# Table of Contents

**Introduction** .................................................................................................................. I

**Classification Display** ................................................................................................... 1-1

  Linked Windows .............................................................................................................. 1-7

**Views** ............................................................................................................................ 2-1

  Classification .................................................................................................................. 2-1
  Array Layout .................................................................................................................. 2-3
  Graph ................................................................................................................................ 2-4
  Bar Graph ....................................................................................................................... 2-5
  Ordered List ..................................................................................................................... 2-6
  Scatter Plot ...................................................................................................................... 2-7
  The Genome Browser Pop-Up Menu .............................................................................. 2-8

**Colors and Interpretations** ........................................................................................... 3-1

  Default colors .................................................................................................................. 3-1
  The “Colors” Menu ......................................................................................................... 3-3
    Color By Primary Experiment ...................................................................................... 3-3
    No Color ....................................................................................................................... 3-4
    Color by Parameter ...................................................................................................... 3-5
  The “Change Experiment Interpretation” window ......................................................... 3-6

**Gene Lists** ..................................................................................................................... 4-1

  The Gene Inspector Window .......................................................................................... 4-3
  A New Gene List ............................................................................................................. 4-7
  Selecting a Different List ................................................................................................. 4-9
  Selecting Many Genes to Make a List ............................................................................ 4-13
    The “Inspect Gene List” window .................................................................................. 4-16

**More Complex Lists** ..................................................................................................... 5-1

  Bookmarks ....................................................................................................................... 5-8

**Functional Classification** .............................................................................................. 6-1
Welcome to GeneSpring Lite.

This tutorial will walk you through some of the features of GeneSpring Lite version 3.0. This tutorial does not cover all features, but should give you a good introduction GeneSpring’s capabilities. This tutorial will concentrate on visualization in the first few lessons, so the results of the analyses are easier to interpret. The full tutorial takes about three hours to go through carefully. It is easier if you print this document out, in order to have it easily accessible.

The demo versions are fully functional, but time limited. The expiration date of every demo program is shown in the “About...” under the “Help” menu. Please contact Silicon Genetics via phone, (650) 591-4459 or email, info@sigenetics.com for more information.

The demonstration version contains some experimental data from an experiment performed on yeast in the Ron Davis laboratory in the department of Biochemistry at Stanford University, with help from Affymetrix and others. The data can be obtained directly from the experiment's web site (http://171.65.21.49/yeast/cellcycle.html). This experiment shows the expression level of most genes in yeast over the course of two cell cycles.

The Tutorial lessons are designed to be done in series, and frequently make use of analyses performed previously. They are saved automatically, and reloaded next time you start GeneSpring Lite.

GeneSpring Lite is a powerful application using a significant amount of memory. At least 64M of RAM is required, and 128M is strongly recommended. Try closing other applications if you are have problems.

If you have trouble with terminology, you can also try looking at the documentation available at http://www.sigenetics.com/GeneSpring/Documentation.htm in the GeneSpring User Manual, or Figure 1.
Figure 1 Names of the different areas in the starting screen of GeneSpring Lite

- **NAVIGATOR WITH FOLDERS**
- **GENOME BROWSER**
- **SLIDER**

**EXPERIMENT PARAMETER SPECIFICATION AREA**
Chapter 1 Classification Display

The setup program for GeneSpring Lite automatically puts GeneSpring Lite onto your desktop or into your windows “Start” menu. Left click the Start button, and find the GeneSpring Lite program in the GeneSpring Lite folder.

When you start GeneSpring Lite, a window will come up asking you to wait while data is being loaded. GeneSpring Lite is currently loading information about yeast, and the experiment you will be working with.

If you have an internet connection, a message window will come up telling you it is caching some files from the MIPS database. This allows you to use gene lists from MIPS for analysis. This message window will go away when the files are downloaded. If you do not have an internet connection, you will get an error message saying GeneSpring Lite was unable to download those files. If you or your company has a firewall and you have not set the firewall preferences under the Preferences menu you may also see this error message. This means you will not be able to use MIPS lists. This tutorial assumes you have an internet connection. Generally, downloaded files are cached, and will not need to be downloaded in the future.

After a few seconds, the “please wait” window should be replaced by a new window shown in Figure 1-1. This window is the main browser display. The main browser display will show the results of most analyses. It, and most other windows in GeneSpring Lite, can be resized at will.

Go to the “File” drop down menu, scroll down to the “New Genome” submenu and select the “Yeast” option. In a moment, the screen will change to show Yeast in an unsorted Classification view.
The main, black window in the upper middle of the screen is the genome browser. It generally displays all the genes in the genome, or at least those of interest.

At this scale it is difficult to make out individual genes. To make out individual genes, you can zoom in on this display. To zoom in, click the left button in one corner of the region you desire to view. Holding the left button down, and move the cursor to the opposite corner of the region you wish to view. The gray rectangle shows you the region you will be zooming in upon. An example of a region being zoomed in upon is shown in the Figure 1-2.
Figure 1-2 How to “Zoom In” on a section

When you release the button, the outlined rectangle will zoom out to fill the whole screen, as displayed in Figure 1-3. Do not worry if you have zoomed into a slightly different area…it does not matter for the purposes of this tutorial. You can zoom out using the Zoom Out button.
Now the genes are much more obvious.

The colors of the genes are determined by the experiment. Genes strongly over-expressed are drawn in red; genes strongly under-expressed are drawn in blue. Genes for which you have very good data are plotted in bright colors; genes for which the experiment produced poor data are drawn in dark colors, so they do not stand out as strongly. The color bar on the right matches colors to relative expression levels.

