User Guide to

GeneQuest

Table of Contents

Welcome to GeneQuest

GeneQuest aids you in identifying the following features in your DNA sequence:

- Open reading frames
- Coding regions
- Splicing junctions
- Translation start and stop sites
- Transcription factor binding sites
- Restriction sites
- Patterns
- Repeats

Note: GeneQuest is one of the "classic" applications in the Lasergene package. Some of its functionality was migrated to SeqBuilder in 2005, and additional functionality was added once SeqBuilder was replaced by SeqBuilder Pro in 2017. In this User Guide, we will point out GeneQuest functions that can be performed more easily in our modern application, SeqBuilder Pro.

To get started with GeneQuest and to see an overview video, go to Getting Started.

These support pages contain written help and tutorials, as well as video tutorials.

- For help **INSTALLING** Lasergene, see our separate Installation Guide.

- For a written GeneQuest **TUTORIAL** with downloadable data, click here.

- Click here for a list of all topics that include **VIDEOS**.

- To **PRINT** the current page of the User Guide, click the printer icon in the top right corner (🖶).

- To download a **PDF** of the entire User Guide, scroll to the bottom of the table of contents on the left, and press **Download as PDF**.

> ✱ **Note:** Throughout this help manual, the use of 'x' in a file path denotes a version number. For example, a folder that is named "Lasergene 11 Data" will be written here as "Lasergene x Data."

Getting Started

GeneQuest aids you in identifying and annotating coding regions and other features of interest in your DNA sequence.

The following video is a general overview of GeneQuest:

Get started with GeneQuest:

1. Create a new project or open an existing project.

2. Choose **Analysis > Show Available Methods** to display a list of methods and choose which of these you would like to apply to the sequence.

3. Move the methods you want from the **More Methods** menu to the Method Curtain. Apply methods to the sequence by using the mouse to drag them from the Method Curtain to the Assay Document, if they are not already there. Once the methods have been applied, the results are displayed graphically on a common horizontal scale. You can Change method parameters at any time. Then add, delete and rearrange methods until you are satisfied with the look of the document.

4. Do any of the following optional tasks.

   • Print your results.

   • Display tabular data summaries (e.g., base content, codon usage).

   • Fold a portion of the sequence as RNA.

- View the predicted electrophoretic separation of restriction fragments through an agarose gel.

- GeneQuest reads and interprets features from the comment or feature panes of GenBank-formatted sequence files. Once you have discovered areas of interest in the sequence, you may create your own annotations and link several features together under a common title and description.

- Search BLAST or Entrez databases for matches to your sequence or use them to locate other sequences of interest.

5. Save your results or export them to another Lasergene application.

6. Exit GeneQuest using **File > Exit** (Win) or **GeneQuest > Quit** (Mac).

GeneQuest Tutorial

In this multi-part tutorial, you will learn how to apply methods to a sequence, predict coding regions, create features for those regions, and corroborate the predictions using BLAST and Entrez searches.

The tutorial begins with [Part A: Apply methods to a sequence](#).

Part A: Apply methods to a sequence

1. [Click here](#) to download the data zip archive, then extract it to any convenient location (e.g., your desktop).

2. Launch GeneQuest and choose **File > New**, then select and open the sequence *WormProblem.seq*. The GeneQuest window appears as shown below:
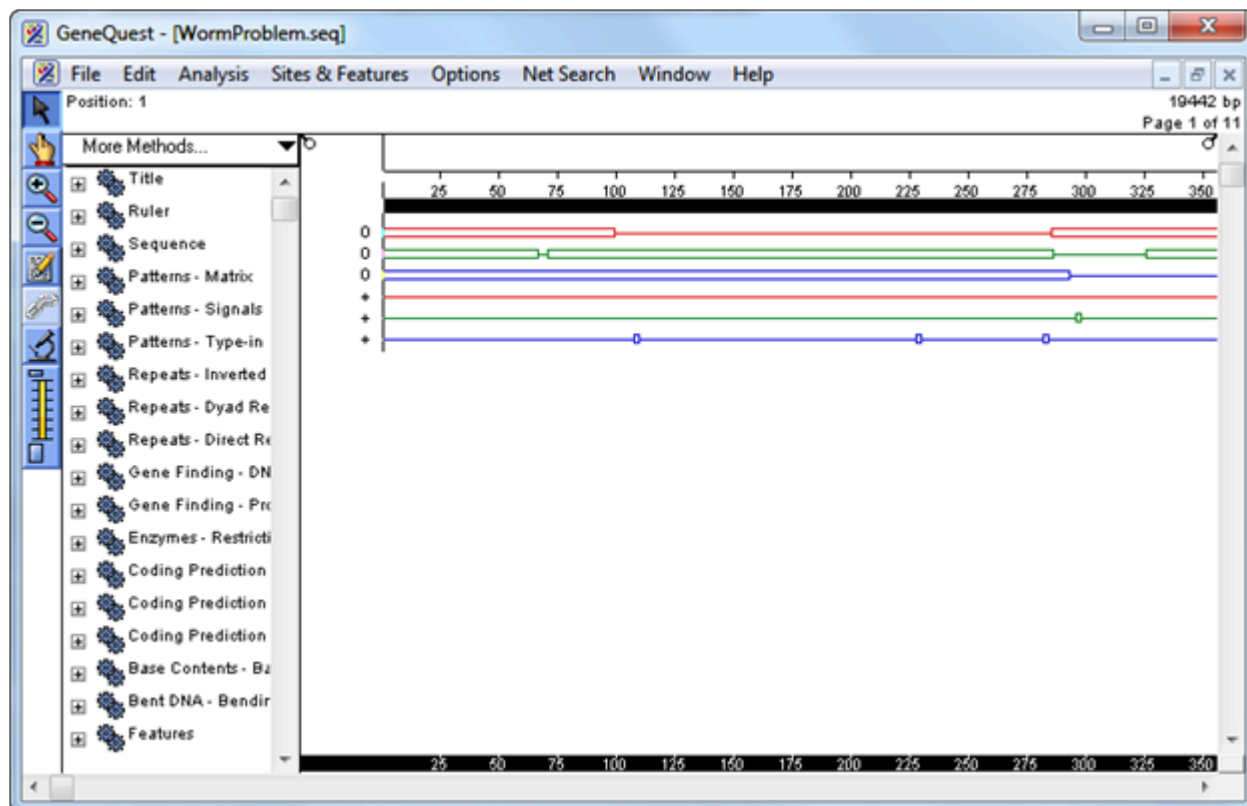
The main portion of the window is the Assay Document. It contains graphic displays of analysis methods that have already been used to analyze the sequence. The [Method Curtain](#) at left displays a list of available analysis methods.

3. Select **Analysis > Show Legend**. The [Legend Curtain](#) is displayed to the right of the Assay Document. This curtain displays the names of the methods that have been applied to the Assay Document. At this point, only the scale, sequence, and the ORFs and Starts for the first 3 reading frames are displayed.

4. In the Method Curtain to the left of the Assay Document, double-click on the method **Coding Prediction-Borodovsky- volvox_0.mat** so that the method parameters dialog is displayed.
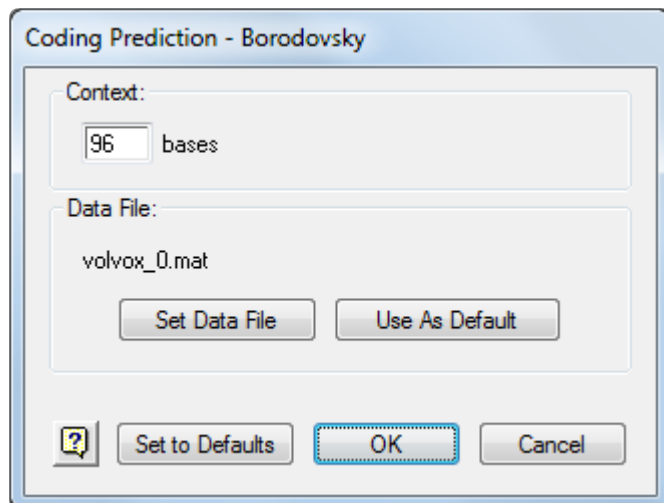
5. Click **Set Data File** and open the [Borodovsky matrix file location for your operating system](). Then click **Open** and then **OK**.

6. Click on the plus sign (Win) or the triangle (Mac) to the left of the method **Coding Prediction – Borodovsky – ce_lo_4.mat** to reveal the list of possible displays. Click the white space to the left of the displays to ungroup them.
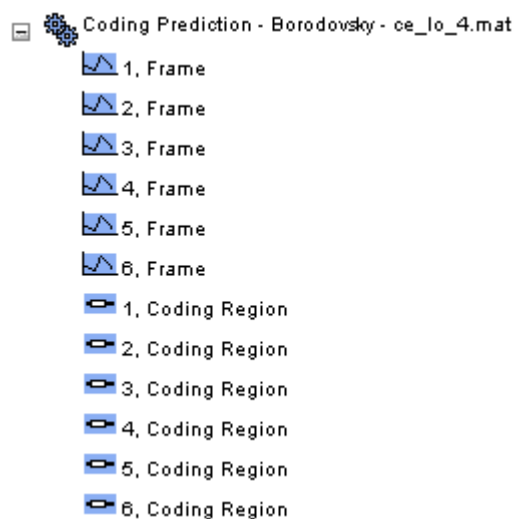
7. Use **Shift+click** to select Frames 1, 2 and 3.

8. Drag the selected displays onto the Assay Document and drop them under the existing displays. GeneQuest calculates potential coding regions and displays them as peaks. The higher the Borodovsky peak, the higher the probability that the area is a coding region.

9. Use colors to differentiate each frame of the newly applied Borodovsky displays from each other. Click the white space underneath the Borodovsky displays to ungroup them. Then click the Frame 1 Borodovsky display to select it. Choose **Options > Line Color > Red**. Repeat this sequence for Frame 2 and Frame 3 displays but select Green for Frame 2 and Blue for Frame 3.

10. **Shift+click** frames 1 and 3 of the Borodovsky methods, then select **Options > Superimpose Objects**.

11. In the Method Curtain, click on the plus sign (Win) or the triangle (Mac) to the left of the method **Patterns – Matrix**. Locate and select the file **Matrix Patterns/celegans_ds_2.mat** and click **Open**. This method will now be listed in the Method Curtain. Select the **Top Strand** of this method and drag it to the bottom of the Assay Document.

12. At the top of the Method Curtain, select **More Methods > Patterns > Matrix**. A new entry called **Patterns – Matrix** will appear at the top of the Method Curtain. Add the new matrix pattern **celegans_as_2.mat** as in Step 11, and again drag the **Top Strand** to the bottom of the Assay Document.

13. Repeat Step 11 for **celegans_atg_2.mat**.

    GeneQuest calculates potential translation start, donor and acceptor sites and displays them as vertical lines. Each apparent line is really a narrow box containing a numeric score denoting the likelihood of that area being a site.

14. To view the scores only, select the Zoom In tool () then click and drag diagonally to draw a box around the area you would like to see in more detail.

15. To view the scores and other information, select the Range Selector tool () then click and hold on a site until an information box is displayed.

    - The cursor snaps to a particular point on the site. The exact base to which the cursor snaps is shown on the header at the top left of the window next to the word "Position."

    - The pop-up information box displays the bases spanned by the pattern and a candidate ranking score.

16. Close the method curtain by selecting **Analysis > Hide Available Method**.

Proceed to Part B: Predict coding regions and create features.

Part B: Predict coding regions and create features

1. To make it easier to locate potential genes, click the object selector tool (). Drag and drop the method names so they appear in the Legend Curtain in the order shown.

2. Continuing to use the object selector, select each matrix pattern result in turn, then use **Options > Line Color** to change the color of the lines. For example, choose green for the donor sites, blue for the acceptor sites, and black for the start sites.

3. Use the scroll bar at the bottom of the window to locate the area of the sequence around 10,000 bp.

4. Select a zoom tool () or () and then click on the document to enlarge or reduce the view until you can see the entire sequence range from 9,000-12,000 bp. Alternatively, select a zoom tool and then "grab" a rectangular area of the document with your mouse to zoom in on that particular region.

5. Click the range selector tool () and then click on the frame 3 ORF centered on base 10,000. Note that a large peak in the frame 3 Borodovsky graph lies entirely within this highlighted ORF. It is likely that this area represents an exon.
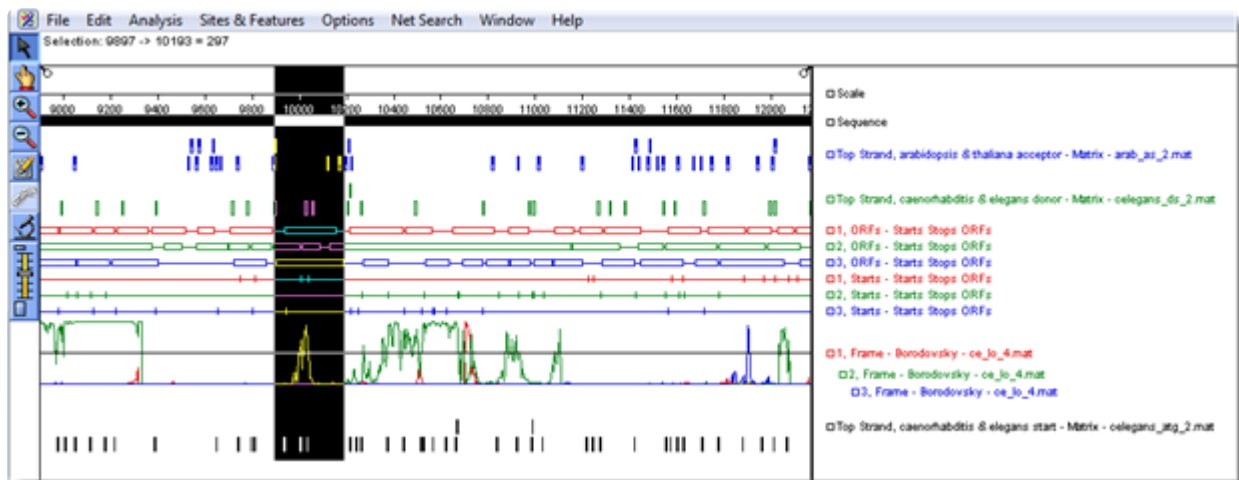
You are now ready to find the beginning of the potential gene.

6. Click the object selector tool (). In the Top Strand…atg_2.mat display, click and hold on the candidate start site at the left end of the ORF. The header shows the position of the candidate, while the pop-up information box shows the score.
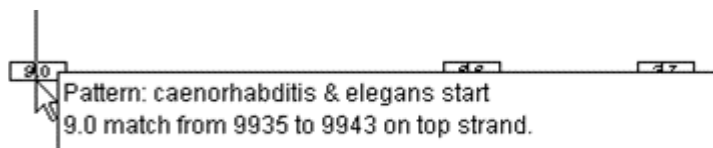
If this is the first exon of a multi-exon gene, it must end with a donor splice site that precedes the stop codon for the ORF in the same frame. In the Top Strand…ds_2.mat display, observe that there are two candidate donor splice sites (shown circled in the following figure) beyond the Borodovsky peak and still within the corresponding frame 3 ORF.
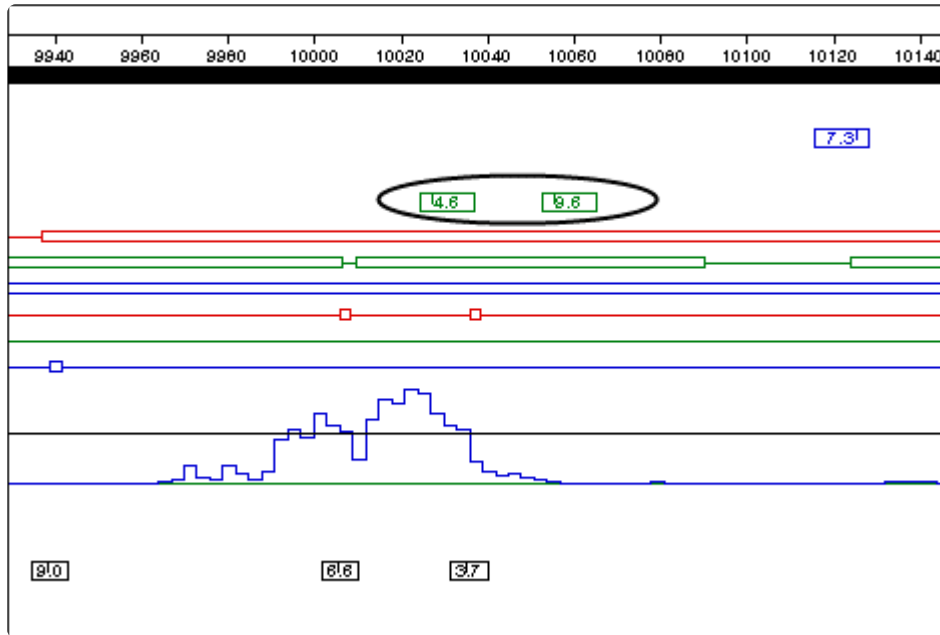
- Exon 1 Donor: Base 10028; score = 4.6

- Exon 1 Donor: Base 10056; score = 9.6

You are now ready to annotate the first exon. The position 10056 indicates the beginning of the donor splice site; therefore, the actual range of the exon extends from 9,939 to 10,055.

7. Select **Edit > Go to Location**. Enter the range 9939,10055 and then click **OK**.

8. Click the Annotation tool ( ) to open the Feature Editor. The location of the gene segment appears in the Segment box.

"]

in the same frame as exon 2, but also in the same ORF. The best candidate acceptor splice site to start exon 3 has a likelihood value of 14.1 out of a maximum 16.7—an excellent candidate. Two candidate donor splice sites are available to define the 3′ end of this exon, at 10,973 and 10,993. The first site scores higher at 6.4.

- Exon 3 Acceptor: Base 10,826; score = 14.1

- Exon 3 Donor: Base 10,973; score = 6.4

The next Borodovsky peak is once more in frame 2, but there is no candidate acceptor site (at the default threshold) under this peak. However, there is a site to the left of the peak, at position 11,021, that scores an impressive 15.1; this is the best candidate start for exon 4. You will note that there is no candidate donor splice site within the corresponding frame 2 ORF. This indicates that Exon 4 is probably the final exon in the gene.

To locate the stop codon:

14. Select **Analysis > Show Available Methods** and expand **Coding Prediction – Start Stops ORFs** by clicking on the plus sign (Win) or the triangle (Mac) next to the method name.

15. Select the **Stops** method for frame 2 and drag and drop it onto the Assay Document. Notice that there is a stop codon at the end of this ORF (base 11,155). Therefore, use the splice site at base 11,021 to start this exon and the stop codon at base 11,155 to terminate this exon.

- Exon 4 Acceptor: Base 11,021; score = 15.1

- Exon 4 Donor: No likely candidates

- Stop Codon: Base 11,155

16. After locating the endpoints for Exons 3 and 4, annotate them by repeating steps 11-14 for each exon and entering the appropriate ranges.

17. Now that all four gene segments have been annotated, click on either the object selector (🖐) or range selector (🖱) and then double-click on the feature to reopen the Feature Editor.

18. In turn, click on the coordinates for the three unnamed segments and type their names into the empty **Segment Name** boxes as **exon2**, **exon3**, and **exon4**.
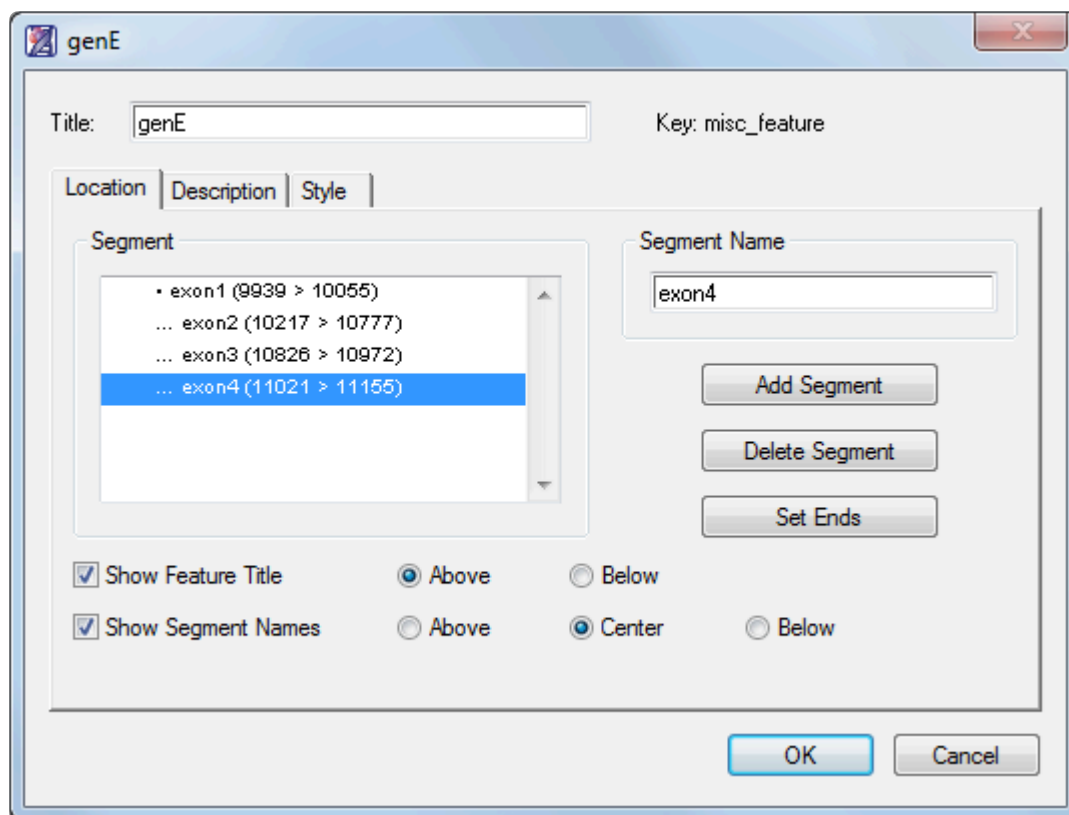
19. Click **OK** to finish. The feature should resemble the figure below:
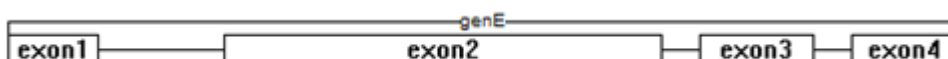
Proceed to Part C: Corroborate predictions via a BLAST search.

Part C: Corroborate predictions using a BLAST search

1. Click on the range selector ( ) then click on the new **genE** feature to select it.

2. Choose **Net Search > BLAST Selection**. Leave all settings at their default values, and then click **OK**. GeneQuest will display search results when they are returned from the server.

3. The first row is the one with the highest Score and is selected by default. Leaving the row selected, press the **Put in Document** tool in the header, then press **OK**.

4. Close the search results using **File > Close** or by clicking the 'x' in the top corner of the window.
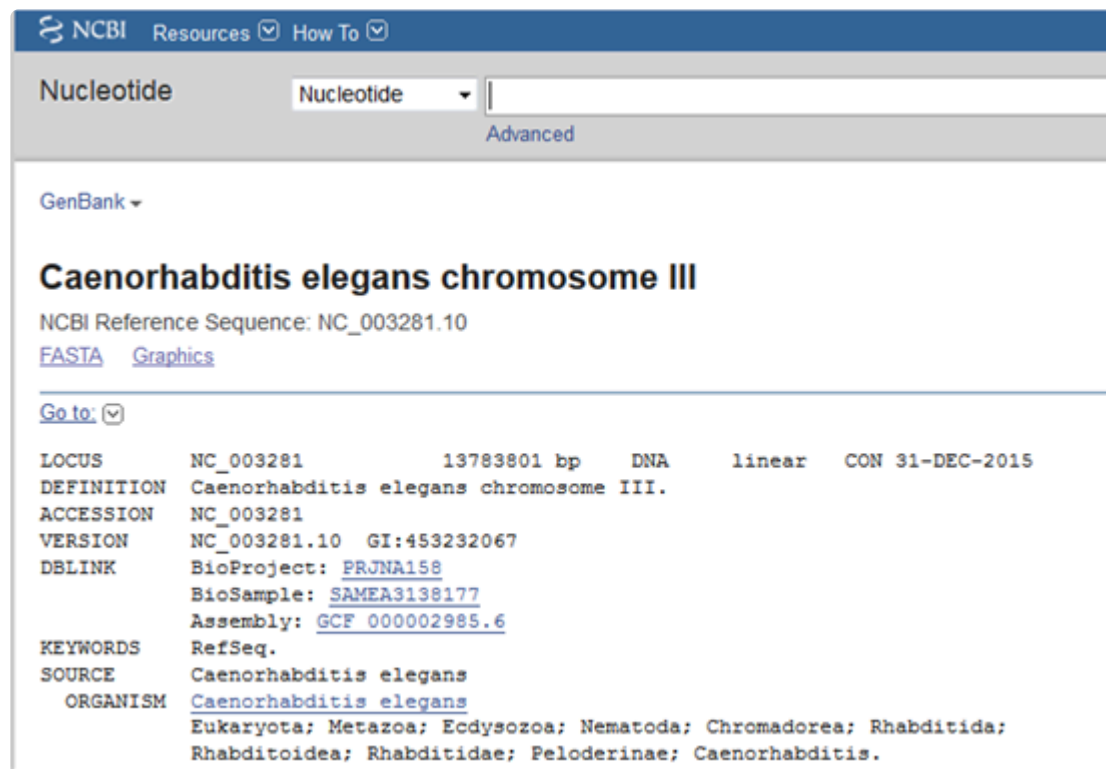
The sequence match now appears on the Assay Document under the **genE** feature, and corroborates the feature created earlier in this tutorial.
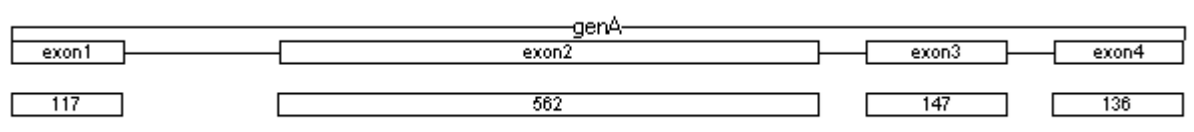
Proceed to Part D: Corroborate predictions using Entrez.

