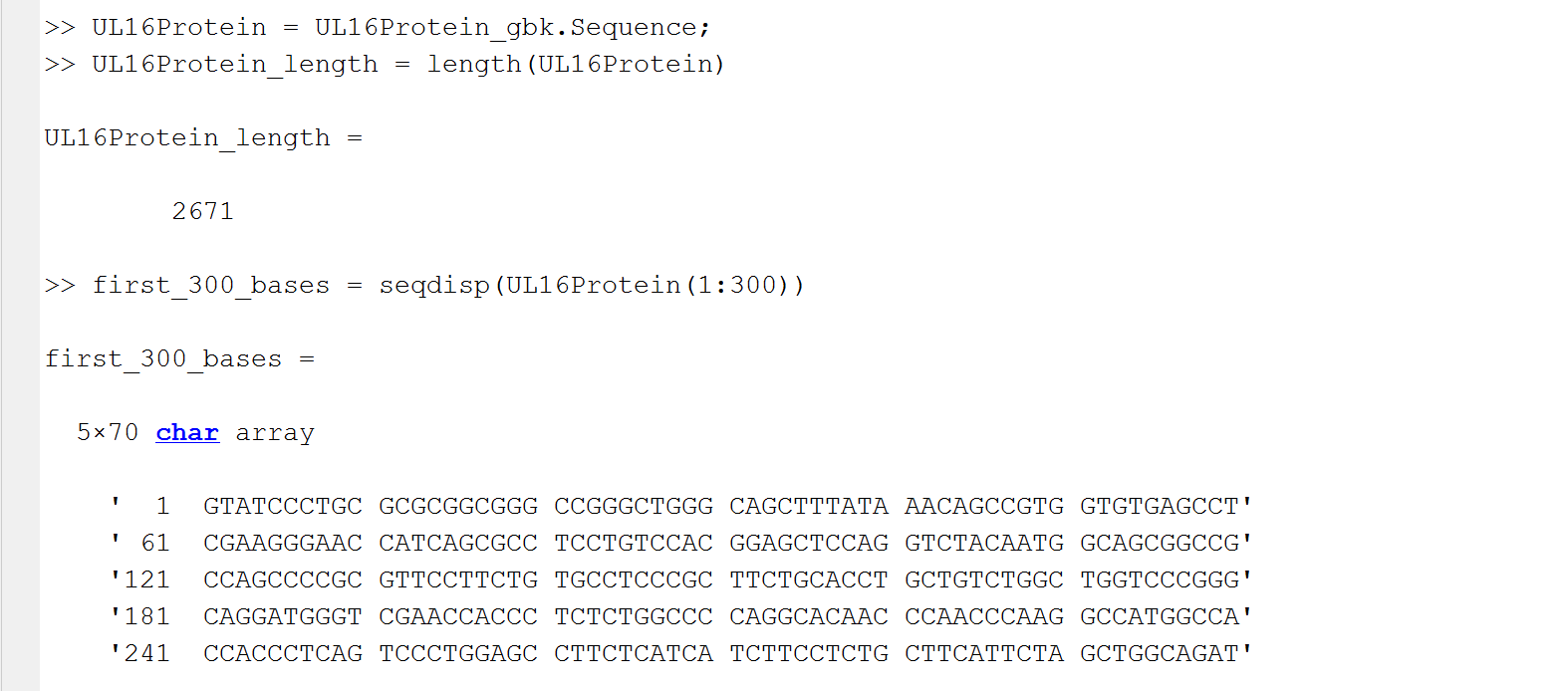
**MATLAB Bioinformatics Toolbox – Calculating and Visualizing Sequence Statistics – Genetic Sequence (Homo Sapiens UL16 Binding Protein 1 Transcript Variant 2 mRNA)**

