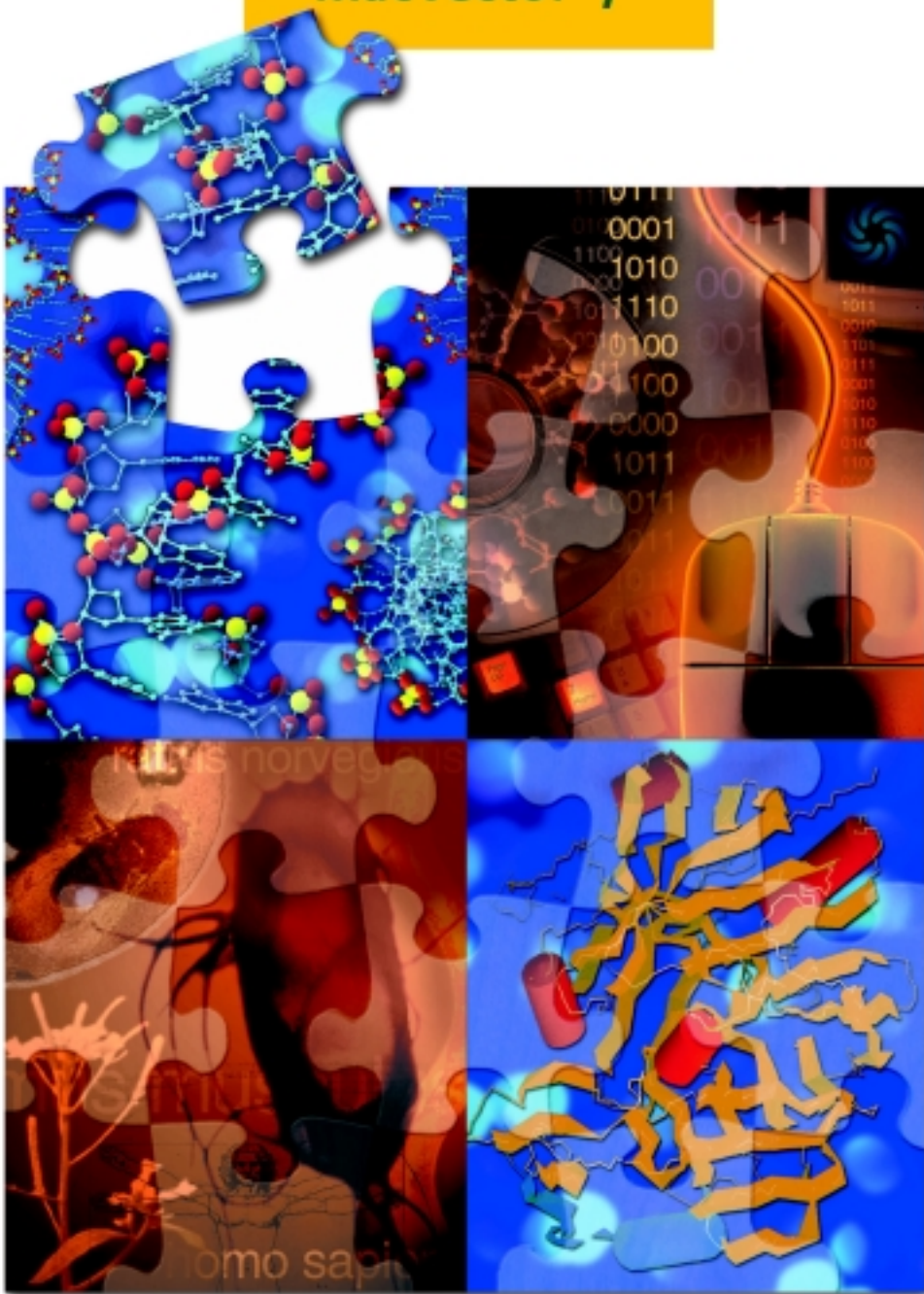


MacVector® 7



Getting Started

Copyright statement

Copyright ©2001 **Accelrys**, a subsidiary of Pharmacoepia Inc. All rights reserved.

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

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.

Accelrys 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 **Accelrys** is provided without warranty or representation of any kind, and **Accelrys** will not be liable for any damages.

This version of Getting Started was published in August, 2001.

Contents

Copyright statement	2
Contents	3
How to use Getting Started	4
Installing MacVector	4
Using this booklet	4
Outline of the experiment	5
After you finish this booklet	6
Starting MacVector	7
Construction of the human beta-globin sub-clone in pUC19	8
Importing sequence files	8
Working with features	10
Searching for restriction sites	12
Inserting the beta-globin gene into pUC19	19
Identifying restriction enzymes for Southern Blot analysis	23
Searching for restriction enzymes	23
Generating fragment predictions	25
Aligning sequences	28

How to use Getting Started

This booklet contains step-by-step tutorials that are designed to introduce new users to the MacVector Sequence Analysis Software.

Installing MacVector


Refer to the Installation Notes for a step-by-step description of how to install MacVector. In this booklet it is assumed that MacVector is already installed.

Using this booklet

We recommend that you work through the tutorials in this booklet sequentially, using the MacVector program. The tutorials are based around a simple experiment and cover the following MacVector operations:

- working with sequence features
- viewing and editing sequence data
- searching for restriction sites
- aligning DNA sequences.

You can find the example files used in this booklet in the `MacVector/Tutorial Files` folder. We recommend that you make a copy of the files contained in the `MacVector/Tutorial Files` folder, and put the copies in a convenient location for your use. Because you modify some of the tutorial files, this allows you and your colleagues to use the tutorials more than once, without the inconvenience of retrieving the tutorial files from the installation disk.

 **NOTE** Throughout the examples, we use the file location `MacVector/Tutorial Files` to refer to the location where your copy of these files is located.

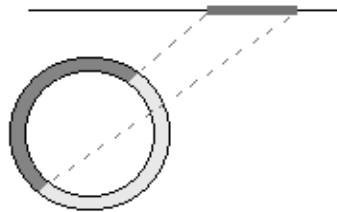
Outline of the experiment

The experiment demonstrates how you might use MacVector to aid identification of polymorphisms in the human beta-globin gene.

