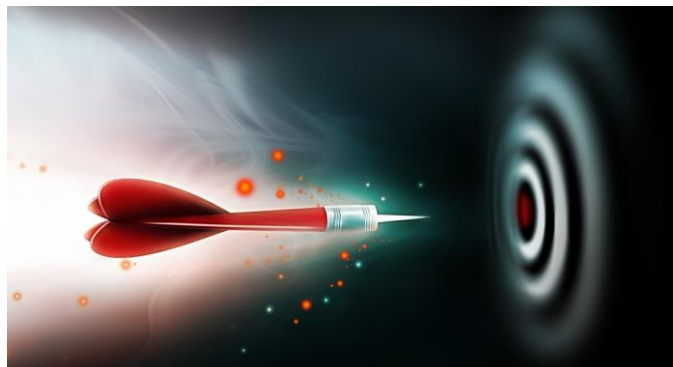


## Mitochondria Amplicon QC Report Targeted Bisulfite Sequencing



SHAKHAWAN K. MAWLOOD

Appendices-5-  
Of  
PhD Thesis

## Contents

<b>1. Sequences Regions of Interest ROI Overview .....</b>	<b>2</b>
1.1. Promoter (-).....	3
1.1.1. QC Data .....	3
1.1.2. ROI Genomic Sequence .....	3
1.2. 12S Ribosomal RNA .....	4
1.2.1. QC Data .....	4
1.2.2. ROI Genomic Sequence .....	4
1.3. Origin of Light Strand .....	5
1.3.1. QC Data .....	5
1.3.2. ROI Genomic Sequence .....	5
1.4. Cytochrome C Oxidase subunit 1 .....	6
1.4.1. QC Data .....	6
1.4.2. ROI Genomic Sequence .....	6
1.5. Cytochrome C Oxidase subunit 2 .....	7
1.5.1. QC Data .....	7
1.5.2. ROI Genomic Sequence .....	7
1.6. T8993G .....	8
1.6.1. QC Data .....	8
1.6.2. ROI Genomic Sequence .....	8
1.7. G11778A .....	9
1.7.1. QC Data .....	9
1.7.2. ROI Genomic Sequence .....	9

## 1. Sequences Regions of Interest ROI Overview

To get mtDNA sequence around interested CpG site, the assembly of genome reference consortium human Build38 (GRCh38 or hg38) was used and for mitochondria (ChrM) use in the assembly. The GRCh38 assembly offers an "analysis set" that was created to accommodate next generation sequencing read alignment pipelines. Some common features were used and applied to all amplicon. Small letter is for repeat sequence and capital letter is unique sequence. **Dark grey** means, regions not covered by target sequencing amplicons and **Light grey** indicates region of interest. The methylated CpG sites were coloured in red and underlined **cg**.

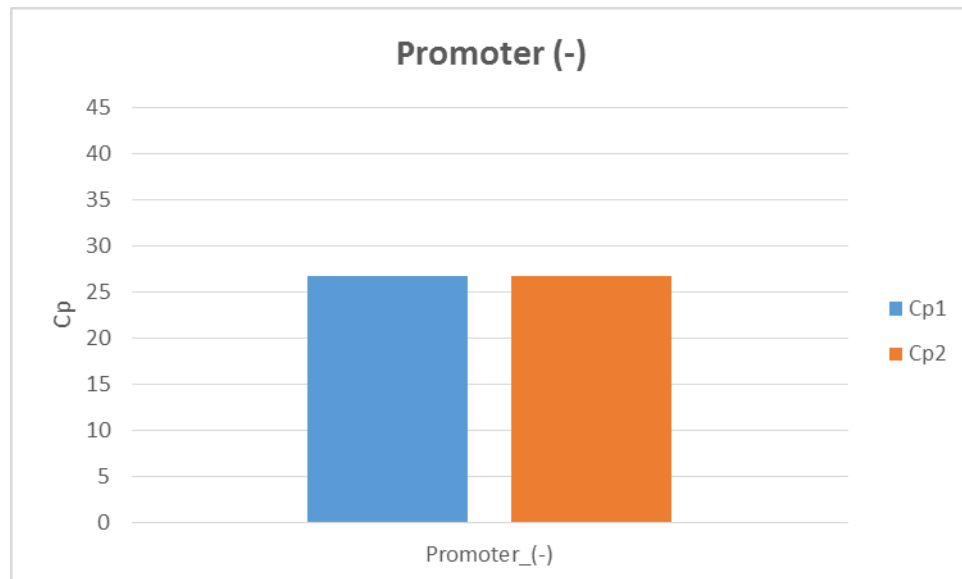
Region of Interest (ROI) ID	Number of Amplicons that passed QC
Promoter_(-)	1
12S_Ribosomal_RNA	1
Origin_of_Light_Strand	1
Cytochrome_C_Oxidase_subunit-1	1
Cytochrome_C_Oxidase_subunit-2_2	1
T8993G	1
G11778A	1