Part D: Corroborate predictions using Entrez

The BLAST search in Part C showed that the gene annotated in Part B is likely to be C. elegans gene enu-3.6. In this section, we will corroborate this finding by comparing our gene to an annotated sequence for C. elegans chromosome III, NCBI Entrez database entry NC_003281.10.

This tutorial is not dependent on the previous tutorials, so you can start from within the currently open GeneQuest project, or begin with a new, empty GeneQuest project.

1. Select **File > Open Entrez Sequence**.

2. Enter **NC_003281.10** in the dialog box, keep the default database selection (nucleotide), and then click **OK**.
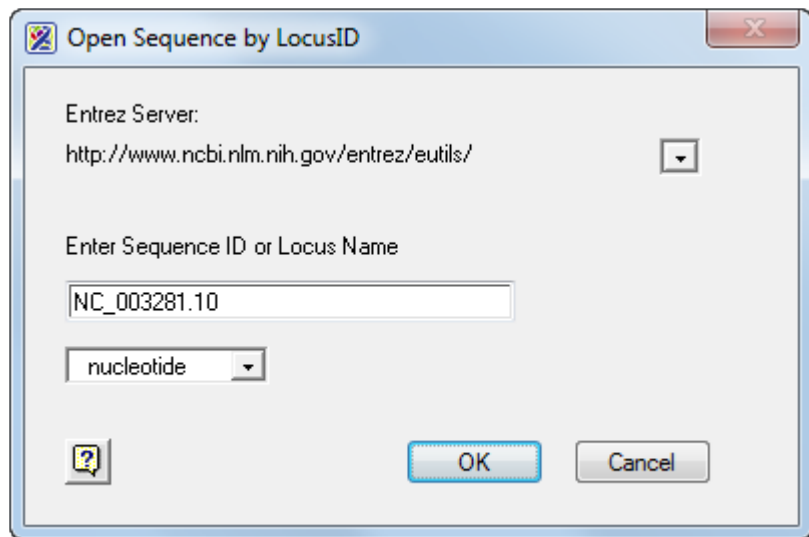
3. When prompted, click **Save**. This saves a local copy of the sequence and opens it as a new GeneQuest document, using the default method set.

4. In the Method Curtain, expand the Features method by clicking the plus sign or triangle to its left. There are over 30,000 features in the list.

5. To sort the features by name, use **Sites & Features > Sort Features > Name**.

6. Locate and select the **enu-3.6:CDS** feature and drag it to the Assay Document

7. Choose **Edit > Go to Location** and search for location **10751000**. Zoom in as needed.
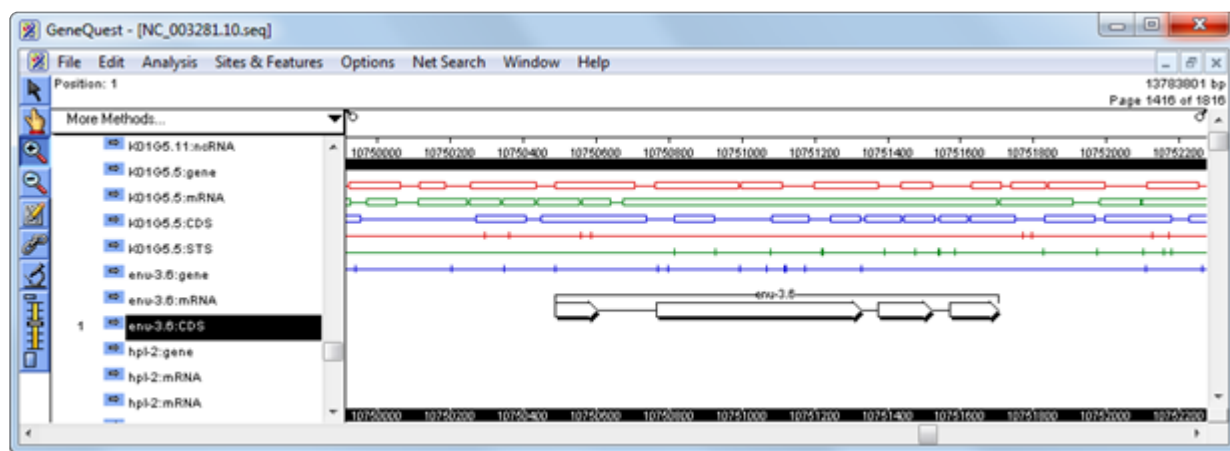
8. Double-click on the feature to open it in the Feature Editor.
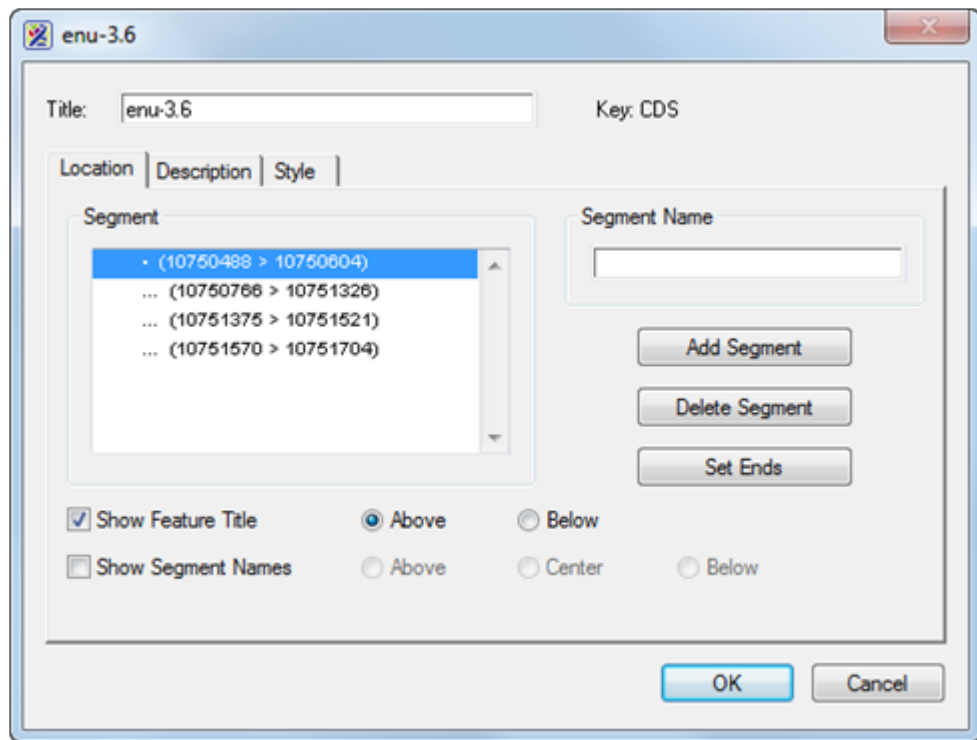
9. Compare this known feature you created in Part C of this tutorial. In both cases, exon 1 is 116 bp in length; exon 2 is 560 bp; exon 3 is 146 bp; and exon 4 is 134 bp. This is good evidence that the gene found and annotated in Part C is the enu-3.6 gene.
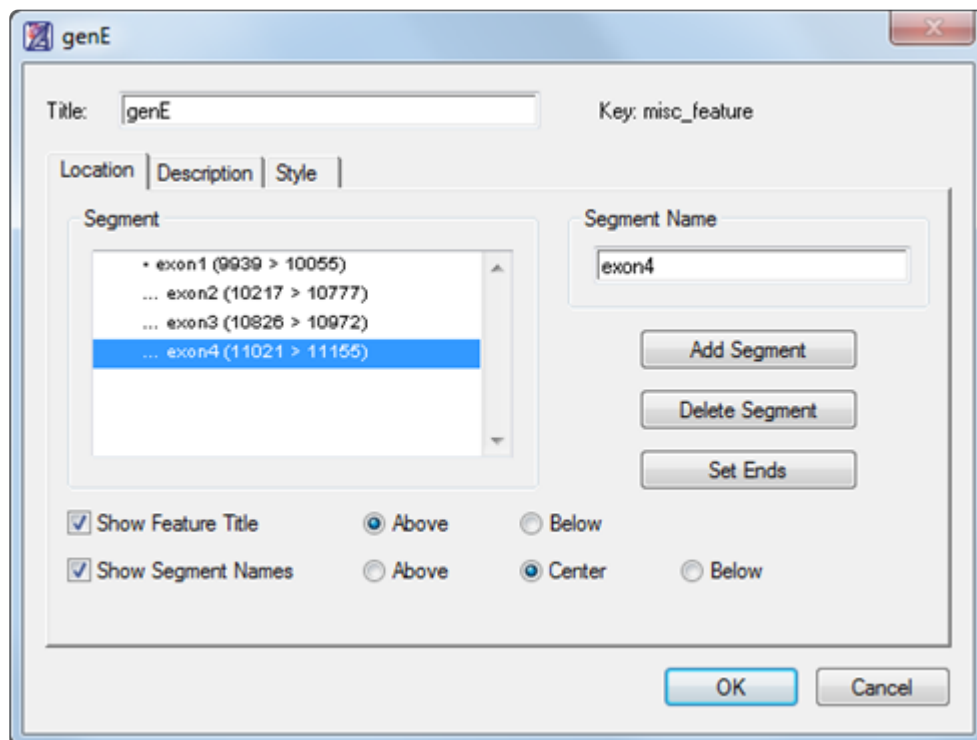
This is the end of the tutorial.

Create and Open Projects

See the following topics for detailed information on creating and opening projects:

- [Open a DNA sequence file](#)
- [Open a protein file or GeneQuest project](#)
- [Open a sequence from another application](#)
- [Open a sequence from an online database like NCBI BLAST or Entrez](#)

Open a DNA sequence file

To create a new GeneQuest project from a DNA sequence file

To open a sequence as a new GeneQuest project, do any of the following:

- Choose **File > New**
- Press **Ctrl/Cmd+N**
- Drag and drop one or more files on the open GeneQuest document
- To reopen a document that has recently been opened, select **File > Recent Documents** and choose it from the list

Next, navigate to and select the desired file.

To analyze a sub-range of a DNA sequence file:

The Set Ends dialog allows you to specify that only a selected sub-range of a sequence be used, or that its complement be used.

After following the steps above, click **Set Ends**.

Choose the sub-range you would like to analyze by typing its endpoints into the **5'** and **3'** text field boxes. Both endpoints are always inclusive, regardless of orientation. Alternatively, you may enter the endpoints by scrolling with the thumbwheels located under the boxes.

- To see the chosen sub-range in cartoon form, look at the box on the top right of the dialog (marked with a yellow arrow, below).
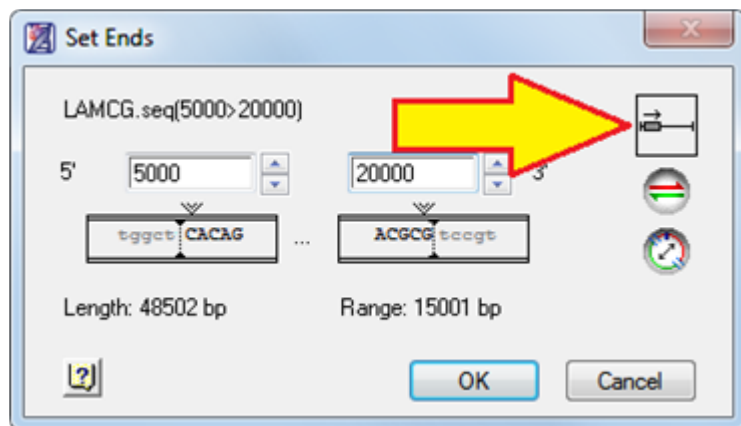
- To excise a polylinker from a nucleotide sequence, enter its endpoints, as above, and then click . GeneQuest will then enter the appropriate endpoints for the remainder of the strand (backbone).

- To choose the complement of the selected nucleotide strand, click .

- If you need to reset the coordinates to their original values, click on the word Length.

When you are finished, click **OK** to return to the file chooser.

Open a protein file or GeneQuest project

To open a GeneQuest project (*.dad*) or protein sequence file (*.pro*):

Do any of the following:

- Choose **File > Open**
- Press **Ctrl/Cmd+O**
- Drag and drop one or more *.dad* or *.pro* files on the open GeneQuest document

Next, navigate to and select the desired file. By default, only *.dad* documents are displayed. To show more file types, use the **Files of Type** drop-down menu.

If you only wish to analyze a sub-range of a DNA sequence file, click **Set Ends**. See Open a DNA sequence file for more information. Choose the coordinates and click **Open**.

If you choose a file that contains multiple sequences, the following dialog opens automatically.
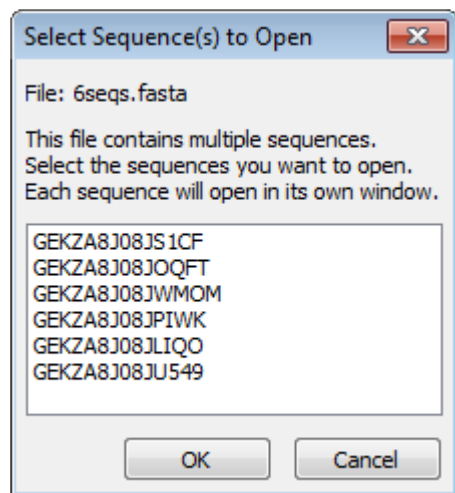
Click on the sequence you want to open. To open multiple sequences, uses **Shift+click** or **Ctrl/Cmd+click** to select them. Then press **OK**. Each selected sequence will open as a separate GeneQuest document.

> ✳ **Note:** If you open an existing GeneQuest project, most methods will appear exactly as they did when the document was last saved. However, GeneQuest will need to recalculate any Repeats or Enzymes > Restriction Map methods each time you open a document.

Open a sequence from another application

Many of the Lasergene applications will import sequences in multiple file formats. If a file type is not listed in your application of choice, you should open it in SeqBuilder Pro and either save it in the Lasergene sequence file format (*.seq*) or select **File > Send Sequence To** to open it in the application of your choice.

In addition, when you perform BLAST or Entrez searches using Lasergene applications, you can enter your choice of sequences from the match list directly to the Lasergene application. When you choose this option, Lasergene applications also save a copy of each sequence as a file in Lasergene format. You may also use the sequence's accession or ID number to directly download the data from NCBI and automatically open it into the Lasergene application.

Open a sequence from an online database

> ✳ **Note:** If you intend to use the enhanced functionality of SeqBuilder Pro for sequence analysis, you should import the online database sequences there. This saves several steps compared to opening them in GeneQuest and then exporting them to SeqBuilder Pro. The search and import functionality is identical in both applications.

GeneQuest's Net Search menu allows you to search online databases (e.g., the National Center for Biotechnology Information's BLAST and Entrez servers), for sequences that match all or part of the current sequence. Once a match has been located, you may add it to the current project or save it for use in other Lasergene applications.

When you perform BLAST or Entrez searches using a Lasergene application, you can enter your choice of sequences from the match list directly to the Lasergene application; the application will also automatically save copies of each file in Lasergene format. You may also use an accession number to directly download sequence data from NCBI and automatically open it in the Lasergene application.

See the following topics for detailed instructions:

- Open an Entrez database file by locus name or accession number

- Open an Entrez database file using a text search

- Open a BLAST database file using a similarity search

> ✳ At the time of each Lasergene release, the GeneQuest Preferences dialog contains current URLs for searching NCBI's BLAST and Entrez databases. If NCBI changes URLs at a later date, you will need to update the Preferences dialog as instructed in Specify URLs for BLAST and Entrez searches.

Specify URLS for BLAST and Entrez Searches

To specify URLs to use for online database searching :

1. Choose **Edit > Preferences** (Windows) or **GeneQuest > Preferences** (Macintosh) and select the **Servers** tab. The upper and lower portions of this dialog let you specify BLAST and Entrez server URLs, respectively. Unless "overwritten" by substituting custom addresses, the **Servers** tab uses the NCBI BLAST and Entrez server URLs that were current at the time of the Lasergene release. At present, these addresses are https://www.dnastar.com/blast/ncbi-blast2.html and https://www.ncbi.nlm.nih.gov/entrez/eutils.
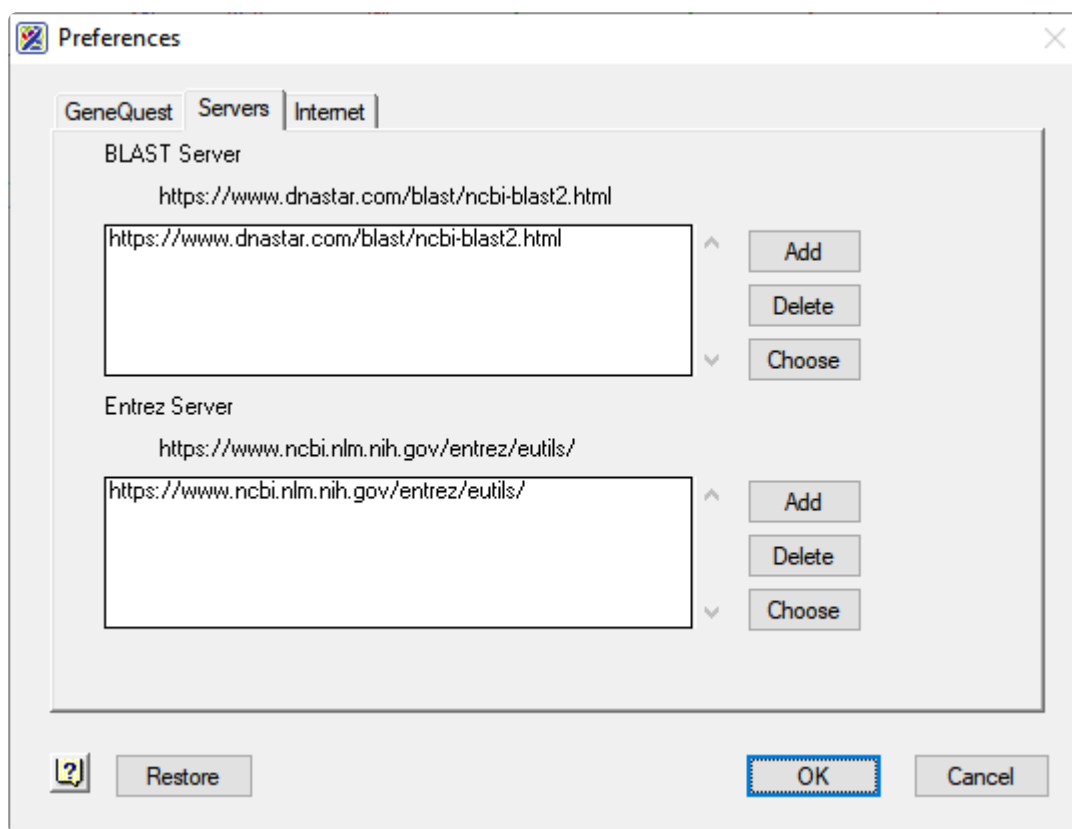
  - To add a new address to either box, click the associated **Add** button and type in the URL.

  - To delete an address, select the URL from the list and click the corresponding **Delete** button.

  - To set the default server for performing a search, select the URL from the list and then click the corresponding **Choose** button.

2. To save changes and edit from the Preferences dialog, click **OK**. To return to the default NCBI server URLs, click **Restore**.

Open an Entrez database file by locus name or accession number

If you have an Internet connection and know the locus name or accession number for a sequence, GeneQuest can retrieve it from the Entrez database at the National Center for Biotechnology Information (NCBI).

✱ **Note:** If you intend to use the enhanced functionality of SeqBuilder Pro for sequence analysis, you should import the online database sequences there. This saves several steps compared to opening them in GeneQuest and then exporting them to SeqBuilder Pro. The search and import functionality is identical in both applications.

1. Select **File > Open Entrez Sequence** or press **Ctrl/Cmd+R** to open the Entrez dialog.
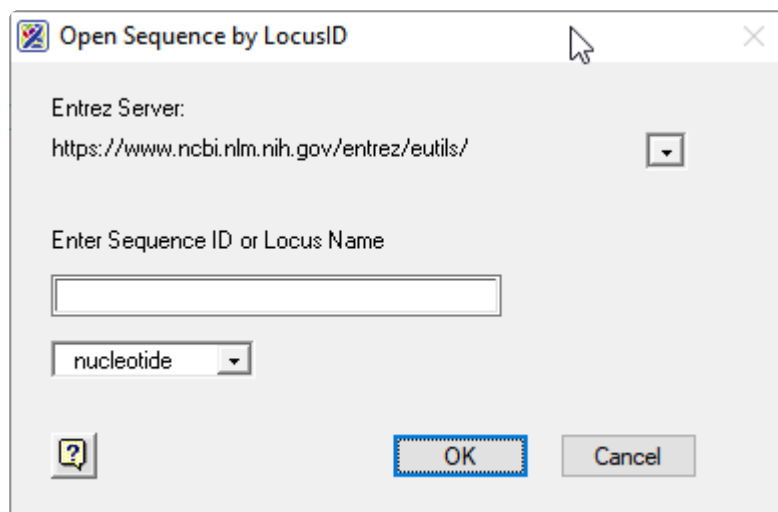
2. The Entrez Server shows the URL specified n the File Preferences dialog (see this topic). In the rare case you want to switch to a different server URL, click the triangle near the top right and choose **A New Server**. Then type in the desired URL.

3. Enter **Sequence ID** (accession number) or **Locus Name** into the text box.

4. Use the drop-down menu to choose the desired database. There may only be one choice available.

5. Click **OK**. If the search was successful, the Save As dialog box will open.

6. Save the sequence under the desired name and location. The default file name is the accession number or locus name followed by the sequence extension (*.seq* or *.pro*). After saving, the sequence will be opened as a new GeneQuest document.

Open an Entrez database file using a text search

GeneQuest provides a graphical interface for constructing Boolean queries of text databases like NCBI's Entrez database. Search results are returned as a list from which you may select any or all sequences, then save them or open them as documents in the current application.

> ✳ **Note:** If you intend to use the enhanced functionality of SeqBuilder Pro for sequence analysis, you should import the online database sequences there. This saves several steps compared to opening them in GeneQuest and then exporting them to SeqBuilder Pro. The search and import functionality is identical in both applications.

To perform a text search of the NCBI Entrez database:

One option is to use **File > Open Entrez Sequence**. Another is described below.