Consider the following scenario:

Genomic DNA samples have been extracted from blood volunteered by a family, in which some individuals are suspected of carrying a form of beta thalassemia. Using sequence data for known forms of beta thalassemia, you are faced with designing a preliminary screening assay to test these samples. The scheme for a simple screening experiment follows:

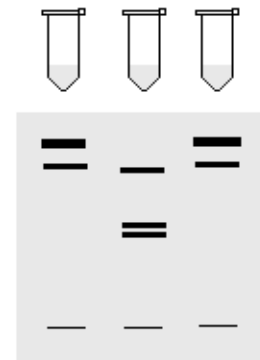
- 1** Sub-clone human beta-globin gene into pUC19.



- 2** Label beta-globin sub-clone to generate a DNA probe.



- 3** Southern Blot genomic DNA samples digested with selected enzymes.




- 4** Hybridize Southern Blot with DNA probe.

The tutorials cover the following activities:

- simulate the construction of the human beta-globin sub-clone in pUC19
- identify candidate restriction reactions for a preliminary screen for beta thalassemia polymorphisms using Southern Blot analysis

- investigate known beta thalassemia mutations in primary protein sequences.

 NOTE Please note that the purpose of these tutorials is to demonstrate the key functionality within MacVector. These tutorials are not intended to provide a recommendation for experimental design.

After you finish this booklet

Once you are familiar with the interface and concepts of MacVector, you can start working on your own projects with MacVector.

Refer to the MacVector User Guide for further information on MacVector functionality and procedural lists of how to use it.

Visit our Website at www.accelrys.com for news of updates and Frequently Asked Questions.

Starting MacVector

1.1 Do one of the following:


- If you have created an alias to MacVector on your desktop, double-click the icon
- Otherwise, open the `MacVector/Tutorial Files` folder by double-clicking its icon, then double-click any file icon in the folder to start the program.

During the startup process, the program will check for the presence of EVE or network copy protection. Keep the mouse still until the MacVector box appears on the screen, so that signals from the mouse do not interfere with this process. If the mouse is moved, you may see the Copy protection error alert box. If this happens, select the **OK** button to close the alert box, then double-click on the file icon again.



The MacVector menu bar and icon appear along the top of the screen.

1.2 If you started MacVector by double-clicking a file icon, close the window that opened by clicking the close button in the top left corner.

 **NOTE** In normal use you will know what sequence you want to work with, and double-clicking the sequence file is an efficient way to start MacVector.

Construction of the human beta-globin sub-clone in pUC19

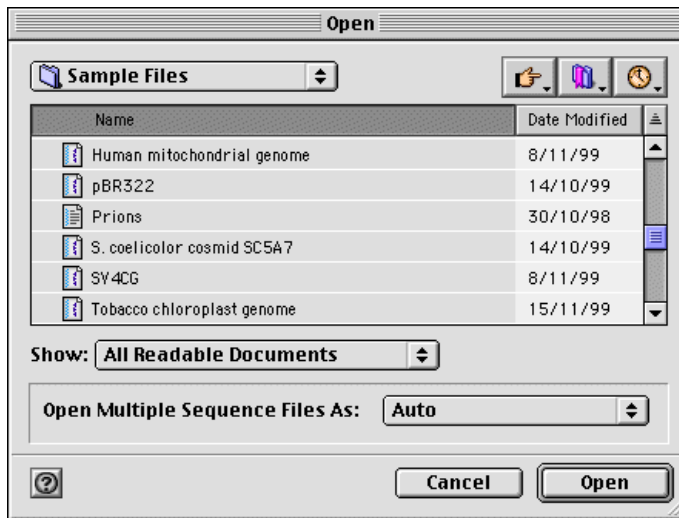
In this first tutorial, you will cover the following operations:

- importing and saving sequences
- using features to identify a region of a sequence
- editing restriction protein files
- finding restriction enzyme sites
- creating and editing sequences.

Importing sequence files

The first step in this tutorial is to import the base sequence of the human beta-globin region of chromosome 11, which is available as HSHBB.gen in the MacVector/Tutorial Files folder.


1.1 Select **File | Open** to display the File Open dialog box.



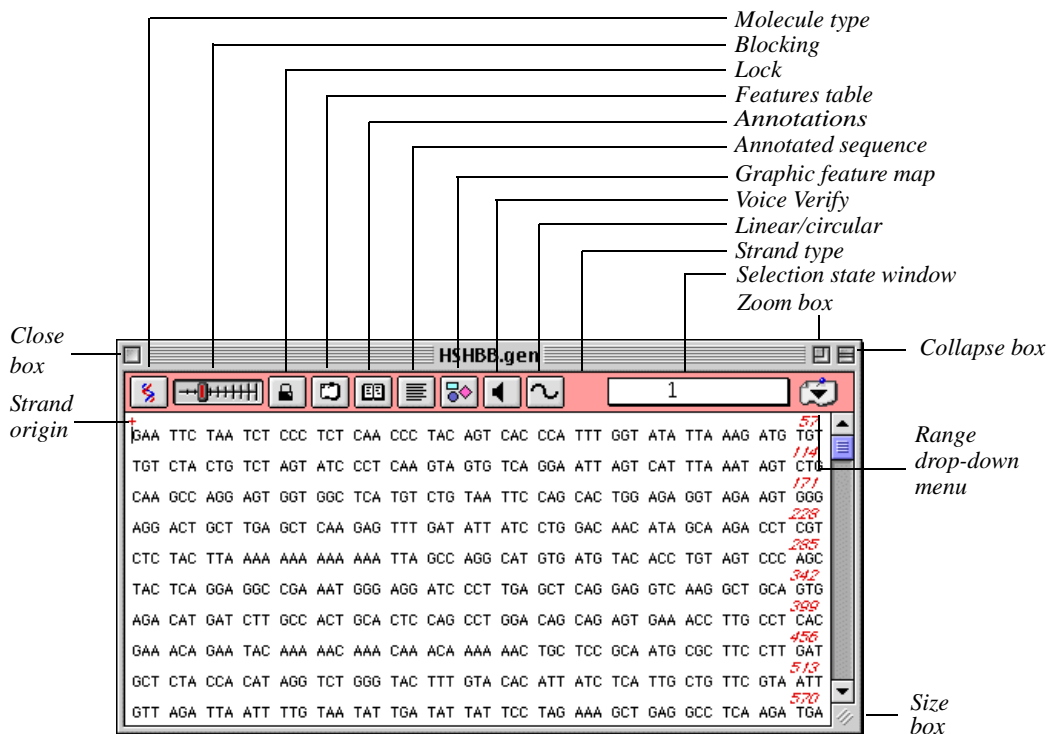
NOTE

The latest (Navigation Services) style of dialog box is illustrated and described here. If your version of MacOS does not have Navigation Services, your file opening dialog box will look different.

