



**CALIFORNIA STATE SCIENCE FAIR  
2003 PROJECT SUMMARY**

<b>Name(s)</b> <b>John P. Mahajan</b>	<b>Project Number</b> <b>S0413</b>
<b>Project Title</b> <b>MicroRNA-Oncogene Proximity and Observed Rates of Cancer: A Statistical Study</b>	
<b>Abstract</b> <b>Objectives/Goals</b> The objective is to determine if a correlation exists between the proximity of microRNAs (miRNAs) to oncogenes related to specific cancers, and incidence rates for each specific cancer. <b>Methods/Materials</b> For this project I needed the list of known miRNAs and their loci, or positions on their respective chromosomes. I obtained the list and sequences of the known miRNAs from the National Center for Bioinformatics (NCBI). I then entered each miRNA into UC Santa Cruz's Genome Browser v17 to find the exact position of each miRNA, of which only a few are known. The list of oncogenes with corresponding cancers was located at the Cancer Gen website. Some oncogenes were specific for one type of cancer, but most are associated with more than one cancer, which complicated the analysis that followed. To obtain the oncogene loci, I also used UC Santa Cruz Genome Browser v17. I did find some discrepancies between the data collected from various web sites. I subsequently had to throw out that data. Next I collected the incidence rates for each cancer that I had an associated oncogene for, from the National Cancer Registry SEER search engine. The analysis of the usable data was done using excel. <b>Results</b> The correlation coefficient for the average distance between all miRNAs and overlapping oncogenes against each type of cancer was 0.3365. The correlation coefficient after removal of overlapping oncogenes was 0.05178. After redistributing the data between high and low incidence rate cancers, the correlation coefficients were 0.5368 and 0.2338, respectively. Again, removing overlapping oncogenes and redistributing the data between high and low incidence rate cancers, the correlation coefficients were 0.9988 and 0.125, respectively. <b>Conclusions/Discussion</b> A correlation between the proximity of miRNAs to oncogenes related to specific cancers, and incidence rates for each specific cancer does not exist; each analysis resulted in low correlation coefficients. Even though there was one strong correlation coefficient of 0.9988, it is not statistically significant because there are only three data points for that category. These findings suggest that the distance between miRNA and oncogenes is not correlated to rates of cancer.	
<b>Summary Statement</b> My project tests whether a correlation between the distance of microRNAs from oncogenes, and incidence rates of cancer exists.	
<b>Help Received</b> My mother and brother helped me glue and mat my board, and my father proofread my writing. My biology teacher helped clarify some of my questions on gene regulation.	