First 300 Bases of Homo Sapiens UL16 Binding Protein 1 Transcript Variant 2 mRNA



1’ GTA|TCC|CTG|CGC|GCG|GCT|GGG|CCG|GGC|TGG|CAG|CTT|TAT|AAA|CAG|CCG|TGG|TGT|GAG|CCT|

GTA = Valine | TCC = Serine | CTG = Leucine | CGC = Arginine | GCG = Alanine | GCT = Alanine | GGG = Glycine | CCG = Proline | GGC = Glycine | TGG = Serine | CAG = Glutamine | CTT = Leucine | TAT = Tyrosine | AAA = Lysine | CAG = Glutamine | CCG = Proline | TGG = Tryptophan | TGT = Cysteine | GAG = Glucine | CCT = Proline

1’ Patterns: 1 Alanine Block, 1 Glycine Sandwich, 1 Proline-Glycine Block, 1 Glutamine Proline Block,

61’ CGA|AGG|GAA|CCA|TCA|GCG|CCT|CCT|GTC|CAC|GGA|GCT|CCA|GGT|CTA|CAA|TGG|CAG|CGG|CCG|

CGA = Arginine | AGG = Arginine | GAA = Glycine | CCA = Proline | TCA = Serine | GCG = Alanine | CCT = Proline | CCT = Proline | GTC = Valine | GAC = Aspartic Acid | GGA = Glycine | GCT = Alanine | CCA = Proline | GGT = Glycine | CTA = Leucine | CAA = Glutamine | TGG = Tryptophan

61 ‘ Patterns: 1 Arginine Block, 1 Proline-Serine Block, 1 Proline Block, 1 Proline Valine Block, 1 Glycine-Alanine Block, 1 Proline Glycine Block, 1 Glutamine Sandwich

181’ CAG | GAT | GGG | TCG | AAC | CAC | CCT | CTC | TGG | CCC | CAG | GCA | CAA | CCC | AAC | CCA | AGG | TGG | CCA

CAG = Glutamine | GAT = Aspartic Acid | GGG = Glycine | TCG = Serine | AAC = Asparagine | CAC = Histidine | CCT = Proline | CTC = Leucine | TGG = Tryptophan | CCC = Proline | CAG = Glutamine | GCA = Alanine | CAA = Glutamine | CCC = Proline | AAC = Asparagine | CCA = Proline | AGG = Arginine | CCA = Proline | TGG = Tryptophan | CCA = Proline

