
Lab 7 - Tyler Bradley

Table of Contents

Lab 7.1.1	1
Lab 7.2.1:	3

Lab 7.1.1

Get the NC_001416 sequence from genbank

```
seq = getgenbank("NC_001416", 'SequenceOnly', 'true');

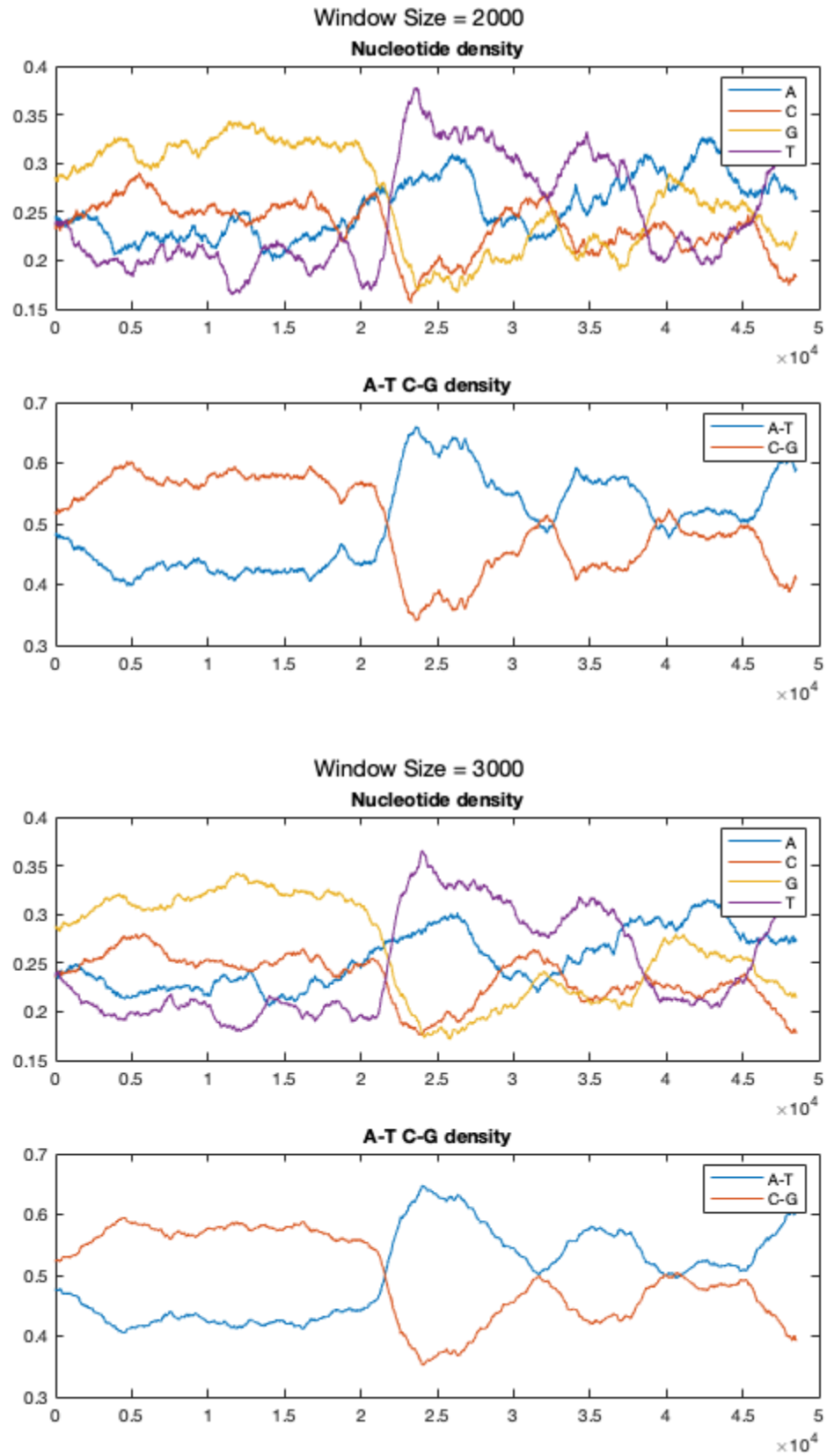
% 1. What is the total length of this sequence?
len_seq = length(seq);
len_seq

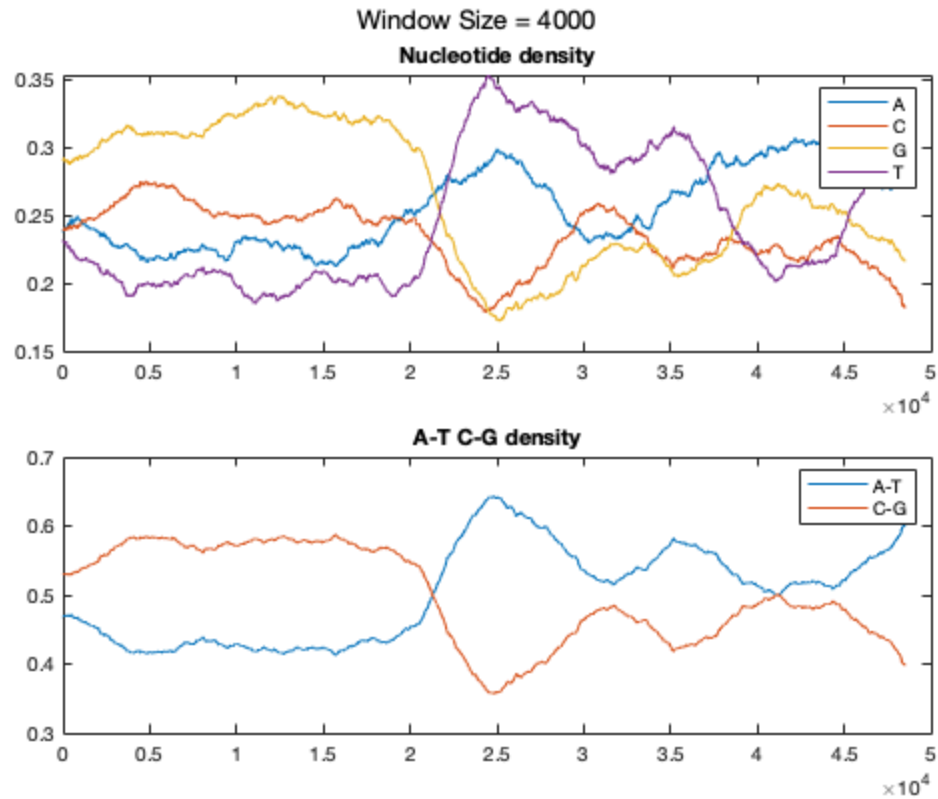
% 2. The local fluctuations in the frequencies of nucleotides provide
% interesting information. The local base composition by a sliding
% window of variable size can be measured. Recall Lab 1.2.1, plot
% density
% of nucleotides along sequence using window size equals to 2000 bp,
% 3000 bp and 4000 bp.
windows = [2000, 3000, 4000];
len_windows = length(windows);

for i = 1:len_windows
    fig_title = "Window Size = " + num2str(windows(i));
    figure(i)
    ntdensity(seq, 'Window', windows(i));
    sgtitle(fig_title);
end

len_seq =

    48502
```