**Total**

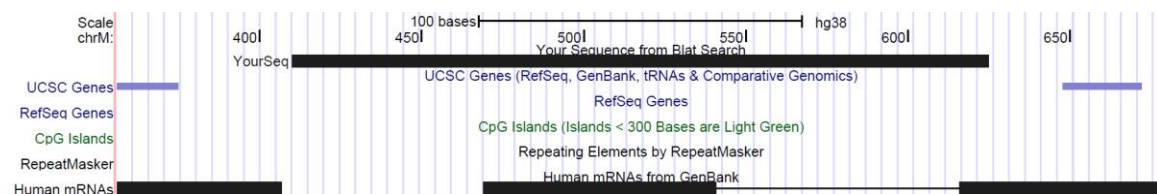
**7**

## 1.1. Promoter (-)

### 1.1.1. QC Data



### 1.1.2. ROI Genomic Sequence

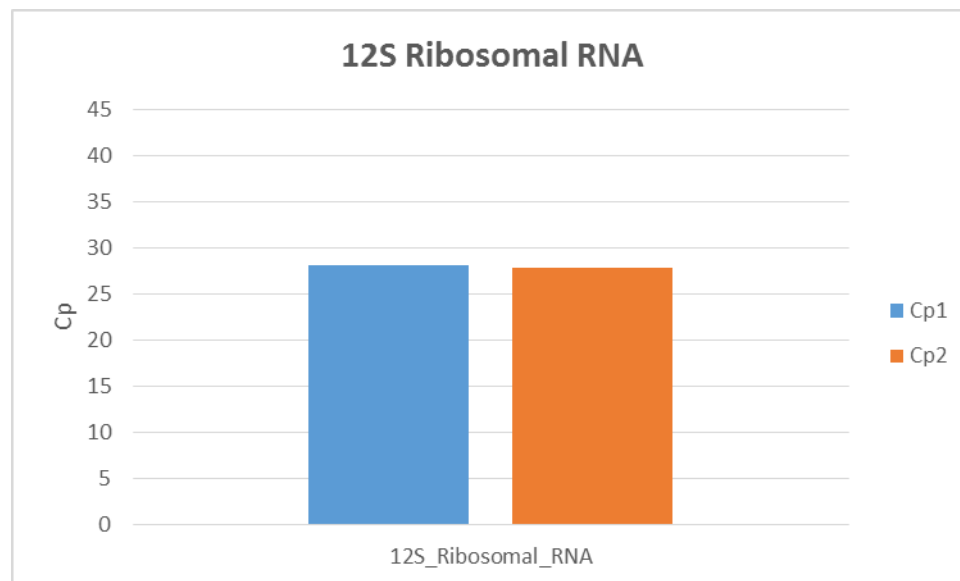


**Light Gray**= Regions of Interest; **Green**= ROI Covered with Amplicons that passed QC; **CG**= CG of interest

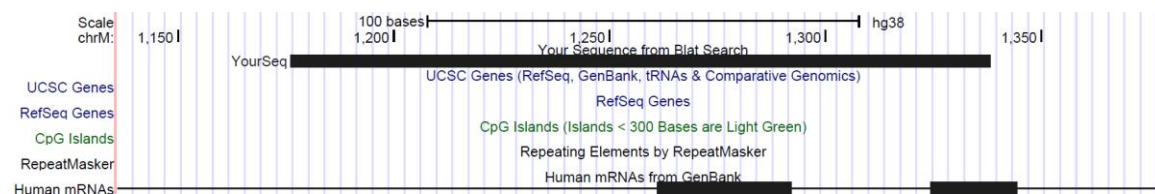
```
>hg38 dna range=chrM:141-700 5'pad=0 3'pad=0 strand=- repeatMasking=lower
TTGCATGTGTAATCTTACTAAGAGCTAATAGAAAGGCTAGGACCAAACCTATTTGTTTATGGGGTGATGTGAGCC
CGTCTAAACATTTTCAGTGTATTGCTTTGAGGAGGTAAGCTACATAAACTGTGGGGGGTGTCTTTGGGGTTTGGT
TGGTTCCGGGTATGGGGTTAGCAGCGGTGTGTGTGTGCTGGGTAGGATGGGCGGGGGTTGTATTGATGAGATTAG