1. Select **Net Search > New Text Search**. The Entrez Query dialog appears.
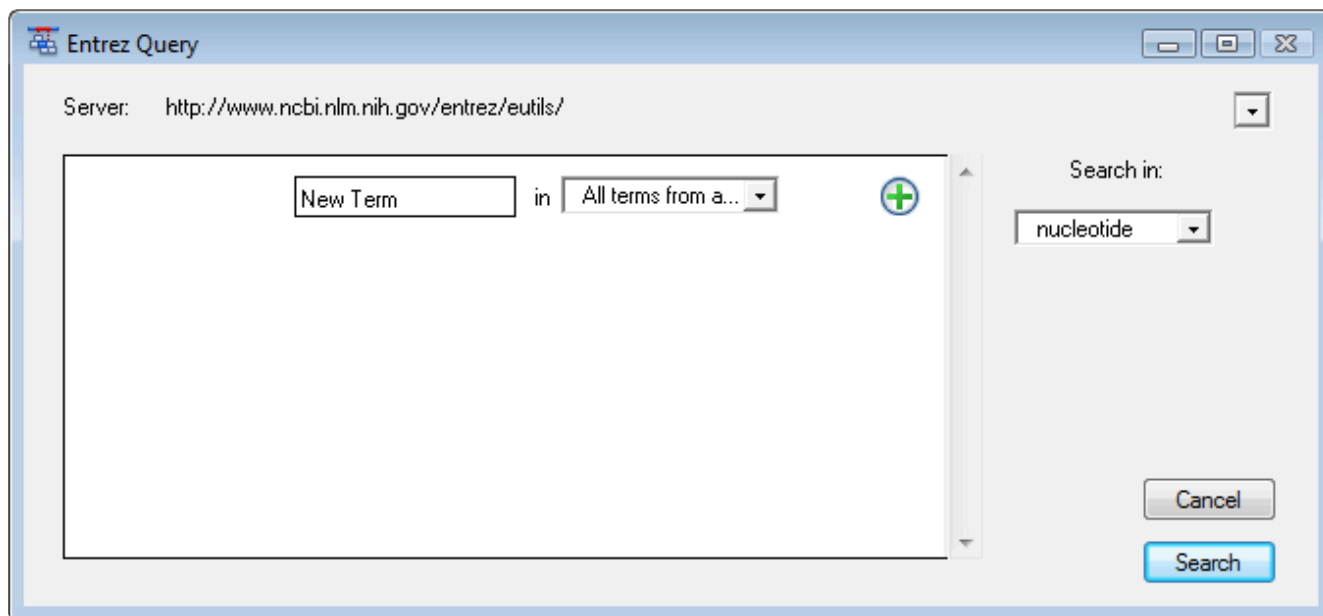
- The **Server** field lists the currently selected Entrez server. You may change to a different server by choosing one from the drop-down menu, or choose **A New Server** and then type a new URL.

- Use the drop-down menu in the lower right of the window to choose the database to search. This menu will reflect the available databases on the selected server.

2. Create the search criteria using the fields provided. To access a context menu of editing commands

(**Cut**, **Copy**, **Paste**, etc.), right-click anywhere in the Entrez Query dialog.
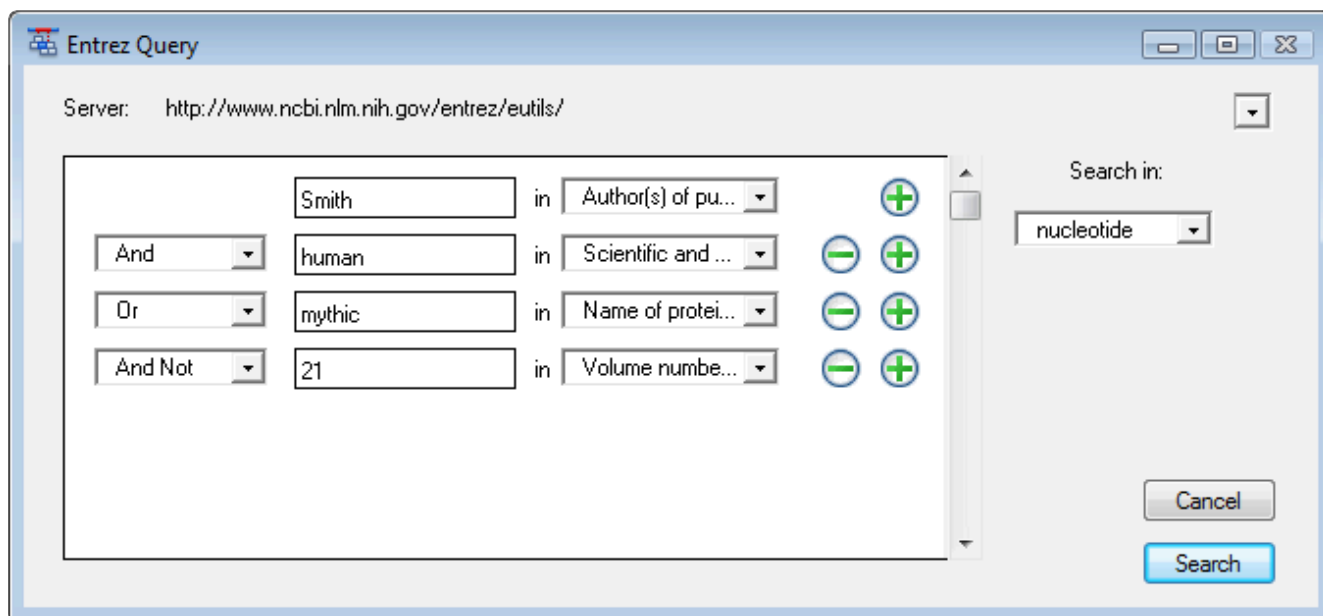
- Enter the first (or only) criterion into the text box on the left.

- You may change the search field heading from the drop-down list to the right of that text box. The default is **All terms from all searches**, which will search every word of every listing for the query term in the database. The field headings vary depending upon the online database that you are using.

- To add a criterion to the query, click the ⊕ button and choose **And**, **Or**, or **And Not** from the drop-down menu on the left of the new row. **And** requires both the first and the second criterion to be present. **Or** will look for matches containing either criterion. **And Not** will search for matches containing the first criterion but not the second.

- To remove a criterion, click the ⊖ button in the row that you want to remove.

3. Click **Search** to send the query to the database. Once the search is completed, any matches are displayed in the Entrez Search Results window.

Note about multiple-term search queries:

All Boolean operators have the same precedence in Entrez. Entrez processes the operators in a left-to-right sequence, so the order of the terms is important.

For example:

- x AND y OR z …is evaluated like… (x AND y) OR z

This means that any results containing "z" are considered a match, even if they don't meet the other criteria.

- y OR z AND x …is evaluated like… (y OR z) AND x

This means that results containing "z" are only considered a match if they also match x.

Entrez Search Results window

When an Entrez database search is complete, the Entrez Search Results window appears, showing a list of sequences that match your query.

If you came to this topic via the BLAST Search Results window, you are in the right place. The buttons near the tops of both dialogs work the same, so the links in the following table apply to both the Entrez and BLAST Search Results windows.
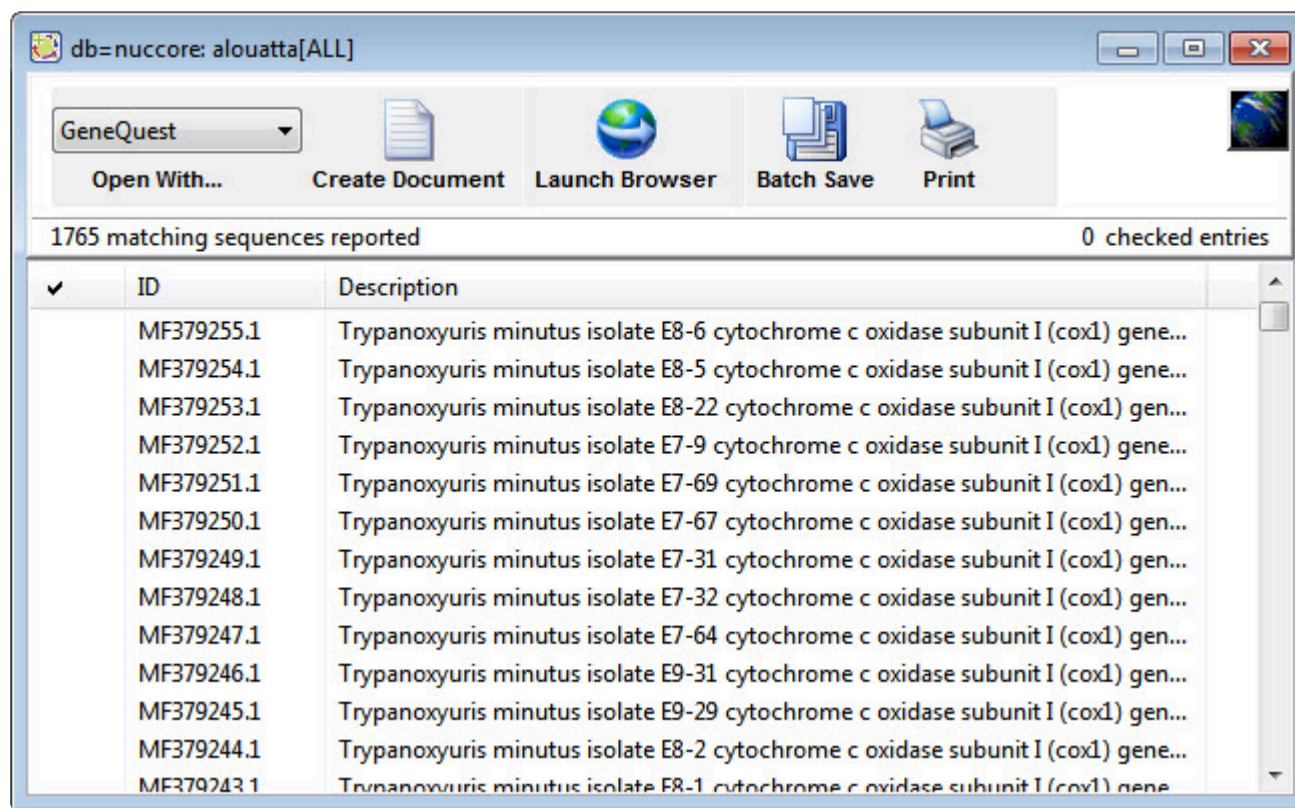
The buttons at the top of the window are used to open, save, copy or print results.

Task	Button to use…
To open one or more results in the selected Lasergene application	**Open With**
To open one or more results in the current application	**Create Document**.
To open one or more results in your default web browser	**Launch Browser**
To save a group of results as sequence files	**Batch Save**
To print one or more results	**Print**

The following table describes tasks that can be performed in the lower portion of the window:

Task	How to…
To sort results	Click on a column header to sort the rows by that column.
To narrow the results of a text query	Make sure the text search results window is active by clicking on it. If you have several text results windows open, select **Net Search > Current Results** to locate the desired window. Select **Net Search > Search These Results** and enter the desired search criteria. Click **Search**.
To change the search query	To modify a text query to obtain different results without having to completely recreate a complex query, select **Net Search > Change This Query**. The Entrez Query dialog appears containing the same search criteria that you last used, allowing you to edit as necessary before performing a new search.

After using **Batch Save** or **Create Document**, selections that have been downloaded or opened are denoted with a black diamond or arrow.
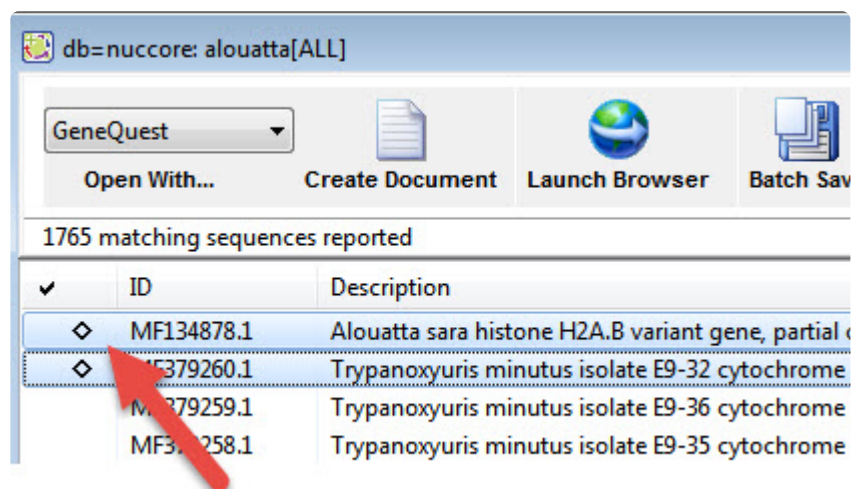
View search results in a web browser

To view one or more results from the Entrez or BLAST Search Results windows in a web browser:

Do either of the following:

- Double-click on any result from the window.

- Select one or more results from the window using click, **Ctrl/Cmd+click**, or **Shift+click**, or by adding a checkbox to the left of the row(s) of interest.Then click the **Launch Browser** button or use the **Net Search > Launch Browser** menu command.

Each selected match will be displayed in a separate browser window.

> ✳ Popup blocking software may prevent Lasergene from opening multiple web browser windows simultaneously. Check the documentation for your popup blocking software for additional information.

Open results in a Lasergene application

To open one or more results from the Entrez or BLAST Search Results windows:

1. Select one or more results from the Search Results window using click, **Ctrl/Cmd+click**, or **Shift+click**, or by adding a checkbox to the left of the row(s) of interest.

2. Do any of the following:

   • To open the results in a Lasergene application other than SeqBuilder, select its name from the drop-down menu above the words **Open With**.

   • To open the results in GeneQuest, click **Create Document**. In the resulting dialog, select one of the following options from the **Save** field:

     ◦ **Top** opens the specified number of matches, starting with the first sequence on the list. Specify a number in the text box (default is 10).

     ◦ **Next** opens the specified number of sequences, starting with the first one currently highlighted. Specify a number in the text box (default is 10).

     ◦ **Checked** opens sequences with a checkmark beside them.

     ◦ **Selected** opens only highlighted sequences.

     ◦ **All** opens all sequences.

   • To save the file(s) in a location other than the default directory, click the **Set Location** button.

3. Click **OK**.

Print search results

To print results from the Entrez Search Results window:

1. Select one or more results from the Search Results window using click, **Ctrl/Cmd+click**, or **Shift+click**, or by adding a checkbox to the left of the row(s) of interest.

2. Click the **Print** button to proceed with printing using the standard Print dialog for your computer's operating system.

To print results from an BLAST Search Results window:

1. Select one or more results from the Search Results window using click, **Ctrl/Cmd+click**, or **Shift+click**, or by adding a checkbox to the left of the row(s) of interest.

2. Click the **Print** button.

3. In the resulting dialog, select one of the following options from the **Print** field:

   - ◦ **Top** prints the specified number of matches, starting with the first sequence on the list. Specify a number in the text box (default is 10).

   - ◦ **Next** prints the specified number of sequences, starting with the first one currently highlighted. Specify a number in the text box (default is 10).

   - ◦ **Checked** prints sequences with a checkmark beside them.

   - ◦ **Selected** prints only highlighted sequences.

   - ◦ **All** prints all sequences.

4. Click **Print** to proceed with printing using the standard Print dialog for your computer's operating system.

Copy search results

Selected results may be copied to the clipboard from the [Entrez](#) or [BLAST](#) Search Results windows.

To copy detailed information:

1. Select one or more results from the Search Results window using click, **Ctrl/Cmd+click**, or **Shift+click**, or by adding a checkbox to the left of the row(s) of interest.

2. Choose **Edit > Copy** to move results to the clipboard. At a minimum, copied results will include an identifier and description for each selected result.

3. Paste into a spreadsheet or text document.

To copy exact text sequence only:

1. In the bottom pane of the BLAST search results window, select the desired text by dragging your mouse.

2. Choose **Edit > Copy** to move the selected text to the clipboard.

3. Paste into a spreadsheet or text document.

✳ When pasting data into a word processing application, choose a mono-spaced font such as Courier so that the alignments display correctly.

Save search results

To save selected results from the Entrez or BLAST Search Results windows as a text file:

1. (optional) Specify a default download folder for online search results.

2. Make the Entrez or BLAST Search Results window the active window.

3. Select **File > Save As**.

   By default, the file will be saved with a text file extension (*.txt*).

To save multiple search results as sequence files:

1. (optional) Specify a default download folder for online search results.

1. Select one or more results from the Search Results window using click, **Ctrl/Cmd+click**, or **Shift+click**, or by adding a checkbox to the left of the row(s) of interest.

2. Click the **Batch Save** button.

3. In the resulting dialog, select one of the following options from the **Save** field:

   - ◦ **Top** opens the specified number of matches, starting with the first sequence on the list. Specify a number in the text box (default is 10).

     ◦ **Next** opens the specified number of sequences, starting with the first one currently highlighted. Specify a number in the text box (default is 10).

     ◦ **Checked** opens sequences with a checkmark beside them.

     ◦ **Selected** opens only highlighted sequences.

     ◦ **All** opens all sequences.

4. (optional) To save the file(s) in a location other than the default directory, click the **Set Location** button.

5. Click **OK** again to save the sequence files to the specified directory.

Specify a download folder for online search results

To specify a download folder for results from an Entrez or BLAST Search Result window:

1. Choose **Edit > Preferences** (Windows), **GeneQuest > Preferences** (Macintosh) or press **Cmd+,**.

2. Select the **Internet** tab.
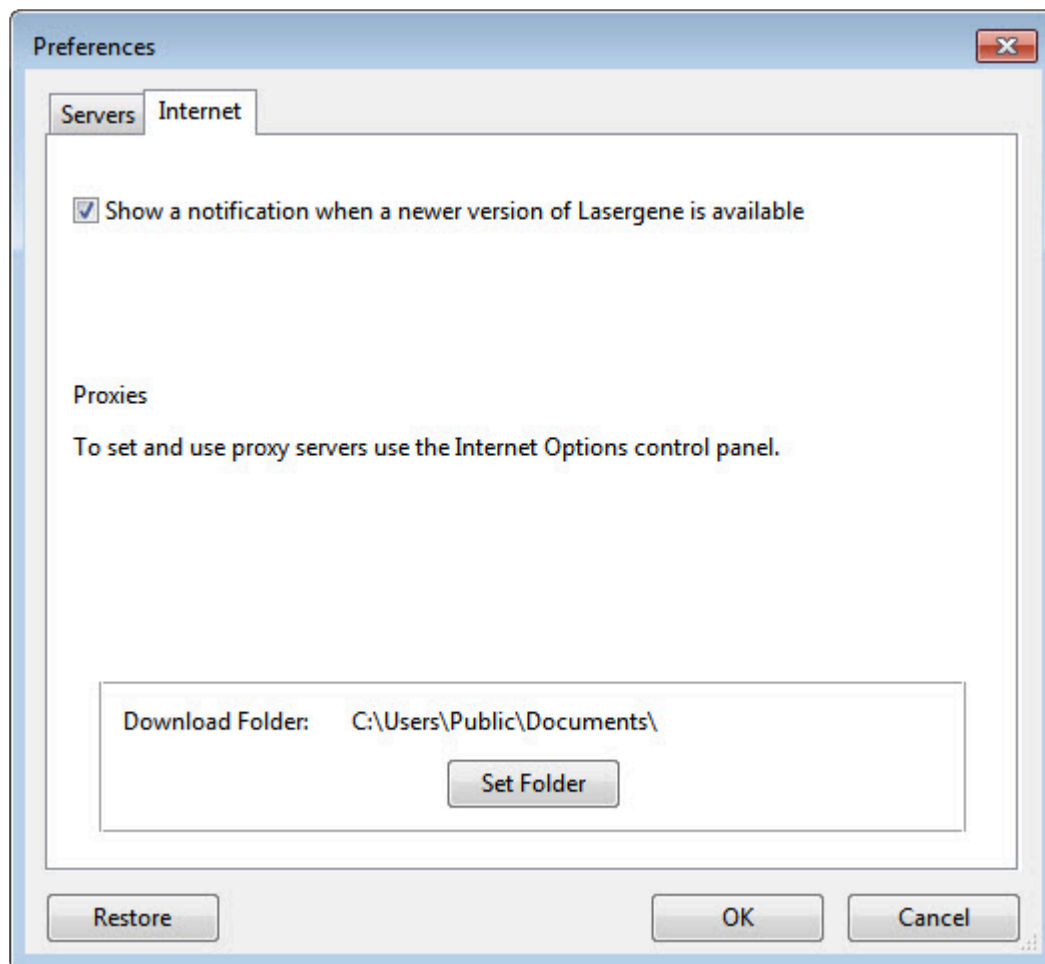
3. (optional) To automatically check for Lasergene software updates, check the box next to **Show a notification when a newer version of Lasergene is available**.

4. To specify the default directory for downloading results from online database searches, press the **Set Folder** button.

5. Navigate to the desired location and press **Select Folder**.

6. To save changes and exit from the Preferences dialog, click **OK**.

Open a BLAST database file using a similarity search

⚹ **Note:** If you intend to use the enhanced functionality of SeqBuilder Pro for sequence analysis, you should import the online database sequences there. This saves several steps compared to opening them in GeneQuest and then exporting them to SeqBuilder Pro. The search and import functionality is identical in both applications.

To perform a BLAST search:

1. Select the sequence or portion of sequence that you want to search for.

2. Choose **Net Search > BLAST Selection** or **Net Search > BLAST ORF**.
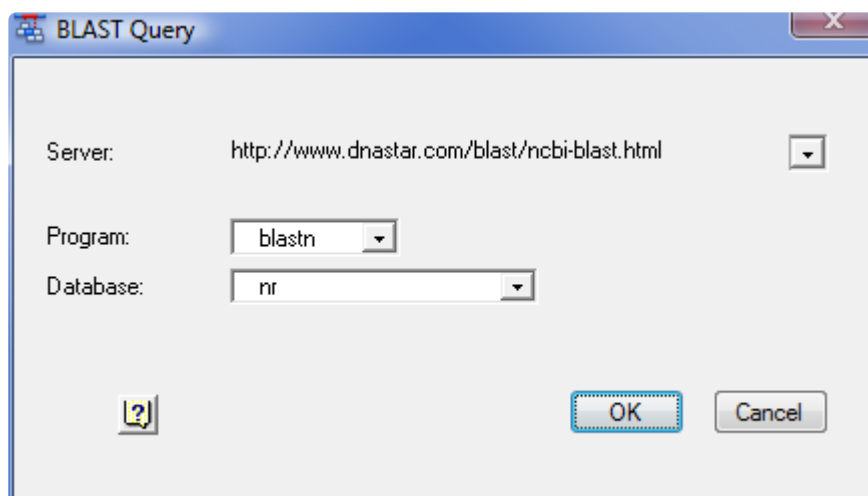
- The **Server** field lists the currently selected BLAST server. You may change to a different server by choosing one from the drop-down list provided, or choose A **New Server** and then type a new URL.

- The **Programs** field lists the currently selected BLAST program. The following programs are available through this list, though some servers may not offer all of these programs:

  - **blastn** searches a nucleotide sequence database with a nucleotide query sequence.
  - **blastp** searches a protein sequence database with an amino acid query sequence.
  - **blastx** searches a protein sequence database with the six-frame translation of a nucleotide query sequence.
  - **tblastn** searches the six-frame translation of a nucleotide sequence database with a protein query sequence.
  - **tblastx** searches the six-frame translation of a nucleotide sequence database with the six-frame translation of a nucleotide query sequence.

- The **Database** field lists the database that will be searched. Choose from the databases available on the server by selecting one from the list provided. Using less common databases may cause a noticeably slower search. According to NCBI, this is because the less frequently used databases may not already be loaded in memory on an available server. In contrast, frequently used databases such as **nr** or **nt** are already loaded in memory on a number of servers. Also note that available databases may include the **refseq_representative_genomes** database. This database contains representative bacterial/archaeal genomes, as opposed to all the genomes (some of which are multiples) now in Genbank or RefSeq. The **refseq_representative_genomes** database can be helpful when matches to *Escherichia coli/ Shigella*, for example, occupy all 100 of the top BLAST results. For example, say you want to find the location of a particular sequence in the curated "model organism version" of the E. coli K-12 MG1655 genome (U00096.3/NC_000913.3). Here is the location of the true sequence match when using different databases: **refseq_genomic** - absent from top 100 hits. **nr** - far down the list of the top 100 hits. **refseq_representative_genomes** – the top hit.

3. After confirming the values in the Query dialog, click **OK** to initiate the search. Once the search is completed, any matches are displayed in the [BLAST Search Results window](). You may continue to work with other Lasergene application activities while waiting for BLAST search results, but you cannot launch another BLAST search until you close the current BLAST window.

