A. The main testing strategy I used was to perform the alignments on 5-50 character substrings of the example TP53 sequences provided. I checked the score and the final output to the screen to make sure that they were reasonable given the length and content of the two strings being aligned. To test my traceback, I aligned two 5-50 character substrings and had my script print out the characters being aligned as the matrix was traversed cell-by-cell. I checked cases where the two sequences had no matching characters, and I also checked cases where one or both of the fasta files passed are empty.
B. To align the BRCA gene and the BRCA exons, I used the parameter of 5 for a match, -4 for a mismatch, -1 for a gap extend penalty and -3 for a gap open penalty. Biologically, these two sequences are identical save for multiple gaps that have been spliced out of one of them. Since we expect many gaps in one of the sequences, it makes sense to penalize gaps minimally, with a small absolute gap extend value such as -1 and a slightly larger gap extend penalty of -3. Since the two sequences are ideally identical in the portions where they do not have gaps, we want to encourage matches with relatively high score of 5 and discourage mismatches with a high absolute penalty of -4. We want to discourage mismatches in a gene-exon alignment because mismatches would imply the existence of indels in addition to gaps, which is less probable when comparing original genes and exons than when comparing the same gene from two different species (for example). The optimal alignment I had was:

```
AATAATCAAGAAGAGCAAAGCATGGATTCAAACTTAGGTGAAGCAGCATC
----------------------------------------------------------------------GTGAAGCAGCATC

TGGGTGTGAGAGTGAAACAAGCGTCTCTGAAGACTGCTCAGGGCTATCCT
TGGGTGTGAGAGTGAAACAAGCGTCTCTGAAGACTGCTCAGGGCTATCCT

CTCAGAGTGACATTTTAACCACTCAGCAGAGGGATACCATGCAACATAAC
CTCAGAGTGACATTTTAACCACTCAGCAG-----------------------------------

CTGATAAAGCTCCAGCAGGAAATGGCTGAACTAGAAGCTGTGTTAGAACA
----------------------------------------------------------------------------------------------

GCATGGGAGCCAGCCTTCTAACAGCTACCCTTCCATCATAAGTGACTCTT
----------------------------------------------------------------------------------------------

CTGCCCTTGAGGACCTGCGAAATCCAGAACAAAGCACATCAGAAAAAGCA
----------------------------------------------------------------------------------------------

GTATTAACTTCACAGAAAAGTAGTGAATACCCTATAAGCCAGAATCCAGA
---TATTAACTTCACAGAAAAGTAGTGAATACCCTATAAGCCAGAATCCAGA
```