TAGTATGGGAGTGGGAGGGGAAAATAATGTGTTAGTTGGGGGGTGACTGTTAAAAGTGCATACCGCCAAAAGATA
AAATTTGAAATCTGGTTAGGCTGGTGTTAGGGTTCTTTGTTTTTGGGGTTTGGCAGAGATGTGTTTAAGTGCTGT
GGCCAGAAGCGGGGAGGGGGGGTTTGGTGGAATTTTTTGTATTGATGTCTGTGTGGAAGTGGCTGTGCAGAC
ATTCAATTGTTATTATTATGTCCTACAAGCATTAAATTAACACACTTTAGTAAGTATGTTCCCTGTAATAT
TGAACTAGGTGCGATAAATAATAGGATGAGGCAG
```

## 1.2. 12S Ribosomal RNA

### 1.2.1. QC Data



### 1.2.2. ROI Genomic Sequence



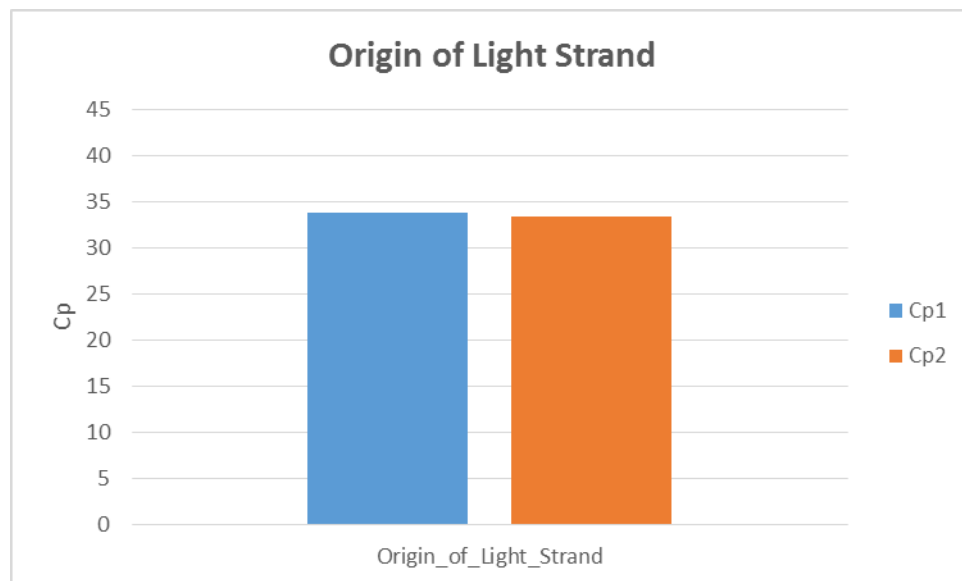
**Light Gray**= Regions of Interest; **Green**= ROI Covered with Amplicons that passed QC; **CG**= CG of interest

```
>hg38_dna range=chrM:501-1700 5'pad=100 3'pad=100 strand=+
repeatMasking=lower
```