- 1.2 Navigate to the MacVector/Tutorial Files folder.
- 1.3 Choose HSHBB.gen from the list of files.
If you are using the older-style dialog box, you will first need to select the **TEXT Files** check box, to display text files.
- 1.4 Select **Open** to import the sequence.

 **NOTE** There may be data within some text file formats that MacVector cannot interpret. This is because the defined file format is not strictly adhered to by the file originators. If this is the case, you will be given the option to end the importing process. In most cases, the data is not relevant to the sequence and it is safe to continue.

Regardless of the file type imported, the sequence is converted internally into MacVector's own format. The sequence is displayed in a sequence window.





For a complete description of the sequence window, refer to the MacVector User Guide, Chapter 5, Working with Sequences.

Now save the sequence as a MacVector file.

- | | |
|-----|--|
| 1.5 | Select File Save As to display the File Save As dialog box. |
| 1.6 | Type HSHBB in the Name text box. There should not be a suffix.
Note: In older-style dialog boxes, this text box is labelled Save As . |
| 1.7 | Select Save to save the sequence in MacVector format. |

The pUC19 sequence is already in MacVector format, so opening the sequence is straightforward.

- | | |
|------|--|
| 1.8 | Select File Open to display the File Open dialog box. |
| 1.9 | Choose pUC19 from the MacVector/Tutorial Files folder. |
| 1.10 | Select Open to import the sequence. |

A new sequence window opens, containing the pUC19 sequence. Each sequence opened will be displayed in its own window.

Move the sequence window so that you can see both sequences at once.

Working with features

This part of the tutorial demonstrates how to use the sequence's features table. A feature is a piece of information about the sequence, associated with one or more residues, and is typically a coding region, exon, promoter or similar item.

- 2.1 Click on the HSHBB sequence window to make it active.
- 2.2 Select the **features table** button at the top of the window to display the features table for HSHBB.
- 2.3 Move the HSHBB Features Table window so that you can see both the table and the sequence.
- 2.4 Select the button at the top left of the Feature Table window, and choose **keyword ascend** from the **Feature sort** drop-down menu. All the features of each type will be grouped together to make searching easier in the following step.
- 2.5 Scroll through the features list, and double-click on the feature `precursor_RNA 62137 63742`.

Feature Type	Start	End	Description
mutation	62238	62238	gene=HBB; a in wt; t in thal
mutation	34275	34275	c in wt; g in persons with s
mutation	34282	34282	c in wt; t in persons with e
mutation	34321	34321	c in wt; t in high G-gamma S
precursor_RNA	54740	56389	hbd mRNA [25]
precursor_RNA	19485	21080	hbe mRNA (alt.) [23], [52], [6
precursor_RNA	19289	21080	hbe mRNA (alt.) [23], [52], [6
precursor_RNA	62137	63742	hbb mRNA [5], [111], [101], [71]
precursor_RNA	39414	40985	hbga mRNA [5], [151], [24], [27
precursor_RNA	34478	36069	hbga mRNA [151], [24], [27]
precursor_RNA	19488	21080	hbe mRNA (alt.) [23], [52], [6
repeat_region	10612	10924	Alu family repeat
repeat_region	10925	10939	Alu flank repeat 3' copy

Double-clicking on a feature has the following effects:

- the sequence window becomes active
- the residues corresponding to the feature are highlighted.

The feature highlighted is the beta-globin gene.



- 2.6 Click on the Features Table window to make it active.
- 2.7 Close the Features Table window by clicking the close button in the top left corner.

Searching for restriction sites

This part of the tutorial demonstrates how MacVector can be used to identify restriction sites for sub-cloning, by identifying restriction enzymes that do not cut within the beta-globin gene region.

- 3.1 Click on the HSHBB sequence window to make it active.
- 3.2 Select **File | Open** to display the File Open dialog box.
- 3.3 Navigate to the MacVector/Restriction Enzymes folder.
If you have the older-style dialog box, ensure that the **Enzyme Files** check box is selected, so that restriction enzyme files are displayed.
- 3.4 Choose Commercial Restriction Enzymes from the list of enzyme files.
- 3.5 Select the **Open** button.

A window opens, displaying the restriction enzymes in the file. By selecting a number of these enzymes, you can limit your search for restriction sites. For the current experiment, we will only use ten of the more commonly used enzymes corresponding to sites in the polylinker region of the pUC19 cloning vector.



3.6 Unlock the file by toggling the **lock** icon to the unlocked padlock.

3.7 Select the **Clear Selection** button to ensure that no enzymes are selected.

3.8 Scroll through the enzyme list and select the following enzymes by clicking on their name column to display a check mark next to the name:

- AccI
- BamHI
- BanII
- EcoRI
- HincII
- HindIII
- PstI
- Sall
- SmaI
- XbaI


 **TIP** To find a name more quickly, type its first letter.

Notice that, as you select the enzymes, a check mark appears next to the name, and the total number of selected enzymes is displayed on the lower bar of the window: make sure that the total number of selected enzymes is 10.

