Biopython Workshop

1. Work with DNA sequences:
    - Sequence: AGTCCCAAGAgTGATG
    - Reverse complement the sequence
    - Transcribe and translate the sequence
    - Define your own stop codon
2. Reading FASTA-files:
    - Read file „e_coli_plasmid_p12579.txt“ using SeqIO (indicate the fasta file format)
    - Use the SeqIO for parsing the contents of the file
    - Which information is accessible?
3. Plot the length of the gene sequences from a FASTA-file using pylab
4. Plot the GC content of the gene sequences from a FASTA-file
5. BLAST over the internet
    - BLAST a nucleotide sequence using Bio.Blast
    - Save BLAST result in a xml-file
    - Access the results using NCBIXML
    - Print alignments
    - Which different species were found?

6. Optional: Take a known sequence and create two primers or use the primers indicated below. BLAST the primers separately and analyze the results:
    i. Identify records from the same species, calculate theoretical product length
    ii. PRIMER fw: TTCGGATTGGGGTCT, PRIMER rw: GTGTGTACAAGC
    iii. How many products would result during the PCR? (consider DNA strand!)
    iv. How many different species would be detected during the PCR? Would there be some false positive results?