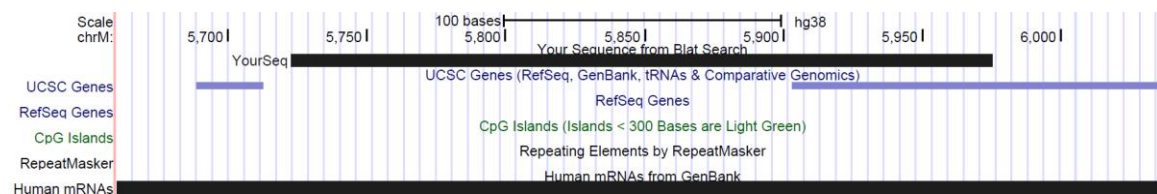
```
CCATCCTACCCAGCACACACACACCGCTGCTAACCCCATACCCCGAACCAACCAAAACCCCAAAGACACCCCCAC
AGTTTATGTAGCTTACCTCCTCAAAGCAATACACTGAAAATGTTTAGACGGGCTCACATCACCCCATAAACAAAT
AGGTTTGGTCCTAGCCTTTCTATTAGCTCTTAGTAAGATTACACATGCAAGCATCCCCGTTCCAGTGAGTTTACC
CTCTAAATCACCACGATCAAAAGGAACAAGCATCAAGCACGCAGCAATGCAGCTCAAAACGCTTAGCCTAGCCAC
ACCCACACGGGAAACAGCAGTGATTAACCTTTAGCAATAAACGAAAGTTTAACTAAGCTATACTAACCCAGGGT
TGGTCAATTCGTGCCAGCCACCGCGGTCACACGATTAAACCAAGTCAATAGAAGCCGGCGTAAAGAGTGTTTAA
GATCACCCCTCCCAATAAAGCTAAAACTCACCTGAGTTGTAAAAAACTCCAGTTGACACAAAATAGACTACGA
AAGTGGCTTTAACATATCTGAACACACAATAGCTAAGACCCGAAACTGGGATTAGATACCCCACTATGCTTAGCCC
TAAACCTCAACAGTTAAATCAACAAAACCTGCTCGCCAGAACACTACGAGCCACAGCTTAAAACTCAAAGGACCTG
GGGTGCTTCATATCCCTCTAGAGGAGCCTGTTCTGTAATCGATAAACCCCGATCAACCTCACACCTCTTGCTC
AGCCTATATACCGCCATCTTCAGCAAACCTGATGAAGGCTACAAAGTAAGCGCAAGTACCCACGTAAAGACGTT
AGGTCAAGGTGTAGCCCATGAGGTGGCAAGAAATGGGCTACATTTTCTACCCAGAAAACCTACGATAGCCCTTAT
GAACTTAAGGGTCGAAGGTGGATTTAGCAGTAACTAAGAGTAGAGTGCTTAGTTGAACAGGGCCCTGAAGCGC
GTACACACCGCCCGTCACCTCCTCAAGTATACTTCAAAGGACATTTAACTAAAACCCCTACGCATTTATATAGA
GGAGACAAGTCGTAACATGGTAAGTGTACTGGAAAGTGCACCTTGGACGAACCAGAGTGAGCTTAACACAAAGCA
CCCACTTACACTTAGGAGATTTCAACTTAACTTGACCGCTCTGAGCTAAACCTAGCCCCAAACCCACTCCACCT
```

