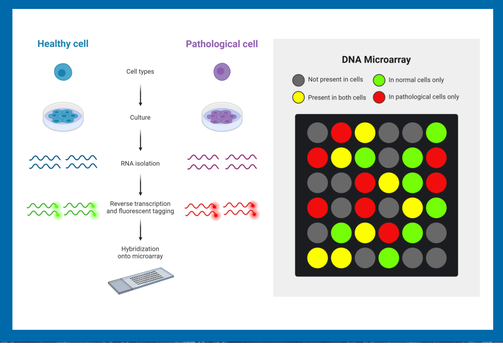


Microarray Analysis: A Real-World Application

Colorado School of Mines

BIOL 301, Introduction to Quantitative Biology II

Emma Benzie, Madilynn Nagy, Morgan Hampton, and Emma Todd

BACKGROUND

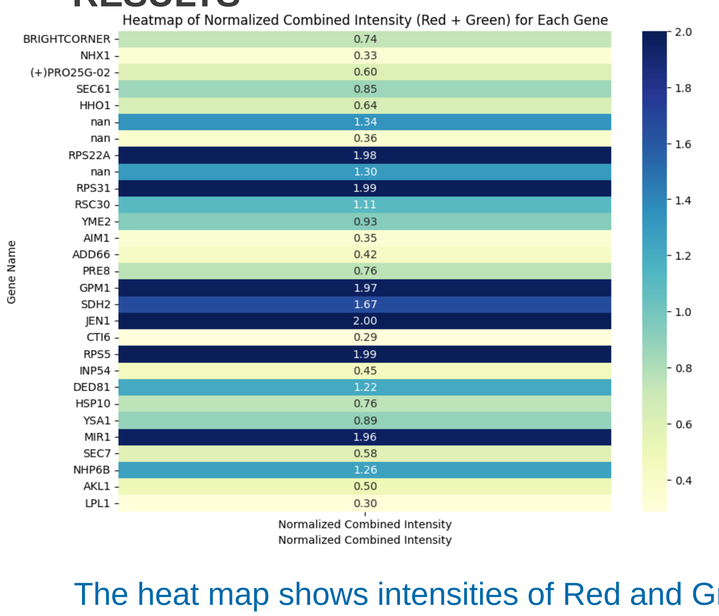
A DNA microarray is a tool that serves several purposes. The primary purpose of a microarray is to measure the level of gene expression, essentially determining the if specific genes are turned “on” or turned “off”. It can also help in detecting a particular mutation in DNA, performing specific disease or personalized medicine diagnostic tests, analyze the DNA across a variety of individuals to find patterns in their DNA which may correlate to a certain phenotype expression, as well as measure the methylation of genes which is important in epigenetic regulation.

Our project aims to apply microscopy image analysis techniques in Python, like noise reduction, contrast enhancement, thresholding, and segmentation to the images produced from the microarray experiments to analyze the results. This was done using data from Princeton.

IMPLEMENTATION

The code takes in a photo that is a 10x10 grid that is then overlayed with computer generated circles. The code then takes the average color of each circle and displays the average color while a helper function computes the average RBG values for each circle which are then organized into a data frame. Background noise and clean visualization is accomplished by using the average RBG and overlayed circles. Red and Green are specifically analyzed via the code. An intensity threshold is then graphed to leave out wells with no binding of cDNA to the oligonucleotide. A heat map was then produced to indicate overall gene expression.

RESULTS



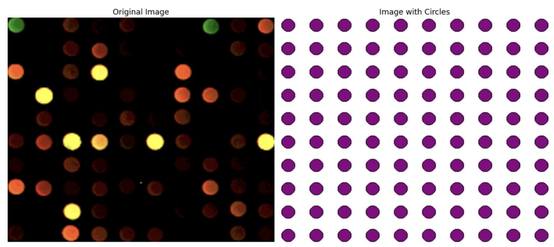
The heat map shows intensities of Red and Green for each of the desired genes. A value of around 1 indicates similar intensities. Low values indicate low intensities and thus low gene expression.