You could leave the restriction enzyme file open and do the analysis now, using the selected enzymes. However, you can also save the selection with its enzyme file for later use:

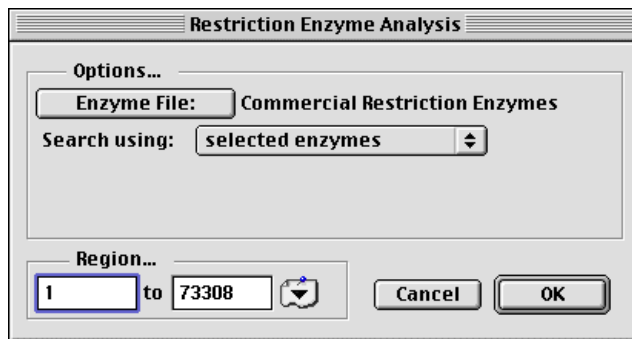
3.9 Select **File | Save** to save the enzyme selection.

3.10 Select **File | Close** to close the file and window.

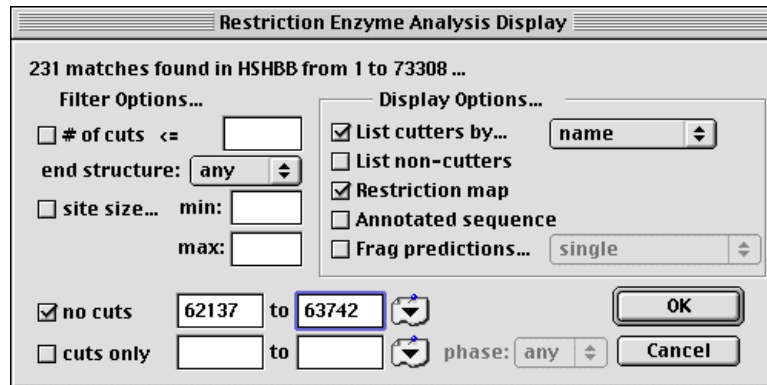
 **TIP** Use **File | Save As** to create a unique file of selected enzymes for restriction site searches.

- ◆ Refer to the MacVector User Guide, Chapter 4, Working with MacVector Files, for a comprehensive description of editing enzyme files.

- 3.11 Click in the HSHBB sequence window to make it active.
- 3.12 Select **Analyze | Restriction Enzyme** to display the Restriction Enzyme Analysis dialog box.
- 3.13 Select the **Enzyme File** button and choose `Commercial Restriction Enzymes` from the `MacVector/Restriction Enzymes` folder.
- 3.14 Select **Choose** to choose the file.
- 3.15 Choose **selected enzymes** from the **Search Using** drop-down menu, to use only those enzymes that were selected in step 3.8.
- 3.16 Choose **1 to 73308: ALL** from the features drop-down menu in the **Region** panel.
- 3.17 Select **OK** to perform the analysis.



When the analysis is complete, the Restriction Enzyme Analysis Display dialog box is displayed. This is used to control the display of results.



You are going to look at two approaches to viewing the results data:

- viewing a restriction map
- listing the cutting enzymes.

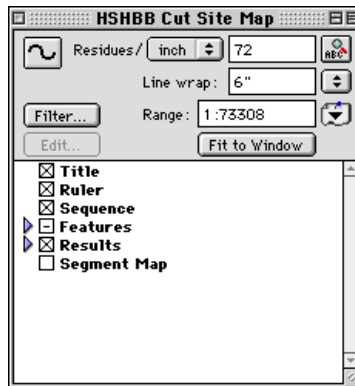
You will filter the results so that only enzymes that do not cut in the beta-globin gene region will be displayed.




- 3.18 Ensure that only the **List cutters by** and **Restriction map** check boxes are selected in the **Display Options** panel.
- 3.19 Ensure that only the **no cuts** check box is selected on the rest of the dialog box, then scroll through the **no cuts** features list, and click on the feature 62137 to 63742:
precursor_RNA,hbb mR...
- 3.20 Select **OK** to display the results.

Two results windows have been generated: a listing of the restriction enzymes, and a cut site map. Both displays show only those enzymes that do not cut in the beta-globin region. Now take a closer look at the region of interest, using the cut site map.

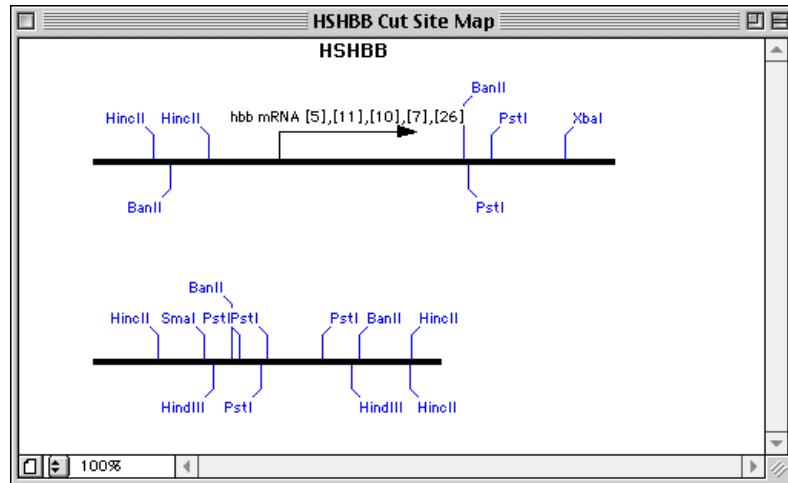
- 3.21 Click in the Cut Site Map window to make it active. The pointer changes to a magnifying glass, and the Graphics Palette dialog box appears.



 **NOTE** If the Graphics Palette dialog box does not appear, select **Windows | Show Graphics Palette**.

You are now going to use the Graphics Palette to emphasize the region of interest on the sequence.