### 1.3. Origin of Light Strand

#### 1.3.1. QC Data



#### 1.3.2. ROI Genomic Sequence



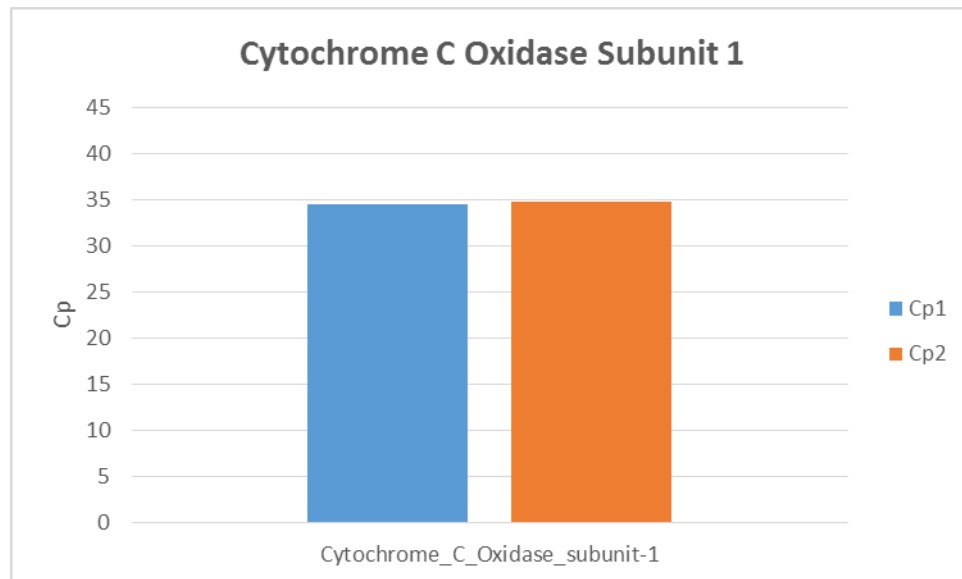
**Light Gray**= Regions of Interest; **Green**= ROI Covered with Amplicons that passed QC; **CG**= CG of interest

```
>hg38_dna range=chrM:5541-6040 5'pad=100 3'pad=100 strand=+
repeatMasking=lower
```

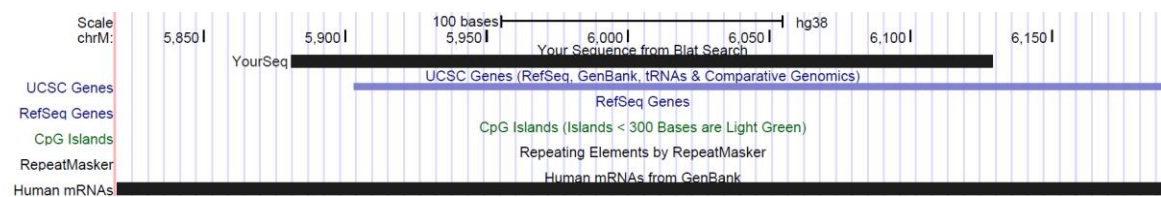
```
CCTTCAAAGCCCTCAGTAAGTTGCAATACTTAATTTCTGTAACAGCTAAGGACTGCAAAACCCCACTCTGCATCA
ACTGAACGCAAATCAGCCACTTTAATTAAGCTAAGCCCTTACTAGACCAATGGGACTTAAACCCACAAACACTTA
GTTAACAGCTAAGCACCCCTAATCAACTGGCTTCAATCTACTTCTCCCGCCCGGGAAAAAAGGCGGAGAAAGCC
CGGCGAGGTTTGAAGCTGCTTCTTGAATTTGCAATTCAATATGAAAATCACCTCGGAGCTGGTAAAAAGAGGCC
TAACCCCTGTCTTTAGATTTACAGTCCAATGCTTCACTCAGCCATTTTACCTCACCCCACTGATGTTGCGCCGAC
CGTTGACTATTCTCTACAAACCACAAAGACATTGGAACACTATACCTATTATTTCGGCGCATGAGCTGGAGTCCTA
GGCACAGCTCTAAGCCTCCTTATTCGAGCCGAGCTGGGCCAGCCAGGCAA
```