BLAST Search Results window

When a BLAST database search is complete, the BLAST Search Results window appears, showing a list of sequences that match your query.
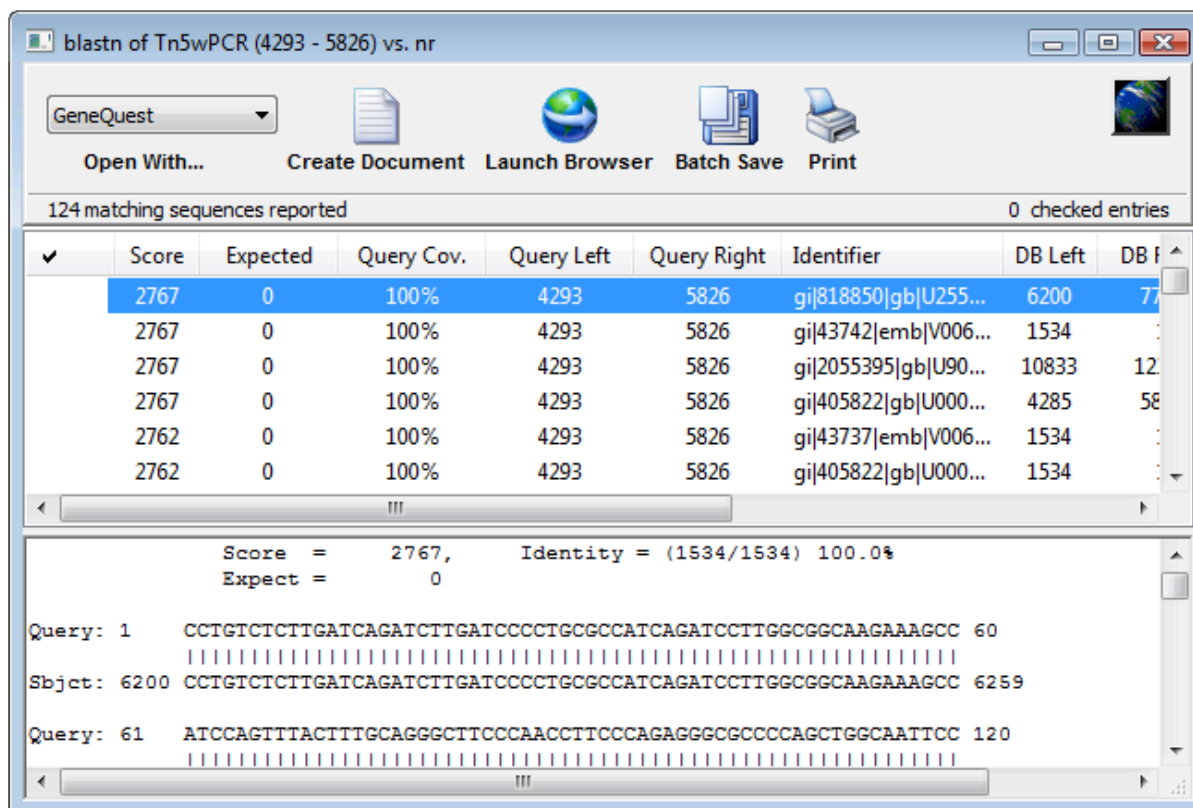
Task	Button to use…
To open one or more results in the selected Lasergene application	**Open With** menu
To open one or more results in the current application	**Create Document**
To open one or more results in your default web browser	**Launch Browser**
To save a group of results as sequence files	**Batch Save**
To print one or more results	**Print**

The lower pane of the BLAST Search Results window contains the alignment of the query (upper sequence) to the highlighted database entry (lower sequence). The upper pane of the window contains the name of possible matches in order of probability. After using **Create Document** or **Batch Save**, selections that have been downloaded or opened are denoted with a black diamond or arrow.
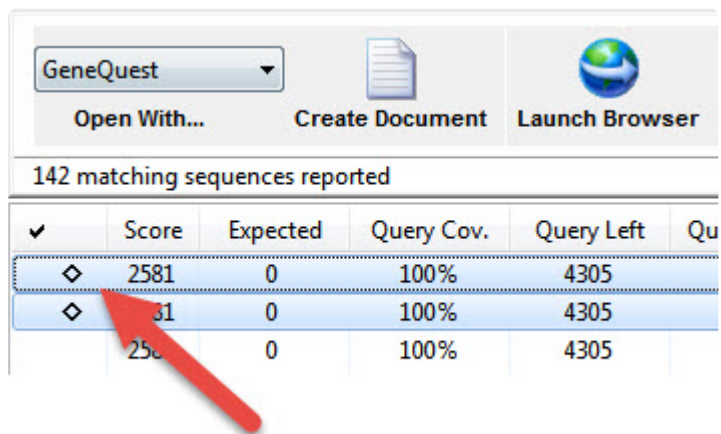
The following columns are available in the upper part of the window. Click on a column header to sort the rows by that column.

Column Name	Description
Score	The similarity score for the result.
Expected	The expectation value. This corresponds to the **Expect** value displayed in the lower pane of the window. **Note:** Detailed information about how **Score** and **Expected** are derived is available at www.ncbi.nlm.nih.gov/BLAST. In general, a higher score and lower expectation connote a better match.
Query Cov.	The percentage of the queried sequence that aligns to a match in the database. For example, if you BLAST a range of 596 bases, and 565 of those bases align to a sequence in the database, then the **Query Cov.** value will be 565/596, or 95%. **Note: Query Cov.** differs from the **Identity** value shown in the lower portion of the BLAST window. Identity is the percentage of bases within the alignment that match. For example, if 565 bases align to a sequence in the database, and there are 3 mismatches within the alignment, then the **Identity** value would be 562/565, or 99.5%.
Query Left	The left end coordinate of the queried sequence that corresponds to the left end coordinate of the matching result (DB Left).
Query Right	The right end coordinate of the queried sequence that corresponds to the right end coordinate of the matching result (DB Right).
Identifier	The sequence identifier for the result.
DB Left	The left end coordinate of the result sequence that corresponds to the left end coordinate of queried sequence (Query Left).
DB Right	The right end coordinate of the result sequence that corresponds to the right end coordinate of

	queried sequence (Query Right).
Description	The description, from NCBI, of the result sequence.

Change the default application for opening Lasergene files

On Windows:

1. Open the **Control Panel**.

2. Go to **Programs > Default Programs > Associate a file type or protocol with a specific program.**

3. Select extension you wish to change in the **Extensions** list and then click the **Change program** button.

4. Click the **Browse** button in the Open With dialog.

5. Navigate to and select the application you want the files to open in. (Lasergene applications are located by default in **C:\Program Files\DNASTAR\Lasergene 'x'** or **C:\Program Files (x86)\DNASTAR\Lasergene 'x'**, where 'x' represents the version number).

6. Click **Open**.

7. Click **OK** to close the Open With dialog, and then click **Close**.

On Macintosh:

1. Select a file of the type you wish to change in Finder.

2. Go to **File > Get Info**.

3. Click the drop-down arrow in the **Open With** field, and navigate to the application you want the files to open in. (Lasergene applications are located by default in **Applications:DNASTAR:Lasergene 'x'**, where 'x' represents the version number).

4. Click the **Change All** button to make the application you selected the default for all files of this type.

To change the default application for opening Lasergene files:

Click here for detailed information.

GeneQuest User Interface

For information about the components of the GeneQuest user interface (window), see:

- Assay Document – the main view in GeneQuest
  - Customize the look of the Assay Document
- Method Curtain
- Legend Curtain
- Toolbar

Assay Document (main view)

GeneQuest's main view is called the Assay Document.
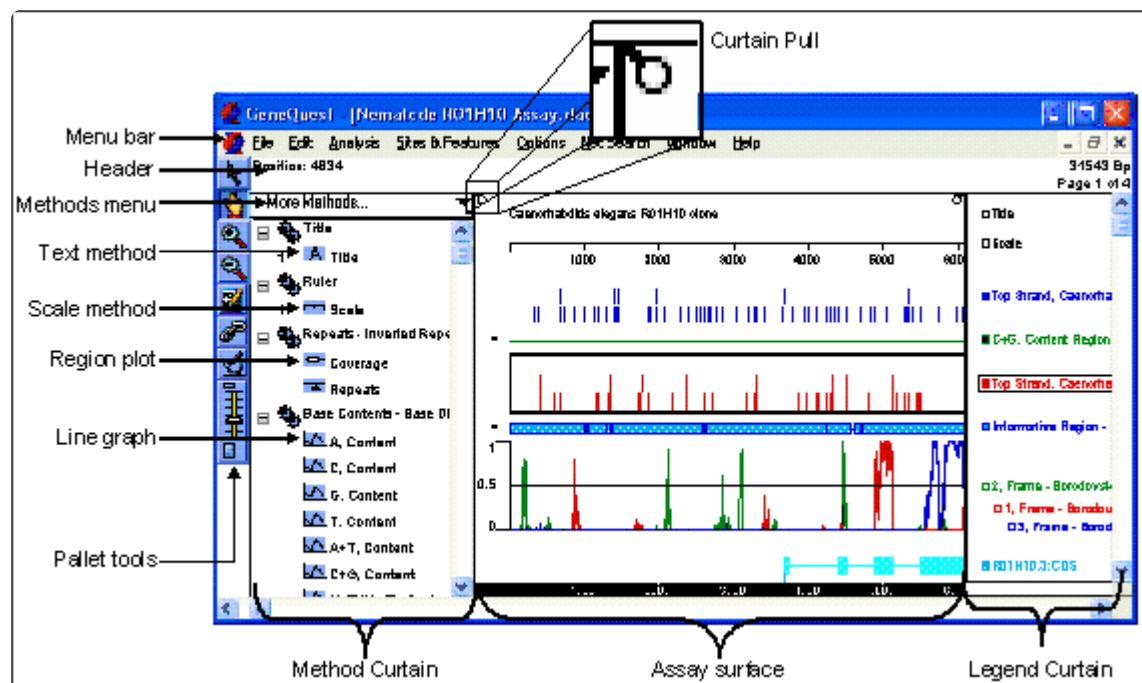
The header contains two entries at the left.
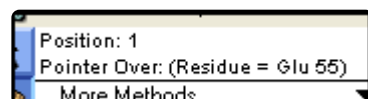
On top is the Site and Sequence display, which shows what has been selected by the range selector.

- If a single nucleotide has been selected, **Position** shows its location.

- If a range has been selected, **Selection** shows the first and last positions of the selected range and the length of the range.

- If a range spanning discontiguous segments of the same feature has been selected, **Selection** shows the first and last positions of each selected segment and the sum of lengths of these segments.

Underneath this is the Pointer Position Display, which shows whether the pointer is over the document, Method Curtain or Legend Curtain. When it is over the document display, the base and its sequence coordinate will appear in parentheses.

There are another two entries on the right.

m.w. 16837.64, ch. -23.87, pl 3.93 [149 AA]
Page 1 of 2
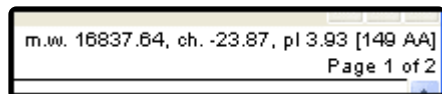
The top number shows the total length of the sequence in base pairs. The bottom display shows the current page and total number of pages in the document. The page numbers may change, depending on amount of zoom, the **Amplitude** control setting, and the **Horizontal Units** setting. Be sure to check the number of pages before printing your document.

GeneQuest offers line graphs and region plots as two ways of visualizing the results of the same method.

- **Line graphs** ( ) are composed of a series of connected points, calculated at each residue and averaged over the window you specify. The Amplitude control palette tool, Options > Plot in Peak Scale command and Options > Plot in Constant Scale command can all be used to change the amplitude of these graphs.

- **Region plots** ( ) use the calculated data to make a yes/no decision, drawing a bar to represent a "yes" answer. The threshold for calling a region a distinct type (e.g., "high GC content") may be adjusted when editing method parameters.
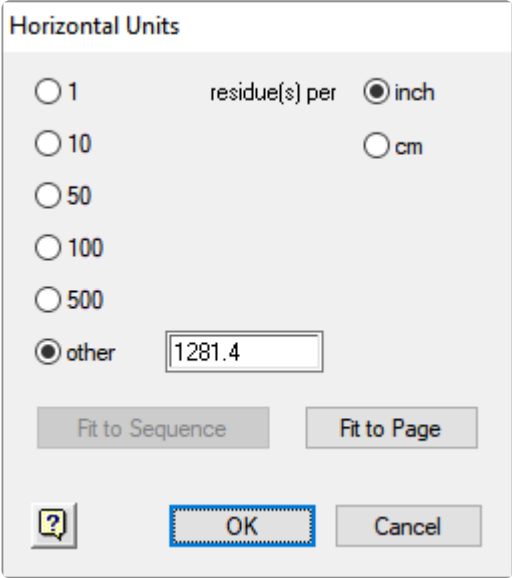
To customize the appearance of the window, see Customize the look of the Assay Document.

Comparing portions of the Assay Document:

The Assay Document has a pane divider to help you compare results from different vertical locations. This allows direct comparison without constantly scrolling up and down the document and without having to drag a method to a new location in order to juxtapose it with another method. Drag the pane divider, located near the bottom right corner of the assay document, up to activate two panes. You will now see two vertical scroll bars that you can use to position method results for comparison. To deactivate the second pane, drag the pane divider back to the bottom of the window.

Customize the look of the Assay Document

The following table describes ways to customize the look of the Assay Document.

Task	How To
Change colors, fonts or other style options	Select **Options > Line Color** (**Line Weight**, **Fill Pattern**, **Fill Color**, **Page Color**, **Font**, or **Size**)
Make the display more compact	Select two or more methods in the assay document and choose **Options > Superimpose Objects**.
Define the number of nucleotides per unit length on the horizontal axis	Use **Options > Horizontal Units**. This command is only enabled if assay document is the active window. • Choose the number of nucleotides using the radio buttons on the upper left. If you choose **Other**, a text field appears for you to enter the number of residues. • Specify units in inches or centimeters using the radio buttons on the upper right. • Use **Fit to Page** to put the entire assay on a single page or **Fit to Sequence** to zoom in/out the amount required to see individual nucleotides. In some cases, the first button may be disabled.
To view a	Click and hold your mouse button down on a method on the Assay Document. The tooltip

tooltip with method information	contains an outline of method properties. Each tooltip varies according to the display type and analysis method, but generally contains information about calculation methods and current parameter settings.
Change method parameters	Double-click on a method on the assay document.
Set the amplitude (height) of a selected line graph	Use **Options > Plot in Peak Scale** to set the amplitude to the maximum range for the current method. Choose **Options > Plot in Constant Scale** to instead to scale the selected line graph to the maximum and minimum possible values for all methods. This option is useful for comparing method results with two different assays. These commands toggle with one another and are disabled unless one or more line graph methods are selected in the Assay Document. Note that Borodovsky methods compute values between 0 and 1 and typically always cover the entire range. Thus, the Peak Scale and Constant Scale for Borodovsky methods will usually be the same.
Change the line color for a selected display	Use the **Object selector** to select one or more displays, then choose **Options > Line Color**.
Change the line weight for a selected display	Use the **Object selector** to select one or more displays, then choose **Options > Line Weight**. Choose from four point sizes: 0.25, 0.5, 1 and 2.
Change the fill pattern for a selected display	Use the **Object selector** to select one or more displays, then choose **Options > Fill Pattern**. The top left pattern, marked by an asterisk (**) is for no pattern. Use this to match the pattern in tandem with the background selection made with *Options > Page Color** (see row below). The solid black pattern to its right is for a completely solid fill.
Change the box and arrow fill color for a selected display	Use the **Object selector** to select one or more displays, then choose **Options > Fill Color**. Note that colors for annotated features are set individually within the [Style tab](#) of the Feature Editor dialog.
Change the color of the assay background	Use **Options > Page Color** to toggle the assay document background between black and white.
Change the	Use **Options > Font** to choose fonts by name. Mono-spaced fonts work best. To change the

Method Curtain

The function of the Method Curtain is to store analysis and annotation methods for inclusion in the Assay Document.

Each analysis method may have one or more of the eight display types: scale, sequence, feature, pattern, repeat, restriction site, line graph, and region plot. For example, Coding Prediction – Local Compositional Complexity has one line graph and one region plot. To see all display types for a given method, click the triangle icon to the left of the method name in the Method Curtain. Any numbers preceding a display type indicate how many times that selection appears on the Assay Document. There are three ways to open the Method Curtain:

- Drag the Method Curtain "pull ring" to the right.

- Select **Analysis > Show Available Methods**. This command toggles with **Analysis > Hide Available Methods**.

- Press **Ctrl/Cmd+[**.

The contents of the Method Curtain are determined initially by the active method outline, usually the default method outline.

The following table shows a list of tasks that apply to the Method Curtain.

Task	How To
Change the method in which analysis methods appear	Use your mouse to drag each method to the desired location.
Apply additional methods or additional copies of methods in use	Use the **More Methods** menu located at the top of the Method Curtain. Click on any method to move it into the main body of the Method Curtain.
Change parameters for a method in the	See Change method parameters or display or hide the results of the method analysis at any time.

Method Curtain	
Display multiple results for the same method with different sets of parameters	See [Compare results for the same method, but using different parameters](#).
Remove a method display from the assay document	Use the [object selector](#) to highlight the method's display on the assay document, then select **Edit > Clear** or press **Backspace** (Win) or **Delete** (Mac).
Remove a method from the Method Curtain	Click on the name of the method, then select **Edit > Clear** or press **Backspace** (Win) or **Delete** (Mac). Any displays created from this method will also be removed from the Assay Document.
Choose which displays to apply to the assay document	Before you drag an entire method from the Method Curtain onto the Assay Document to display its results, you should note that some methods contain a dozen or more displays. To choose a subset of the available displays, click the expand triangle to the left of the method name in the Method Curtain. All options will be displayed as already selected. Click in the white space just to the left of the subtype to toggle between adding/removing the selection. Use **Shift+click** or **Ctrl/Cmd+click** to change multiple items.
Apply a method to the assay document	Use your mouse to drag the method from the Method Curtain to the assay document. Use **Shift+click** or **Ctrl/Cmd+click** to select multiple methods so that you can drag them as a group.
Format a method that has been applied to the assay document	Select it with the [object selector](#) then choose **Options > Line Color** (or **Line Weight**, **Fill Color**, or **Fill Pattern**). Use **Shift+click** or **Ctrl/Cmd+click** to select multiple methods so you can format them as a group.
Remove methods from the assay document	Use your mouse to drag the method back to the Method Curtain or select the method and select **Edit > Clear** or press **Backspace** (Win) or **Delete** (Mac). The method will be removed from the document but will remain in the Method Curtain.

Locate a method in the Method Curtain	Select the method on the Assay Document, then use **Edit > Find in Curtain**.
Locate a method on the Assay Document	Select the method in the Method Curtain, then use **Edit > Show on Display**.

Legend Curtain

Method legends appear on the far right of the sequence on the [Assay Document](#).

If you are [zoomed in](#), you won't be able to read the legends without scrolling to the right. The Legend Curtain gives you a way to see these labels without leaving the region of interest. There are three ways to open the Legend Curtain:

- Drag the Legend Curtain "pull ring" (the small circle in the upper right corner of the assay document screen) to the left.

- Select **Analysis > Show Legend**. This command toggles with **Analysis Hide Legend**.

- Press **Ctrl/Cmd+]**.

Toolbar

The GeneQuest toolbar consists of eight buttons arranged on the left side of the [Assay Document](#). They are used to manipulate the position and scale of methods and to examine particular regions in greater detail.

Tool	Name	Function
	Range Selector	Used to select a range of bases in the nucleotide sequence. When the range selector is active, the cursor appears as an arrow. When using the range selector, you can select a sub-range of the entire assay by dragging anywhere on the Assay Document, or by clicking distinct elements on region plots. Clicking a region plot will select the entire region, with the proper strand orientation. For example, selecting a region of a Borodovsky plot of the 4th, 5th, or 6th frame will select a region on the complement strand (but note that you cannot select the corresponding region if you display the data as a graph instead of a region plot). Alternatively, you can select a region on the complement strand by holding down the **Alt** key as you select the region. For all selections, the Assay Document header shows what region has been selected. The **Range Selector** creates a vertical selection, spanning horizontal coordinates for all methods on the Assay Document. When you click with the **Range Selector** on a discrete element, GeneQuest selects the precise sequence range corresponding to the region. When you **Shift+click** multiple such elements, GeneQuest selects the whole range of sequence between the left end of the first element and the right end of the last element. Contrasting highlighting marks the range you have selected. **Note:** If you select an ORF using the range selector, the terminator will be added to your selection, making GeneQuest consistent with NCBI and with other Lasergene® applications. If you wish, you may observe the terminator portion by making the selection, then using the **Zoom In** tool.
	Object Selector	Used to select and manipulate entire method displays or to do such things as change the color and pattern of a display. When the object selector is active, the cursor appears as a hand. If you would like to select multiple objects, hold down the Shift key while you select the desired objects using the mouse button. You may also use the **Object Selector** to drag method displays to new vertical locations, and to superimpose method displays. **Note:** You may use either the range selector or the object selector to double-click on any method display to launch the corresponding parameter dialog.
	Zoom In/ Out	After you have selected **Zoom In** or **Zoom Out**, the cursor changes to a magnifying glass. You may control zooming by using either a clicking or dragging technique. Each click expands or reduces the view by 150%, whereas dragging defines an area for zooming. For zooming in, dragging defines the area you want to expand to fill the window. For zooming out, the area defines the size the current view will shrink to. If you were to print the assay document, it would print at the zoomed scale. The number of pages needed is

		indicated in the top right of the document window. To scale the view to a set horizontal scale, select **Options > Horizontal Units**.
		Note: Some methods (like "Coding Prediction – Borodovsky":#coding-prediction-borodovsky) cannot provide useful information once you have zoomed out past a certain threshold. These methods temporarily disappear from the Assay Document until you zoom in again.
	Annotation Tool	Opens the Feature Editor dialog and allows you to create a feature or edit an existing feature. Highlight the area of interest with the **Range Selector** before using the annotation tool. The three panes that are accessed through the Feature Editor are" Location":#location-tab, "Description":#description-tab and "Style":#style-tab.
	Joining Tool	Joins the selected region as a new segment in an existing feature annotation. Select a range for the new segment using the range selector prior to choosing this feature. Once you have activated the Join Current Selection To dialog, select a feature to join to and click Join. Instead of using the joining tool, you may also select **Sites & Features > Join to Feature**.
	Microscope Tool	Displays the sequence for the region corresponding to the current horizontal location of the mouse pointer (whether **Range Selector**, **Object Selector** or **Zoom In/Out** tool). It doesn't matter which method, if any, is under the pointer. The **Microscope Tool** lets you stay zoomed out for the overall big picture, but still read the sequence. As the cursor's position changes over the Assay Document, GeneQuest updates the microscope window to display residues underneath. The microscope views have two display methods. The Method Plot View, activated by clicking the magnifying glass at the upper left corner, displays the nucleotide sequence and zoomed-in views of the graphs currently underneath the cursor. The Site & Sequence view, activated by clicking the restriction site icon (the lower of the two), shows both strands of the nucleotide sequence, all six translated reading frames and the restriction sites for every restriction enzyme within the Enzymes file. You can increase the size of the microscope window by dragging on its lower right corner.
	Amplitude Control Tool	Lets you globally define the amplitude for all graphs on the Assay Document. Drag the slide control upward to reduce the vertical area each graph occupies. Drag the slide control downward to increase the vertical area each graph occupies.

GeneQuest Methods

The following table lists methods you can drag onto the Assay Document, including many methods designed to find potential genes.

Task	Apply this method
Display the sequence on the document	Apply Sequence.
List known features for the sequence.	See Annotate Features.
Add a document title	Apply Title.
Add a ruler	Apply Ruler.
Plot the local frequency of particular bases	To plot the local frequency of the four bases and A+T and G+C pairings and the local strand skew of the complements AT and GC, apply Base Contents – Base Distribution.
Predict the bending of free DNA	To predict the bending of free DNA by showing localized angles of helix trajectory computed for a specified window size, apply Bent DNA – Bending Index.
Find genes using coding region recognition algorithms	Three methods use coding prediction algorithms to find potential genes. • To use Borodovsky's Markov method to identify potential coding regions, apply Coding Prediction – Borodovsky. • To use the Shannon information theory to find regions rich with information, apply Coding Prediction – Local Compositional Complexity. • Use open reading frames (ORFs), apply Coding Prediction – Starts Stops ORFs. This method locates and summarizes open reading frames longer than a user-specified minimum length. You can designate whether or not a start codon is required. Starts and stops are displayed independently for all frames.
Apply restriction	To create restriction maps using any of the enzymes from the DNASTAR enzymes catalog, apply Enzymes – Restriction Map.

enzymes to the sequence	
Find matches to a user-specified sequence file	To identify regions that match those of a user-specified DNA sequence file, use Gene Finding – DNA Finder. This method displays results separately for the top and bottom strands. To identify regions whose translation matches a user-specified protein sequence file, use Gene Finding – Protein Finder. This method's display combines results for all six reading frames.
Find genes using patterns	Three methods use patterns to locate potential genes: • To use log odds or positional frequency matrix patterns to locate starts, stops, and intron & exon splice sites, apply Patterns – Matrix. • To use a text-based pattern database of known DNA transcription factor sites, apply Patterns – Signal. • To use patterns you type in, apply Patterns – Type-In Pattern.
Find genes using repeats	Three methods use repeat locations to locate potential genes: • To locate dyad repeats and palindromes, apply Repeats – Dyad Repeats. • To locate direct repeats, apply Repeats – Direct Repeats. • To locate inverted repeats, apply Repeats – Inverted Repeats.