- 3.22 Move the Graphics Palette dialog box so that you can see the entire HSHBB Cut Site Map window.
- 3.23 Highlight the **Range** text box, then type in the new range 60000:70000 and press **enter**.
- 3.24 Choose **cm** from the **Residues/** drop-down menu.
- 3.25 Type 500 in the **Residues/** text box and press **enter**.
- 3.26 Type 12cm in the **Line wrap** text box and press **enter**.
- 3.27 Ensure that the **Ruler** check box is clear so that the ruler is not displayed on the map.
- 3.28 De-select the **Features** check box to remove the feature symbols from the display. You will need to click it twice to make the check box clear.
- 3.29 Click on the triangle next to the **Features** check box to show an expanded tree of feature symbols.
- 3.30 In the expanded list, click on the triangle next to the **precursor_RNA** check box.
- 3.31 Select the **62137:63742** check box to display the beta-globin gene sequence symbol.

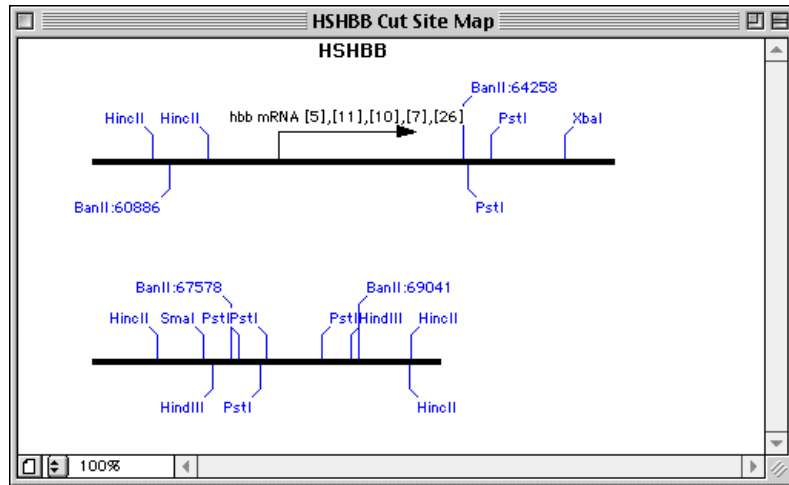


The display shows the feature of interest, plus where each of the restriction enzymes cuts the sequence. You can see from the map that BanII provides the smallest fragment that includes the intact beta-globin gene sequence.

You can label the exact position of the cut on the map.

- 3.32 In the Graphics Palette dialog box, click on the triangle next to the **Features** check box to collapse the Features tree.
- 3.33 Click the triangle next to the **Results** check box to expand the Results tree.
- 3.34 Click on the BanII text to highlight it, then select the **Edit** button. The Sequence Symbol Editor dialog box is displayed.
- 3.35 In the Label text box, enter the text
<Description> : <Start>, using the capitalization shown. These are special symbols which label the result feature with its name and starting residue number.
- 3.36 Select **OK** to make the change and close the dialog box.

The BanII cut sites are now labelled with the cut position.



The exact cut positions are also displayed in the Enzyme Cutters window, which shows that BanII cuts at positions 60,886 and 64,258.

The screenshot shows a window titled "HSHBB Enzyme Cutters" with the text "Sequence Range: 60000 to 70000". It contains a table with the following data:

Enzyme	#Cuts	Positions
BanII	53	60886 64258 67578 69041
HincII	28	60695 61322 66730 69624 69655
HindIII	19	67377 68952
PstI	31	64301 64560 67668 67913 67982 68631
Smal	4	67269
XbaI	26	65421

You can save this information as a feature, to use in the next part of the tutorial.

3.37 Close the cut site map window.

You cannot edit features while a map window generated from them is open.



- 3.38 Click in the HSHBB window to make it active.
- 3.39 Unlock the sequence.
- 3.40 Select the **features table** button to display the features table.
- 3.41 Select the **add (+)** button.
- 3.42 Select **misc_feature** from the **keyword** scrolling list.
- 3.43 Type 60887 in the **From** text box.
- 3.44 Type 64258 in the **To** text box.
- 3.45 Type a description in the **Description** text box, for example BanII cut sequence.
- 3.46 Select **OK** to add the feature. The **misc_feature** is added to the list.
- 3.47 Click the **Close** button in the top left of the features table window to remove the window from the display.

Now you need to save the sequence.

- 3.48 Select **File | Save** to save the sequence with its new feature.
- 3.49 Finally, tidy up the screen by removing the cutters list windows.

Inserting the beta-globin gene into pUC19

This part of the tutorial shows how MacVector can be used to simulate a cloning experiment, by excising the beta-globin gene from the HSHBB sequence and inserting it into pUC19.

The first thing to do is create a new sequence and copy pUC19.



- 4.1 Select **Windows | pUC19** to make the pUC19 sequence active.
- 4.2 Select **1 to 2686: ALL** from the sequence window features chooser to select the entire sequence.
- 4.3 Select **Edit | Copy** to copy it to the clipboard.
- 4.4 Select **File | New | Nucleic Acid** to open a new nucleic acid sequence window.
- 4.5 Select **Edit | Paste** to copy pUC19 into the new sequence window.

The copied sequence retains all the feature information from the original sequence.

- 4.6 Select **File | Save** and save the sequence as pHBG, to a folder of your choice.
- 4.7 Select **Windows | pUC19** to make the pUC19 sequence active.
- 4.8 Select **File | Close** to close the pUC19 sequence window.

Now you need to find out where BanII cuts the pHBG sequence. You can use the restriction site search and results display options that you set in “*Searching for restriction sites*”, p12.

- 4.9 Click in the pHBG window to make it active.
- 4.10 Select **Analyze | Restriction Enzyme** to display the Restriction Enzyme Analysis dialog box.
- 4.11 Select **OK** to perform the analysis.
- 4.12 When the Restriction Enzyme Analysis Display dialog box is displayed, remove the check from the **no cuts** check box, then select **OK** to display the results.

The restriction sites all occur in one area of the sequence. This is the polylinker region of the cloning vector. BanII cuts at position 406.