AGGCCTTTCTGCTGACAAGTTTGAGGTGTCTGCAGATAGTTCTACCAGTA
AGGCCTTTCTGCTGACAAGTTTGAGGTGTCTGCAGATAGTTCTACCAGTA

AAAATAAAGAACCAGGAGTGGAAAGGTCATCCCCTTCTAAATGCCCATCA
AAAATAAAGAACCAGGAGTGGAAAG----------

TTAGATGATAGGTGGTACATGCACAGTTGCTCTGGGAGTCTTCAGAATAG

AAACTACCCATCTCAAGAGGAGCTCATTAAGGTTGTTGATGTGGAGGAGC

AACAGCTGGAAGAGTCTGGGCCACACGATTTGACGGAAACATCTTACTTG

CCAAGGCAAGATCTAGAGGGAACCCCTTACCTGGAATCTGGAATCAGCCT
----------AGGGAACCCCTTACCTGGAATCTGGAATCAGCCT

CTTCTCTGATGACCCTGAATCTGATCCTTCTGAAGACAGAGCCCCAGAGT
CTTCTCTGATGACCCTGAATCTGATCCTTCTGAAGACAGAGCCCCAGAGT

CAGCTCGTGTTGGCAACATACCATCTTCAACCTCTGCATTGAAAGTTCCC
CAGCTCGTGTTGGCAACATACCATCTTCAACCTCTGCATTGAAAGTTCCC

CAATTGAAAGTTGCAGAATCTGCCCAGAGTCCAGCTGCTGCTCATACTAC
CAATTGAAAGTTGCAGAATCTGCCCAGAGTCCAGCTGCTGCTCATACTAC

TGATACTGCTGGGTATAATGCAATGGAAGAAAGTGTGAGCAGGGAGAAGC
TGATACTGCTGGGTATAATGCAATGGAAGAAAGTGTGAGCAGGGAGAAGC

CAGAATTGACAGCTTCAACAGAAAGGGTCAACAAAAGAATGTCCATGGTG
CAGAATTGACAGCTTCAACAGAAAGGGTCAACAAAAGAATGTCCATGGTG

GTGTCTGGCCTGACCCCAGAAGAATTTATGCTCGTGTACAAGTTTGCCAG
GTGTCTGGCCTGACCCCAGAAGAATTTATGCT----------

AAAACACCACATCACTTTAACTAATCTAATTACTGAAGAGACTACTCATG

TTGTTATGAAAACAGATGCTGAGTTTGTGTGTGAACGGACACTGAAATAT
----------GAGTTTGTGTGTGAACGGACACTGAAATAT

TTTCTAGGAATTGCGGGAGGAAAATGGGTAGTTAGCTATTTCTGGGTGAC
TTTCTAGGAATTGCGGGAGGAAAATGGGTAGTTAGCTATTTCT----------

CCAGTCTATTAAAGAAAGAAAAATGCTGAATGAGCATGATTTTGAAGTCA
--CATGATTTTGAAGTCA

GAGGAGATGTGGTCAATGGAAGAAACCACCAAGGTCCAAAGCGAGCAAGA
GAGGAGATGTGGTCAATGGAAGAAACCACCAAGGTCCAAAGCGAGCAAGA

GAATCCCAGGACAGAAAGATCTTCAGGGGGCTAGAAATCTGTTGCTATGG
GAATCCCAGGACAGAAAGATC----------

GCCCTTCACCAACATGCCCACAGATCAACTGGAATGGATGGTACAGCTGT
----------AACTGGAATGGATGGTACAGCTGT

GTGGTGCTTCTGTGGTGAAGGAGCTTTCATCATTCACCCTTGGCACAGGT
GTGGTGCTTCTGTGGTGAAGGAGCTTTCATCATTCACCCTTGGCACA---

GTCCACCCAATTGTGGTTGTGCAGCCAGATGCCTGGACAGAGGACAATGG

CTTCCATGCAATTGGGCAGATGTGTGAGGCACCTGTGGTGACCCGAGAGT
----------CAATTGGGCAGATGTGTGAGGCACCTGTGGTGACCCGAGAGT

GGGTGTTGGACAGTGTAGCACTCTACCAGTGCCAGGAGCTGGACACCTAC
GGGTGTTGGACAGTGTAGCACTCTACCAGTGCCAGGAGCTGGACACCTAC

CTGATACCCCAGATCCCCCACAGCCACTACTGA
CTGATACCCCAGATCCCCCACAGCCACTACTGA

C. Global alignments:

	Global	Local
Cat-Rat	558	691
Human-Rat	597	735
Human-Cat	657	894

Parameters: 5 for match, -1 for mismatch, -2 for gap extend, and -3 for gap open. As the original parameters provided worked well for the gene-exon alignment, I changed them to penalize gaps more and mismatches less. Since we are comparing a likely well-conserved gene (due to its functional importance) among various mammals, I gave gaps a big penalty; at least in the local

alignment, gaps indicate indels and thus secondary protein sequence differences. Mismatches are to be expected, so I penalized them less under the assumption that they are less likely to indicate secondary sequence differences (a mismatch causes the biggest changes when it's between first nucleotides in a codon). Of course, this assumption leads much to be desired, since my alignment doesn't consider codons or even the effects of mismatches of different type of characters (a purine-purine vs. a purine-pyrimidine mismatch, for example). However, based on the combined results from my alignments, it seems to me that the human and cat TP53 sequences are most similar.