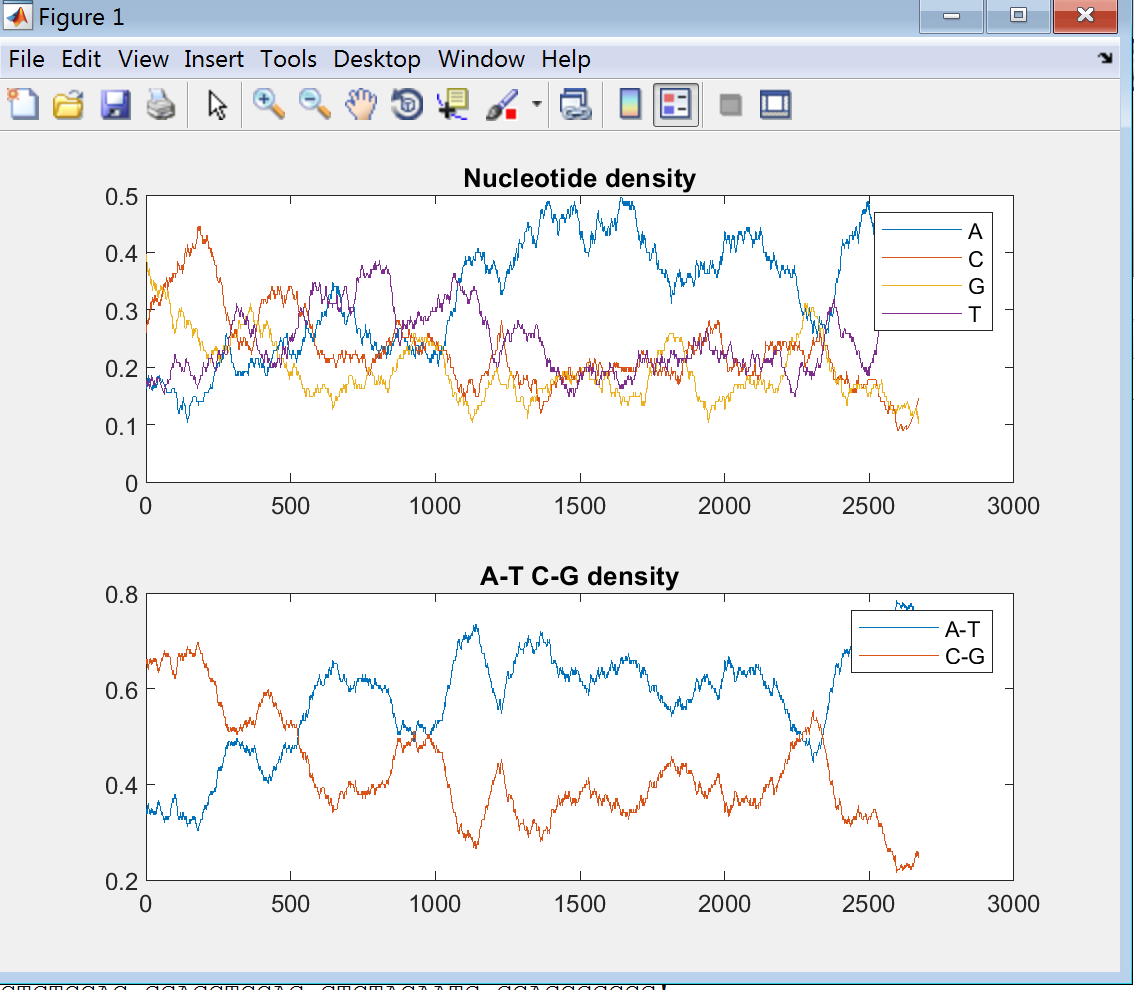
181’ Patterns: 2 Glutamine-Proline Blocks, 1 Glutamine Sandwich Pattern, 2 Proline Sandwiches

241 ‘ CCA | CCC | TCA | GTC | CCT | GGA | GCC | TTC | TCA | TCT | TCC | GCT | TCA | TTC | TAG | CTG | GCA | GAT |

CCA = Proline | CCC = Proline | TCA = Serine | GTC = Valine | GGA = Glycine | GCC = Alanine | TTC = Phenylalanine | TCA = Serine | TCT = Serine | GCT = Alanine | TCA = Serine | TTC = Phenylalanine | TAG = STOP Codon | CTG = Leucine | GCA = Alanine | GAT = Aspartic Acid

241’ Patterns: 1 Proline Block, 1 Proline-Serine Block, 1 Proline-Valine Block, 1 Glycine-Alanine Block, 1 Serine Block, 1 Serine-Alanine Block, 1 Serine Sandwich

Nucleotide Density Figure 1:



X-Axis = Position on the Genetic Sequence (Base Pairs)

Y-Axis = Density of the Amino Acids or Amino Acid Pairs

Nucleotide Density Chart Inference:

* Adenine has the greatest collective nucleotide density, followed by Thymine, Cytosine, and Guanine.
* Adenine has its most widespread range of increasing nucleotide density from the 1000th Base Pair Position to the about the 1700th Base Pair Position.
* Thymine has its most widespread range of increasing nucleotide density from the 200th Base Pair Position to about the 600th Base Pair Position.
* Cytosine has its most widespread range of increasing nucleotide density from the start of the sequence to the 200th Base Pair Position. It is very short.
* Guanine has its most widespread range of increasing nucleotide density from about the 2000th Base Pair Position to the 2400th Base Pair Position.

**A-T – C-G Density Chart**

The Adenine-Thymine Pairing has an overall higher nucleotide density than the Cytosine-Guanine Pairing.

