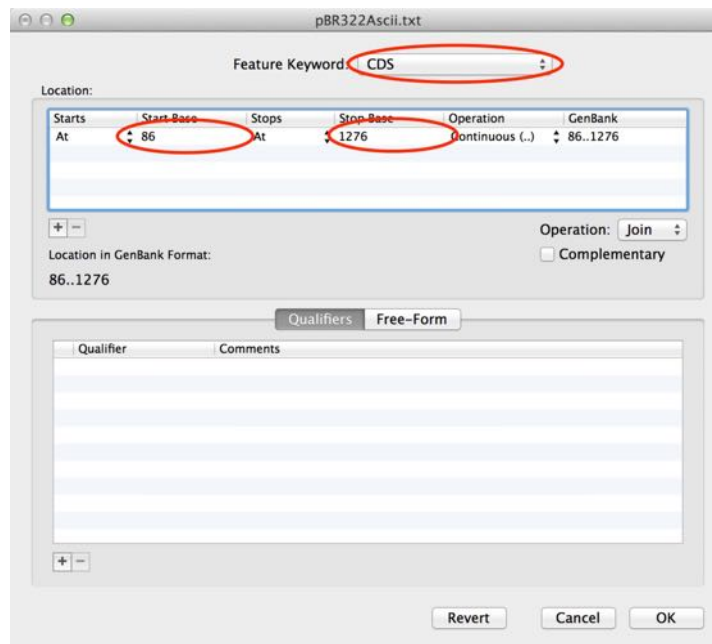
We can confirm this by switching to the **Features** tab.



Lets create a new feature!

Click on the **Create** button (circled above).

A drop down dialog sheet appears. This is where define what type of feature we want to create, along with the start and stop location and some comments.

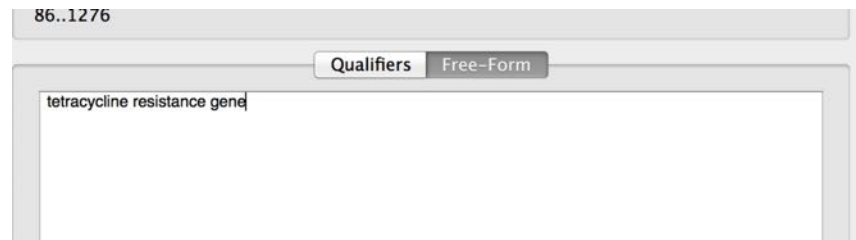


Select **CDS** from the drop down **Feature Keyword** menu, then double-click on the **Start Base** entry and enter “86”. Finally, double-click on the **Stop Base** entry and enter “1276”.

If your feature is something that is translated into protein, you should always use **CDS** as the **Feature Keyword**. You will find that any sequence you download from Genbank will always have protein-coding regions annotated as CDS features. In

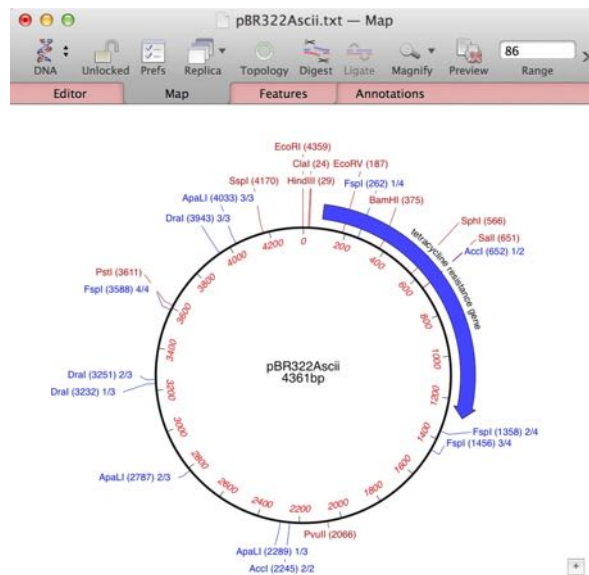
addition, there are many parts of MacVector that key off CDS features to display translations along with the DNA sequence.

In the lower pane, click on the **Free-Form** tab and type “tetracycline resistance gene” into the edit box and click on the **OK** button.



Switch back to the **Map** tab.

