

```

function [] = efetchHW5(accession_number)
% Automatically download all 16S sequences from a bacterial genome given the
accession number
%
% Test with:
% efetchHW5('NC_000913');
% efetchHW5('NC_005296');
% efetchHW5('NC_003888');

% Open the file and get name and path to open and write file
% [Name, Path] = uigetfile('.txt');
% filename = strcat(Path, Name);
filename = '/Users/kyleking/Documents/Developer/__Matlab/Matlab-for-
Bioninformatics/Week 10 (Nov 6 - No 11)/efetchTest.txt';
fid=fopen(filename, 'w');

% Get feature table
URL=['http://eutils.ncbi.nlm.nih.gov/entrez/eutils/efetch.fcgi?
db=nucleotide&strand=1&id=' accession_number '&rettype=ft'];

% Write to file
ft_str = urlread(URL); % Read URL
% fprintf(fid, '%s\n', [ft_str]); % Write to file

% Use regexp to look for:
% #####      #####      rRNA
%          product      16S( ribosomal RNA) - removed for less specific search
% Save start and stop locations of the start and end number for annotated sequence
start_start_number = regexp(ft_str, '\d*\t\d*\trRNA\n\t\t\tproduct\t16S');
start_end_number = regexp(ft_str, '\d*\trRNA\n\t\t\tproduct\t16S');
end_start_number = regexp(ft_str, '\d\t\d*\trRNA\n\t\t\tproduct\t16S');
end_end_number = regexp(ft_str, '\d\trRNA\n\t\t\tproduct\t16S');

% Search through sequences found to save to file
for i = 1:length(start_start_number)
    % Should be from 223771 to 225312 given NC_000913
    URL = ['http://eutils.ncbi.nlm.nih.gov/entrez/eutils/efetch.fcgi?
db=nucleotide&id=' accession_number '&seq_start=' ft_str(start_start_number(i):
end_start_number(i)) '&seq_stop=' ft_str(start_end_number(i):end_end_number(i))
'&rettype=fasta'];

    % Write each Fasta sequence to file
    fasta_str = urlread(URL);
    fprintf(fid, '%s', [fasta_str]);
end

fclose(fid); % Close the file
end

```