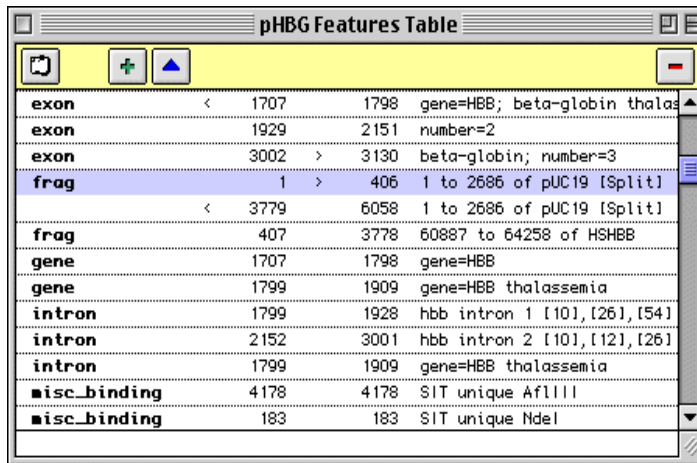
- 4.13 Select **File | Close** to close the pHBG Cut Site Map window.

The next stage is to copy the BanII fragment from the HSHBB sequence and paste it into pHBG.

- 4.14 Select **Windows | HSHBB** to make the HSHBB sequence active.
- 4.15 Select **60887 to 64258: misc_feature** from the feature table drop-down menu to select the BanII restriction fragment containing the beta-globin gene. The feature should appear near the top of the feature list, although the list order differs depending on the sort options that were set in the features table window.
- 4.16 Select **Edit | Copy** to copy the fragment.
- 4.17 Select **Windows | pHBG** to make the pHBG sequence active.
- 4.18 Select the text in the Selection State window, then type 407 and press the **enter** key to position the cursor just before position 407.
- 4.19 Select **Edit | Paste** to paste the gene into pHBG.
- 4.20 Select the **features table** button to open a features table window.
- 4.21 Select the button at the top left of the Feature Table window, and choose **keyword ascend** from the **Feature sort** drop-down menu.

2687

If you now scroll through the features table for pHBG, you will see that the features from both pUC19 and the HSHBB fragment have been inherited. There are also two MacVector *frag* features, documenting that the original pUC19 sequence is now split, and that residues 407 - 3778 originated from residues 60,887 - 64,258 of the HSHBB sequence.



Feature	Start	End	Description
exon	1707	1798	gene=HBB; beta-globin thalas
exon	1929	2151	number=2
exon	3002	3130	beta-globin; number=3
frag	1	406	1 to 2686 of pUC19 [Split]
frag	3779	6058	1 to 2686 of pUC19 [Split]
frag	407	3778	60887 to 64258 of HSHBB
gene	1707	1798	gene=HBB
gene	1799	1909	gene=HBB thalassemia
intron	1799	1928	hbb intron 1 [10],[26],[54]
intron	2152	3001	hbb intron 2 [10],[12],[26]
intron	1799	1909	gene=HBB thalassemia
misc_binding	4178	4178	SIT unique AflIII
misc_binding	183	183	SIT unique NdeI

4.22 Close the features table window by clicking on the **Close** button in the top left corner.

4.23 Select **File | Save** to save the modified pHBG sequence.

You have successfully inserted the beta-globin gene into the polylinker of pUC19, and stored the sequence as the file pHBG.

You might want to practice displaying a circular map, showing the insertion that you have just done.

4.24 Click in the pHBG sequence window to make it active.



4.25 Click on the feature map button to display a map. The Graphics Palette dialog box is also displayed.



4.26 Click the linear/circular button on the Graphics Palette dialog box to toggle the map display to a circular one.

You can now experiment with the feature display settings. Expand the features tree by clicking on the triangle to the left of the Features check box, then add or remove features from the display until you have the picture that you want.

4.27 Finally, tidy up the display by closing all sequence and result windows, clicking the close button in the top left corner of each window in turn.

Identifying restriction enzymes for Southern Blot analysis

The pattern of fragments from an enzyme digest can be used as a preliminary screen. For our experimental scenario, we need to be able to identify known mutations of the beta-globin gene. For this tutorial, we have three sequences available:

- beta-globin gene sequence from HSHBB, called beta_globin
- var_48
- var_64.

The mutant gene var_48 contains six base-pair mutations, whereas var_64 has a single base-pair deletion.

Searching for restriction enzymes

In this part of the tutorial, you are going to look for restriction enzymes for each variant.

- 1.1 Select **File | Open** to display the File Open dialog box.
- 1.2 Choose beta-globin from the MacVector/Tutorial Files folder.
- 1.3 Select **Open** to import the sequence.
- 1.4 Select **Analyze | Restriction Enzyme** to display the Restriction Enzyme Analysis dialog box.
- 1.5 Select the **Enzyme File** button and choose Southern Blot Enzymes from the MacVector/Tutorial Files folder.
- 1.6 Select **Choose** to choose the file.

This enzyme file is a copy of the Commercial Restriction Enzymes file, with two enzymes selected. For the initial search, however, you will use all enzymes in the file.



- 1.7 Choose **all enzymes** from the **Search Using** drop-down menu.
- 1.8 Choose **1 to 1606: ALL** from the drop-down menu in the **Region** panel.
- 1.9 Select **OK** to perform the analysis.

The Restriction Enzyme Analysis Display dialog box is displayed.

- 1.10 Ensure that only the **List cutters by** check box is enabled in the **Display Options** panel.
- 1.11 Ensure that the **no cuts** check box is not selected.
- 1.12 Select **OK** to display the results.

Now you are going to repeat this analysis for the sequences var_48 and var_64. Do not close any windows at this stage.

- 1.13 Select **File | Open** to display the File Open dialog box.
- 1.14 Choose var_48 from the MacVector/Tutorial Files folder.
- 1.15 Select **Open** to import the sequence.
- 1.16 Select **Analyze | Restriction Enzyme** to display the Restriction Enzyme Analysis dialog box.
- 1.17 Select **OK** to perform the analysis.
- 1.18 When the Restriction Enzyme Analysis Display dialog box appears, select **OK** to display the results.
- 1.19 Repeat steps 1.13 through 1.18 for the var_64 sequence.

Take a few moments to scan the results.

- 1.20 Arrange the displayed windows so that the three Enzyme Cutters lists are side by side.
- 1.21 Scroll through each list so that the **AvaII** and **AxyI** enzymes are visible in each.