Lab 7.2.1:

1. Suppose we have two hidden states, N and M, and four possible observations: A, T, G, and C, generate the transition matrix and emission matrix randomly.

```
% We can generate a transition matrix (4x4) and a emission matrix (4x2)
% all with random values between 0 and 1 with the rand function
trans_guess = rand(2, 2);
em_guess = rand(2, 4);
```

```
% 2. Encode the nucleotide ?A?, ?C?, ?G?, and ?T? by 1, 2, 3
% and 4 using nt2int.
int_seq = nt2int(seq);
```

```
% 3. Use hmmtrain to update the transition and emission matrix.
[est_trans, est_em] = hmmtrain(int_seq, trans_guess,
    em_guess, 'Maxiterations', 5000);
```

```
% 4. Use hmmviterbi to infer the hidden state of the observations.
states = hmmviterbi(int_seq, est_trans, est_em);
```

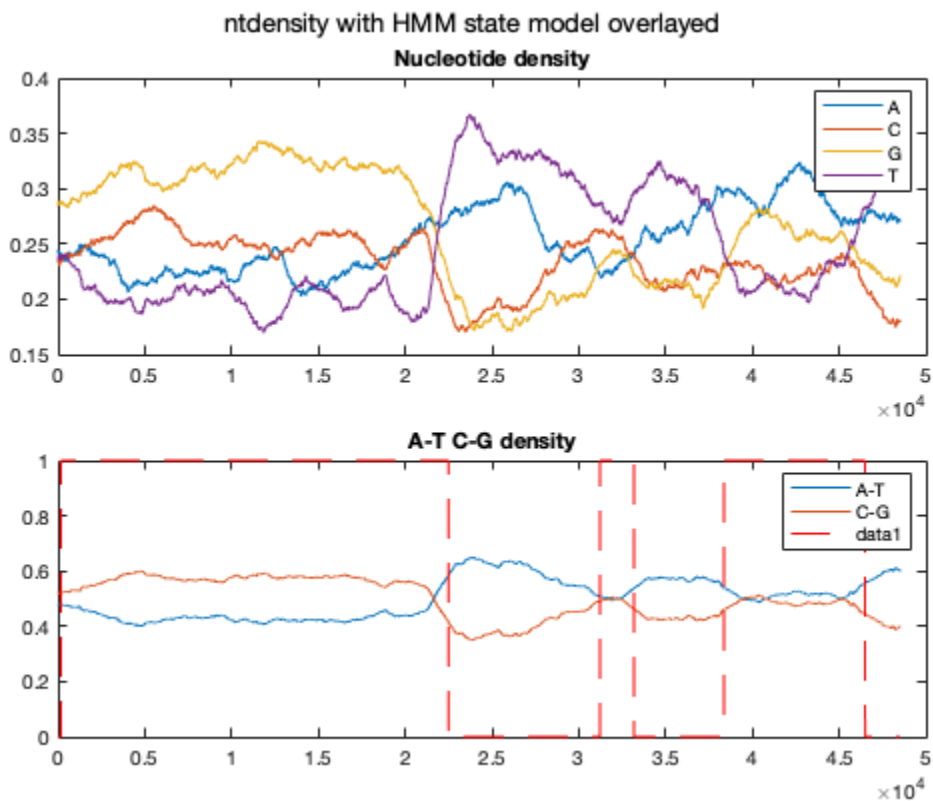
```
figure(4)
hold on
```

```

ntdensity(seq)
line(1:length(states),states-1, 'Color', 'red', 'LineStyle', '--')
sgtitle('ntdensity with HMM state model overlayed')
hold off
% 5. Plot nucleotide density and change points together.

% Hint: Expected result is shown below. If you can't get a similar
% result, your model is probably a bad one. Try to think about what
% causes this problem (did you initialize your model correctly? Did
% you
% read the instruction of hmmtrain and hmmviterbi? Did you use those
% two functions correctly? Also refer to hmm lecture notes for help.).

```



Published with MATLAB® R2018b