

#	Lines of Code	Meaning of each line
1	fastafilename="p53sequence.fasta"	Sets the name of your sequence file to a variable
2	pos = -1	This says we didn't find anything yet
3	listArray = [],[]	each gene will become a list in this list
4	newdir=r"C:\Users\John\Desktop"	new directory as a string
5	print "Total: "+ str(total)	
6	fastafile=file(fastafilename,"r")	"r" reads file
7	print os.getcwd()	prints location of new directory
8	for i in range(1, len(listArray[0])): for gene_item in range(len(listArray[0][i])): if listArray[0][i][gene_item] != listArray[1][i][gene_item]: mismatch += 1 total+=1	#for anything within the length of one of the lists #for each letter in each string in each list #if the two items are not equal  #increase mismatch counter
9	flines=fastafile.readlines()	reads the entire file as strings (so line by line)
10	print "Mismatch: "+ str(mismatch)	
11	total = 0.0	total base pair counter
12	entryPlace = 0	preparing to set a location in our array
13	for line in flines: if line[0] == '>': pos = pos + 1 entryPlace = 0 listArray[pos].append(line) entryPlace += 1	#for each string in the entire aligned file #if the line starts a new gene #move it into the array #in this position #add all of the lines after the carrot into the specified list #add the lines in one after another
14	mismatch = 0.0	Mismatch counter
15	os.chdir(newdir)	changes directory to new directory

16	print "Total: "+ str(total)	
----	-----------------------------	--