Since this is a time series experiment, the expression level, and thus color, is determined by time in the experiment. The time used for the colors at the moment is the 0 minute time point. This time is written in black just below the genome display window, on the left. You can change this time by moving the slider on the bottom of the display. Do so now. Notice, the colors and time change accordingly. GeneSpring Lite interpolates for time points between experiments.
The picture in the lower right of the screen has changed. This picture (also taken from the Stanford web site, http://171.65.21.49/yeast/cellcycle.html) is the yeast culture from the nearest time point to the point currently being displayed in the genome browser. Generally, this picture can be anything associated with an experiment to help you remember what was going on with that experiment. You can remove this picture from the screen to make more room by going to the “View” drop down menu and selecting the “Visible” submenu. Deselect “Picture” from the options presented.

Selecting the “Animate” box will make time advance automatically.

There is a gray colored rectangle which looks like a gene in the genome browser in Figure 1-4 (top line, 14 genes from the right). This is a genomic element for which the experiment did not produce data. These are generally non Genes like centromeres, tRNA or rRNA.

Zoom in further. Select a thin box as you did before, encompassing several genes, and zoom in. When there is space, the systematic name (and common name, should there be one) of the genes will be displayed, as in Figure 1-5. You may need to zoom in several times.
Figure 1-5 The physical position display at high magnification, displaying gene names

Experiment with zooming in and out. Clicking the “Zoom Out” button or menu (under “View”) will zoom out by a factor of two; clicking the “Zoom Fully Out” menu (also under “View”) will return you to the full genome view. You can also find these menu choices by right clicking in the black genome browser.

If you click a gene’s rectangle, it becomes white. This indicates the gene has become selected. Clicking elsewhere will deselect the previously selected gene, and, if you are choosing another gene, you will select that gene. You can select multiple genes by holding down the shift key while selecting. You can also deselect a gene by clicking it again. You can select all genes in a region by holding down the shift key while creating a rectangle, as is done for zooming, when you finish all the genes passing through the rectangle will be selected. You can use selected genes to make a gene list; this will be discussed later.

Congratulations. You can now navigate around the physical position display. All the displays in GeneSpring Lite can be navigated in the same way.
1.1 Linked Windows

A new feature in GeneSpring Lite 3.0 is the New Linked Window. With this feature you can open a new window linked to your current window. You can change the view in the second window to another option, such as graph. You can use this feature to look at the same list at different time-points simultaneously. In Figure 1-6, the highlighted genes (from a list of genes similar to CLN1) is simultaneously displayed in the Classification, Array, Graph, and Bar Graph views.

Figure 1-6 Linked windows

The linked windows can also be used to view different experiments.
Chapter 2  

2.1 Classification

For those genomes about which you have some knowledge of the genes, the Classification View can be used to divide up the genes into sensible categories. The knowledge can come from a list of classifications, one per gene, or it can come from a set of gene lists. In the latter case, a gene can be in more than one classification.

When you initially choose “Classification” from the “View” drop down menu, you will see a “no classification chosen” message at the top of the Genome Browser, just like Figure 1-1. To choose a classification, open the Gene Lists folder by clicking it. Position your cursor over the folder (with sublists) you would like to use as classifications and right click. A pop-up menu will appear. Select the “Use as Classification” option.

In a moment the screen will update to show the classifications.
You may want to enlarge your screen to make the classification more easily seen. You can do this by double clicking on the blue banner bar, or by dragging the edges of the screen.

**Figure 2-2 The Classification Display**

GeneSpring Lite 3.09f yeast Genes: all genomic elements

- **CELL GROWTH, CELL DIVISION AND DNA SYNTHESIS (787 ORFs)**
- **CELL RESCUE, DEFENSE, CELL DEATH AND AGEING (354 ORFs)**
- **PROTEINS, GENES, DNA (proteins are not localized to the corresponding organelle)**
- **CELLULAR COMMUNICATION/SIGNAL TRANSDUCTION (126 ORFs)**
- **PROTEIN LOCALIZATION (proteins are localized to the corresponding organelle)**
- **CLASSIFICATION NOT YET CLEAR-CUT (148 ORFs)**
- **ENERGY (246 ORFs)**
- **INTRACELLULAR TRANSPORT (450 ORFs)**
- **IONIC HOMEOSTASIS (121 ORFs)**
- **METABOLISM (1044 ORFs)**
- **PROTEIN DESTINATION (539 ORFs)**
- **PROTEIN SYNTHESIS (346 ORFs)**
- **RETRORTRANSPORTERS AND PLASMID PROTEINS (113 ORFs)**
- **TRANSCRIPTION (744 ORFs)**
- **TRANSPORT FACILITATION (304 ORFs)**
- **UNCLASSIFIED PROTEINS (2536 ORFs)**

Time: 0 minutes
Magnification: 1

Copyright 2000 Silicon Genetics
2.2 Array Layout

Another way to look at your data is the “Array Layout” in the “View” menu. The first time, you will get a message stating no array is selected. To select an array, go to the Navigator and open the Array Layout folder (currently at the bottom of the list). Choosing Pat Brown’s yeast Layout will result in a screen like this:

![Figure 2-3 Array Layout](image-url)
2.3 Graph

You can view experimental data as line graphs or histograms. Generally, the vertical axis is relative gene expression, and the horizontal axis is experiment type. If experiments are repeated, error bars can be plotted. If there are several categories of experiments, there can be several lines per gene.

![Graph Layout](image)

Figure 2-4 The Graph Layout

This display allows you to visualize one experiment or a set of experiments by plotting the relative expression of each gene over the experimental parameter, such as time.
2.4 Bar Graph

The “Bar Graph” view is similar to the “Graph View”, except bars are used instead of lines.