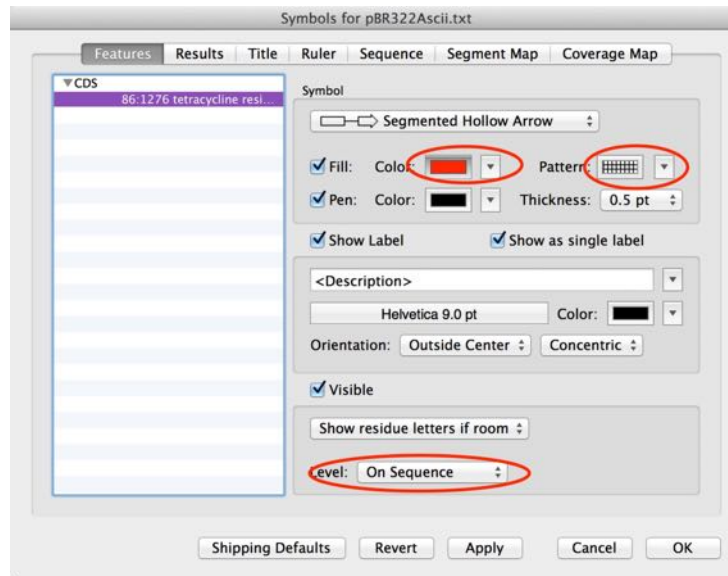
You should see that the new CDS feature is displayed as a blue arrow.



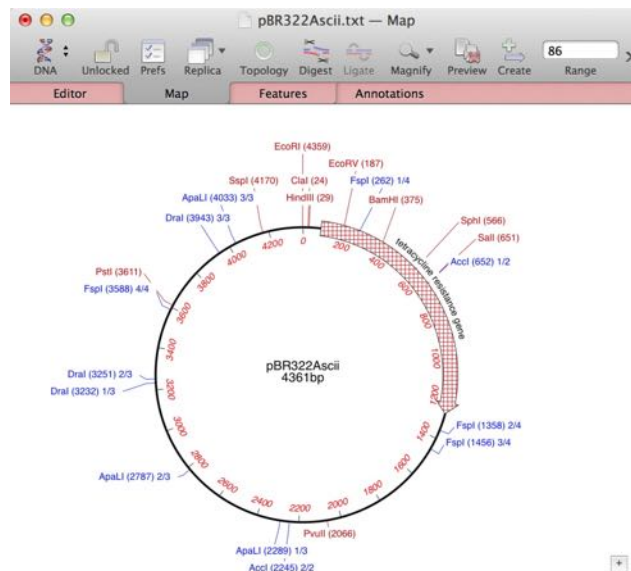
Changing The Appearance Of A Feature

Now that we have created a feature, lets make it look a bit different!

Double-click on the blue arrow to open up the **Symbol Editor**. Set the controls circled in red to the settings shown below and click **OK**.



The display should update to reflect the new values;

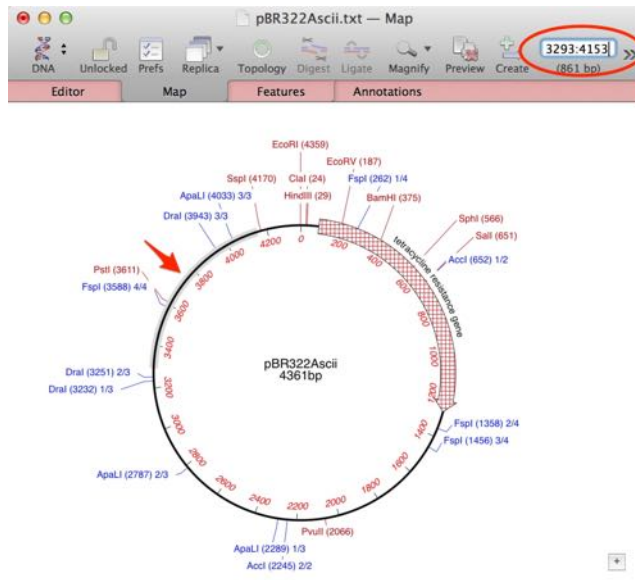


Complementary Strand Features

Now let's create a feature that is on the "minus" or "complementary" strand.

Click in the **Range** edit box and enter "3293:4153".

Note how a pale grey selection appears on the sequence line in that region of the circle;



Click on the **Create** button and edit the Feature as shown below

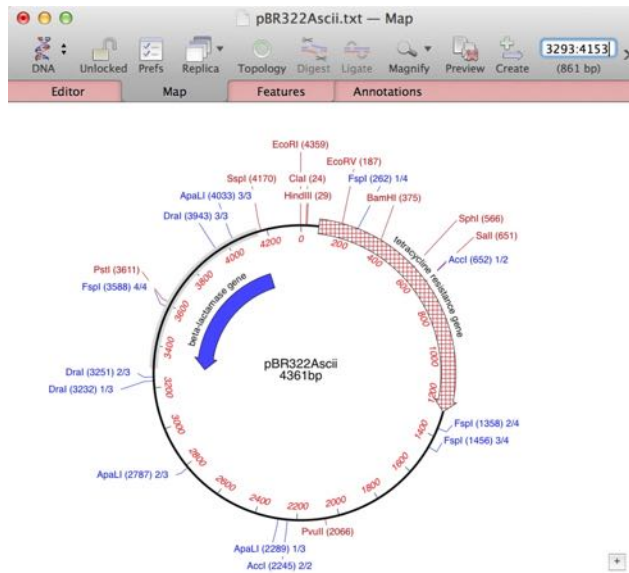
Note how the **Start Base** and **Stop Base** entries are already filled out with the selection and the **Feature Keyword** remembers the last value we used.

