Project Proposal

John Palmer

- retrieve lots of HCV sequences containing core protein

- retrieve a random sample of sequences from Africa and Middle East, and

- read in the fasta files (using biopython NOT seqUtils)

- pairwise align the sequences to E1 reference sequence from NC\_004102 reference using gotoh2

- translate nucleotide sequences into amino acid sequences using a MANUALLY CODED translat

- perform multiple sequence alignments between all sequences

-

make phylogenetic trees describing the HCV infection of two different countries in Africa

- compare these phylogenetic trees to