Each experiment point has a set of bars. Each bar in a set corresponds to a gene. The height of the bar is proportional to the expression value for that gene in that experiment.

All of the usual zooming, coloration, animation, movement (panning), selection and inspection commands work in the Bar Graph view.

Figure 2-5 The Bar Graph display

Yeast cell cycle time series (no 90 min)

Relative intensity

Expression

Trist

time (minutes)

0.0 20.0 40.0 60.0 80.0 100.0 120.0 140.0 160.0

0.0 1.0 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0

0.0 0.2 0.4 0.6 0.8 1.0 1.2 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0 6.5 7.0 7.5

time 0 minutes
Magnification : 1

Animate
Zoom Out
2.5 Ordered List

You can also use the “View” menu option “Ordered List” to view your data. If you have a list of genes in an order indicated by numbers associated with the genes, you can display them in that order. Vertical green lines proportional to the associated number are attached to each gene. The numbers can be entered by you, or a by-product of GeneSpring Lite analyses, like similarity level or distance from the beginning of a gene or a promoter. The figure below shows a zoomed view of an ordered list.

![Image of an ordered list view](image-url)
2.6 Scatter Plot

Similarly, choose “Scatter Plot” from the “View” menu, producing the picture below. (If need be, reset your colors to “Color by Primary Experiment”.) Detail on scatter plots is out of the scope of this tutorial, however, all the normal manipulations (zooming, selecting, double clicking) work, and you can graph any experiment on any axis, with either normalized, raw, or control measurements.

One other method worth mentioning involves classifying each gene into one of a number of categories. It is possible to assign each gene to a category. Then each category will be assigned a color, and the genes are plotted through that color. This allows the use of proprietary or hypothesized functional information, or results from other types of assays.

Return now to coloring by the experimental expression levels by choosing the “Color by Primary Experiment” option from the “Colors” menu. Return to the “Classification” display found in the “View” menu. This should return the screen to a display similar to the original display at the start.

Figure 2-7 The Scatter Plot display
of lesson one, except centromeres, tRNAs, etc. are not being shown, because you are displaying the “all ORFs” list rather than the “all genomic elements” list.

2.7 The Genome Browser Pop-Up Menu

There are many aspects of the Genome Browser you can change. Right click in the Genome Browser to see the menu shown in Figure 2-8.

![Figure 2-8 The pop-up menu](image_url)
Chapter 3  Colors and Interpretations

3.1  Default colors

You can change the default colors used by GeneSpring Lite through the “Preferences” box under the “Tools” drop-down menu.

![GeneSpring Preferences](image)

**Figure 3-1 The Preferences Box**

Select the + sign next to Show Color details.
Use the drop-down menus to select new colors for the default settings. Click the “OK” button to accept your changes.
3.2 The “Colors” Menu

You have three options under the Color menu.

3.2.1 Color By Primary Experiment

“Color by Primary Experiment” is the default colors shown throughout most of this manual (see Figure 3-3). This command initiates the coloring scheme allowing you to view the expression levels of the genes at differing points within the primary experiment. To color the genes by primary experiment, go to the “Colors” menu and select the “Color by Primary Experiment” command. To change the coloring scheme, go to the same menu and select a new color scheme.

This coloring scheme allows you to view the expression levels of the genes at differing points within the primary experiment being viewed.
The coloring scheme of the display is shown in the Color Bar on the right, in Figure 3-3, “The “Colors” Menu”. The color bar indicates what the colors mean. The vertical axis indicates the gene’s relative intensity on a continuous scale; red indicates overexpression, purple average expression, and blue indicates underexpression. The horizontal axis indicates the trust or reliability of the experimental data for a particular gene, also on a continuous scale; a dark shade indicates a low confidence in the data regarding that gene, lighter shades indicate higher confidence in the gene’s data. This coloration has been chosen so your attention will not be distracted by low quality data. You can change the default settings of these colors through the “Preferences...” pop-up under the “Tools” drop-down menu.

In series experiments, the genes are colored by their expression level at the particular point in the series indicated by the scroll bar and noted in written form directly beneath the genome browser. The particular experiment illustrated in Figure 3-3 is a time series. In such cases it is the time, as the series experiment parameter, specified in the gray box at the bottom, directly beneath the genome browser. The time-point in the series shown may be changed by moving the slider in the scroll bar at the bottom edge of the browser display. As you do this, the colors of the genes change according to their expression at the new time point. The series variable can be made to automatically move forward by using the “Animate” option. This will scroll through all of the expression changes and therefore the color changes in the experiment. Times between experiments are interpolated for continuous data. For different experiments the parameter in the gray box might indicate the temperature, drug amount, nutrient supply or any other relevant parameter or combination of parameters.

3.2.2 No Color

“No Color” marks all genes in gray (see an example in Figure 3-4). Unlike in the other coloring displays there is no right-hand legend to indicate what the colors mean, because all the genes are colored exactly the same, at every point. No color can be useful when you want to look at all the genes equally without losing the underexpressed genes in the background.

This command initiates the coloring scheme allowing you to view all of the genes as equal. To color the genes in this manner, go to the “Colors” menu and select the “No Color” command. To change the coloring scheme, go to the same menu and select a new color scheme.
3.2.3 Color by Parameter

"Color by Parameter" requires parameters be set in the initial loading of the experiment or changed in the “Change Experiment Interpretation” under the “View” menu (see an example in Figure 3-11).