The screenshot shows the 'Feature Editor' window for pBR322Ascii.txt. The 'Feature Keyword' is set to 'CDS'. The 'Location' tab is active, displaying a table with the following data:

Starts	Start Base	Stops	Stop Base	Operation	GenBank
At	3293	At	4153	Continuous (..)	3293..4153

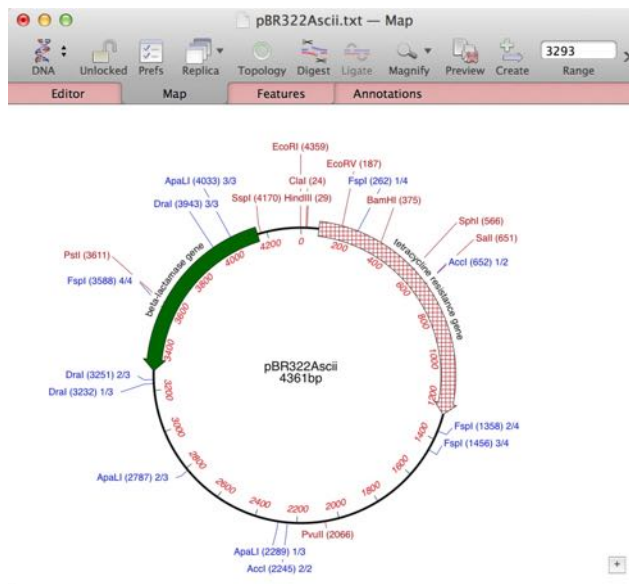
Below the table, the 'Location in GenBank Format' is shown as 'complementary(3293..4153)'. The 'Operation' dropdown is set to 'Join', and the 'Complementary' checkbox is checked. The 'Qualifiers' tab is also visible, showing the text 'beta-lactamase gene'.

Make sure you have the **Complementary** box checked and click **OK**



Note how the blue arrow points in the “other direction” to indicate the CDS is translated anti-clockwise on the molecule.

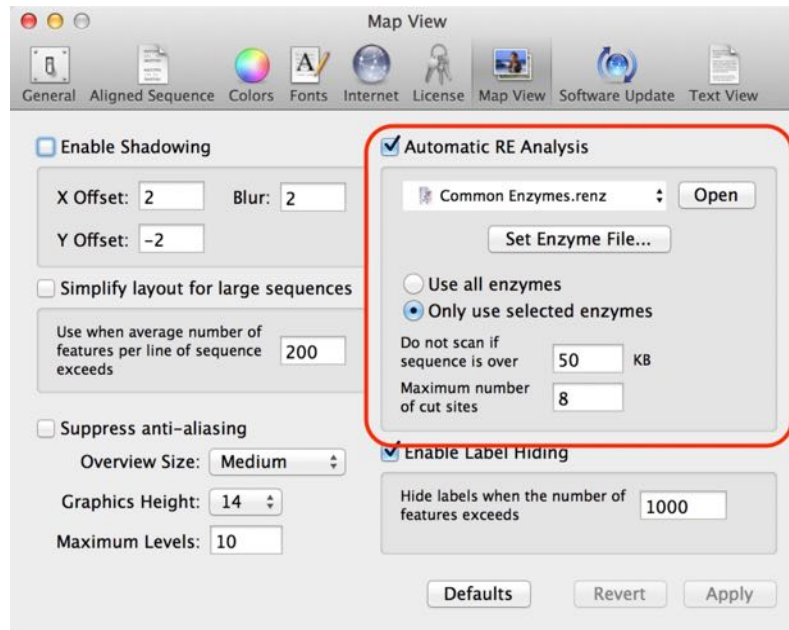
Double-click on the blue arrow and change the color to green and make sure the **Symbol** is placed “On Sequence”



Restriction Enzymes

Every time you open a DNA sequence, MacVector automatically runs a restriction enzyme analysis on the sequence and displays the results in the **Map** tab. You can control which enzymes are used in the analysis through the **Map Preferences** tab.

Click on the **Prefs** button on the window toolbar



The area outlined in red controls how the restriction enzyme sites are calculated and displayed. There are a few things to note;

Restriction enzymes names and recognition sequences are stored in special files with a .renz extension. You can find a selection of these files in /Applications/MacVector/Restriction Enzymes/. By default we use the file called Common Enzymes .renz, but you can choose any file you like by clicking on the **Set Enzyme File** button.