Base Contents – Base Distribution

This method examines the distribution of nucleotides and groups within a sequence. The overall patterns of nucleotide composition in the sequence can be described by the %GC, GC-skew and AT-skew. You can plot the frequency of single bases, base pairs and base skew.

The **Window** value is the number of bases over which the composition is calculated. For example, to calculate the frequency of C in a window of 100 bases, GeneQuest would count the number of Cs within the given 100 bases and divide the sum by 100. Values are plotted at the window center. Wider windows smooth out transient effects; narrower windows allow some small peaks to show through.

GeneQuest plots values for all nucleotides 5' to 3' on the sense strand. The complement is not used, nor are wobble frequencies taken into account.

To change method parameters:

1. Double click on the method name in the Method Curtain; or select the method display and then **Analysis > Method Parameters**.

2. The default **Window** size, 100 bases, is optimal for sequences averaging about 2000 bp. Increase this number for larger sequences and decrease it for smaller sequences.

3. Change the significance thresholds, if necessary. The significance thresholds for the bases have default settings of: A=0.233, C=0.253, G=0.272, T=0.242. These thresholds were chosen based on global frequency values for coding regions of the *Escherichia coli* genome.

Bent DNA – Bending Index

This method predicts the bending of free DNA. If short stretches of a particular bent sequence repeat with a period close to the 10.5-bp pitch of the B-DNA double helix, it follows that local bending of a helix will sum to produce a persistent curvature of the whole sequence. According to Trifonov and Sussman (1980), certain dinucleotides, such as AA, when repeated with a period close to 10.5 bp, form a curvature in DNA as a result of a "wedge" angle inherent in AA dinucleotides.

Bent DNA – Bending Index computes the helical trajectory over the arc length. At the same time, the angle formed by two vectors parallel to the helical midpoint at the ends of the arc is measured. Using an arc length that exceeds the helical periodicity eliminates this periodicity bias for longer DNA segments. The method makes no attempt to predict the super coil structure of DNA. Nor does it take into account involvement of DNA-binding proteins on chromatin structure. **Bent DNA – Bending Index** should be used with the assumption that the DNA is free of proteins.

To change method parameters:

1. Double click on the method name in the Method Curtain or select the method display and then **Analysis > Method Parameters**.

2. Specify **Arc Length**, the window size used for each trajectory point determination.

3. Specify **Threshold**, the minimum value for calling a region "bent," using values ranging from $0°$ to $180°$.

Coding Prediction – Borodovsky

Coding Prediction – Borodovsky plots are used to evaluate coding potential independent of identifying ORFs. This method, located in the More Methods section, finds potential coding regions by analyzing the sequence against data contained in statistical matrix files. Combining this method with Coding Prediction – Starts Stops ORFs is particularly informative.

The method is demonstrated in the following video.

See Installed Lasergene file locations for the location of the Borodovsky matrix files after a normal Lasergene installation.

You should choose which Borodovsky matrix file to use based on the organism from which the DNA was obtained. Files are further differentiated by model order, which is the maximum mer length included in the matrix. For eukaryotes, separate matrix files are provided for sequences with different ranges of GC content.

The naming convention for Borodovsky matrix files makes it easy to choose an appropriate file. The first few characters of the file name denote the organism, the middle characters describe GC content (high, low, or a percentage) and the last digit tells the order of the mer. It is usually best to start with the largest mer size available because the method will be more selective of coding regions. If results appear as flat lines (no peaks or bars), just switch to a matrix with a smaller order and try again. For organisms with less sequence data available, you may have no choice but to use lower order matrix files.

For a list of species names that correspond to the Borodovsky matrix files, see the file *Borodovsky Species Names.txt* also located in the Borodovsky Matrices folder. (See above for the directory path.)

To generate a summary table of codon usage for any selected segment of the sequence, which may also be

useful in identifying coding regions, choose **Analysis > Codon Usage**.

Results of the Borodovsky method may be displayed as line graphs and/or region plots, with separate plots for each of the six frames. The algorithm used to generate Borodovsky plots is Lukashin and Borodovsky (1998). A potential coding region will appear as a bar in a region plot or as a sharply rising, flat topped peak in a line graph. Coding regions will have a value over the threshold of 0.5, and preferably, near 1.

Tips for using the Borodovsky method:

- Before analyzing the Borodovsky graphs, it often helps to color-code each of the reading frames. For instance, you could choose red for reading frame one for both ORFs and Borodovsky plots, blue for frame 2 and green for frame 3 etc.

- It may also be helpful to superimpose the three Borodovsky plots for each strand by dragging them with the mouse or by highlighting them and selecting **Options > Superimpose Objects**. One layout that works well is to position the superimposed Borodovsky plots just below the corresponding ORFs plots for the same reading frames. You may then quickly establish whether any Borodovsky peak corresponds to an ORF in the same frame, as it should in genuine coding regions.

- In addition to correlating Borodovsky peaks with ORFs in the same frame, you may seek evidence for the starts of genes. The Coding Prediction – Starts Stops ORFs method allows you to pinpoint all the ATG triplets in all six frames.

- More powerfully, the Patterns – Matrix method identifies the subset of ATG triplets that are in a sequence context typical of start sites in specific organisms. For the **Patterns – Matrix** method, select the appropriate start site matrix pattern for your organism, and then apply region plots for the start sites for both strands.

- If you are working with a eukaryote, you might continue your evaluation of ORFs by locating candidate splice sites (see Look for exons).

To specify the default Borodovsky matrix:

1. Select **Edit > Preferences**.

2. Choose the **GeneQuest** tab.
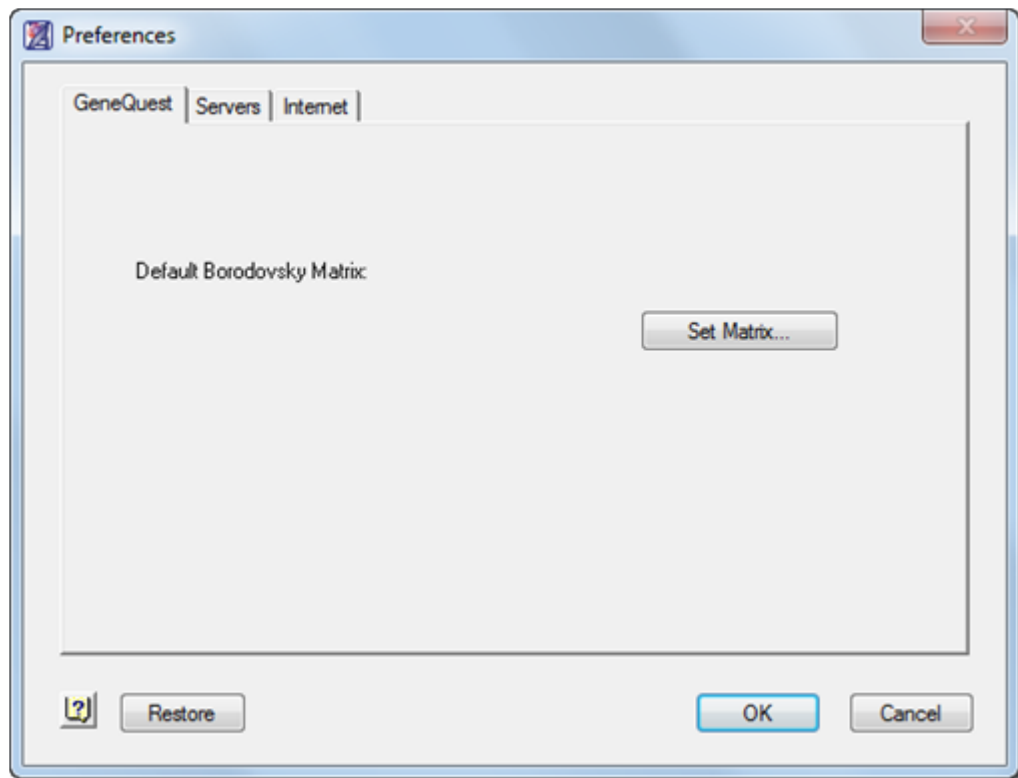
3. Click the Set Matrix button and navigate to the location.

To change method parameters, including specifying the default Borodovsky matrix:

1. Double click on the method name in the Method Curtain or select the method display and then **Analysis > Method Parameters**.

2. Specify **Context**, the window size used in the prediction. The number you choose should be a multiple of three.

    • In most cases, the default (96 bases) is sufficient.

    • Smaller values produce more false positives but also increase the chance of finding small exons.

    • Higher values may miss small exons but clarify the signal for coding in longer exons.

    • You may restore the default window size, if desired, by pressing the **Set to Defaults** button.

3. Select the desired Borodovsky matrix file by clicking **Set Data File**. GeneQuest comes with a set of matrix files from over two dozen genomes and can use new matrices created for GeneMark. If you use a particular matrix often, you may want to set it as your default Borodovsky matrix by clicking **Use as Default**.

Coding Prediction – Local Compositional Complexity

This method, located in the More Methods section, identifies regions rich in information. These regions often correlate with coding regions and regulatory elements. Results can be presented as line graphs and/or region plots. Information content is based upon the Shannon information theory formula.

$$H = -\Sigma\ p(i)\ \log\ p(i)$$

Where:

i = {A G T C}

$p(i) = N_i/L$

N_i = count of base i in the context length and L = context length.

The method provides a quick way to evaluate the information content of a region of DNA. While it is not reliable for distinguishing coding from non-coding sequence, it can distinguish random or repetitive DNA from biologically interesting DNA.

To change method parameters:

1. Double click on the method name in the Method Curtain; or select the method display and then **Analysis > Method Parameters**.

2. Specify **Context**. This field determines the value of "L," the number of bases examined at a single time.

3. Decide whether or not to select **Smoothing Window**. This feature averages the context peaks over the specified range.

4. Specify **Threshold**, the level at which a line plot is called informative or not. In most cases, the default values are appropriate.

You may wish to raise or lower **Context** or **Smoothing Window** depending upon the length of your sequence and the preponderance of repetitive sequence within the region. Low values may distinguish repetitive from non-repetitive DNA more accurately, but at the cost of increased noise.

To generate a summary table of codon usage for any selected segment of the sequence, which may also be useful in identifying coding regions, choose **Analysis > Codon Usage**.

Coding Prediction – Starts Stops ORFs

One way to locate potential coding regions is to identify and to apply Open Reading Frame (ORF) plots for all six reading frames to your sequence. Although only a subset of ORFs are genuine coding regions, knowing their locations is a helpful adjunct to application of the more powerful methods of gene finding.

> ✸ **Note:** ORF detection can also be done in SeqBuilder Pro, our modern application that is installed alongside GeneQuest. To learn about finding and displaying ORFs in SeqBuilder Pro, click here.

The Coding Prediction – Starts Stops ORFs method, located in the **More Methods** section, predicts coding regions in the traditional way, by locating start codons and stop codons. The region between these codons is defined as the open reading frame (ORF). Results are presented as a series of region plots for each reading frame. Plots for starts, stops and ORFs are displayed individually. You can help define the boundaries of coding regions by displaying this method alongside (i) Borodovsky peaks for each frame and (ii) (in eukaryotes) candidate splice sites (see Patterns – Matrix).

Start and stop codons are determined by the active genetic code. Before applying the **Coding Prediction > Starts Stops ORFs** method to your sequence, you may select the appropriate code from the **Options > Genetic Codes** list. The code can be edited, if necessary, by selecting **Options > Edit Selected Code**.

If you choose to display an ORF region plot (a type of graph with bars representing "yes" areas and lines representing "no" areas), the plot will not include the terminator. If you select the ORF using the range selector, however, the terminator is added, making GeneQuest consistent with NCBI and with other Lasergene applications. If you wish, you may observe the terminator portion by making the selection, then using the **Zoom In** palette tool.

To change method parameters:

1. Double click on the method name in the Method Curtain; or select the method display and then **Analysis > Method Parameters**.

2. Specify the **Minimum ORF Length**, the minimum number of codons required to call a region an ORF. The shortest exons rarely fall under 25 codons (75 nucleotides).

3. Decide whether or not to check **ORFs Must have a Start Codon**.

    • When checked, this forces GeneQuest to begin an ORF with a start codon and end with a stop codon.

    • When unchecked, an ORF begins immediately after a stop and runs until the next stop codon, assuming there are at least 75 bp between stops.

To generate a summary table of codon usage for any selected segment of the sequence, which may also be useful in identifying coding regions, choose **Analysis > Codon Usage**.

To investigate further, consider using the [Borodovsky method for evaluating coding potential](#).

Enzymes – Restriction Map

This method identifies restriction sites and displays them as a mini-map, showing recognition sites as vertical bars. GeneQuest can also superimpose mini-maps to show restriction patterns from multiple digests. To view information about a particular enzyme, double click on its name in the Method Curtain or on its map display on the Assay Document to launch a text box that includes a list of isoschizomers, the enzyme recognition sequence, cut sites, site complexity rating, class (precise, random or unknown cutter), and vendor information, including cost/unit.

> ✳ **Note:** Restriction map functionality and useability is greatly improved in SeqBuilder Pro, our modern application that is installed alongside GeneQuest. To learn about working with enzymes in SeqBuilder Pro, click here.

For information on using this method, including a video demonstration, see Analyze restriction sites.

To change method parameters:

1. Double click on the method name in the Method Curtain; or select the method display and then **Analysis > Method Parameters**.

2. Choose whether you want to treat the sequence as **Linear** or **Circular**.

3. Check **Examine Context at Ends of Sub-range** if you are using a sequence sub-range for your assay. If checked, GeneQuest will examine 25bp upstream and downstream beyond your chosen coordinates. This will determine if any recognition sites located outside the range cut bases inside it.

4. Check **Report Sites that Might Cut Degenerate Sequence** to display sites with degenerate bases potentially cut by a given endonuclease. For example, Acc65I, which recognizes GGTACC, will recognize the degenerate sequence GRTACC only if this box is checked.

5. Choose a **Selector** from the list box to use a subset of restriction sites from an existing enzyme file. Selectors are created using DNASTAR's SeqBuilder Pro application. Each selector will identify only the specific subset of sites specified by the selector.

6. Choose **Show Filter Result** from the **Selector** list to open a window that displays picked and remaining enzymes in the checked selector. You may double-click the name of any restriction site in that window or in the Method Curtain to view information regarding the enzyme's recognition sequence, cut site, complexity, and other properties.

Features

GeneQuest reads and interprets features from the comment or feature panes of GenBank-formatted sequence files and displays them in the Assay Document. Once you have discovered areas of interest in the sequence, you may create your own annotations and link several features together under a common title and description. For information on creating your own features, see Annotate Features.

Gene Finding – DNA Finder

This method, located in the More Methods section, finds DNA regions which match those in the sequence file(s) you specify. GeneQuest uses the selected files to search both the top and bottom strands of your sequence.

To apply the method:

1. Double click on the method name (**Gene Finding – DNA Finder**) in the Method Curtain to open the Add Sequences dialog.

2. Locate and select the appropriate sequence files. Note that you may select multiple sequences, and of a variety of file types, including Lasergene *.seq files*, Lasergene project files from SeqBuilder Pro or GeneQuest, FASTA files, ABI or SCF trace files, SFF flowgram files, Phred files *.phd*), zip archives, DNA Multi-Seq Files (*.mseq*), and File of Filenames (*.fof*). Click **Done**.

The top and bottom strand for each sequence you selected will now appear in the Method Curtain below the **Gene Finding** method. If you selected multiple sequences, a **Composite** top and bottom strand will also be listed. The **Composite** strands are created by superimposing all of the top/bottom strands for your specified sequences.
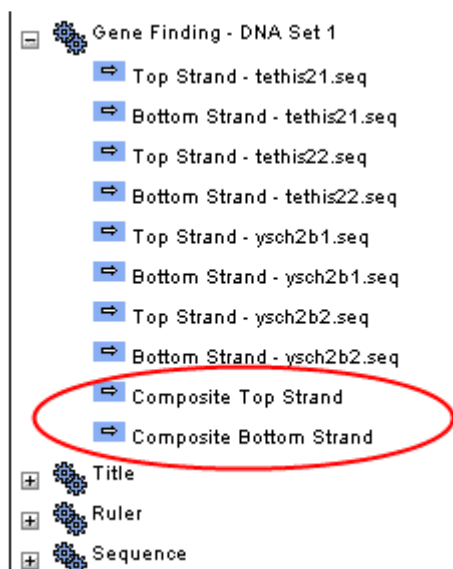
A hit is recorded when the match reaches or exceeds the minimum match size parameter (see below). Gap characters are not inserted in the identical region to extend an alignment. With this method, you can locate multiple matches for a sequence, as long as the areas are at least as long as the minimum match size.

The score printed on the region plot indicates how many bases in the two sequences matched up. Using the **Zoom In** palette tool, zoom in until you see the numbers. Clicking once on the plot and holding the mouse button reveals the matching coordinates of the two sequence stretches and tells whether the match occurred on the top or bottom strand (see Customize the look of the Assay Document). If you wish to see

the residues for the match, continue using the **Zoom In** tool until the sequence is displayed.

To change method parameters:

1. Double click on the Gene Finding method name in the Method Curtain or select the method display and then **Analysis > Method Parameters**.
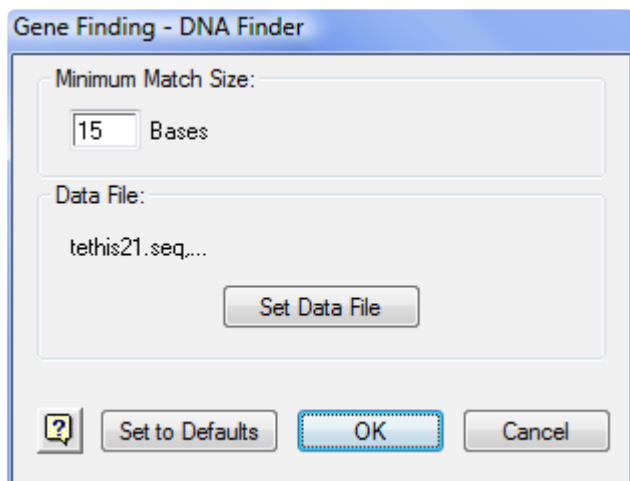
2. Specify the Minimum Match Size, the minimum number of identical bases required to define a match. Note that GeneQuest will not accept a value less than 12 for this parameter.

3. To choose different files, click Set Data File to return to the Add Sequences dialog.

Add Sequences dialog

The Add Sequences dialog allows you to select multiple sequences for use with the Gene Finding – DNA Finder method. Access the Add Sequences dialog by double clicking on **Gene Finding – DNA Finder** in the Method Curtain.
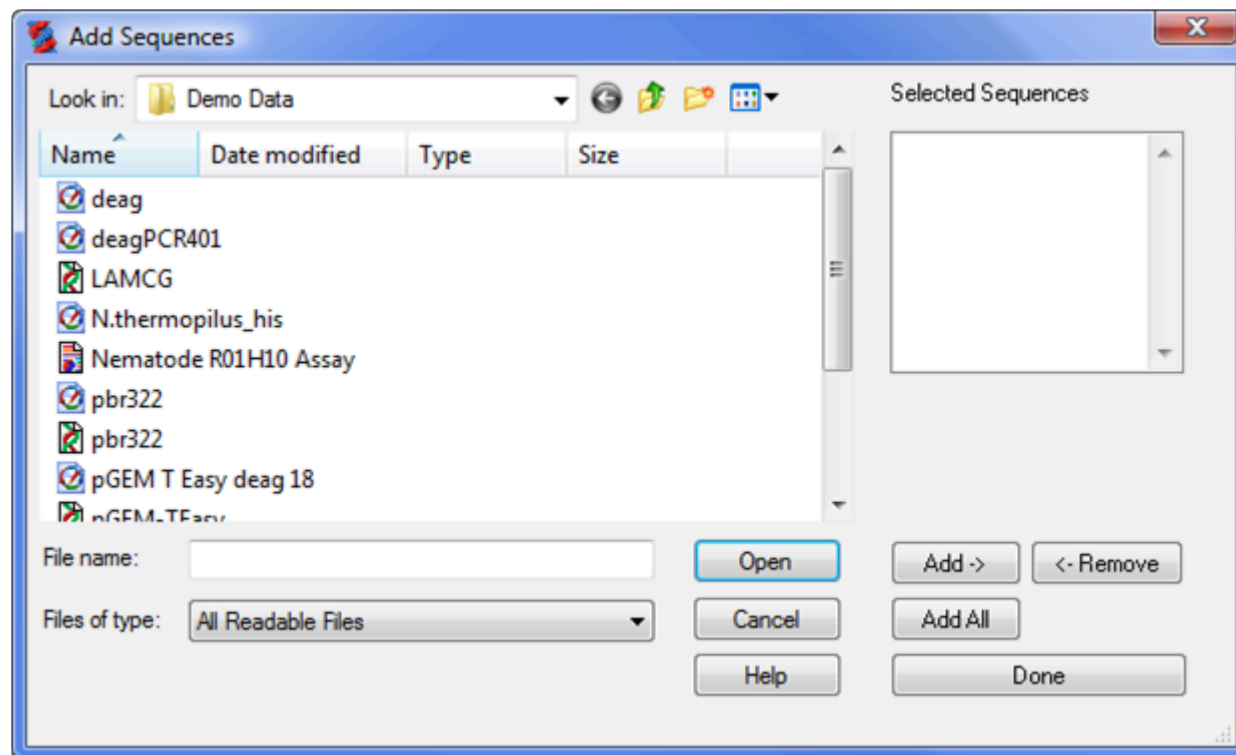
1. Locate and select the desired file(s).

2. Click **Open** to move the selected sequence(s) to the **Selected Sequences** (Win) or **Chosen Files** (Mac) list.

3. You may also use the following buttons to add sequences to the **Selected Sequences/Chosen Files** list.

    •   To add a single sequence, select its name from the top/left list and then click **Add** (Win) or **Add File** (Mac).

    •   (Windows only) Select multiple sequences via **Ctrl+click** or **Shift+click**, and then click **Add**.

    •   (Macintosh only) Select the folder in the top window and click the **Add Folder** button.

    •   To add all the sequences in a folder, click **Add All**.

    •   To delete a sequence from the **Selected Sequences/Chosen Files** list, select its name and click **Remove**.

4. When all the desired sequences are listed in **Selected Sequences/Chosen Files**, click **Done**.

Gene Finding – Protein Finder

This method, located in the More Methods section, finds protein regions which can code for a given Lasergene format protein sequence. GeneQuest uses the file you select to search all six reading frames of the nucleotide sequence.

To apply the method:

1. Double click on the method name in the Method Curtain to open the Protein Sequence dialog.

2. Locate the appropriate sequence file and highlight it. Acceptable sequence formats are Lasergene protein files (.pro) or text files (.txt).

3. Click **Open**.

A hit is recorded when the match reaches or exceeds the minimum match size parameter. Gap characters are not inserted to extend an alignment. With this method, you can locate multiple potential coding areas for a specified sequence, as long as the areas are at least as long as the minimum match size. This method can help locate introns and exons within a stretch of DNA.

The score printed on the region plot indicates how many nucleotides matched the protein. Use the **Zoom In** tool to enlarge the display until you see the numbers. If you want to know more about a particular match, select it using the object or range selector and continue to hold down the mouse button. GeneQuest will display the matching coordinates of gene and protein, the reading frame of the match, and the twenty residues at the 5' and 3' ends (see Customize the Display).

To change method parameters:

1. Double click on the method name in the Method Curtain or select the method display and then **Analysis > Method Parameters**.

2. Specify the **Minimum Match Size**, the minimum number of identical codons required to define a match.

3. To choose a different file, click **Set Data File** to return to the Protein Sequence dialog.

Patterns – Matrix

This method, located in the More Methods section, finds sequence features such as start sites, splice acceptor sites and splice donor sites. The matrix patterns are defined in separate data files, which are normally installed in the Matrix Patterns folder, found in the following location:

- **Windows:** *C:\Users\Public\Public Documents\DNASTAR\Lasergene 'x' Data*

- **Macintosh:** *Hard Drive:Applications:DNASTAR:Lasergene 'x'*

Whereas the start codons identified by the codon prediction method are simply in-frame ATGs, the matrix pattern method takes into account local sequence context with the aim of predicting the subset of ATGs that function as start codons.