Some experiments have more than one measurement taken at a given time point. For instance, you could do a time series experiment both with and without the presence of a disease. Color by Parameter allows you to color code these lines by which category the measurements are in (such as diseased or healthy). This is different from all of the other coloring schemes (which are per-gene) because it is per-measurement.
3.3 The “Change Experiment Interpretation” window

Click the “View” drop-down menu and then select “Choose Experiment Interpretation...” to get the “Choose Experiment Interpretation” window.

The “Choose Experiment Interpretation” window allows you to change several aspects of your experiment. There are four formulas used to determine how the gene’s relative intensity data is plotted on the y-axis. In the “Choose Experiment Interpretation” window, when you click the arrow at the right side of the white box labeled “Mode”, a menu appears with the four formulas to choose from:

- **Ratio [signal/control]** is the default display for graphing relative intensity values on the y-axis. It displays the normalized expression value as the relative intensity viewed on the y-axis, without modifying that number for visual reasons. Using this algorithm to display your data means the y-axis can be numbered from zero to infinity.

- **Log ratio [log base 2 of ratio]** makes under-expressed genes visually take up as much space as over-expressed genes. The difference between this function and [log10ratio] is scale. Using a logarithm of the ratio can lose detail in the way data is displayed, especially for genes that are highly over- or under-expressed. Using this algorithm to display your data means the y-axis can be numbered from negative to positive infinity.

- **Log ratio [log base 10 of ratio]** makes under-expressed genes visually take up as much space as over-expressed genes. The difference between this function and [log2ratio] is scale. A logarithm has the most detail around “normal” expression levels, but it allows you to easily look at genes that are highly over- or under-expressed. Using this algorithm to display your data means the y-axis can be numbered from negative to positive infinity.
• Over/under expression amount \[+/- \ (\text{ext}(\abs{\log(ratio)}) - 1)\] makes under-expressed genes visually take up as much space as over-expressed genes. The difference between this and the two log ratio options is that both logarithms lose some detail, especially away from “normal” expression levels, but this algorithm retains this detail. Using this formula to display your data means the y-axis can be numbered from positive to negative infinity.

You can also choose another option in the “time” box in the center of the “Choose Experiment Interpretation” window.

For example, the “non-continuous” option will result in the figure below, experiment points unconnected by lines.

![Figure 3-6 The “Like CLN1” list, displayed as non-continuous](image)

Other options presented for parameters include “Replicate”, for many experiments averaged together and “Color Code” which permits you to color by parameter, in this case, minutes. To use the color code you must rename the interpretation and save it. The resulting picture (in graph view) looks like Figure 3-7.
Colors and Interpretations

GeneSpring Lite

The “Change Experiment Interpretation” window

Figure 3-7 The “Like CLN1” list, displayed as non-continuous and colored by parameter

Figure 3-8 Saving an interpretation
To save, change the default name from “Default Interpretation” and click the save button. You can save several interpretations relevant to the same experiment. The interpretations are accessible under the “Experiments” folder in the Navigator. Interpretations are sorted by experiment, click the plus (+) icon next to your experiment name to see the interpretations available.

For now, return the “Choose Experiment Interpretation” to its original configuration.

For an experiment with many parameters, you can reconfigure the interpretation to make some aspects of it more obvious. For example, with four different parameters, such as in the fictitious “Yeast Extraterrestrial Studies” experiment, you will get a Change Interpretations box like Figure 3-10.
### Continuous Element
- A continuous variable is when each parameter-value of the experimental parameter exists on a continuum with the other parameter-values in that experimental parameter, rather than as discrete points. More loosely, each parameter-value is related to the parameter-values on either side of it and it is sensible to draw lines between them. In this example, Kryptonite Concentration (ppm) is left as the Continuous Element, so the lines are drawn from the least to the most concentrations.

### Non-Continuous Element
- A non-continuous (or set) variable is when each parameter-value of the experimental parameter exists independent of each other, as discrete points. Making the parameters “Variety of yeast” into a Non-continuous element, splits the display into two parts, one for yeast type A, and one for yeast type B.

### Replicate
- A parameter defined as a replicate is graphically a hidden variable; no visual distinction is made based upon this parameter or its parameter-values. The “Test Repeat Number” is defined as a replicate, so all repeats are averaged together. This saves clutter on the screen.

### Color Code
- A color code is used for experimental parameters whose parameter-values exist independent of one another, but not unrelated to one another. When a color code is graphed, separate gene points or gene lines are drawn; one per each gene per each parameter-value defined by a color. Defining “Andromeda strain infection” as the color code means the differences in behavior are more easily seen, the infected strains are drawn in red and the healthy ones in blue.
Figure 3-11 Graph display with a non-continuous element and colored by parameter

As each gene is shown twice, once for the infected strain and once for the healthy, the Gene Inspector will bring up the following screen.
Figure 3-12 The “Gene Inspector” window
Chapter 4  Gene Lists

Making and comparing lists is one of the most powerful tools within GeneSpring Lite. In order to provide maximum flexibility there are many, many ways to make lists.

It is very important to give your lists sensible names. If you try to give a new list the same name as an already existing list, GeneSpring Lite will bring up an error message asking you to rename the list. Resist the temptation to name the lists sequentially (List1, List2, List3) as while sequential naming convention may be perfectly obvious now, but in two weeks with 130 new lists in your Navigator you won’t remember which is which easily. It is also helpful to give your lists short names, as they will be more easily viewed in the Navigator.

Lists are accessible under the “Gene Lists” folder. Lists are sorted alphabetically each time you start the program. Lists have an icon of DNA-on-a-page next to their names.

