



# CALIFORNIA STATE SCIENCE FAIR 2013 PROJECT SUMMARY

<b>Name(s)</b> <b>Jazz Pouls</b>	<b>Project Number</b> <b>S0521</b>
<b>Project Title</b> <b>Analyzing Genomic Divergence in the Swainson's Thrush</b>	
<p style="text-align: center;"><b>Abstract</b></p> <p><b>Objectives/Goals</b> My objective was to find if peaks of divergence between the two subspecies of Swainson's Thrush were randomly distributed in the genome. Then, I could determine if certain regions of the Swainson's Thrush genome appeared to be more divergent and resistant to introgression. My hypothesis predicted that the divergence peaks were non-randomly distributed in the Swainson's Thrush genome.</p> <p><b>Methods/Materials</b> I used a data set containing 154,123 single nucleotide polymorphisms in the Swainson's Thrush genome. Using the programming language R (and some C code), I coded a series of programs that calculated statistics like the density of fixed differences (df), nucleotide diversity (<math>\pi</math>) and <math>F_{st}</math>, a measure of fixation between species. Then, using a homemade kernel smoother, I plotted the data and identified the peaks in divergence. Comparing the genetic coordinates of these peaks to the peaks found in <i>Ficedula</i> flycatchers, I looked at whether or not the peaks in passerine species seemed to be randomly distributed in the genome.</p> <p><b>Results</b> I found evidence that the peaks of the Swainson's Thrush were not randomly distributed in the genome when compared with peaks in <i>Ficedula</i> flycatchers. To do this, I found the distances between peaks of the Swainson's Thrush and the <i>Ficedula</i> flycatchers and added them together to get the "distance index" and compared this to the distance indices of the 500,000 sets of random peaks. The Swainson's Thrush distance index was definitely smaller than the distance indices of the random sets of peaks.</p> <p><b>Conclusions/Discussion</b> The results of my analysis suggest that certain regions of the passerine genome are more divergent and resistant to introgression and gene flow due. This is important because it sheds light on the physical process of speciation and how genomic divergence creates it. A question for further research is where these regions tend to be (i.e. in the telomeric or centromeric regions) and why these parts of the genome are more diverged between species.</p>	
<b>Summary Statement</b> My project used code in R to calculate the genomic divergence between two subspecies of the Swainson's Thrush and I found that the regions in the genome with the most divergence were non-randomly distributed in the genome.	
<b>Help Received</b> Eric Anderson, mentor, helped with the coding; Kristen Ruegg = original collector of the data set used	