The naming convention for matrix pattern files makes it easy to choose an appropriate file. The first few characters of the file name denote the organism; the middle characters describe whether the pattern searches for acceptor splices (as), donor splices (ds) or start sites (atg); and the last digit tells whether the pattern looks for frequencies of nucleotides (1) or dinucleotides (2). Dinucleotide pattern files are less likely to generate false positive results. However, if you fail to locate the patterns you seek using a dinucleotide pattern file, you may observe candidates by using the corresponding mononucleotide pattern file. Also, while the pattern files tend to be species-specific, if there is no pattern file for your organism, you may be able to obtain reasonable results using a pattern file from a similar species.

Note: For a list of species names that correspond to the matrix pattern files, see the file *Species Names.txt* also located in the *Matrix Patterns* folder. (See above for the directory path.)

For additional information on using this method, along with a video demonstration, see [Look for exons](#).

To apply the method:

1. Double-click on the method name in the Method Curtain to open the Matrix Pattern dialog.

2. Locate the appropriate matrix pattern file and highlight it.

3. Click **Open**.

To change method parameters:

1. Choose a matrix pattern by performing steps 1-3, above.

2. Double click on the method name in the Method Curtain; or select the method display, then select **Analysis > Method Parameters**.

3. Specify **Threshold**, taken as half of the average of the log-odds scores of sequences that were used to train the pattern. Increase the threshold for more stringency, decrease for more matches.

4. To select a different pattern, click **Select Pattern File** to return to the Matrix Pattern dialog.

Patterns – Signal

This method, located in the **More Methods section**, finds transcription factor binding sites using species-specific sequence patterns. To enable you to search your sequence for known transcription factor binding sites, the *TFD.dat* file is provided. The data in this file is summarized in the file *explanation.html*.

The following video shows the method in action.

To apply the method:

1. Double-click on the method name in the Method Curtain to open the Transcription Factor dialog.

2. Locate the appropriate transcription factor file and highlight it.

3. Click **Open**.

This method does not have any parameters that can be adjusted.

To view the sequence pattern for a particular binding site:

Hold the mouse pointer down on one of the sites displayed on the Assay Document; in a second or two, a text box displays the name of the site, the sequence that was matched, the sequence coordinates, and whether the match to the site is on the top or bottom strand.

To search for one specific transcription factor binding site:

Type in the sequence pattern for the site using the Patterns – Type-in method.

To create your own database of transcription factor binding (or other) sites:

Create a fixed-width text file with the following format and content:

Column	1	2	3	4	5
Width	16	21	31	3	7
Content	gene name	site name	sequence pattern	number of bases	index

For example, one row in the in the fixed-width text file, would appear as:

GR	GRE_CS5	SGGTWCAANNTGTYCT	16	S00117

The "width" is the field width in characters. The text string for the field should be a maximum of one character fewer than the width, and the last character must be a space. Where text strings are shorter than (width – 1), fill in the fixed-width field with spaces.

Look up information about the Transcription Factors Database:

Use **Sites & Features > Look Up Site**. The command is only enabled after you add the **Patterns – Signal** method to the curtain, then apply the method to the sequence. The sub-menu item **Description** looks up the site information on the Ghosh database, hosted on the DNASTAR website. The sub-menu choice **Abstract** calls up the reference abstract at NCBI for the paper in which the transcription factor was published. Both choices open a web browser, and multiple selections will open multiple browser windows.

Patterns – Type-In

This method, located in the More Methods section, finds transcription factor binding sites using species-specific sequence patterns. To enable you to search your sequence for known transcription factor binding sites, the *TFD.dat* file is provided. The data in this file is summarized in *File Explanation.html*.

To apply the method:

1. Double click on the method name in the Method Curtain.

2. Type in a **Name** for the pattern.

3. Type in the **Pattern** itself in the appropriate fields.

    • To incorporate inclusive expressions in your pattern, use square brackets around whichever bases are acceptable. Example: "[CG]" means C or G is acceptable.

    • For exclusive expressions, use curly brackets around whichever bases are unacceptable. Example: "{AT} " means neither A nor T is acceptable.

    • To avoid typing repeated expressions, type the base, followed by the number of repeats in parentheses. Example: "AG(3)T" would search for the string AGGGT;.

    • For gapped expressions, type in the symbol X followed by the acceptable maximum number of gaps. Example: "AX(1,3)G" would recognize the following sequence strings: AXG, AXXG, AXXXG.

4. Set the **Threshold** to specify how many mismatches you will permit.

    • This method will not gap a pattern unless it is part of the syntax.

    • If you have used ambiguous characters or inclusive expressions, keep the **Threshold** high to avoid false matches.

To change method parameters:

1. Double click on the method name in the Method Curtain or select the method display

2. Select **Analysis > Method Parameters**. This returns you to the Type-In Pattern dialog.

Repeats – Direct

This method, located in the More Methods section, finds all direct repeats that exceed a specified length and are located within a defined proximity. This proximity is defined in the method parameters. A direct repeat exists where one region of DNA is followed by an identical region shortly thereafter. An example is

TAGCGTCCAgaaaatgga…gtcaattctcTAGCGTCCA

To change method parameters:

1. Double click on the method name in the Method Curtain or select the method display and then **Analysis > Method Parameters**.

2. Type the **Minimum Length** for your repeats in the appropriate window.

3. Click **No Limits** if you don't want to limit the distance between direct repeats or click **Less Than** to open a text box where you can specify a maximum repeat distance.

Note that this method searches only one strand. The complement of any direct repeats found occurs on the opposite strand.

To view a tabular results report that categorizes repeats by both type and length:

Choose **Analysis > Repeat Summary**.

Repeats – Dyad

This method locates all dyad repeats that exceed a specified length and are located within a defined proximity. This proximity is defined in the method parameters. A dyad repeat exists when one region of DNA is followed shortly thereafter by its mirror image. An example is:

TAGCGA…AGCGAT

To change method parameters:

1. Double click on the method name in the Method Curtain; or select the method display and then Analysis > Method Parameters.

2. Type the Minimum Length for your repeats in the appropriate window.

3. Click No Limits if you don't want to limit the distance between dyad repeats or click Less Than to open a text box where you can specify a maximum repeat distance.

Note that this method searches only one strand. The complement of any dyad repeats found occurs on the opposite strand.

To view a tabular results report that categorizes repeats by both type and length:

Choose **Analysis > Repeat Summary**.

Repeats – Inverted

This method, located in the **More Methods** section, finds all inverted repeats that exceed a specified length and are located within a defined proximity. This proximity is defined in the method parameters. An inverted repeat exists where one region of DNA is followed by its complement shortly thereafter. An example is TGCCATGCGCATGGCA.

To change method parameters:

1. Double click on the method name in the Method Curtain; or select the method display, then select **Analysis > Method Parameters**.

2. Type the **Minimum Stem Length** for the repeats in the appropriate window.

3. Click **No Limits** if you don't want to limit the distance between inverted repeats or click **Less Than** to open a text box where you can specify a maximum repeat distance.

Note that this method searches only one strand. The complement of any inverted repeats found occurs.

To view a tabular results report that categorizes repeats by both type and length:

Use **Analysis > Repeat Summary**.

Ruler

The **Ruler** method, initially located in the **More Methods** menu, can be used to add a ruler to the Assay Document. The ruler's appearance is determined by following the steps below. You may add more than one ruler to your document by moving multiple copies of the **Ruler** method into the Method Curtain and from there to the Assay Document. Note that if you simply place a single ruler method at multiple locations on the Assay Document, each ruler will be identical.

1. Apply the **Ruler** method to the Assay Document .

2. In the assay document, double-click on the **Ruler** method.

3. Click **Stroke Up** or **Stroke Down** to specify the placement of ticks.

4. If you want to use GeneQuest's coordinates, check **Set Divisions Automatically**. To set your own coordinate system, uncheck the box and enter a value in the **Number Every** text box.

5. Click **OK**.

Sequence

The **Sequence** method, initially located in the **More Methods** menu, adds the DNA sequence to the Assay Document. You may add more than one sequence to your document by moving multiple copies of the **Sequence** method into the [Method Curtain](#), and then to the Assay Document.

1. Apply the **Sequence** method to the Assay Document.

2. In the Assay Document, double-click on the **Sequence** method.

3. (optional) If you want to temporarily suppress the sequence (e.g., if there are too many horizontal units per page) enter zero or another low value for **Suppress When Bases Overlap More Than**. This will hide a sequence in circumstances under which it will be illegible.

4. Click **OK**.

> ✳ **Note:** If you are zoomed in to the point where the overlap threshold is exceeded, the sequence will be displayed as a solid black bar. Zooming in will cause the sequence to reappear.

Title

The **Title** method, initially located in the **More Methods** menu, adds a title to the Assay Document. Text for the title is entered by following the steps below. You may add more than one title to your document by moving multiple copies of the **Title** method into the Method Curtain. You may change the text and other parameters for each **Title** method before or after applying the methods to the Assay Document. Note that if you simply apply a single title method to multiple locations on the Assay Document, each title will be identical.

1. Apply the **Title** method to the Assay Document.

2. In the assay document, double-click on the **Title** method.

3. Enter a title in the **Title Display** field.

4. Select one of the three **Justify** buttons to set the horizontal position.

5. To establish a sequence-specific position (optional), check **Over Residue** and enter a coordinate for the beginning of the title.

6. Click **OK**.

Work with Methods

For information on working with the GeneQuest methods, see the following topics:

- Look for exons
- Analyze restriction sites
- Change method parameters
- Compare results from different methods
- Compare results for the same method, but using different parameters
- Apply a saved set of methods to the project

The following video shows some of these methods in action:

Change method parameters

To change these method parameters	Use this command
Analysis > Fold As RNA	Use **Options > Folding Parameters**.
Analysis > Repeat Summary	Use **Options > Repeat Summary Options**.
Analysis > Base Composition	Use **Options > Base Composition Settings**.
Sites & Features > Restriction Fragment Summary	Use **Options > Fragment Summary Options**.
Methods on the Assay Document	Use the **Object selector** to select the method result (graph, plot, etc.) from the Assay Document or its name from the Method Curtain, and then choose **Analysis > Method Parameters**. Method parameters may also be changed by double clicking on the method display or on its name in the Method Curtain.

Compare results from different methods

Since all GeneQuest methods share the same horizontal axis, nucleotide positions are comparable across analysis methods. It is easy to juxtapose and superimpose graphs and plots in order to compare results from different methods.

To juxtapose two results, drag one beside the other with the object selector. To superimpose multiple results, select two or more displays, and then select **Options > Superimpose Objects**. Alternatively, you may simply drag the methods on top of one another using the mouse. When two or more graphs are superimposed, GeneQuest aligns them vertically by their threshold lines, shown at the far left of each graph on the Assay Document. Superimposed graphs will be appropriately offset vertically based on these threshold values. The method at the top of the stack has its title indented furthest right, while the title of each layer down in the stack is slightly less indented.

To separate methods that were previously superimposed, use **Options > Expand Objects** or drag the methods apart using your mouse.

To move a method display, select the object selector to select the display and then drag it to a new location. If you don't want to move similar methods near each other for comparison, you can still view them side by side using the [pane divider](#).

Later, if you wish to separate methods that you previously superimposed, simply select the superimposed methods and choose **Options > Expand Objects**.

Compare results for the same method, but using different parameters

You may compare the results of a method using different parameter settings. This requires that at least two separate copies of the method first be applied to the Assay Document. The steps are as follows:

1. Move two or more separate copies of the method from the **More Methods** menu into the Method Curtain.

2. Apply each copy to the Assay Document by dragging it and dropping it.

3. Double-click on each method and set the parameters as desired.

4. (optional) For ease of comparison, move the methods next to one another or superimpose them. See Compare results from different methods.

Apply a saved set of methods to the project

When you create a new project, GeneQuest selects a file called the default method outline (named *default.dao* on Windows and *Default DNA Outline* on Macintosh) to determine which methods to place in the Method Curtain, which of those to apply to the Assay Document, and how they will be decorated.

In the course of revising your project, you may decide to delete some of the default methods, add different methods, and optimize parameters and display options. GeneQuest allows you to save your optimized method outline and apply it to selected Assay Documents or as the default outline for all new Assay Documents.

Task	How To
Save the current method setup to apply to new projects by default	Choose **Analysis > Save as Default Method Outline**. The method outline is saved as part of a project file. If you save a GeneQuest project, close it, and reopen it later, it will look as it did when you saved it.
Save the current method setup to apply to new projects on demand	Choose **Analysis > Save Method Outline**.
Apply a method setup that you saved earlier	Choose **Analysis > Apply Method Outline** and select the desired *.dao* file.
Apply the same methods you used in another GeneQuest document	Choose **Analysis > Apply Method Outline** and select the GeneQuest document.
Apply the original GeneQuest default method outline	Choose **Analysis > Apply Method Outline** and open *C:\Users\Public\Documents\ DNASTAR\Lasergene x Data\Templates\Default 8 DNA Outline.dao* (Windows) or *DNASTAR:Lasergene x Data:Templates: Default 8 DNA Outline.dao* (Mac).

Look for Exons

If you are working with a eukaryote, you may wish to see if two or more neighboring exons are part of the same gene. If this is the case, you should be able to locate a donor splice at the 3' end of one exon and an acceptor splice at the 5' end of the next. You can search for acceptor and donor splices using the [Patterns – Matrix](#) method. Matrix pattern files are available in both "donor" and "acceptor" varieties for a number of organisms. This is the same method used for predicting start sites, except that distinct matrix pattern files are used for start sites vs. donor splice sites vs. acceptor splice sites.

To jump to sites containing the highly conserved core residues of the splice pattern:

First select the range of the sequence you want to work in. Then choose **Edit > Move Selection**, **Edit > Find/Move Next** and **Edit > Find/Move Previous** to jump through the sites. Once you have located splice sites, you may annotate the feature and join related exons together to form a putative gene.

To confirm the annotated feature has no internal stop codons:

Select the feature and choose **Analysis > Codon Usage**. A table summarizing the frequency of occurrence of all codons in the selected feature appears.

To seek evidence for the terminus of a coding region:

Do any of the following:

- Seek stop codons.

- Correlate the falloff in the last of a series of Borodovsky peaks with the stop codon for the corresponding ORF. In eukaryotes, you may also seek an ORF and corresponding Borodovsky peak for which there are no compelling splice sites available for a subsequent exon.

- Use the [Patterns – Type-in Pattern](#) method to seek nearby potential polyadenylation signals on the appropriate strand.

The next step is to attempt to ascertain the identity of your gene via a [BLAST search](#).

To align exons to a reference sequence for corroboration:

This procedure is demonstrated in the following video:

Analyze Restriction Sites

To create restriction maps in GeneQuest, apply the Enzymes – Restriction Map method to the Assay Document. Use the **Microscope** tool to view restriction sites on an expanded portion of the sequence.

The following video shows the method in use:

Once you have applied the **Enzymes – Restriction Map** method, you can do any of the following:

Task	How To
Display a table of restriction analysis results for **Enzymes – Restriction Map** methods applied to the Assay Document	Choose **Sites & Features > Restriction Fragment Summary**. The restriction enzymes selected from the Assay Document will appear in the tabular summary. If none are selected, then all will appear in the tabular summary. The summary lists the following data for single, double, and triple digests; the number and position of cuts; the enzyme responsible for each cut; and the length of the fragment. To change the parameters for this method, use **Options > Fragment Summary Options**.
View the predicted electrophoretic separation of restriction fragments	Hold down the **Shift** key while you use the object selector to select the desired enzyme plots. When the appropriate enzymes have been highlighted, choose **Sites & Features > Agarose Gel Simulation**.

While viewing the simulated gel, you may move the cursor over any fragment to display a header showing its range, size, and the restriction enzymes cutting each end. The far left and |

through an agarose gel	far right lanes are molecular size standards from bacteriophage lambda. The other lanes show all possible combinations of single, double and triple digests. Above each lane is the name of the enzyme(s) or molecular size standard. If you move the cursor over a fragment in the simulated gel, the header displays its range, size, and the restriction enzymes cutting each end. Fragments in red represent fragments within the current range selected in the assay document. You can change the selected range by double clicking on any DNA fragment in the gel simulation.
Toggle between treating a sequence as linear or circular	Choose **Sites & Features > Linearize Sequence** to treat the sequence as linear. Choose the command again to treat the sequence as circular. The ability to linearize sequences is especially important in restriction site analysis, since GeneQuest may find fewer cut sites on a linear strand than on a circular one.
Toggle between searching for restriction sites inside or both inside & outside the specified sequence sub-range	Choose **Sites & Features > Sub-range Context** to search both inside and outside the specified sequence subrange. Use the command again to restrict the search to the portion within the subrange. Note that this option does nothing unless you have previously used **Set Ends** to specify a sub-range of a sequence.
Toggle between including or not including ambiguous DNA residues when searching for restriction sites	Choose **Sites & Features > Include Uncertain Sites** to include ambiguous bases when searching for restriction sites. Use the command again to look only for literal matches. **Example:** When this option is active (i.e. checked), the sequence ANGCTT, an uncertain site for HindIII (AAGCTT), would be labeled as a cut site.

> ✱ **Note**: DNASTAR's SeqBuilder Pro application maintains a list of restriction enzymes that is used in all Lasergene applications, including GeneQuest. Refer to the SeqBuilder Pro User Guide for information on updating, deleting, or modifying enzymes or enzyme selectors. That User Guide also describes how to access restriction enzymes with unusual cleavage properties.
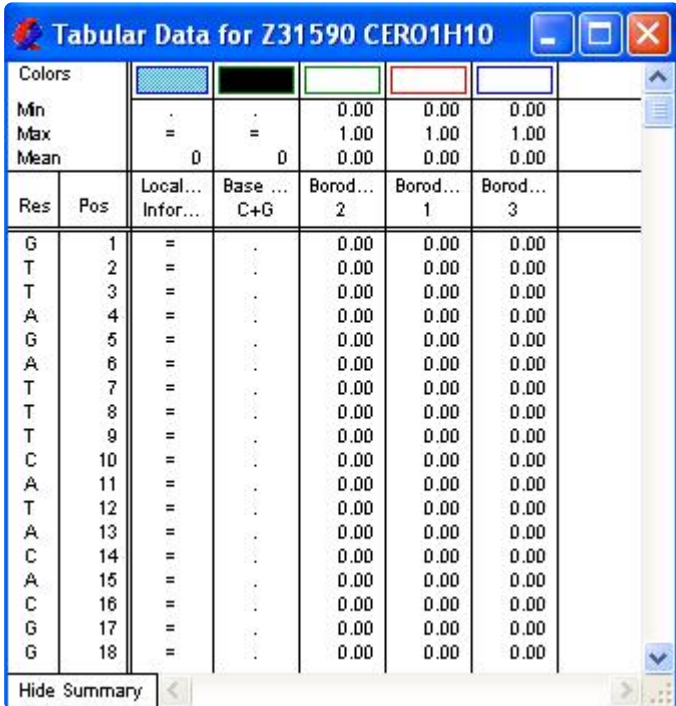
Search the Sequence

The cursor may be moved by using the mouse, or by selecting one of several cursor placement commands from the menus, as described in the table below. Refer to the assay document header to learn the precise position of the cursor.

Task	How To
Go to a numerical location on the sequence	Choose **Edit > Go to Location** or press **Ctrl/Cmd+J**. Type the desired position or range into the box in the dialog. Inclusive ranges for the top strand are typed with the symbol > between the two positions (e.g., 2500 > 3000). To exclude a particular range, write the higher number first (e.g., 3000 > 2500) to highlight everything except for that range. GeneQuest accepts a comma (,) or two periods (..) in place of the greater-than symbol (>). Use the less-than symbol (<) to select ranges on the bottom strand. Click **OK** to move the pointer. If you entered a number that is outside the range of the sequence, GeneQuest will position the pointer at the end of the sequence.
Find a nucleotide sequence string	Choose **Edit > Find** or press **Ctrl/Cmd+F**. Type a search string (e.g., gctttag) up to 256 characters in length into the **Find** box. Use the radio buttons to search **Either Strand** or only the **Top Strand** or **Bottom Strand**. Click **Next** or **Previous** to find the next/previous appearance of the search string in the sequence. Both options cause you to exit from the Find dialog and return to the Assay Document. The pointer will be positioned at the beginning of the sequence region that matches the search string. To continue jumping to the next candidate site of the same type in the highlighted range, select **Edit > Find/Move Next**. To move to the previous candidate site, select **Edit > Find/Move Previous**. **Note:** If the candidate start, donor or acceptor splice sites you initially locate are in the reverse strand, **Find/Move Next** moves the selection to the left and **Find/Move Previous** moves the selection to the right, consistent with the orientation of the reverse strand.
Find translation starts, donor splice sites or acceptor splice sites	Use the **Range selector** to select a sequence range, then choose **Edit > Move Selection** or press **Ctrl/Cmd+L**. Specify whether you want to find a potential **Translation Start (ATG)**, **Acceptor Splice (AG)** or **Donor Splice (GT)**, then click **Next** or **Previous** to find the next/previous appearance of the search string in the sequence. Both options cause you to exit from the Find dialog and return to the Assay Document. The pointer will be positioned at the beginning of the sequence region that matches the search string. To continue jumping to the next candidate site of the same type in the highlighted range, select **Edit > Find/Move Next**. To move to the previous candidate site, select **Edit > Find/Move Previous**.
Highlight a region beginning with an acceptor splice and ending	Use the **Range selector** to select a sequence range, then choose **Edit > Move Selection** or press **Ctrl/Cmd+L**. Select **Acceptor Splice** and click **Next**. Press **Ctrl/Cmd+.** until the left side of the highlighting is aligned with some method feature. Select **Edit > Move Selection**, choose **Donor Splice**, and click **Previous. Press *Ctrl/Cmd+,** repeatedly, until the right side of the highlighting is also aligned appropriately.

with a donor splice	
Remove a selection made using the Object selector	Use **Edit > Clear**.

View Data Tables

The table below shows how to generate statistical tables in GeneQuest:

| Task	How To
View a summary of restriction fragments	Choose **Sites & Features > Restriction Fragment Summary**. To specify whether fragments in the table should be sorted by length or by location, choose **Options > Fragment Summary Options**. The command is enabled only if the Restriction Fragment Summary is the active window.
View results for relevant methods currently on the Assay Document	Choose **Analysis > Tabular Data**. The first column shows a nucleotide, the second its position and the remaining columns show the value assigned to each nucleotide for a given analysis. Methods which calculate a numerical value (e.g., line graphs) display that value, while those making simple yes/no decisions (e.g., region plots) show a letter or symbol in positions having a value of "yes." Pattern matches are not shown in the Tabular Data Summary, as there is not always one value per residue. To show the parameter window for the method in question, double-click anywhere on a column. Parameters may be changed by entering the new value(s) and clicking OK. GeneQuest recalculates the results. If you prefer to return to the default settings, click the Set to Defaults button in the parameter window.

	If you plan to print or copy part of the selection, drag the cursor horizontally across column headers or vertically across residues. To highlight non-adjacent columns, hold down the Shift key while individually selecting the names of the columns that you want. To view an expanded header at the top of each method column, click Show Summary in the lower left corner of the Tabular Data Summary. The new fields list the minimum, maximum and mean values for the plot and show the color scheme used on the assay document. To rearrange the order in which columns appear in the summary, first change the order of the graphs on the Assay Document and close, then reopen, the tabular summary. You can copy tabular data to your clipboard in the form of tab-delimited text or as an image. If you only want to copy certain columns for export, hold down the Shift key while you click each of the desired method titles. When the appropriate columns have been highlighted, select Copy. Note, however, that popular spreadsheet programs may be limited to 65,536 rows of data and would not be able to accommodate all the tabular data for sequences longer than 65,536 bases. Select File > Print Selection to print the highlighted portion of the table.
View the composition of nucleotides, dinucleotides, trinucleotides, tetramers, pentamers, and hexamers in the sequence	Choose **Analysis > Base Composition**. All nucleotide percentages are shown for mononucleotides, dinucleotides and trinucleotides, while longer strings are sorted into frequent and infrequent species. GeneQuest examines the entire nucleotide sequence unless a sub-range is selected. Method parameters for this method can be changed using **Options > Base Composition Settings**. To specify how many of the most frequent and least frequent mers you would like to see displayed in the table, and to learn how many hexamers are present in the sequence, choose **Options > Base Composition Settings**. The command is only enabled if the Base Composition summary is the active window.
View the distribution of codons in the nucleotide sequence, or selected nucleotide sub-range	Choose **Analysis > Codon Usage**. Distribution is shown for a single reading frame, beginning with the first nucleotide in the highlighted portion of the sense strand. The Codon Usage window contains a nucleotide "Base Composition" table and a "Codon Preferences" table.
View all repetitive portions of a sequence in a single window	After you have applied a repeat method (Repeats – Inverted, Repeats – Dyad, or Repeats – Direct), use **Analysis > Repeat Summary** to view all repetitive portions of a sequence in a single window. The command is disabled until at least one repeat method has been applied to the sequence. After applying a repeat method, however, the summary will remain available as long as the repeat method remains in the Method Curtain, even if it is removed from the Assay Document. The summary is divided into four areas – direct repeats, short inversions,

palindromes and dyad repeats – depending upon which repeat methods have been applied to the sequence. GeneQuest shows the residues involved and the length of each repeat, as well as the starting position of the repeat in the sequence. The amount of data displayed is limited only by your computer's memory. If your computer lacks enough memory, however, it is possible that the list of repeats may be truncated. Repeat Summary remains dimmed until at least one repeat method has been applied to the sequence. After applying a repeat method, the summary will remain available as long as the repeat method remains in the Method Curtain, even if it has been removed from the Assay Document.