Once a gene is visible, you can double click it to get further information about the gene. To ensure consistency, use the “Find Gene” menu command (in the “Edit” menu). Type “cln1” in the “Find Gene” window, and press enter (or click “OK”). This will center the display around the cyclin 1 gene, as shown in Figure 4-2.
Both the systematic name (YMR199W) and the common name (CLN1) are visible. Double click the rectangle representing the CLN1 gene (white in Figure 4-2 because it is already selected). This will bring up a new window—the “Gene Inspector” window—telling you more about the selected specific gene. A picture of the window is shown in Figure 4-3.
4.1 The Gene Inspector Window

The black graph in the center of this window shows a graph of the relative expression of cyclin 1 plotted against time for this experiment. There were two cell cycles, which are quite visible from the two peaks shown. The 90 minute time point is missing from this experiment because of experimental problems.
The text at the top of the box contains a brief description of what is known about this gene:

- common name
- ID number in the Saccharomyces Genome Database
- signal strength in the current experiment (the median of the data points for the gene)
- brief description of the gene (from the Saccharomyces Genome Database)

At the bottom right, are buttons causing a browser to come up on your computer, searching respectively the Saccharomyces Genome Database, the MIPS yeast database, or GenBank. Try clicking these—they will bring up information from these databases for the CLN1 gene. After finishing with the browser windows, you may close the browser normally.

Perhaps the most interesting feature in this window is the “Find Similar” button. This will find all of the genes in the current genome whose expression profile is similar to the CLN1 profile. Similar is defined as having a correlation coefficient of at least 0.95 (or whatever number you type in the “minimum correlation” box). Click “Find Similar”. This brings up the window shown in Figure 4-4.

![New gene list (117 genes)](image)

**Figure 4-4 The New Gene List window containing the list of genes meeting the correlations specified in Figure 4-3**

The list of genes shown in Figure 4-4 is a list of the 118 genes having a similar expression profile to CLN1. All of this was done on the fly, in a fraction of a second. You can also do analyses across experiments, or you can draw your own gene profile. For more information see Chapter 7, “More Powerful Correlations”.

As well as this window, another window will come up advising you some files from the MIPS yeast database are being downloaded. These are the actual gene lists from MIPS. They are cached, so you will not be bothered by having to download them in the future. You do not have to wait for this if you do not wish to…it is just preparing additional functionality.
Once these lists have all been downloaded (this will take a few minutes with a good internet connection), GeneSpring Lite will compare each list to the list just created, to see if there are any statistically significant matches. That is, GeneSpring Lite is seeing if anyone it knows of has made a list of genes substantially similar to the list of genes which you have just made. These matches will then be displayed, and can help give you some idea of the functionality of the gene list you have just created. These lists will be displayed at the bottom of the window just show, together with a P-value.

The P-value shown is the probability such a resemblance of lists is accidental, taking into account all the lists compared. Not surprisingly, the lists shown are associated with the cell cycle.

You have the possibility at this point to rename this list by changing the line in the “name” box. Do not do so for the purposes of this tutorial; just click “OK”. This makes a new list called “like YMR199W (CLN1) (0.95)”. 

Figure 4-5 The New Gene List window including the “Similar lists” box
The creation of the new list has several effects:

- The list is saved to disk as an HTML file, containing information about the genes in the list, and the method used to generate them. When saved, it looks like Figure 4-6. Saved files will be automatically reloaded when you start GeneSpring Lite again.

- The genes shown in the main screen are restricted to just that list.

- The list appears in the menu system (to be shown later).

Figure 4-6 Format of the CLN1 Gene List as it would be saved in a HTML file (this figure only illustrates the first page of that file)

The restriction of displayed genes can be demonstrated by returning to the main screen, and zooming fully out (either with the “Zoom Fully Out” menu choice, or by pressing on “Zoom Out” several times). You may close the Gene Inspector window at this stage. Returning to the main screen will result in a screen like Figure 4-7.
4.2 A New Gene List

Figure 4-7 The physical position display showing a restricted gene list

Fewer genes are visible, and they all have roughly the same color, except for CLN1 which is still highlighted in white. Try varying the time (using the slider along the bottom or clicking the animate box), and noticing all the genes change colors together. This is expected, as they were selected by their similar expression patterns.

The similarity of the expression patterns can be seen explicitly by changing the view type. Choose the menu option “Graph” under the “View” menu. This will convert the main display into a graph of relative expression level against time, for the 118 genes behaving like CLN1.

If needed, you can increase the width of the window for clarity. To resize the window, move the cursor to the edge of the window, wait until the cursor becomes a double-headed arrow, click, and drag the border to where you wish it to be. You can also double click the blue banner at the top of the window to resize it to fit your screen. You can also alter the proportions of Navigator to Browser by grabbing and dragging the grey line between them.
Figure 4-8 The graph display of the “like CLN1” list’s relative expression levels

All the genes peak and dip at the same places. The CLN1 gene is still highlighted in white—the other genes are shown in the colors corresponding to their expression at time 119.6 minutes. Again, you can change the time with the scroll bar at the bottom of the screen. The green vertical TimeLine marks off the time chosen by the scroll bar’s slider.

The name of the experiment (Yeast cell cycle time series (no 90 min)) appears in the upper right corner of the Genome Browser. The name of the selected gene (CLN1) appears just below it.

You can change the selected gene by clicking another gene’s graph line. Double clicking any gene’s graph will bring up the Gene Inspector window this lesson started with, for the selected gene. Clicking in a place where there are lots of nearby genes will be considered ambiguous, and will not result in a gene being selected. You can solve this problem by zooming in on a gene as done in a previous lesson.
4.3 Selecting a Different List

You are currently displaying a gene list you created. You can see which gene lists are available by selecting the “Gene Lists” icon in the Navigator at the left of the main screen. Do this now. The “Gene Lists” section will expand to show the currently available lists.

