

```
#!/usr/bin/perl -w
# GC_Analyzer.pl (C) Ulrich Krauss (2013)
#
# DESCRIPTION
# Reads in a DNA Sequence file (only DNA sequence no header) and generates
# a "out-file.txt" containing a list with mean GC values per user defined window
#
# USAGE: GC_Analyzer.pl INPUT.txt
#
# NOTE1: The sequence window over which the script averages GC content has to
# be adjusted in the script (see below)
#
# NOTE2: The input DNA sequence file (e.g. INPUT.txt) has to be located in the same
# directory as the script
```

```
use strict;
use warnings;
```

```
#Declare variables and Input Data
```

```
$/ = '\777'; # entire input to be read in one slurp
my $ORFs = <>; # read input, assigning to single string
my $window = 10; # define sequence window, e.g averaging over 10 bp
my $length = length($ORFs);
my $codoncount = $length/$window;
```

```
#Print Input Data Statistics
```

```
print " \n\n";
print "Input data: \n\n";
print "sequence length: $length \n";
print "number codons: $codoncount \n";
```

```
#Initialize Counts
```

```
my $count_of_g = 0;
my $count_of_c = 0;
my $count_g = 0;
my $count_c = 0;
my $i = 0;
my $pos = 0;
```

```
#Split Sequence in Codons
```

```
my @sequence = split( ' ', $ORFs );
my @codons = ($ORFs =~ m/...../g);
```

```
#Count Codons
```

```
foreach my $nuc (@sequence) {

    if ($nuc eq 'G' or $nuc eq 'g') { ++$count_of_g;}
```

```
    if ($nuc eq 'C' or $nuc eq 'c') { ++$count_of_c;}
}

# derive GC content

my $gc_content = ((( $count_of_g + $count_of_c ) / $length ) * 100);
$gc_content = sprintf ( "%.2f" , $gc_content);

#Count over Window and analyse Local GC

foreach my $block1 (@codons) {
    $i=$i+1;
    $pos=$i*$window;
    my @seqBlock = split(' ', $block1 );
    print "@seqBlock \n";

    foreach my $base (@seqBlock) {

        if ($base eq 'G') {++$count_g}
        if ($base eq 'C') {++$count_c}
        print "$base \n";

    }
}

my $local_gc = ((( $count_g + $count_c ) / $window ) * 100);
print "block # $i Seq-Position: $pos local_gc: $local_gc \n";
open (MYFILE, '>>out-file.txt');
print MYFILE "$i $pos $local_gc \n";
close (MYFILE);
$count_g = 0;
$count_c = 0;
}
print "GC content: $gc_content \n";

exit;
```