## 1.4. Cytochrome C Oxidase subunit 1

### 1.4.1. QC Data



### 1.4.2. ROI Genomic Sequence

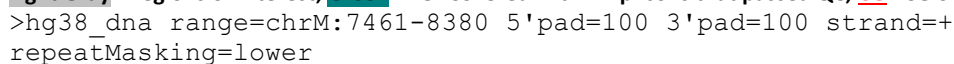


**Light Gray**= Regions of Interest; **Green**= ROI Covered with Amplicons that passed QC; **CG**= CG of interest

```
>hg38_dna range=chrM:5781-6340 5'pad=100 3'pad=100 strand=+
repeatMasking=lower
```

```
CTGCTTCTTCGAATTTGCAATTCAATATGAAAATCACCTCGGAGCTGGTAAAAAGAGGCCTAACCCTGTCTTTA
GATTTACAGTCCAATGCTTCACTCAGCCATTTTACCTCACCCCACTGATGTTTCGCCGACCGTTGACTATTCTCT
ACAAACCACAAAGACATTGGAACACTATACCTATTATTCCGCGCATGAGCTGGAGTCTTAGGCACAGCTCTAAGC
CTCCTTATTCCGAGCCGAGCTGGGCCAGCCAGGCAACCTTCTAGGTAAACACCACATCTACAACTTATCTCACA
GCCCATGCATTTGTAATAATCTTCTTCATAGTAATACCCATCATAATCGGAGGCTTTGGCAACTGACTAGTTCCC
CTAATAATCGGTGCCCCGATATGGCGTTTCCCCGCATAAACAACATAAGCTTCTGACTCTTACCTCCCTCTCTC
CTACTCCTGCTCGCATCTGCTATAGTGAGGCCGAGCAGGAACAGGTTGAACAGTCTACCTCCCTTAGCAGGG
AACTACTCCACCCCTGGAGCCTCCGTAGACCTAAC
```