The windows have been resized for easy viewing of the list names. You can use the slider at the bottom of the Navigator to look at all of the names.

The list currently being displayed is highlighted in the Navigator. There is another folder inside the “Gene Lists” folder containing the MIPS gene lists. At this point you will change which genes are being viewed. Click the “all ORFs” list. This produces a screen like Figure 4-10. It may take a second or two for the graph to update, as it is drawing so much.
Gene Lists

GeneSpring Lite

Selecting a Different List

Chapter 4-10

What a mess! Lines everywhere! Part of the aim of GeneSpring Lite is to avoid having to look at such things. Next, look at some MIPS lists. Click the “MIPS - Saccharomyces Cerevisiae functional categories” folder.
You can see more of the Navigator by using the scrollbar at the right or bottom, or by resizing the Navigator. You can resize the Navigator by clicking the gray vertical bar between the Navigator’s vertical scroll bar and the browser, and then dragging the cursor to where you want the border to be.

You are now seeing some more folders, as well as some lists (with the DNA-on-a-page-of-text icon). Select the seventh of these lists, “ENERGY”.

Figure 4-11 Graph display of all ORFs, with the MIPS folder in the Navigator open
GeneSpring Lite is now displaying those genes known to be involved with energy production. Unsurprisingly, they are fairly constant in this cell growth experiment. Nevertheless, there are a few especially active near the start. Next, you will make a list containing just those genes.
4.4 Selecting Many Genes to Make a List

If you want to make a list of all the genes passing through a certain area of the Genome Browser, you can do so with a combination of the cursor and the shift key. First, click in a black section of the browser window to deselect CLN1, which does not appear in the “energy” list, but is still selected. (In the upper right corner of the genome browser is the name of a single selected gene.) Secondly, *hold down the shift key* and drag a rectangle (as if you were zooming) over the genes which you wish to select (upper left), as shown in Figure 4-13. Let go of the shift key only *after* you have completed your rectangle.

![Figure 4-13](image)

*Figure 4-13 Selecting the genes in the “Energy” list that are active at the beginning of the experiment*

When you release the cursor, the genes passing through this rectangle will be highlighted.

Alternatively, you can use the “Make Gene List” command in the Tools drop-down menu to create a list, see Chapter 5, “More Complex Lists” for more information.
Figure 4-14 The genes in the “Energy” list selected in Figure 4-13 are highlighted to indicate they were selected.

You can now make a new list of these genes by right clicking the (black) genome browser, and choosing the option “Make List From Selected Genes”. This comes up with a “New Gene List” window, containing the genes highlighted in white. You have just performed a rather sophisticated analysis.
Change the name to “initially high” by deleting the selected text and typing in “initially high”, and click “OK”. Now you have added another gene list to the choices in the Navigator, and the main screen will look like Figure 4-16.
Now, click elsewhere in the black genome browser to deselect all the genes, returning them to their default colors.

4.4.1 The “Inspect Gene List” window

The choice of genes displayed in the genome browser can be limited by choosing a gene lists. Creating gene lists can be done in a number of different ways. The “Inspect Gene List” command, in the “Tools” menu, opens the “Inspect Gene List” window. The “Inspect Gene List” window displays the common and systematic names of all the genes in the gene list currently being displayed in the genome browser. Listed in the bottom of the window are the names of other lists significantly similar to the displayed list. See Figure 4-17. In this window you may utilize any of the commands listed on the right-hand buttons, currently “Print” and “Copy to Clipboard”.

Figure 4-16 Graph of the genes in the “Initially high” list

*GeneSpring Lite*  Selecting Many Genes to Make a List

*Gene Lists*
Figure 4-17 “Inspect Gene List” window

This window results from the “Inspect Gene List” command, found in the “Tools” menu. Alternatively, you can right-click the list and select the “Properties” option. The name of your inspected gene list is in the banner bar and is the first list listed in the “Similar lists” box. In Figure 4-17, the list of genes being viewed is the “like YMR199W (CLN1) (0.95)” list. The lower left-hand box, titled “Similar lists”, contains the lists GeneSpring Lite knows of resembling the inspected list. *Resembling* means the lists contain a statistically significant number of overlapping genes. How statistically significant the similarities are is given in the left column of the bottom box. Statistical significance is listed as the P-value (the probability of a false positive) for each of the lists in the
right column. Both of the left boxes have scroll bars at their far right edges, to let you move up and down the list of items displayed in these boxes. In the “Inspect Gene List” window you may utilize any of the commands listed on the right buttons, or double click any of the listed genes or gene list for further details.

Close the “Inspect Gene List” window by clicking the “Cancel” button or the small “x” in the upper right corner.

Finally, select the “like YMR199W (CLN1) (0.95)” list in the Navigator to get back the first list you made.

Congratulations! You can now manipulate lists, and have done some simple but powerful analyses.

Clean up after this lesson by closing the window titled “Gene Inspector YMR199W” if you have not already done so. Do this by clicking in close box in the upper right hand corner of the Gene Inspector window or by clicking on the “Cancel” button.
Chapter 5  More Complex Lists

In Chapter 4, “Gene Lists”, you made a list of genes like CLN1. This was a very easy and quick operation, but there was not much flexibility. In this lesson, you will make lists in some other ways.

Choose the menu option “Make New Gene List...” in the “Tools” menu. The following window will come up.

