MAGIC Tool Tutorial

The purpose of this tutorial is to provide the new MAGIC Tool user with a few simple exercises to help you learn how to discover what MAGIC Tool can do. In order to make the most of this tutorial, you must have internet access in order to download the MAGIC Tool software as well as the MAGIC Tool User's Guide. All MAGIC Tool resources can be found at the MAGIC Tool web site <www.bio.davidson.edu/MAGIC/>.

The first section of the User's Guide called "Getting Started" has the step by step directions for a beginner. However, the User's Guide assumes the user has data files already available. This tutorial will show you how to obtain some simple data files that can be used as a way learn about MAGIC Tool.

Tutorial

1. Go to the **MAGIC Tool** web site: <www.bio.davidson.edu/MAGIC/>.

2. Scroll down to the Download link. Click on the text and you will find the GNU license that is required of all open source software. If you do not intend to alter the software or sell it, you can scroll to the bottom and click on "I agree". From the resulting page, you will find a few options for downloading the software. Choose the format that is appropriate for you.

3. Go back to the MAGIC Tool main page and scroll down to find the **User's Guide**. This PDF file is designed to show you how to get started and go deeper into MAGIC Tool. The instructor's Guide provides more details about the algorithms used in MAGIC Tool. The PDF link called "Additional Math Background for Understanding Data Analysis" provides Math Minutes to explain the basic math behind DNA microarray data analysis.

4. Now download the **tiff files** that are the raw data. There are four levels of options. The simplest is the one grid set which is only one small section from a much larger microarray. The next level up is a four grid set of spots. The next level up has two similar types of data: Whole yeast genome (Y01 Version from Stanford) and Whole yeast Genome (ISB Version). These are much larger sets of spots from two different DNA microarrays. These are single experiments from a larger set of experiments. The final and most complex set of tiff files are the ones used by Joe DeRisi in the landmark paper examining the diauxic shift of yeast metabolism (*Science*. 1997. 278: 680-686).

5. The last two sets of files you will need are the **gene lists** (some times referred to as GodLists) and the **gene info files**. At this time, GCAT has compiled only gene info files for yeast, but other model organism files are under development. Gene lists identify each spot on the microarray so you will know which genes you are looking at on the chip. Gene info files contain the chromosomal location and functional information for each annotated gene in the genome. This information can be imported to enhance your interpretation and searching of the data.

6. Read the User's Guide to see the best way to launch MAGIC Tool. Note that software written in Java automatically request a limited amount of RAM from your computer. This is not a problem to

launch the program but since DNA microarrays generate large data sets, you should boost your RAM if possible. Therefore, we have created scripts to boost the allocation of RAM to MAGIC Tool.

7. Although it is not critical for the practice data since you are not interested in the biology behind the data, it may be a good exercise to download the excel files and the jpeg file to understand how the chips were printed. Understanding the chip's layout is critical to analyzing real data (either your own or the DeRisi set of data). Knowledge of the chip's layout is important for addressing the chip (telling MAGIC Tool how the chip was printed) so the software will know how the spots relate to the gene lists.

8. A good starting place is to work with the one grid set of data. Keep in mind that MAGIC Tool running on Mac OSX or Linux uses Unix and so you will want to eliminate any folders or file names that contain spaces. Check your file names and folders at this time to ensure no surprises later.

9. Follow the Getting Started directions using the one grid set of tiff files. In this project, you will grid, segment and explore the data. You cannot cluster single time points.

10. Perform similar steps for the four grid set of data. The only difference is the ability to apply the grid patter from one group of spots to additional groups of spots.

11. The next level of complexity is a whole-genome chip. You can choose either the Y01 set where genes were printed once or the ISB chp where each gene was spotted in duplicate (side by side). The duplicate printing allows you to examine consistency and to utilize the MAGIC Tool averaging capacity.

12. Finally, you can use the DeRisi tiff files to create a data set that has multiple time points. This is a way to generate an expression file that can be clustered by MAGIC Tool in several ways.

13. We also have pre-fab/hypothetical data sets for you if you don't want to endure the tedious steps necessary to generate your own time course data. You can use the 1 grid, four grid, or DeRisi expression files for use with clustering. Clustering is explained at increasingly detailed information in the User's Guide, Instructor's Guide and the "Additional Math Background for Understanding Data Analysis" file.

14. Instructors can create expression files using a spreadsheet program and then saving the file as a tab delimited text file (one of the "save as" options). These tab delimited files can be imported and used for further practice by students.

15. MAGIC Tool is public domain and open source which means you and your students can download the software as many times as you want and load on as many computers as you want. We hope that MAGIC Tool will continue to improve over time, so you may want check the download page periodically to see if any new versions have been posted.