### 1.5.1. QC Data

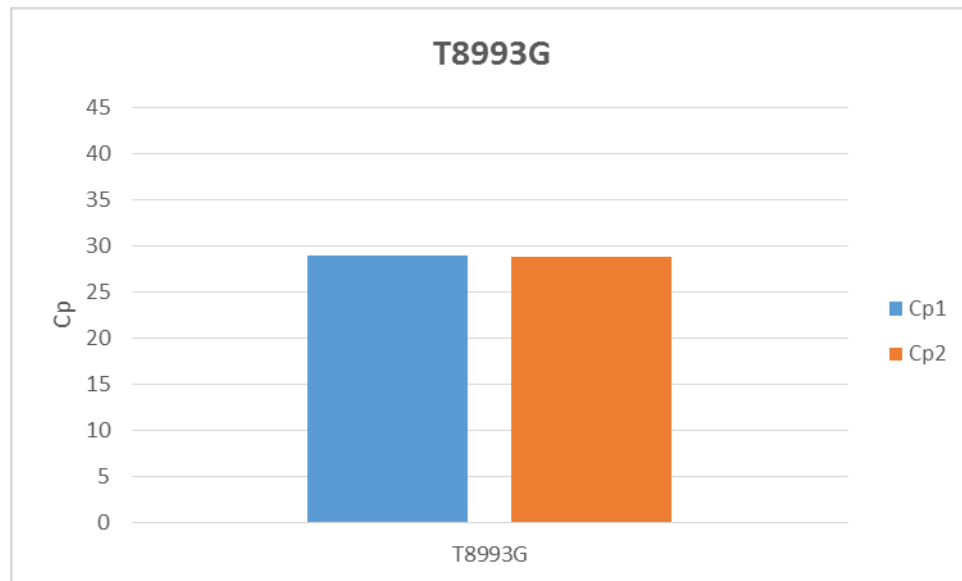


tcgaacccccccaagctggtttcaagccaaccccatggcctccatgactttttcAAAAAGGTATTAGAAAAACCA  
TTTCATAAAGTTTGTCAAAGTTAAATTATAGGCTAAATCCTATATATCTTAATGGCACATGCAGCGCAAGTAGGTC  
TACAAGA CGCTACTTCCCCTATCATAGAAGAGCTTATCACCTTTTCATGATCACGCCCTCATAATCATTTTTCCCTTA  
TCTGCTTCCTAGTCCTGTATGCCCTTTTCCCTAACACTCACAACAAAACCTAACTAA **TACTAACATCTCAGACG**CTC  
AGGAAATAGAAAC **CG**TCTGAACTATCCTGCC **CG**CCATCATCCTAGTCCTCAT **CG**CCCTCCCATCCCTA **CG**CATCC  
TTTACATAACAGA **CG**AGGTCAA **CG**ATCCCTCCCTTACCATCAAATCAATTGGCCACCAATGGTACTGAACCTA **TCG**  
AGTACAC **CG**ACTA **CG****CG**ACTAATCTTCAACTCCTACATACTTCCCCCATTTATTCCTAGAACCCAGGCGACCTGC  
GACTCCTTGACGTTGACAATCGAGTAGTACTCCCGATTGAAGCCCCCATTCGTATAATAATTACATCACAAGACG  
TCTTGCACTCATGAGCTGTCCCCACATTAGGCTTAAAAACAGATGCAATTCCCGGACGTCTAAACCAAACCACTT  
**TCACCGCTACACGACCGGGGGTATACTACG**GTCAATGCTCTGAAATCTGTGGAGCAAACCCACAGTTTCATGCCCA  
TCGTCTTAGAATTAATTCCCCCTAAAAATCTTTGAAATAGGGCCCCGTATTTACCCTATAGCACCCCTCTACCCCC  
TCTAGAGCCCATGTAAAGCTAACTTAGCATTAACCTTTTAAAGTTAAAGATTAAAGAGAACCAACACCTCTTTACA  
GTGAAATGCCCACTAAAT

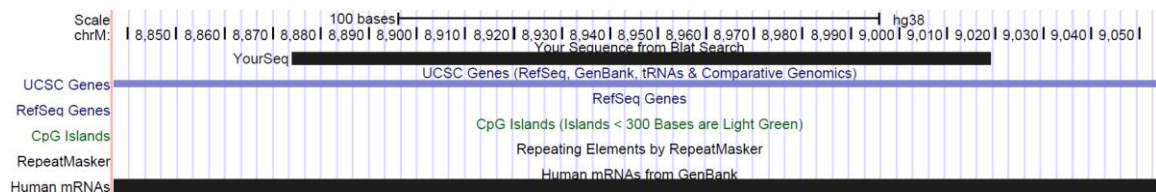


## 1.6. T8993G

### 1.6.1. QC Data



### 1.6.2. ROI Genomic Sequence



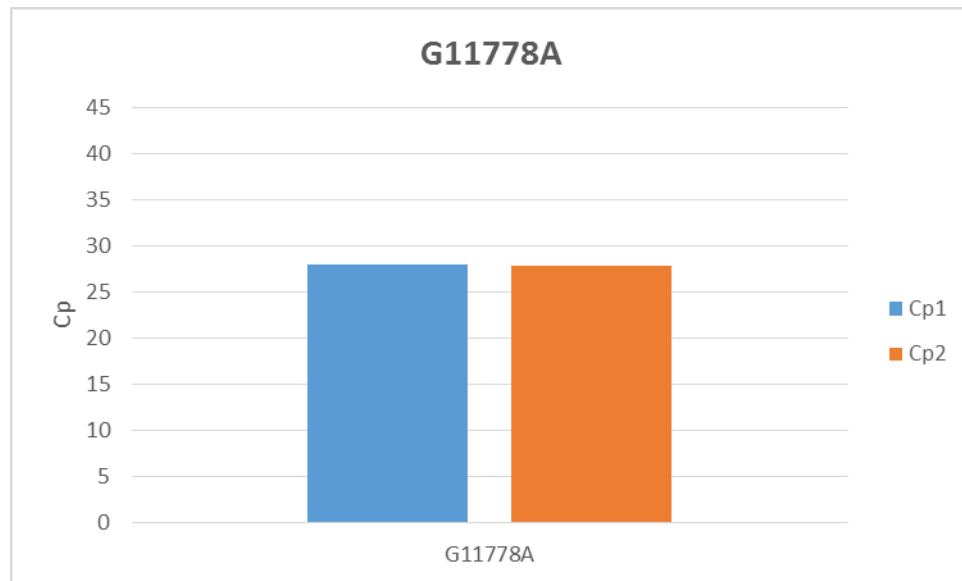
**Light Gray**= Regions of Interest; **Green**= ROI Covered with Amplicons that passed QC; **CG**= CG of interest

```
>hg38_dna range=chrM:8781-9160 5'pad=100 3'pad=100 strand=+
repeatMasking=lower
```