Even though the Common Enzymes.renz file has just a subset of all of the commercially available enzymes, there are still a huge number of enzymes that would completely saturate the Map if all of them were shown. So there are a few filters that can be applied to help reduce the clutter;

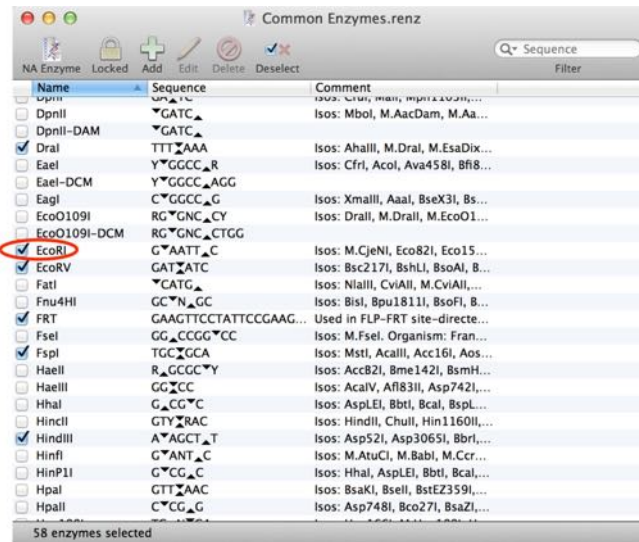
Use all enzymes vs **Only use selected enzymes**. You can select your favorite enzymes in any .renz file and just use those in the analysis. We'll talk about this more below.

Do not scan if sequence is over. Short sequences can be scanned very quickly, but large genomes might take some time to scan and display. This lets you restrict the search to just shorter sequences.

Maximum number of cut sites. This suppresses the display of sites if there are more than this number of hits.

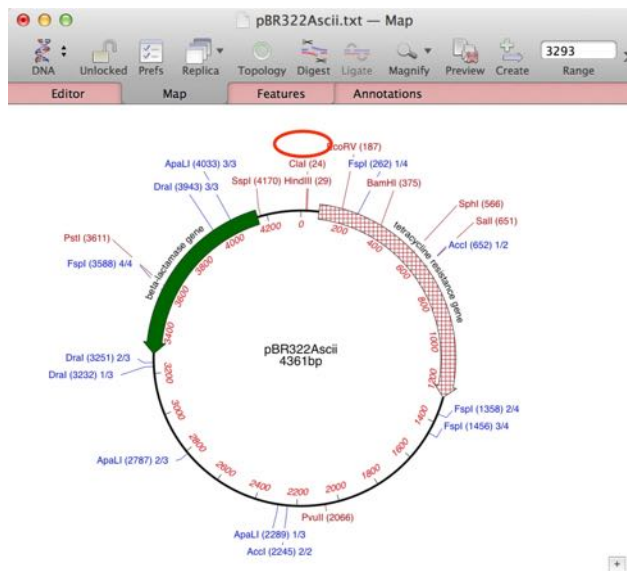
Click on the **Open** button.

This opens the selected enzyme file – in this case the Common Enzymes .renz file.



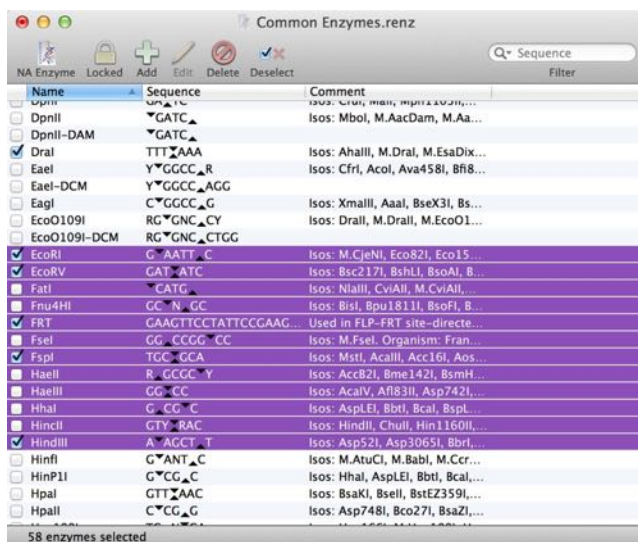
Note how a number of the sites have a checked box next to them. These are the enzymes that are being used in the analysis.

Uncheck the box next to EcoRI.

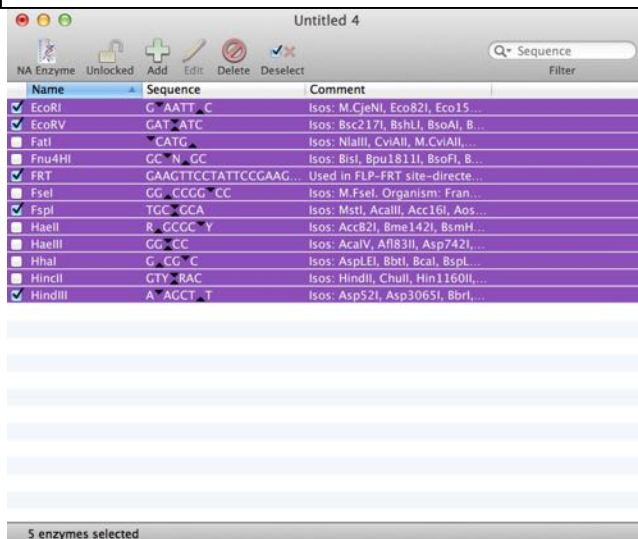


Note how the EcoRI site at the 12 o'clock position of pBR322 disappears. You can use this dynamic update approach to turn enzymes on and off in all of the open windows.