To designate whether repeats in this table should be sorted by length or by location, use **Options > Repeat Summary Options**. The command is only enabled if Repeat Summary is the active window.

Annotate Features

When you import a Lasergene or Genbank sequence into GeneQuest, features from the document are placed as annotation methods at the bottom of the Method Curtain. GeneQuest also lets you create new annotations to label areas of interest in your DNA sequence. A feature can be composed of one or more segments.
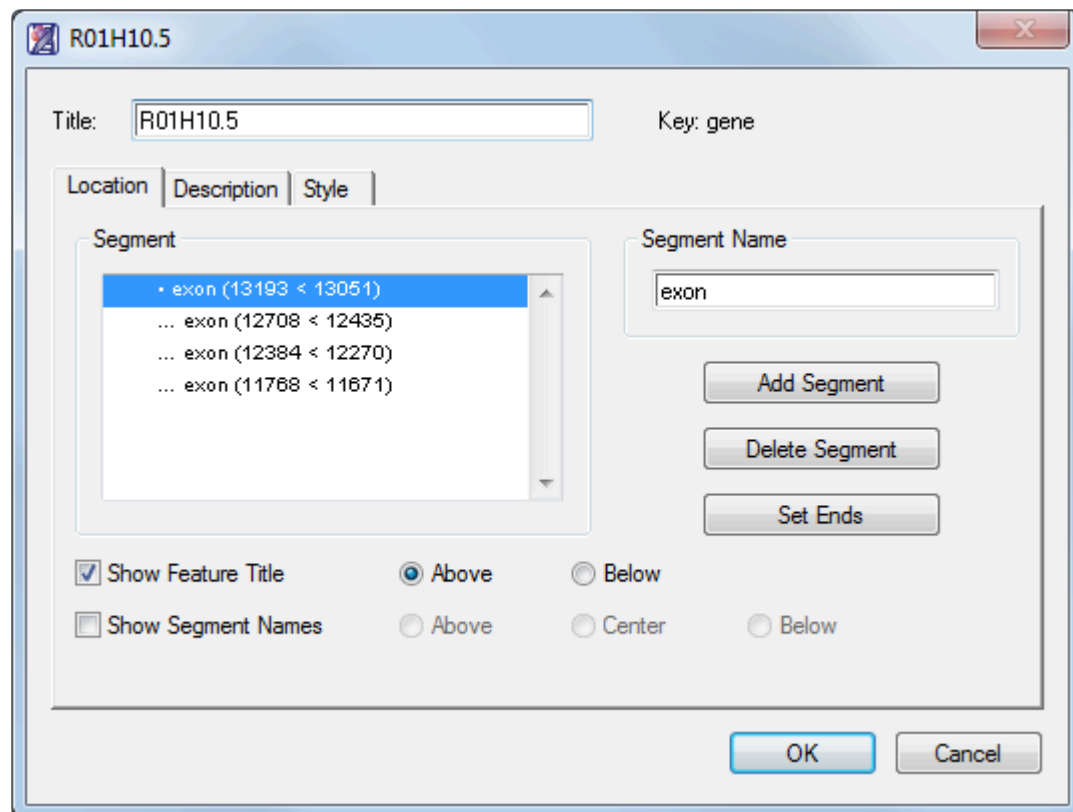
Note: Feature annotation functionality and useability is greatly improved in SeqBuilder Pro, our modern application that is installed alongside GeneQuest. To learn about annotating features in SeqBuilder Pro, click here.

A feature can be composed of one or more segments. Using the Feature Editor, both the feature and its segments can be annotated independently or joined under a common heading and description.

To create a new feature:

1. Activate the **Range selector** ( ).

2. Select a range or starting residue. To use the currently highlighted range of sequence as the first segment, highlight an area of interest. To use the residue that the cursor was last resting on (even if the cursor line is not visible at present), position the cursor at the desired point.

3. Do one of the following:

- Click on the **Annotation** tool ((![icon])).

- Select **Sites & Features > New Feature**.

- Press **Ctrl/Cmd+K**.

- Right-click anywhere on the document window and choose **New Feature**.

4. Use the Location, Description, and Style tabs to assign a design style, font, size and color for the feature. You can also make notes about the feature and assign a keyword description.

To edit an existing feature:

1. Do either of the following:

- Double-click on the feature.

- With the **Range selector** (![icon]) active, click on the feature, and then select **Sites & Features > Edit Feature**.

2. Use the Location, Description, and Style tabs to assign a design style, font, size and color for the feature. You can also make notes about the feature and assign a keyword description.

To link two or more features together:

Use this option to join a selected region as a new segment in an existing feature annotation.

1. Activate the *Range selector*(![icon]).

2. Select a range that you wish to add to an existing feature.

3. Choose **Sites & Features > Join to Feature** or click on the **Joining Tool** tool (![icon]).

4. Specify which feature to append your region onto and click **Join**.

Except for having a different range, the new segment will have all the same Location, Description and Style characteristics as the segment it is appended to.

To generate a feature based on a region plot:

1. Apply one or more methods whose results are displayed as region plots () .

2. Select the region plot(s) using the **Object selector** ().

3. Select **Sites & Features > Generate Features**.

4. If you have selected different region plots for the same method (e.g., two **ORF** plots from **Coding Prediction – Starts, Stops, ORFs**), a dialog box opens.

5. Choose whether you wish to combine related decisions or make all separate features. The former causes a single joint feature display to be created, while the latter creates separate displays for each of the varieties.

6. Click **OK**. Results appear as one or more new displays at the bottom of the GeneQuest assay document.

To sort features:

Use the **Sites & Features > Sort Features** command to sort the features listed in the method curtain.

- **Type** – Sorts by feature type, or key.

- **Name** – Sorts by feature name.

- **Position** – Sorts by position on the sequence.

By default, features are sorted by **Position**.

Location tab

The Location tab, located in the [Feature Editor](#), lets you determine what sequence coordinates are associated with a given feature, where the feature appears and how the feature is labeled. The feature can be one contiguous stretch of residues or it can be a series of segments.

To access this tab, open the Feature Editor. This dialog always opens with the **Location** tab active. For a list of ways to open the Feature Editor dialog, see [Annotate Features](#).

- Use **Title** to name the feature or edit the current name.

- **Segment** lists all segments in a feature. Double clicking a segment lets you set its range by selecting the **Set Ends** button. Selecting a segment makes it active in the **Segment Name** field.

- The **Segment Name** field displays the name, if any, of the selected segment in the **Segment** window to the left. Under segment name, type in a name for the segment (e.g., the name or number of the exon) if desired.

- Click **Add Segment** to add a new segment after the selected segment if desired. (Note that you may instead find it more convenient to return to the assay document, select the next segment you want to annotate, and then select **Sites & Features > Join to Feature**.)

  - If the last segment in the list is highlighted, the default range begins with the first residue

following the last segment and ends at the last residue in the sequence.

  ◦ If you have highlighted a segment higher in the list, **Add Segment** uses the range from the end of the previous segment to the beginning of the succeeding one.

- Click **Set Ends** to modify the segment end points and to set the range and orientation of the segment.

- Click **Delete Segment** if you wish to remove the selected segment from the segment window.

To annotate the feature on the bottom strand rather than on the top strand, click the **Set Ends** button, then click the white button with two arrows on it. If you have chosen the "arrow" style in the Style tab, the arrow will now point in the opposite direction.

- Check the *Show Feature Titl*e box to display the feature title when the feature is displayed on the assay document. After placing a check, click Above or Below to determine the position of the title display. If there is no check mark present, GeneQuest will not display the feature's title when it displays the feature on the document.

- Check **Show Segment Name** to show the segment name when the segment is displayed on the assay document. After placing a check, click **Above**, **Center**, or **Below** to determine the position of the name display. If there is no check mark present, GeneQuest will not display the segment's name when it displays the feature on the document.

After you have set the feature's location, you may wish to use the [Description tab](#) to modify the class of a feature or notes about it, or the [Style tab](#) to change its graphic appearance. Or click **OK** to save changes and close the dialog.

Description tab

If you imported features from Genbank, the Description tab of the Feature Editor shows the Genbank keyword(s) already associated with each feature. You may also use this dialog to record notes about the feature.

To access the dialog, open the Feature Editor and click on the **Description** tab. For a list of ways to open the Feature Editor dialog, see Annotate Features.
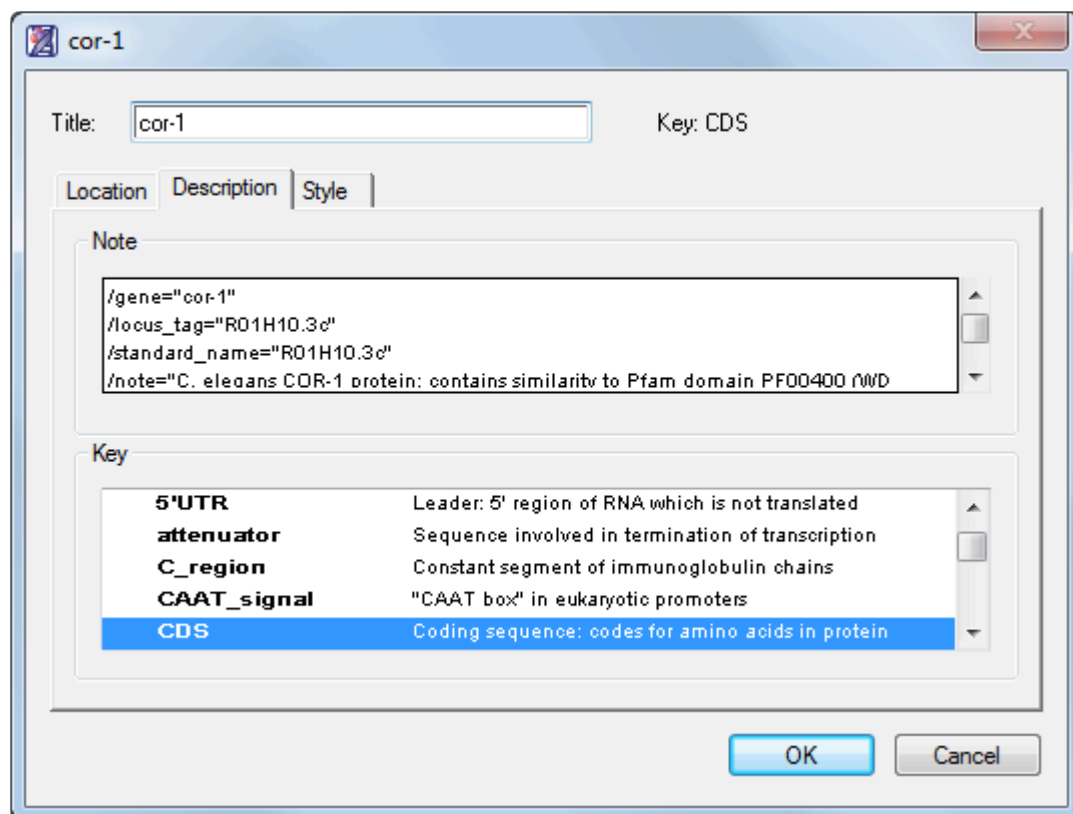
- Type a name for the feature in the **Title** window, or edit the current name, if desired. This is the title displayed on the Assay Document if you check the **Show Feature Title** box in the Location tab.

- Scroll down the **Key** list and click on the appropriate key descriptor for your feature. The selected **Key** will appear at the right of the **Title** text box in all future **Feature Editor** sections.

- Use the **Note** box to type in a note about the feature, if you wish.

After you have specified a description, you may wish to use the Location tab to modify the feature's range, or the Style tab to change its graphic appearance. Or click **OK** to save changes and close the dialog.

Style tab

The Style tab of the [Features Editor](#) lets you determine the graphic appearance of your feature. It is accessed by opening the Feature Editor and clicking on the **Style** tab. For a list of ways to open the Feature Editor dialog, see [Annotate Features](#).
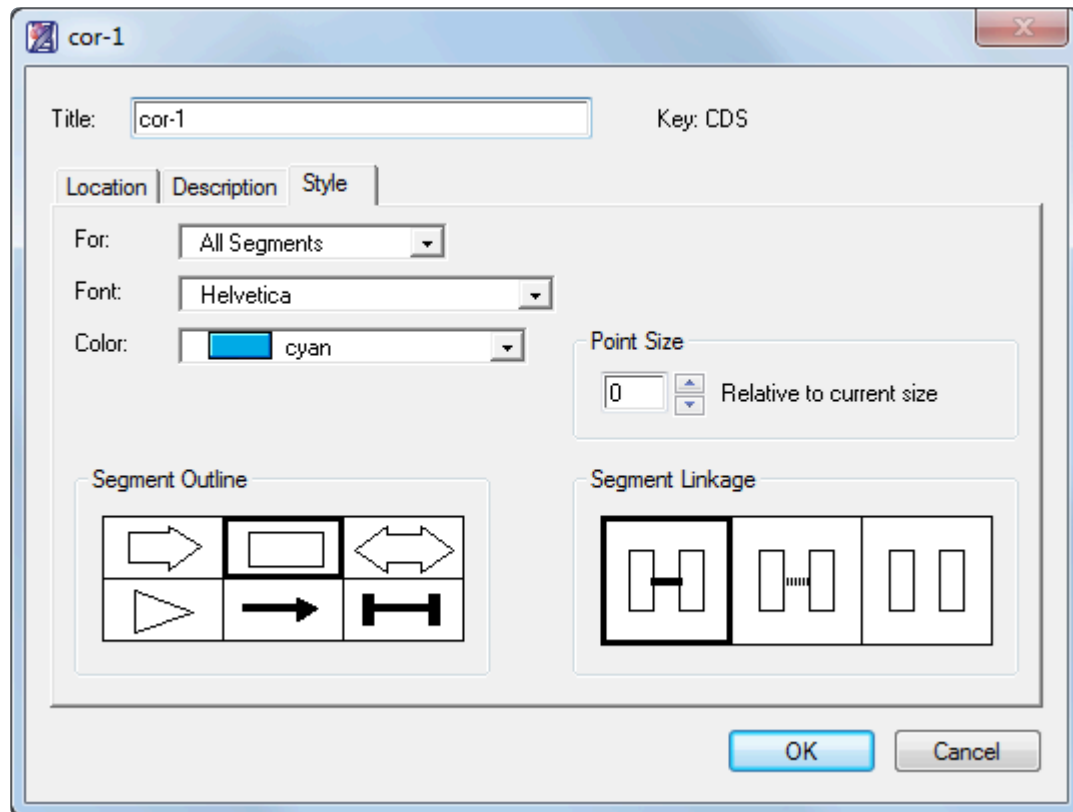

- Type a name for the feature in the **Title** window, or edit the current name, if desired. This is the title displayed on the Assay Document if you check the **Show Feature Title** box in the Location tab.

- Select the segments for which these styles will apply in the **For** list. To have all segments of this feature use the same style, select **All Segments**.

- Use the **Font** list to select the font that the feature will use.

- The **Color** menu lets you choose a color for your feature or segment.

- If the feature has multiple segments, click the appropriate boxes to select graphics for **Segment Outline** and **Segment Linkage**.

- The **Size Relative to Current Point Size** box lets you pick the font size for your feature or segment. Fill in a size in the text field or use the up and down arrows. The point size is relative to the **Size** chosen under the **Options** menu.

> ✳ **Note:** To modify the fill pattern for the graphic display of the feature, highlight the feature on the Assay Document then select **Options > Fill Pattern**.

After you have specified a description, you may wish to use the Location tab to modify the feature's range, or the Description tab to modify the class of a feature or notes about it. Or click **OK** to save changes and close the dialog.

Run a Simulated Agarose Gel

GeneQuest helps you predict electrophoretic separation of restriction fragments through an agarose gel. The following video demonstrates this.

> ✳ **Note:** Agarose gel functionality and useability is greatly improved in SeqBuilder Pro, our modern application that is installed alongside GeneQuest. To learn about running an agarose gel simulation in SeqBuilder Pro, click here.

Start by applying one or more enzymes from the Enzymes – Restriction Map method to the Assay Document. Then select **Sites & Features > Agarose Gel Simulation**.

The restriction enzymes selected from the Assay Document will appear. If none are selected, then all will appear in the simulation.

The far left and far right lanes of the Agarose Gel Simulation are molecular size standards from bacteriophage lambda. The other lanes show all possible combinations of single, double and triple digests. Above each lane is the name of the enzyme(s) or molecular size standard. Fragments in red represent fragments within the current range selected in the Assay Document.
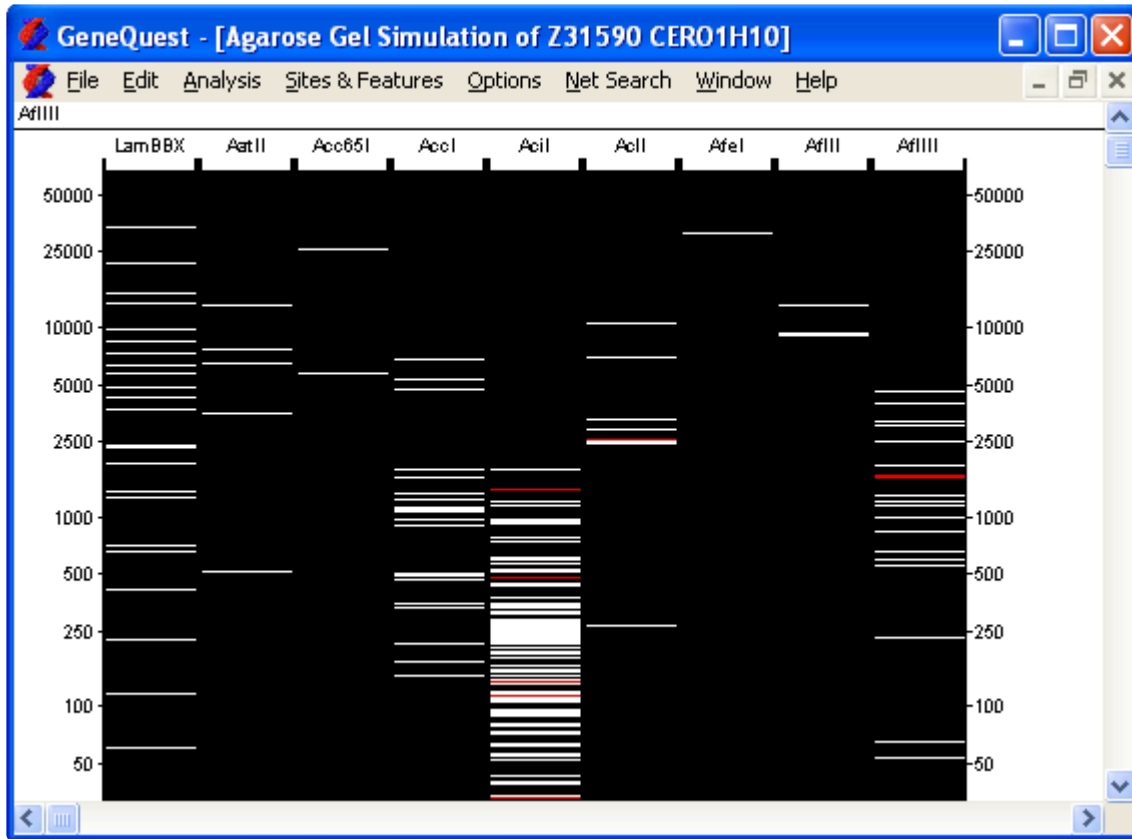
Move the cursor over a fragment in the simulated gel to see its range, size, and the restriction enzymes cutting each end in the header.

Change the selected range by double clicking on any DNA fragment in the gel simulation.

View a Portion of the Sequence as Folded RNA

You can view a portion of a sequence as folded RNA. GeneQuest uses the Vienna RNA folding procedure, taken from Zuker's optimal RNA folding method, to fold the sense strand of selected DNA regions as RNA.

Use the **Range selector** to select a portion of sequence. The selection must be 1500 bases or less in length. Then choose **Analysis > Fold As RNA**. The folding procedure may take up to several minutes.
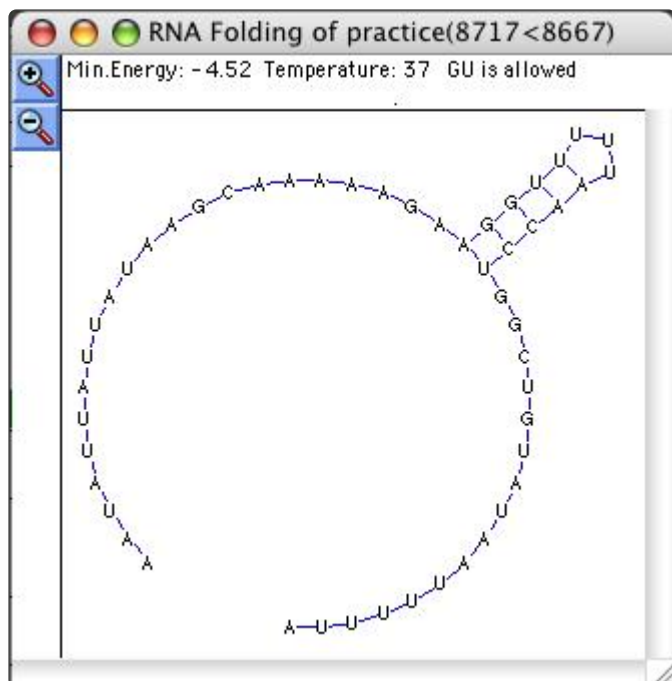
The folding model is an interactive two-dimensional plot. At the top of the RNA folding window, GeneQuest lists the minimum energy (in Kcal/mol) of the structure, the temperature used in folding and the status of G-U pairing. Use **Zoom In** ( ) and **Zoom Out** ( ) to magnify or reduce the display by 150%. To copy the image as a Windows picture, use **Edit > Copy**. To instead save it as a Connect file (.ct), use **File > Save As** and choose that option from the drop-down menu.

Editing RNA folding parameters:

Parameters can be changed either before or after folding RNA. To set parameters before folding the selection, hold down the **Alt** key while selecting **Analysis > Fold As RNA**. To alter the parameters after viewing the folded RNA, select **Options > Folding Parameters**. The command is hidden unless the RNA folding window is the active window. In both cases, the following dialog opens.

- Enter the **Temperature** that you want to be used in the folding algorithm. The default is 37°C. Altering temperature will modify the nearest neighbor free energy values ([Frier et al., 1986](#)) used to calculate the minimal free energy of the folded structure.

- The next three buttons determine the state of guanine to uracil base pairing. You can allow G-U pairing everywhere, allow it only inside helical structures, or forbid it entirely. Allowing G-U everywhere will increase the complexity of secondary structures and forbidding it will decrease complexity.

- Checking either of the two checkboxes, **fold with specially stable tetra loops** or **use dangling end energies**, will cause GeneQuest to use alternate nearest neighbor free energy values while within these respective structures.