Enzyme	1	2	3	4	5
AvaII	2	292	506		
AxyI	2	267	355		
BamHI	1	476			
BanI	2	387	429		
BbrPI	1	473			
BbsI	1	213			
BbvI	5	259	439	454	
Bbv12I	2	58	385		
Bbv16I I	1	213			
BfaI	3	29	818	1296	
Bme18I	2	292	506		
BmyI	3	58	385	432	
BpiI	1	213			

The pattern of cuts suggests that the two enzymes AvaII and AxyI could form the basis of a digest to distinguish the three sequences.

Generating fragment predictions

In this part of the tutorial, you will take the two enzymes found in the previous part, and look at their predicted digest fragment pattern.

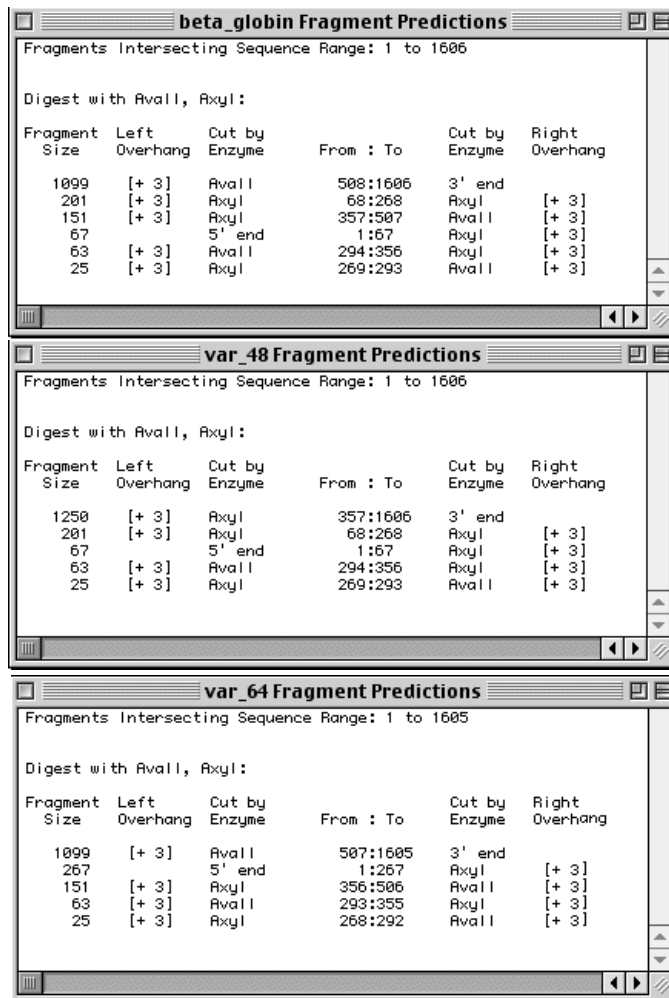
- 2.1 Select **Windows | beta-globin** to make the beta-globin sequence active.
- 2.2 Select **Analyze | Restriction Enzyme** to display the Restriction Enzyme Analysis dialog box.
- 2.3 Ensure that Southern Blot Enzymes from the MacVector/Tutorial Files folder is the **Enzyme File** in use.
- 2.4 Choose **selected enzymes** from the **Search using** drop-down menu to use only selected enzymes.

AvaII and AxyI are the only selected enzymes in this file.

- 2.5 Select **OK** to perform the analysis.
- 2.6 When the Restriction Enzyme Analysis Display dialog box appears, select **Frag predictions** from the **Display Options** panel, and select **double** from the adjacent drop-down menu.

This display option shows the fragments produced for each pair of enzymes used in the digest.

- 2.7 Select **OK** to display the results.
- 2.8 Repeat steps 2.1 through 2.7 for the var_48 and var_64 sequences.
- 2.9 Arrange the windows so that the Fragment Predictions lists are all visible.



The fragment predictions show that these enzymes can provide the basis for a Southern Blot analysis. This restriction data could be used in conjunction with restriction data for other mutations to modify the Southern Blot for a broader screening analysis.

2.10 Tidy up the display by closing all result windows, clicking the close button in the top left corner of each window in turn. Leave the sequence windows open for the next tutorial.

Aligning sequences


In this final tutorial, you are going to perform a simple ClustalW sequence alignment so that you can visualize the mutations easily.

- 1.1 If you have not carried on from the previous tutorial, you need to open the three sequences `beta_globin`, `var_48`, and `var_64`. These are in the `MacVector/Tutorial Files` folder.
- 1.2 Select **Windows | beta-globin** to make the sequence active.
- 1.3 Select **Analyze | ClustalW Alignment** to display the ClustalW Alignment dialog box.
- 1.4 Select **Use Defaults**, then select **OK** to perform the alignment.

The Multiple Sequence Alignment Editor window and the Alignment Views dialog box are displayed.

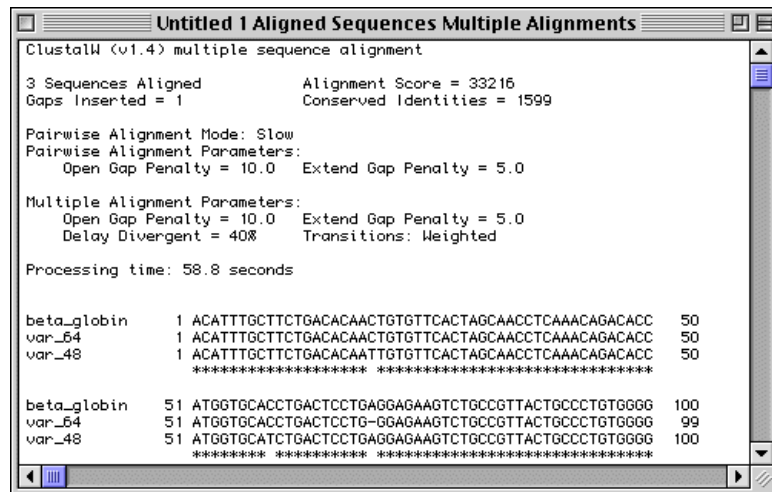
- 1.5 In the Alignment Views dialog box, enable the **Multiple Alignment** check box in the **Text Display** panel, and also the **Multiple Alignment** check box in the **Picture Display** panel.
- 1.6 Select the **Set Options** button.
The Multiple Alignment Options dialog box is displayed. It has five tabbed sections.
- 1.7 Select the **Text Display** tab, and ensure that the **Show identities** check box is enabled.
In the text display, this will mark positions in the alignment where all sequences match.
- 1.8 Select the **Consensus** tab, and ensure that the **Show consensus line** check box is enabled.
In the picture display and MSA editor windows, this will display a consensus line under the alignment.