Click on the EcoRI name so that the entire row selects, hold down the <shift> key and click on the HindIII site so that all of the rows in between highlight.



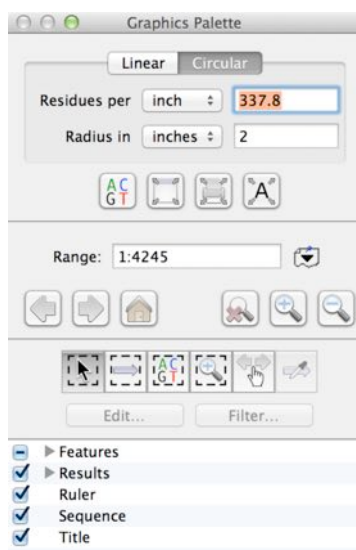
Choose **Edit | Copy** to place the selected enzymes on the clipboard. Now choose **File | New | Restriction Enzyme** to create a new, empty, Restriction Enzyme document. Finally choose **Edit | Paste** to add the selected enzymes to the new document.



You can use this approach to create your own personalized restriction enzyme files containing just those sites of most interest to you. Don't forget to save the file and then use the **Set Enzyme File** button in the **Map Preferences** tab to select your personal file.

Using The Graphics Palette

Whenever a Map view is active in the front most MacVector window, a floating Graphics Palette window appears. (If this is not displayed, choose **Window | Show Graphics Palette** to make it visible).



General Functionality

You can use this to turn on and off sequence, features and results objects in the Map display using the lower tree view. However, there is a lot of additional functionality controlled by the palette;

Linear/Circular – the tabs at the top let you toggle the Map between linear and circular views of the sequence. The circular option is disabled if the underlying sequence does not have circular topology.

Residues per inch/Line wrap – this section controls the scale of the Map and how “wide” a line is.



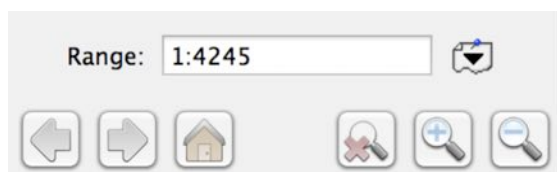
A row of 4 buttons provides “quick layout” functionality:

Zoom to Sequence – this adjusts the Residues per Inch so that individual sequence residues are visible. It does this without affecting the current zoomed section.

Fit to Window – adjusts the Residues per Inch and Line Wrap so that the current zoomed section fits in the current window.

Fit to Page – adjusts the Residues per Inch and Line Wrap so that the current zoomed section fits in the current printed page.

Fit Residues – similar to Zoom to Sequence, but this resets the zooming so that the entire sequence is displayed.



The range section lets you zoom to display a specific section by typing in the box or selecting a feature from the popup features menu.

Left Arrow/Right Arrow – these let you “nudge” the zoomed section to either side. You can also use the left/right keyboard arrow keys. For unzoomed circular sequences, these will rotate the graphic on screen.

Home – if you ever get “lost” in a sequence, this will center the sequence on the screen.

Reset zoom – restores the display to show the full sequence. Equivalent to a double-click in a blank area of the Map.

Zoom In/Zoom Out – this pair of buttons zooms in and out of the sequence in two-fold increments. You can also use the up/down keyboard arrow keys.



There is a row of six “mode” buttons that controls what happens when you click and drag with the mouse in the Map window.

Select Zoom – this is the default and the mode used by older versions of MacVector – you can click on features or sites to select them and if you click, hold and drag, the display resets to “zoom in” to the segment you selected.

Select Features – clicking and dragging selects all the features or sites that are touched by the selection rectangle.

Select Sequence – clicking and dragging selects just the sequence touched by the selection rectangle.

Magnify – in this mode a click magnifies the display 2-fold. Hold down the <option> key to reduce the magnification 2-fold.

Slide – this mode lets you drag the current zoomed region to the left or right. You can also use this to rotate circular sequences so that any arbitrary location is set to the 12 o’clock position.