Print, Save and Export

See the following topics for details:

- [Print the Assay Document](#)
- [Save a sequence file or GeneQuest project](#)
- [Export a sequence to another Lasergene application](#)
- [Export a GeneQuest project to GenVision](#)

Print the Assay Document

Task	How To
Print the active (topmost) window	Choose **File > Print**. Depending upon your printer, you may first be able to set parameters such as the print range, the number of copies, whether to collate copies and whether to print the current view to a separate file.
Print the selected portion of the Tabular Data Summary	Choose **File > Print Selection**. This command is only active when a selection has been made in the Tabular Data window. The ensuing dialog is identical to the File > Print dialog.
Change the location of page breaks when printing the assay document	Use **Options > Show Page Breaks** to see where breaks between pages fall on assay surfaces larger than one page. Changes made in the **File > Print Setup** and **Options > Horizontal Units** (see Customize the look of the Assay Document) and scaling using the **Amplitude control**, **Zoom In** and **Zoom Out** tools all affect where page breaks occur. **Options > Show Page Breaks** is inactive unless the assay document is the active window.
Specify printing parameters such as which printer to use, page orientation, paper size, and paper source	Choose **File > Print Setup** (Win) or **File > Page Setup** (Mac).
Optimize the number of pages and location of page breaks	Choose **Options > Show Page Breaks** to see where breaks between pages fall on assay surfaces larger than one page. Avoid bisecting important features during printing by manipulating the size of the display until page breaks fall in more convenient locations. The following all have an effect on the number of pages and/or locations of page breaks: using **Options > Horizontal Units**, the **Zoom In**, **Zoom Out**, and **Amplitude control** tools and the legend printing option you choose. To review how many pages long a document is, refer to the display at the right of the Assay Document header. However, the exact length of the printed document depends upon

which legend printing option you choose. To find the exact length, use the **Print** command, click on the desired legend option, then click **Cancel**. The assay document header will be updated to reflect the exact number of pages required to print the document using the chosen settings.
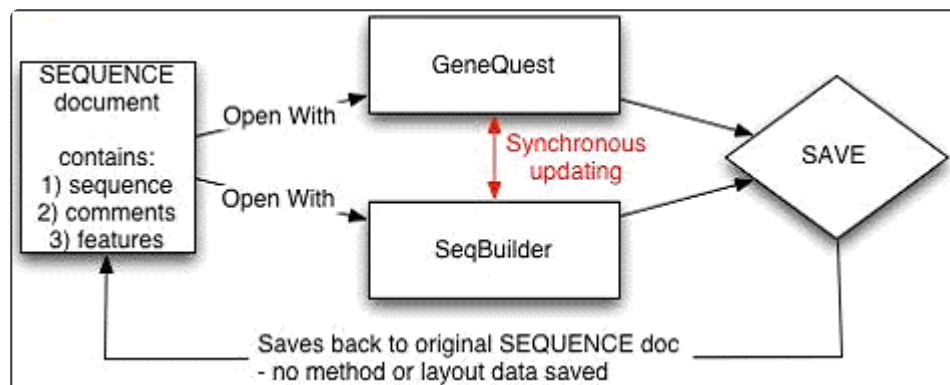
Save a sequence file or GeneQuest project

To save an assay document, select **File > Save** or use **Ctrl/Cmd+S**. If you have already saved a document but want to save a copy of it under a new name or location, use the **File > Save As** command. In either case, the current method outline will be saved as part of the project file. Because of this, all methods that have been moved into the Method Curtain or dragged onto the Assay Document will still be there when you reopen a saved document. Similarly, edits made to parameter settings are saved, as are any visual enhancements made to the Assay Document (like line color).

Saving a modified sequence file using **File > Save** :

GeneQuest and SeqBuilder Pro can open project files created by the other applications.

Use **File > Save** to save sequence information, features, and comments back to the original sequence file (.*seq* or .*pro*). Application-specific information, such as layout and formatting, is not saved.

As long as you only use **File > Save**, these three programs can continue to use synchronous updating, as described in Export a sequence to another Lasergene application.

Saving a modified sequence file using **File > Save As** :

GeneQuest and SeqBuilder Pro can open project files created by the other applications.

Use **File > Save As** to save the application-specific data into a project file (e.g., .*sbd*, .*dad*, or .*pad* projects). Once you use **File > Save As**, you can no longer use synchronous updating to the original sequence file because the newly-saved project file has copied the information from that sequence file into the new project file along with additional project information.

You can re-establish a synchronous updating link between SeqBuilder Pro and GeneQuest by opening the newly-created project file in the other application(s). You do not need to close the original project file. However, because **File > Save As** has been used in another open application, that file will no longer be updated when you make changes in the application where you last used **File > Save As**.
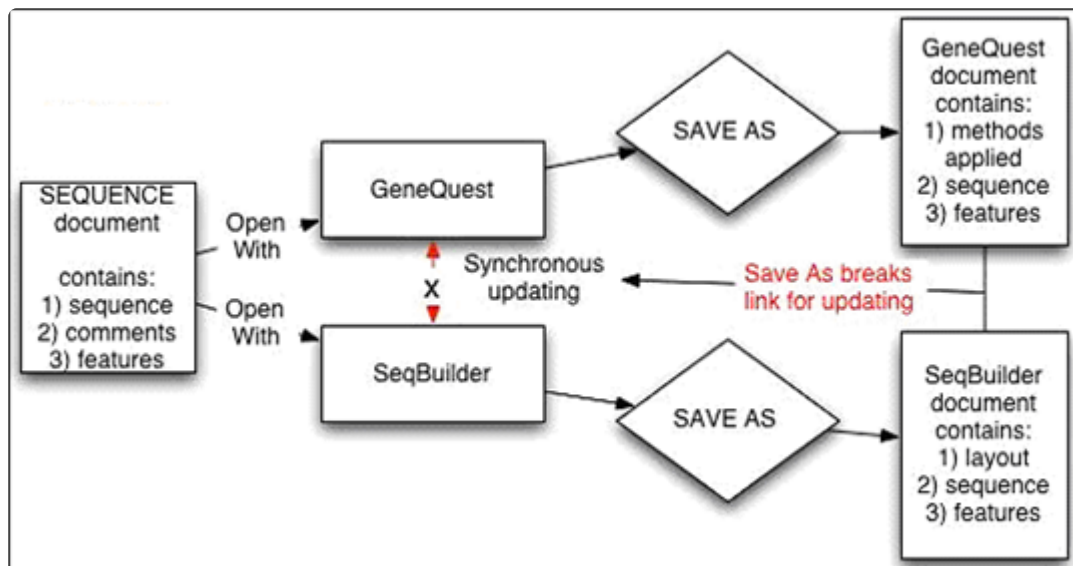
Saving a modified project file:

GeneQuest and SeqBuilder Pro allow one sequence per project and allow you to open and modify a project file from the other applications' project files. These two applications can save the project file in the format of any other Lasergene application file.

Just as when working with a sequence file, **File > Save** and **File > Save As** work somewhat differently. For example, if you have opened a GeneQuest project file in SeqBuilder Pro:

- Selecting File > Save in SeqBuilder Pro saves the project in the original GeneQuest document, including the modified sequence information, GeneQuest methods, and SeqBuilder Pro layout information.

- Selecting **File > Save As** in SeqBuilder Pro saves the project file as a SeqBuilder Pro (.sbd) document, and it contains the modified sequence information, the GeneQuest methods, and the SeqBuilder Pro layout and formatting information. However, it breaks the current synchronous updating link between the two applications.

To reconnect the synchronous updating link between the two applications, you need to open the new project file in GeneQuest.

- Selecting **File > Save** maintains the connection between the two applications so that synchronous updating can continue.

- Selecting *File > Save A*s creates a new document that breaks the current connection between applications and prevents synchronous updating until both applications once again have the same project file opened.

Export a sequence to another Lasergene application

Synchronous updating allows you to work on the same file in multiple Lasergene applications at the same time, and is available in SeqBuilder Pro and GeneQuest, and—on a limited basis—in Protean 3D. Changes made to a sequence in one application are automatically applied in the other shared applications.

To export a sequence to another Lasergene application with synchronous updating:

1. Open the sequence of interest.

2. Send the sequence to 1-3 additional applications using **File > Send Sequence To > [Lasergene application]**. For example, you could open the same document in:

   - SeqBuilder Pro – to edit the sequence
   - GeneQuest – to further analyze a region of conservation

For information about how **File > Save** and **File > Save As** affect synchronous updating, see Save a sequence file or GeneQuest project.

Note that Protean 3D does not allow the editing of sequences or features, and synchronous updating is not supported for *.pdb* files. As a result, Protean 3D has limited, unidirectional synchronous updating capacity compared to some other Lasergene applications. Changes made in another application may be synchronized to Protean 3D, but not the reverse, as in these two scenarios:

- A file is open in another Lasergene application and you have made unsaved changes to it (e.g., trimming, adding a feature). When you initially open the same file in Protean 3D, any changes made to the original sequence or its features will also appear there.

- The same file is already open in both Protean 3D and another application. You make changes to the file in the other application. Only feature edits will synchronize to Protean 3D; sequence edits are ignored.

Export a GeneQuest project to GenVision

GeneQuest assay documents can be exported into GenVision format by selecting **File > Export as GenVision Project**. This function is disabled if **Set Ends** (see Open a DNA sequence file) was previously used.

In the resulting dialog, specify a name and location for your GenVision project, then click **Save**.

You will be prompted with the following dialog:
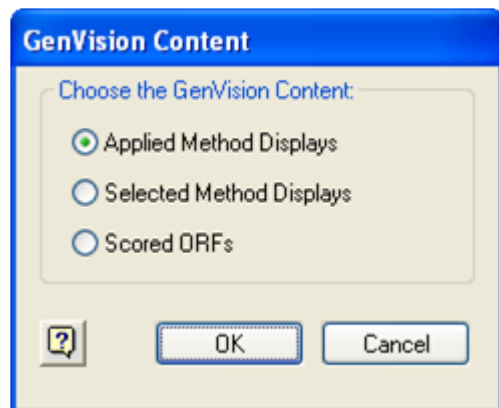
- **Applied Method Displays** will export all analysis methods currently applied to the Assay Document.

- **Selected Method Displays** will export only those analysis methods currently selected on the Assay Document.

- **Scored ORFs** will export ORFs that correspond to a Borodovsky method applied on the Assay Document for the same frame.

Select the GenVision content you wish to export, and then click **OK**. A GenVision project (*.gnv*) will be exported, in addition to a text (*.txt*) file for each exported panel. A new folder containing the exported text files will be created within the same directory as your GenVision project and will be named the same as your GenVision project with a data (*.data*) extension.

Additional considerations when exporting from GeneQuest to GenVision:

- The figure will be exported as linear.

- All exported text files will be prefixed with **GQ** to denote that they were exported from GeneQuest.

- Features are grouped into panels by feature type. One panel will be created for each feature type.

- For each feature panel created, a corresponding panel will be created for the feature labels.

- Non-feature analysis methods are exported as one panel per selected/applied item.

- Fill and line colors will be preserved for each exported method.

- An End Legend file is created for all of the panels within the export and is named the same as your GenVision project with _table inserted. For example, *GenVisionProject_table.txt*.

Close and Exit

To close the active assay document:

Select **File > Close** or use **Ctrl/Cmd_W**. If the document has never been saved, or has changed since the last save, you will be asked whether you wish to **Save** the document with its changes or **Discard** the changes before closing.

To leave GeneQuest:

Select **File > Exit** (Win) or **GeneQuest > Quit** or **Cmd+Q** (Mac). GeneQuest will prompt you to save any unsaved projects before exiting. Click **Save** to preserve a document. Click **Discard** to close the document without saving recent changes or to throw away a document that has never been saved.

Appendix

Choose the genetic code

To select a genetic code:

To choose from a list of genetic codes, select **Options** and hover over the **Genetic Codes** menu item. The active genetic code is preceded by a checkmark. Click the desired code to select it.

When selecting a genetic code whose code file is absent from Lasergene's GeneticCodes folder, the Find dialog will open automatically. Navigate to the folder containing the missing genetic code file (which will have the extension .rtc), and then click **Open**.

Genetic code files are installed, by default, in the following locations:

- **Windows:** *C:\Users\Public\Public Documents\DNASTAR\Lasergene 'x' Data\GeneticCodes*

- **Macintosh:** *Hard Drive:Applications:DNASTAR:Lasergene 'x':GeneticCodes*

… where 'x' represents the version number.

To change the list of codes displayed in **Options > Genetic Codes**:

Select **Options > Genetic Codes > Modify Menu**. This launches the Genetic Codes dialog, which lists all of the available codes on the left, and all of the codes current listed in the menu on the right. By default, all available codes are displayed.
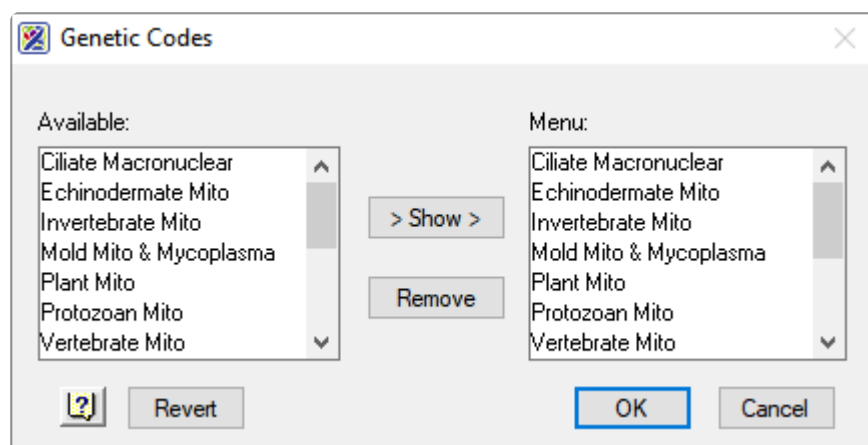
- To add a code to the menu, select it in the **Available** list and then click **Show**.

- To remove a code from the menu, select it in the **Menu** list and then click **Remove**.

- To save changes, click **OK**. Alternatively, to change all code assignments to the last saved version, click **Revert** and then **OK**.

Modify the genetic code

Changes made to a genetic code in GeneQuest will affect the same genetic code in all other Lasergene applications.

To modify or rename a genetic code for all Lasergene applications:

Open the two-tabbed Genetic Code Editor using **Options** > **Edit Genetic Code**. To assign a name for the code, type a name into the **Description** field. If you do not make changes to this field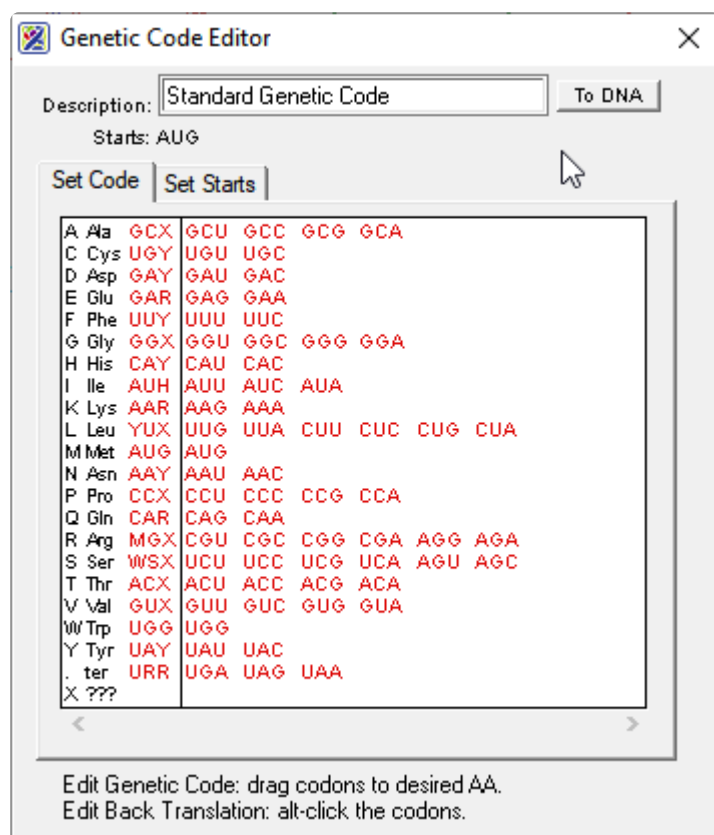, the original genetic code will be overwritten with any changes made below.To toggle between RNA and DNA, use the button in the upper right of the dialog .

Tasks in the Genetic Code tab	How to…
Modify codon assignments, including stop codons	Use the mouse to drag a codon to the appropriate amino acid row. The color of the codon will change from red to black to show it has been moved. SeqBuilder Pro only considers red codons when it calculates the aggregate codon for the set.
Include the newly moved	**Alt+click** (Win) or **Option+click** (Mac) on the codon to change it back to red. The sum of all codons in the row is recalculated to give the revised reverse translation codon. The same

codon when calculating the aggregate codon	recalculation occurs instantly when a codon is removed from a row. Reverse translation codons, shown in the third column from the left, are the codons used in reverse translating a protein sequence.
Change the codons that will be used for back translation without changing the forward translation	**Alt+click** (Win) or **Option+click** (Mac) on codons to exclude. For example, in the Standard Genetic Code, there are two codons for Cysteine (UGU and UGC) and the default reverse translation codon is UGY (Y means pyrimidine, i.e. C or U). To reverse translate Cysteine residues as UGC, **Alt+click** (Win) or **Option+click** (Mac) on the UGU codon to change its color to black so it is not considered for reverse translation. The reverse translation codon for Cysteine will now read UGC instead of UGY.

Tasks in the Start Codons tab	How to…
Edit start codons	Select the desired start codons visually by finding the row containing the first base (on the left), following that over to the column belonging to the second base (in the center) and selecting the row in the table that aligns with the third base (on the right). Click an amino acid to turn it green,

	and cause > to appear beside it. The codon representing the amino acid now appears at the top of the screen to the right of the word **Starts**.
Remove a start codon from the list at the top of the page	Click on the corresponding amino acid in the body of the table.

Once you have made the desired changes, click **OK**. If you have changed the description, a confirmation dialog appears. Choose **Save As** to save the code, or **Discard** to discard changes and begin editing anew.
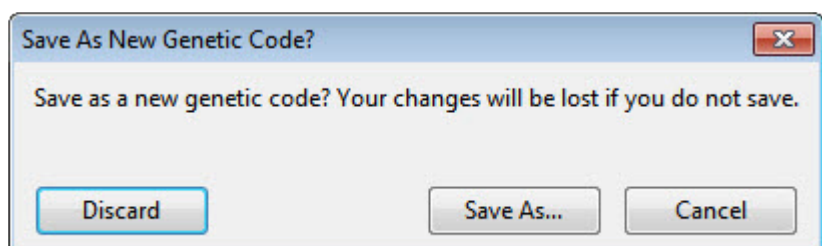
Use editing commands

The following editing commands can be used in multiple views, panels and dialogs.

Task	How to…
Remove the current selection from the project and place it on the clipboard	Choose **Edit > Cut** or use keyboard shortcut **Ctrl/Cmd+X**. Among other things, you can cut range and text selections, comments in the Comment view, features in the Features view and translation rulers. You can also use **Edit** > **Select All** within the Sequence, Circular, Linear, or Minimap and then use **Edit > Cut** to remove the contents. **Note:** Cut data does not remain strand-specific. If you copy data from a forward strand, you may copy it to the reverse strand and vice versa.
Paste the contents of the clipboard into an insertion point	Choose **Edit > Paste** or use the keyboard shortcut **Ctrl/Cmd+V**.
Delete the selection from the project without placing the contents in the clipboard	Choose **Edit > Delete** or press the **Delete** key
Select the entire contents of the active window	Choose **Edit > Select All** or use the keyboard shortcut **Ctrl/Cmd+A**. Neither option is available within the Features view.
Reverse the effect of the previous operation(s)	Choose **Edit > Undo** or the keyboard shortcut **Ctrl/Cmd+Z**. You may use these options multiple times to undo multiple actions.
Reverse the effect of the last Edit > Undo command	Choose **Edit > Redo** or the keyboard shortcut **Ctrl/Cmd+Shift+Z** You may use these options multiple times to redo multiple actions.

Supported file types

For a list of file formats supported by this DNASTAR product, please see the File Formats page of our website.

IUPAC codes

IUPAC codes for amino acids:

1-Letter Code	3-Letter Code	Amino Acid
A	Ala	Alanine
B	Asx	Aspartic Acid or Asparagine
C	Cys	Cysteine
D	Asp	Aspartic Acid
E	Glu	Glutamic Acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
J	Xle	Leucine or Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
O	Pyl	Pyrrolysine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
U	Sec	Selenocysteine
V	Val	Valine
W	Trp	Tryptophan
X	Xaa	Unspecified or unknown
Y	Tyr	Tyrosine

Z	Glx	Glutamic Acid or Glutamine

IUPAC codes for nucleotides:

Symbol	Description	Bases Represented			
A	adenosine/adenine	A			
C	cytidine/cytosine		C		
G	guanosine/guanine			G	
T	thymidine/thymine				T
U	uridine/uracil				U
W	weak	A			T
S	strong		C	G	
M	amino	A	C		
K	keto			G	T
R	purine	A		G	
Y	pyrimidine		C		T
B	not A		C	G	T
D	not C	A		G	T
H	not G	A	C		T
V	not T	A	C	G	
N or -	any base (not a gap)	A	C	G	T

✳ The non-standard nucleotide codes "I" (Inosine) or "X" (xanthine) are not supported.

Turn off usage logging

By default, usage logging is enabled. To opt out of usage logging, launch the DNASTAR Navigator and go to **View > Preferences > Lasergene** (Win) or **DNASTAR Navigator > Preferences > Lasergene** (Mac). Uncheck the box and click **Apply** and then **OK**.

Troubleshoot the error "A different version is running"

If you receive the error "A different version of Lasergene is running" when you try to launch a Lasergene application, but another version of Lasergene actually isn't running, there may be a lingering .state file from another version that needs to be deleted.

The solution is to locate and delete files named *STARDM*.state*, …where '*' represents the version number.

These files may exist in the following directories:

Windows 7 and Windows 8:

- *C:\Users\username\AppData\Local\DNASTAR\DataManager*

- *C:\ProgramData\DNASTAR\DataManager*

Macintosh:

- *Hard Drive:Library:Preferences:DNASTAR:DataManager*

- *Hard Drive:Users:username:Library:Preferences:DNASTAR:DataManager*

If you find any of the *.state* files listed above, delete them. Once the files are deleted, the error should no longer appear.

Installed Lasergene file locations

The following file names use 'x' to represent the version number.

File Category	Application	Path
Application[2]	ArrayStar	Windows: C:\Program Files (x86)\DNASTAR\Lasergene x\ArrayStar
	SeqNinja (command line)	Windows: C:\Program Files (x86)\DNASTAR\Lasergene x\SeqNinjaCL
	All others	Windows: C:\Program Files (x86)\DNASTAR\Lasergene x Mac: /Applications/DNASTAR/Lasergene x
Borodovsky Matrix files	GeneQuest	Windows: C:\ProgramData\DNASTAR\Lasergene 'x' Data\Borodovsky Matrices Mac: ~/Library/Applications/DNASTAR Data\Lasergene 'x' Data\Borodovsky Matrices
Data Manager[2] (*DMx, DMx.exe*)	SeqBuilder Pro, Protean 3D, GeneQuest, MegAlign Pro	Windows: C:\Program Files (x86)\DNASTAR\Lasergene x Mac: /Applications/DNASTAR/Lasergene x
Data Manager State File	SeqBuilder Pro, Protean 3D, GeneQuest, MegAlign Pro	Windows: C:\Program Data\DNASTAR\DataManager Mac: ~/Library/Application Support/DNASTAR/DataManager, ~/Library/Preferences/DNASTAR/DataManager
License Manager[2]	All	Windows: C:\Program Files (x86)\DNASTAR\License Manager

		Mac: /Applications/DNASTAR/ License Manager
Server License File (*lservrc*), Server License Manager [2], *Server Executables (_lserv, lservnt.exe*)	All	Windows: C:\Program Files (x86)\DNASTAR-LicenseServer\ Server Mac: ~/Library/DNASTAR-LicenseServer
Standalone & Trial Licenses (*.license)*, License Server Client License (*.lshost)*, Key Server Client License (*.keyhost*)	All	Windows: C:\Program Data\ DNASTAR\Licenses Mac: ~/Library/Application Support/ DNASTAR/Licenses
Preferences	Protean 3D, Navigator, SeqNinja (DNA*), GenVision Pro	Windows: C:\Users\<user>\DNASTAR Mac: ~/Library/DNASTAR
	ArrayStar	Windows: C:\Users\<User>\AppData\ Roaming\DNASTAR\ArrayStar
	All others	Windows: C:\Users\<user>\AppData\Local\ DNASTAR\ Mac: ~/Library/Preferences

✱ AppData is a hidden folder in Windows. To unhide the folder, go to **Organize > Folder** and **Search Options > View > Show Hidden files and folders**.

Research references

Bolshow A, McNamara P, Harrington RE and Trifonov EN (1991). "Curved DNA without A-A: Experimental estimation of all 16 DNA wedge angles." Proc. Natl. Acad. Sci. 88: 2312-2316. [View on PubMed](#).

Frier SM et al. (1986). "Improved free-energy parameters for predictions of RNA duplex stability." Proc. Natl. Acad. Sci. 83: 9373-9377.

Lukashin A and Borodovsky M (1998). "GeneMark.hmm: new solutions for gene finding." NAR Vol. 26, No. 4, pp. 1107-1115.

Trifonov EN and Sussman JL (1980). "The pitch of chromatin DNA is reflected in its nucleotide sequence." Proc. Natl. Acad. Sci. 77: 3816-3820.