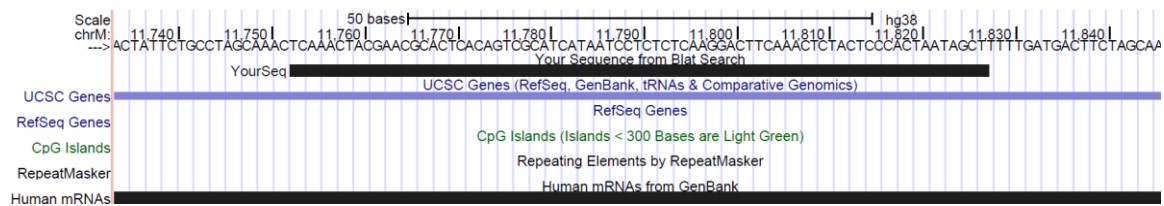
```
CGGACTCCTGCCTCACTCATTTACACCAACCACCCAACCTATCTATAAACCTAGCCATGGCCATCCCCCTTATGAGC
GGGCACAGTGATTATAGGC TTTTCGCTCTAAGATTAAAAATGCCCTAGCCCACTTCTTACCACAAGGCACACCTAC
ACCCCTTATCCCCATACTAGTTATTATCGAAACCATCAGCCTACTCATTCAACCAATAGCC CGTGGCCGTACGCCT
AACCGCTAACATTACTGCAGGCCACCTACTCATGCACCTAATTGGAAGCGCCACCCTAGCAATATCAACCATTAA
CCTTCCCTCTACACTTATCATCTTCACAATTCTAATTCTACTGACTATCCTAGAAATCGCTGTCGCCTTAATCCA
```

## 1.7. G11778A

### 1.7.1. QC Data



### 1.7.2. ROI Genomic Sequence



**Light Gray**= Regions of Interest; **Green**= ROI Covered with Amplicons that passed QC; **CG**= CG of interest

```
>hg38_dna range=chrM:11541-12040 5'pad=100 3'pad=100 strand=+
repeatMasking=lower
```

```
TCCTTGTAATCCCTATGAGGCATAATTATAACAAGCTCCATCTGCCTACGACAAACAGACCTAAAAATCGCTCA
TTGCATACTCTTCAATCAGCCACATAGCCCTCGTAGTAACAGCCATTCTCATCCAAACCCCTGAAGCTTCACCG
GCGCAGTCATTCTCATAATCGCCACGGGCTTACATCCTCATTACTATTCTGCCTAGCAAACTCAAACTACGAAC
GCACTCACAGTCGCATCATAATCCTCTCTCAAGGACTTCAAACTCTACTCCCACTAATAGCTTTTGTGACTTC
TAGCAAGCCTCGCTAACCTCGCCTTACCCCCACTATTAACCTACTGGGAGAACTCTCTGTGCTAGTAACACGT
TCTCCTGATCAAATATCACTCTCCTACTTACAGGACTCAACATACTAGTCACAGCCCTATACTCCCTCTACATAT
TTACCACAACACAATGGGGCTCACTCACCCACCACATTAACAACATAAAA
```