- 1.9 Select the **Picture Shading** tab, and ensure that the **Outline**, **Bold** and **Shade** check boxes in the **Identities** panel are enabled.
In the picture display, this will clearly highlight the residues at positions where all the sequences match.
- 1.10 Select **OK** to close the Multiple Alignment Options dialog box, and **OK** to display the results.

 **TIP** In the Alignment Views dialog box, you can use the **Create Consensus Sequence** check box to display the consensus sequence in a new sequence window. You can then save it in the normal way as a sequence.

Look at the displayed results in turn.

The text display



```

ClustalW (v1.4) multiple sequence alignment
3 Sequences Aligned           Alignment Score = 33216
Gaps Inserted = 1             Conserved Identities = 1599

Pairwise Alignment Mode: Slow
Pairwise Alignment Parameters:
  Open Gap Penalty = 10.0      Extend Gap Penalty = 5.0

Multiple Alignment Parameters:
  Open Gap Penalty = 10.0      Extend Gap Penalty = 5.0
  Delay Divergent = 408       Transitions: Weighted

Processing time: 58.8 seconds

beta_globin   1 ACATTTGCTTCTGACACAACCTGTGTTCACTAGCAACCTCAAACAGACACC 50
var_64        1 ACATTTGCTTCTGACACAACCTGTGTTCACTAGCAACCTCAAACAGACACC 50
var_48        1 ACATTTGCTTCTGACACAATTGTGTTCACTAGCAACCTCAAACAGACACC 50
*****

beta_globin   51 ATGGTGCACCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGG 100
var_64        51 ATGGTGCACCTGACTCCTG-GGAGAAAGTCTGCCGTTACTGCCCTGTGGGG 99
var_48        51 ATGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGG 100
*****

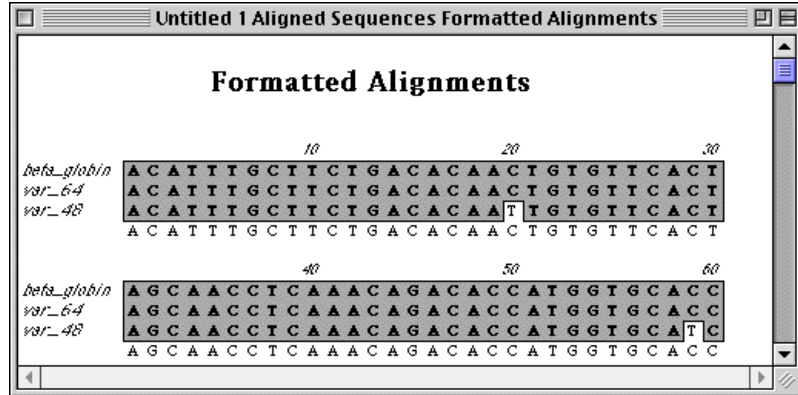
```

The text display has the following features:

- the sequences are displayed aligned in the standard way
- a consensus line appears at the bottom of the sequences, with an asterisk (*) for a match and a space for a mismatch.

The format of the display can be changed: try experimenting with the **Text Display** tab of the Multiple Alignment Options dialog box (accessed in step 1.6 on page 28).

The picture display

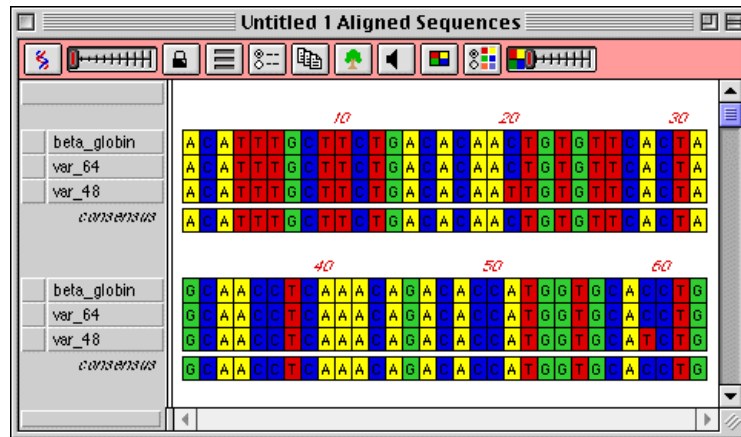


The picture display has the following features:

- the sequences are displayed aligned in the standard way
- matching bases are shown shaded grey
- mismatches and inserted gaps are shown shaded white
- a consensus sequence appears at the bottom, giving the best match at each position.

The format of the display can be changed: try experimenting with the **Picture Fonts** and **Picture Display** tabs of the Multiple Alignment Options dialog box (accessed in step 1.6 on page 28).

The Multiple Sequence Alignment Editor



The Multiple Sequence Alignment (MSA) Editor has the following features:

- the sequences are displayed aligned in the standard way
- each type of base is displayed in a different colour, enabling mismatches to be seen easily
- a consensus line is displayed under the alignment
- full formatting controls are provided for the display
- full editing capabilities are provided for renaming, reordering, moving, changing, adding and deleting sequences.

A full description of the MSA Editor appears in the MacVector User Guide, Chapter 6, Working with Multiple Sequences.

Using the displays, it is easy to identify the gap in var_64, at position 70 (loss of A), and the six substitutions in var_48, at positions:

- 20 (C to T)
- 59 (C to T)
- 511 (C to G)
- 569 (G to T)

- 576 (C to T)
- 1161 (T to C).