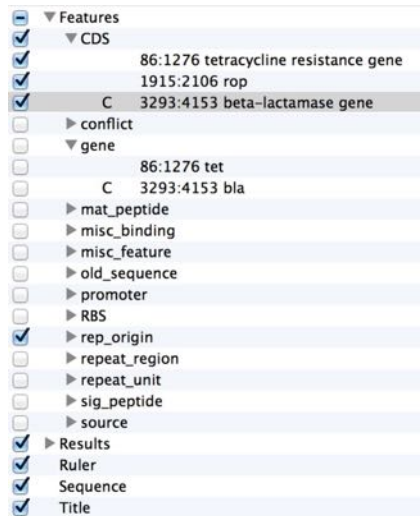
Copy Feature Appearance - this mode is only available if you have one or more features selected. Once selected, if you then click on a different feature, all the selected features will change appearance to match that feature

Visibility List View

The bottom section of the **Graphics Palette** controls the visibility of all of the different features.

Choose **File | Open** and locate and open the file
/Applications/MacVector/Sample Files/pBR322.nucl.

This is a fully annotated copy of pBR322. Look at the list area when this **Map** is active;



Most of the Features are hidden in this sample file.

Turn some of the checkboxes on and off again and watch how the **Map** display updates. Also watch what happens when you turn off and on the **Ruler**, **Sequence** and **Title** checkboxes.

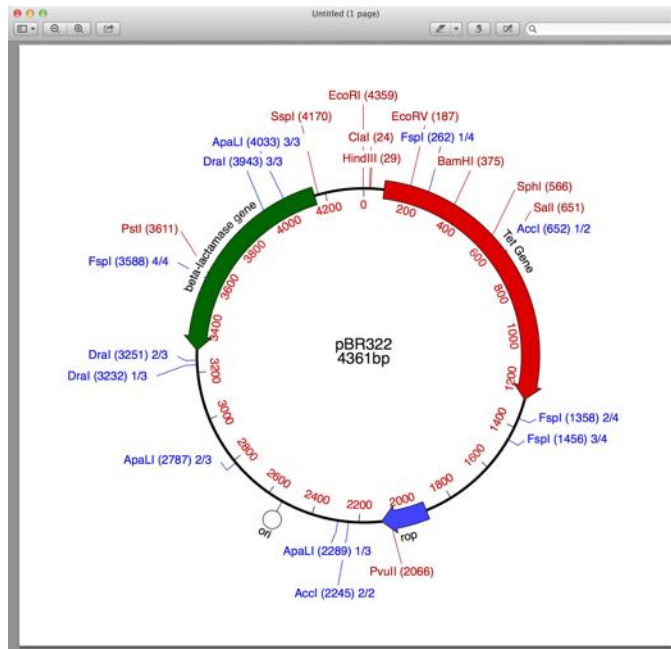
Getting Graphics Out Of MacVector

There are several ways that you can get high-resolution vector-based graphics out of MacVector and into other applications. Lets take a look at the two easiest approaches.

Choose **Edit | Copy**. This will copy the graphics from the **Map** tab. It is not necessary to have anything selected in the **Map** tab – the entire view will always be copied.

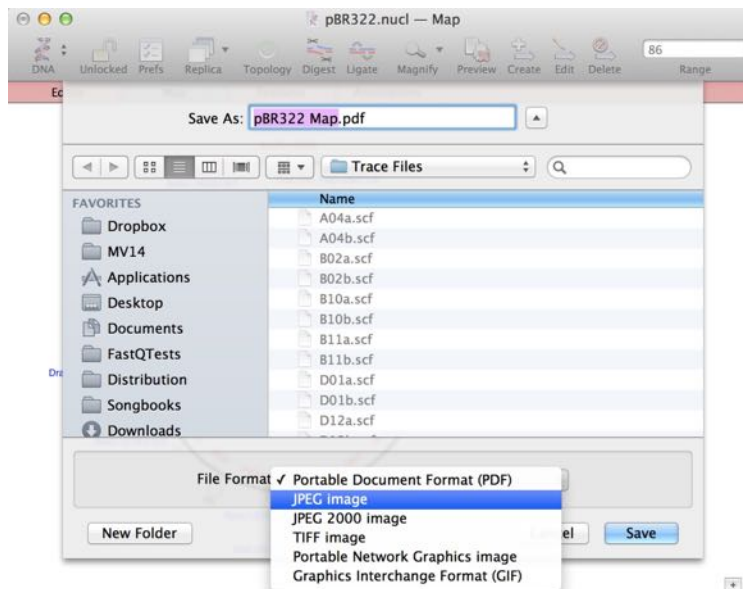
Open the application `Preview.app`. You can find this in the `/Applications/` folder. Choose **File | New From Clipboard**.