![Figure 5-1 Make New Gene List window](image)

You will use this window to make a list of all the genes with a strong signal and are over-expressed at the 40 minute time point in the “Yeast cell cycle time series” experiment. If you have already loaded your data into GeneSpring Lite, you can use any experiment. In this window you will create a list of restrictions, and GeneSpring Lite will create a list satisfying all of those restrictions. At the moment, the status box at the top of the window says you are choosing genes from the “all ORFs” list, and of those 6188 genes, all 6188 pass the (nonexistent so far) restrictions. Later, you can change the gene list used at the start by opening the “Gene Lists” folder at the left side of this window.
Create a new signal strength restriction by using the navigator. Open the Experiment folder within the “Make New Gene List” window by clicking its icon at the left. In the default version there are two experiments, select the “Yeast cell cycle time series” experiment in this folder. Right click the “Yeast cell cycle time series” experiment, and choose the “Add Signal Strength Restriction” option from the pop-up menu. This will bring up a small window as shown in Figure 5-2.

![Figure 5-2 The Signal Strength Restriction window](image)

Change the entry in the “Strength” box to 200. The units are the units of the raw data coming out of the instrument - in this case, the units from the Affymetrix scanner. Click “OK”, and the new constraint will appear in the “Make New Gene List” window, as shown in Figure 5-3.
Figure 5-3 The Make New Gene List window with a signal strength restriction of 200 for the yeast cell cycle time series experiment

Now only 3465 of the 6188 genes pass the restrictions. Which means only 3465 genes have a signal strength of at least 200 in the “Yeast cell cycle time series” experiment. For this experiment, the signal strength is the median of the readings for each gene from the raw data. For two color experiments, the signal strength generally comes from the control signal.

Add in another constraint—this time on the relative expression level at the 40 minute time point. In the navigator, open the “Yeast cell cycle time series” experiment with a left click on the gray plus icon. The navigator will resemble the one pictured in Figure 5-3. Right click the “time 40.0 minutes” sub-experiment, and choose the “Add Expression Restriction” option in the pop-up menu that appears. This will bring up a screen like Figure 5-4.
You want over-expressed genes, so enter “1.5” in the “Minimum” box, and leave the maximum box blank. Click “OK”. This will add a new constraint in the “Make New Gene List” window, as shown in Figure 5-5. The window displayed in Figure 5-5 has been manually made wider for clarity (to make your window wider, just select the edge and drag it, while holding the mouse button down).
Now there are only 90 genes passing restrictions. You could go on and add more restrictions, change existing restrictions or remove them, but this will be enough for this tutorial. Click the “Make List” button at the top of the window. This brings up a window listing the genes, and letting you set a name. This is the same “New Gene List” window you encountered in Chapter 4, “Gene Lists”.

Figure 5-5 The Make New Gene List window with both a signal strength restriction and an expression level restriction
Figure 5-6 A Make New Gene List window containing all of the genes that met the restriction stated in the Make New Gene List window shown in Figure 5-5

Change the name to “High 40”, and click “OK”. This new list will now be shown in the main browser. To make this obvious, return to the main browser, and choose the “Graph” menu option in the “View” menu. A screen like Figure 5-7 should result.
When you are working with more experiments, you can do more complex correlations finding which genes are similar across multiple experiments. This can be done in conjunction with weighting of experiments, time shifts in comparisons, and the restrictions just demonstrated. Such analyses are not shown in this tutorial.

Clean up after this lesson by closing the “Make New Gene List” window.
5.1 Bookmarks

If you ever need to pause in the midst of an analysis, you can use the “Bookmark” feature to hold your place. If, for example, you were looking at the “before CLN1” list suddenly decided to make the following changes:

- change the interpretation to Log ratio [log base 2 of ratio]
- change the interpretation to non-continuous
- change color to color by parameter
- select YOL011W

You could save all of these changes by clicking “Bookmark” and selecting “Save Bookmark”. This brings up the following (hopefully familiar) dialog box.

![Save Bookmark Window]

Name your bookmark and click save. If you have not changed the default settings, the bookmark will be saved under C:\Program Files\SiliconGenetics\GeneSpring\data\yeast\Bookmarks\BookmarkName. The bookmark saves all your current settings. You can change the settings again or shut down your computer. Later, when want to resume at this spot, go to “Load Bookmark from File”, select your bookmark and click “Open”.

Copyright 2000 Silicon Genetics
Chapter 6 Functional Classification

It is possible to display genes according to some classification. This is the default way to display unsequenced organisms. Generally, the classification can come from either proprietary data that has assigned a label to each gene, or it can come from a set of lists, such as the MIPS lists. For the purposes of this tutorial, the MIPS lists will be used for classification.

Position the cursor over the “MIPS—Saccharomyces Cerevisiae functional categories” gene list in the navigator, and click the right button, getting a pop-up menu as shown in Figure 6-1.

Figure 6-1 Selecting a set of lists to use as classifications
Select the “Use as Classification” option. This makes the gene lists in that folder the classifications for the genes being displayed. The result should look like Figure 6-2.

![GeneSpring Lite](image)

Figure 6-2 Display of all ORFs using the MIPS lists as classifications

This is a bit of a mess, since there are too many categories to fit nicely onto the screen. Make the screen bigger by dragging the border outwards. In particular, make the window taller.
Each gene is divided up according to the gene lists in the MIPS folder, with the genes listed below their classifications. There are too many genes to see the genes clearly at the moment. Select the gene list “like YMR199W (CLN1) (0.95)” by clicking it in the navigator.
Here you can more easily see the individual genes as small, distinct rectangles. You can zoom in to see some genes in more detail. The gene names will appear when there is enough space. Here you see the genes of yeast categorized by the MIPS functional categories into several groups. It is possible for a single gene to be in more than one group. Genes not mentioned in any of the gene lists end up in the “unclassified” section on the bottom. The “UNCLASSIFIED PROTEINS” classification is a list of genes MIPS has actively specified as unclassified. Some classifications like “PROTEIN SYNTHESIS” contain no genes. It is not surprising, given the source of the list, that there are many “CELL GROWTH, CELL DIVISION AND DNA SYNTHESIS” genes.