DISCUSSION

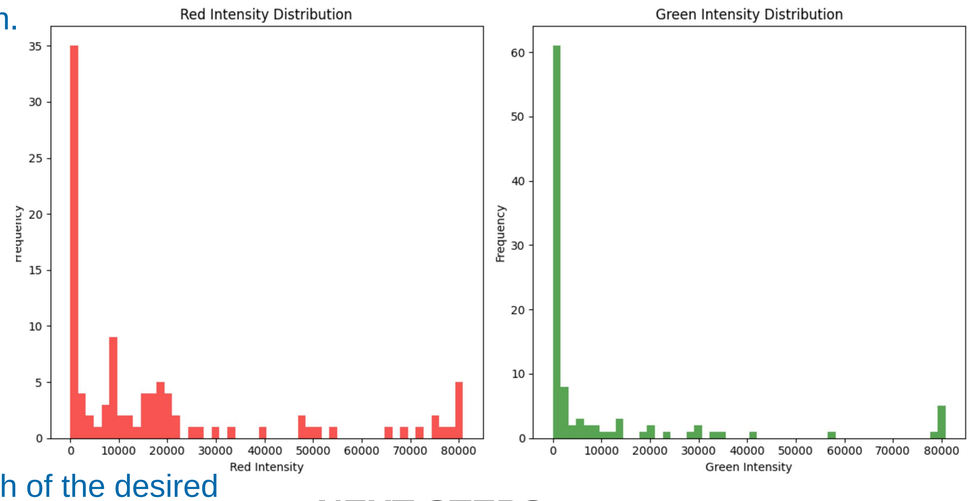
This project has the potential to significantly benefit both research and medicine by enhancing the accessibility and efficiency of DNA microarray analysis. This is critical for advancing personalized medicine and disease diagnostics through genomics. By developing an image analysis tool that simplifies the interpretation of microarray results, we aim to empower and educate both researchers and clinicians who may not have extensive coding expertise. This could facilitate more widespread adoption of DNA microarray technology in various research and clinical settings. It could ultimately lead to better understanding and treatment of genetic disorders and diseases as well as simplify and speed up the analysis process.

Out[17]:

	Position	R	G	B	Red Intensity	Green Intensity	Gene/Control Name	Sequence Type
0	(0, 0)	58	131	35	18386	41527	BRIGHTCORNER	CONTROL
1	(0, 1)	0	0	0	0	0	(-)3XSLV1	CONTROL
2	(0, 2)	64	21	5	20288	6657	NHX1	OLIGO
3	(0, 3)	10	0	0	3170	0	SKI3	OLIGO
4	(0, 4)	42	5	2	13314	1585	POL5	OLIGO
...	...	...	...	...	...	...	...	...
95	(9, 5)	54	5	3	17118	1585	GIN4	OLIGO
96	(9, 6)	0	0	0	0	0	NaN	OLIGO
97	(9, 7)	0	0	0	0	0	SLD7	OLIGO
98	(9, 8)	57	20	4	18069	6340	LPL1	OLIGO
99	(9, 9)	23	0	0	7291	0	NaN	OLIGO



By taking in microarray data and imaging, analysis can be run to better understand and analyze gene expression. A heat map is created in order organize and visualize gene expression using RBG values from the original image. A CSV file of results is given to the researcher once the code runs to completion.



NEXT STEPS

This project paves the way for several future research opportunities in genomics, bioinformatics, and medicine. One potential direction is the application of advanced machine learning and artificial intelligence techniques that can be used to enhance the accuracy and efficiency of microarray data interpretation and analysis. This would enable predictions of gene interactions in clinical outcomes. It may also eliminate potential human error that could occur during analysis and interpretation. Additionally, expanding the analysis tool to incorporate other genomic data, such as RNA-seq and whole-genome sequencing, could provide deeper insights into disease mechanisms. This could aid in identifying biomarkers. This project may also inspire studies focusing on genetic expression patterns in specific populations or diseases. This could ultimately inform targeted therapies and personalized treatment plans. This project aims to make microarray simple and accessible.