The graphics will get displayed in a Preview window. Note that from here you can save the document in a number of different formats by choosing **File | Save...**



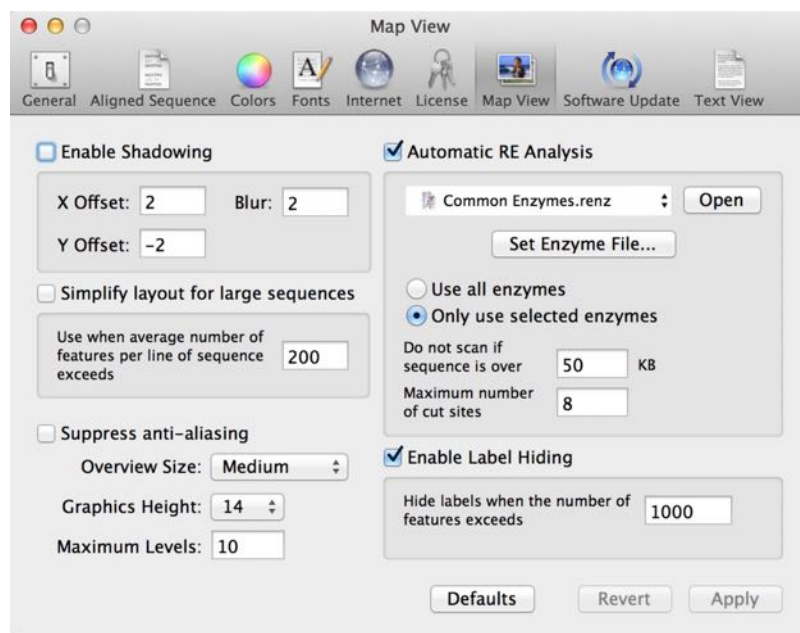
Choose **File | Export Tab Contents As...** This will open a **Save As...** dialog sheet that lets you choose the file format that would like to save the graphics in.

The permissible formats vary depending on the version of Mac OS X you are running. This is from OS X 10.8.



Map View Preferences

We took a look at the Preferences pane earlier when we learned how to control the appearance of Restriction Enzymes. Lets look at some other settings;



Enable Shadowing

If you like your images to have a 3D drop shadow, this is where you can turn this on and off.

Simplify layout for large sequences

Large sequences with many features can generate extremely crowded maps that can take some time to layout and display. Normally, MacVector tries to layout the features so that no graphical objects overlap. When “simplify layout” is turned on, MacVector will force all of the features onto the same level, above or below the sequence depending on the strand of the feature. The edit box controls how dense the features should be before the simplification kicks in. One nice effect of this is that you can “zoom in” and as the density of the features thins out, the display will switch from the crowded “simple” layout to the more carefully laid out non-overlapping display.

Enable Label Hiding

Like the “Simplify layout” setting, this helps to tidy up crowded maps and speed up layout and display. Laying out a lot of labels, particularly on circular maps, can be fairly slow, so this can really speed up the display of circular bacterial genomes. If you routinely work with moderately large sequences (e.g. BACs), you might want to increase the “number of features” cutoff to ensure labels are generally shown.

Other Settings

Suppress anti-aliasing. Checking this can also help to speed up display, at the expense of the quality of the graphics. However, with modern “retina” displays, this is less critical than in years past.

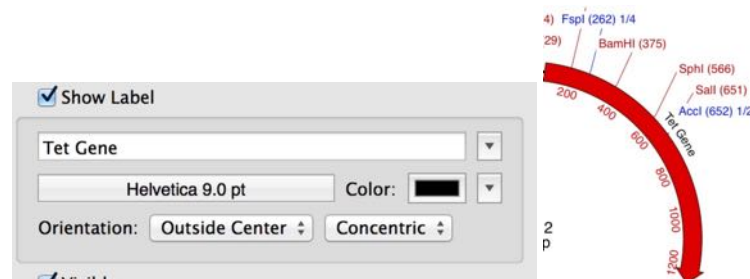
Overview Size. This controls the size of the small overview pane in the lower part of the Map display.

Graphics Height. This is the number of pixels used by the feature objects. If you like them “thick” try setting this to 16 or more. If you like them thin, try 10 or less. Like all of these Preference settings, this affects all Map windows.

Maximum Levels. With crowded maps, MacVector will try to avoid letting features overlap. However, if any feature has to be moved more than “maximum levels” away from the sequence line, MacVector gives up and just lets the features overlap on the top level. If you have the need to display a large number of features on the same level, increase this value appropriately.

Understanding How Labels Are Displayed

You can type any text into the **Label** field of the **Symbol Editor** and that will be displayed exactly as typed into the box;



However, MacVector has a much more sophisticated “meta-tag” language that can be used to format the labels. Meta-tags have a specific format with angled brackets (<>) surrounding the tag keyword e.g. <Description>. These are very useful for showing specific information in the Map view labels. By default, MacVector uses <Description> for all new features, but you can click on the popup menu to the right of the label text box to display a list of other valid meta tags;

<Description> or <Desc> - normally displays the entire description associated with a feature. However, if the description contains a /label qualifier, then the value of that qualifier will be used instead. Alternatively, if a /gene qualifier or a /note qualifier is present, then the values of one of those will be used.

<Start> or <start> - substitutes the co-ordinate of the start location of the feature.