With the full version of GeneSpring, you can make new classifications like this by simply making a directory, and putting the lists you want to classify by into that directory.
Chapter 7 More Powerful Correlations

In Chapter 4, “Gene Lists”, you found all the genes behaving like CLN1. This is very powerful, but sometimes you will want to look for genes with other shapes. Or, you might want to find all genes whose behavior is like CLN1, except offset in time.

7.1 Finding Offset Genes

Begin in the same way as “Gene Lists”, zoom in on the “CLN1” gene using the “Find Gene” command in the “Edit” menu and double click the CLN1 gene to get the “Gene Inspector” window. This time, enter “-10.0” in the box labeled “time offset”. The screen should now look like Figure 7-1.

Figure 7-1 The Gene Inspector window with a time offset of -10
Click “Find Similar”. GeneSpring Lite will now look for genes with a similar shape, except starting ten minutes earlier than CLN1. Their graphs will be shifted to the left by ten minutes.

![New gene list (10 genes)](image)

**Figure 7-2** The New Gene List window containing the genes meeting the correlations set in the Gene Inspector window as shown in Figure 7-1

GeneSpring Lite found ten such genes. Change the name to “before CLN1” and click “OK”. Close the “Gene Inspector YMR199W” window, as it is not needed any more. Change the main browser window to graph mode (“Graph” under the “View” menu) so you can see what you have produced. Indeed, you have produced a set of genes with shapes similar to CLN1 (which peaks at about 20-30 minutes and 100-110 minutes), except ten minutes to the left. This can be used if you want to see what genes might have triggered CLN1 production.
Figure 7-3 Graph of the new gene list, “before CLN1”
7.2 Pseudo Genes

Now you will do something more powerful still … you will draw a picture of a profile, and find genes with a similar profile.

First draw the picture. Go to the “View” menu, and choose “Show Pseudo Genes”. A horizontal green line appears on the screen. This green line is the pseudo gene whose shape you will modify.

Figure 7-4 The Graph display showing a pseudo gene
To change the shape of this new gene, click where you want the graph to go, while holding down the control key. For instance, click in the middle of the screen with the control key held down, and you will get Figure 7-5.

![GeneSpring 3.00 yeast Genes: before CLN1](GeneSpring Lite)

Figure 7-5 Altering the profile of the pseudo gene

The green line has gone up at the 70 minute time point to where you placed the cursor. Now, draw a gene similar to the one in Figure 7-6, by clicking the cursor at the desired height for each time point. In this example, try to find genes active early in the experiment, but at no other time.
The entire line beyond the 20-minute time point is below the relative intensity level where the pseudo-gene started. Now, find similar genes in the same way as you did for CLN1 in “Gene Lists”. Double click the pseudo gene (without the control key this time), and you will get the “Gene Inspector” window.

In the “Gene Inspector” window you can save your drawn gene (or any gene profile) by clicking the “Save Profile” button in the upper left hand corner. This will bring up the “Choose Profile Name” box. The default name for such profiles are “Gene name in Experiment name”.
Figure 7-7 The Gene Inspector window displaying the pseudo gene created in Figure 7-6

Click “Find Similar”, and GeneSpring Lite will find roughly eight genes from all of the genes in this experiment. The exact number will vary depending upon how you drew your picture.

If you want only the genes in your current list to be considered, click the empty circle next to “Consider displayed Genes” rather than accepting the default, “Consider All Genes”.
Figure 7-8 The New Gene List window containing a list of genes similar to the pseudo gene (created in Figure 7-6)

Clicking any of the individual genes names in your new list will bring up a new “Gene Inspector” window for that gene. If clicking “Find Similar” for your drawn gene had listed only one gene, try double clicking that gene (in the list) and trying the “Find Similar” again.

Click “OK” in the New Gene List box. In the main screen you will see your genes; they are very close to the green line.
With a combination of drawing genes, and selecting rectangles, you can make very sophisticated queries with very little effort, and instantly see the results.

The Pseudo Gene feature also works in the Bar Graph and Scatter Plot views.
7.3 Conclusion

Congratulations. You have finished the GeneSpring Lite tutorial. While the tutorial has not covered many significant sections of GeneSpring Lite’s functionality, it will have given you an idea of the powerful and easy-to-use analyses and visualizations available in GeneSpring Lite.

Close GeneSpring Lite by clicking in the close box in the upper right hand corner of the main window.
Index

A
animate 1-5
B
Bar Graph view 2-5
Bookmark 5-8
C
contact Silicon Genetics I
D
drawn gene 7-6
drawn gene. See also pseudo gene
E
eperiment parameter
condition 3-10
continuous 3-10
non-continuous 3-10
replicate 3-10
Expression Restriction 5-3
F
Find Gene 4-1
Find Similar 4-4, 7-2
Functional Classification 6-1
G
gene colors
over-expressed 1-4
under-expressed 1-4
Gene Inspector 4-2
GeneSpring requirements I
L
Linked Window 1-7
lists 4-11
M
Make List 5-5
Make List From Selected Genes 4-14
P
pseudo gene 7-4
pseudo gene. See also drawn gene
P-value 4-5, 4-17
R
resize the navigator 4-11
S
Save Profile 7-6
select a gene 1-6
selecting several genes 4-13
Show Pseudo Genes 7-4
Signal Strength Restriction 5-2
T
TimeLine 4-8
U
Use list as Classification 6-2
V
View
Bar Graph 2-5
Z
zoom in 1-2
zoom out 1-6