<Stop> or <stop> - substitutes the co-ordinate of the stop location of the feature.

<Type> - substitutes the name of the feature type (e.g. CDS or misc_feature).

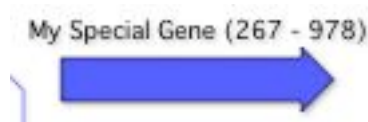
<Size> or <Length> - substitutes the length of the feature in residues.

<Total> - substitutes the total number of cuts for this Restriction Enzyme

<Index> - substitutes the number of this Restriction Enzyme cut.

In addition, you can include Genbank qualifiers as a meta-tag. So, e.g., <gene> will substitute the value of the /gene= qualifier associated with a feature, if one exists. More details on the Genbank feature format are included in the next section.

You can combine meta-tags along with free text in the Show Label field to display a lot of useful information. So, for example, typing the text <gene> (<start> - <stop>) might show a feature as;



Similarly using the text <Description> (<Start>) [<Index>/<Total>] for Restriction Enzyme Results will create labels like this;



GenBank Feature Format Primer

GenBank is the primary US repository of DNA and Protein sequences, curated by the NCBI (National Center for Biotechnology Information) at the National Institutes of Health (NIH) in Bethesda Maryland. Sequences maintained by GenBank have two main types of annotations. MacVector arbitrarily splits these up into "Features", which we define as annotations that have a defined location on the sequence (such as a gene or a site), and "Annotations" which are general data associated with a sequence such as an accession number, publications or authors.

GenBank features have a defined type such as CDS, mRNA, promoter etc - only a limited number of types are allowed. You can find a full description of the GenBank file format at;

<http://www.ncbi.nlm.nih.gov/collab/FT/>

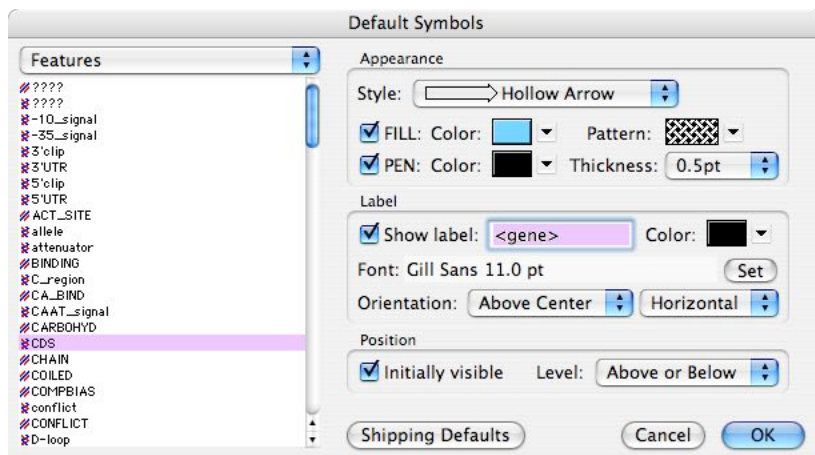
Each feature can have one or more qualifiers associated with it. The permissible qualifiers depend on the type of feature. For example, here is the definition for the gene feature type;

Feature Key	gene
Definition	region of biological interest identified as a gene and for which a name has been assigned;
Optional qualifiers	<pre> /allele="text" /citation=[number] /db_xref="<database>:<identifier>" /experiment="text" /function="text" /gene="text" /inference="TYPE[(same species)][:EVIDENCE_BASIS]" /label=feature_label /locus_tag="text" (single token) /map="text" /note="text" /old_locus_tag="text" (single token) /operon="text" /product="text" /pseudo /phenotype="text" /standard_name="text" /trans_splicing </pre>
Comment	the gene feature describes the interval of DNA that corresponds to a genetic trait or phenotype; the feature is, by definition, not strictly bound to it's positions at the ends; it is meant to represent a region where the gene is located.

Note how the qualifiers always start with a forward slash. Most are followed by an equals sign and text in quotes. Some qualifiers have no additional information (e.g. the `/pseudo` qualifier in this example).

Setting Symbol Defaults

If you consistently use the same set of qualifiers for your sequence, or if you primarily use sequences downloaded from the NCBI, where most sequences are consistently annotated, you might want to always show features labeled by their appropriate qualifier. To do this, hold down the **<option>** key and select **Options | Default Symbols** – this opens the default Symbol Editor. (Without the **<option>** key, this opens the Symbol Editor for the current sequence). For our example, make sure you have the Features popup selected, then scroll down to the CDS feature. Finally, type the new meta-tags you want in the Show Label edit box and use the opportunity to change any other default appearance information you want;



Click **OK**. From now on all NEW sequences you create (including any sequences you download from the NCBI, or read from disk in non-MacVector format) will use this information to display the feature in the **Map** tab. Existing sequences saved in MacVector format will override this information with their own default settings.