BIOINFORMATICS WORKSHOP

Sometimes genetic mutations in an organism can lead to it being more or less susceptible to certain drug treatments. This is particularly important when we are looking at the effectiveness of drug treatment against parasitic worms as there may be resistant and susceptible strains of worms. In this workshop we will be using an example sequence to design primers for and then we will analyse. To do this we will be using bioinformatics. In this workshop we are aiming to do the following:

1. Understand how the primers are designed for the PCR process.
2. Look for mutations in our sequence to identify our resistant and susceptible strains.

When we are designing primers for our PCR we need to think about a few things in order to get a good primer designed for our sequence. Good primer design is essential for a successful PCR. In order to design your primer there are some important considerations, which include but are not exclusive to the following:

a. Primer length
b. Primer melting temperature
c. Primer annealing temperature
d. GC content

1. **Primer length** – Optimal primer length is generally considered to be between 18 to 22bp long. This is long enough for the primer to anneal to our DNA easily and to be specific for our DNA region.

2. **Primer Melting Temperature (Tm)** – This is the temperature where the double stranded DNA dissociates to become single stranded DNA. Temperatures generally range between 52 and 58°C. Secondary annealing can occur with temperatures above 65°C.

3. **Primer Annealing Temperature (T_a)** – This is the temperature required for the primer to anneal to the DNA sequence. Too low a T_a and the primers may anneal to sequences that we are not targeting therefore leading to a reduction in the product yield. Too high a T_a may also lead to a low product yield as the primer annealing may be reduced.

4. **GC content** – This is the number of G's and C's in the DNA sequence. The total of which should be between 40 – 60%.

We will design primers using a primer design software tool called PRIMER 3 (Rozen and Skaletsky, 2000) for the example sequence below:

TTTCCGACGACGACGACAGCAAGGGCAGCTTCGACATCGTCTGACGAAATACGGTAGTGACCTCAGAAAGTCAGAATACAAGTAGTAGCAATAAGGTACAGTAGCAAGAG
TTGTGCAGTGCAGTTTCTTAGGGAGAGCAGAGCAGATAAAAATGGTGGAATCGACACTGTGAGCCATCATGCAGGTGTATAGGCAACTAAGTTTGATCAGATAGTATCATTT
TGAAAACAAGATTGTGCATTCAGACATATATATAGTGGACAAATGAGGTAAACACAGCCAAAAGACGGTTAAATTGGACTATGTCGATCCCCATCTTTATTGGGAAGAGAAA
AATACAAAAGGATTAACAACAAAATGCTTGGATTACAACACCAGACAAGGGAAAAAGAATCTGGCAGAACAGGAATGGCATTGTTCAAACCTCCATCTCCATTGTTCCTGC
CATCCAAACGTTTATGCATGCCATTAGATAAACCTCCAAACTCTTTGGCGCTATACTAGCTATTATCCCGCTGCAAAGATGTCTGCAGAATAGCCGAAGGAAGGTAAAATTCC
ACCACCAGGGGAGCCCAGAATATTAGGAAGATTTGAAGAGAAAACGAAGAGGAGAGGCTACACTCAAGAGTTCAGTCCATCACGGCAGTTATCAAAAGACTAGTCGACCT
CCGCATACTGGGAGTGCGAGGCCGCAAGGAGAGCGGCTCCGATGCCAGATCCGTCGTTGGCCAGCTTAGCAACCAGCGACGAAGAGGCTTCCTCCCCGAGGAGATCTGTA
AGAGTTGCTTCTAAGCAACTGCTGAACTTCTTGTAATGTTCATAGAGCCCACCATCCAAGGCAACGACGGTCCTTGGCATCGTGCTGCCATCAGTTGGCACTTTGTCACGGC
CTAGCTTCTTCAGGATGCCATATATACCAGCAGCGGCCAAGCGTGCTCCGCGCTCAGCAACAATGTCACATATGTGACGAGTAATGTATCTTACTTCCAGGGAAGTATCTGCA
ACCTGAACATGAAATCGAGATGTCATACAAGTAATCAATGGTCATTGAACATTTCTGGTATCCAGTTAGGTTCACGCAAGTCATACCCCAACGATATCCTTTAGCTTAGTTCCC
AGAACTTTAAGGTCATGTGATGCATCATGGTGCATGGCTGACATATCCGGGGTCCTGCAATGCATAAGATTCAGTAGTTGAATTTATACGGAGTAAGAATACACTACAAAACT
TCATACAGATAAACGCTTTAATCGTAATTGAGCTCATGTTCAGAAAACATTTGTCAAGATTTAACACAGATTTCTGCTACTAATGGCATGGGCTCTTCATTAGTTCTCAAGAAT
TCTATGCTATTCAATACACATATCAATCAACAAATAGTATATCTACTTAGCCTCAAGTAAATATTCGATTAGTTCCAAGAAGAGGAGCATAACACTATTCACCATCACTCTTTTG
AGCCACTTGAAAAAACAAAAGAAAAAGGAAAACAAACATTTGGTTTTGAATTAGAATATTAAGGTGTGCACGAAACAAACTCTATCCAAGTAAGCAATTGCACACGCCAT
CCACCAAATTATTCTAATACACCAGAGAAAAAAGATACCAAGCTTTGAATTGATCGCAGTTGAGGACAAAACAGAGCATTGACCAGATGTAAAGTATTTCATCACCAGTTAT
GATGAATCGGCATTAAATTTTGATTTTGCCACCTCACATAGATTACTTGACTGAAGTATATGTTAACTTCAGAAGGCCCGAAATTTGCCAATTCTTTTATAGCTTTGGGCTTTG
GCCTTACACGGACTAGTCAACTGATGTAACATGATAAATTCGGACGAGAATCTGTTTCCTTGCGCAGTTTTGAAAGCTCAATATTAGTGTGTAAAAATTCTTGATCACCGTTT
AGCTACAGCAAAGATGAAGAGTTCAATGACTAATGGGTCAAGAAAGTATACCTCAGTACAAACGGCTGTTCCAACTTAGGTGGAACAACATCTCCAAACAATGCAGCATCA
TGAGCCAGTTTCAGTAAGATTCTTCGCACAATCTCTCCAAGATACATGCCAGAAATCATCTTTTCATATATCTGTTACCCAGAAGATATTCGCATATATTAGTGCAGTTTAGCAA
CAAAACATAAATGGAAGTGTAAATACCACCGAGATAAACACCTGTTCTCCAGGATTCAGACTTTCAAAATCCATAGCTTTATCGTATTCTGAAAGTGGAAGCTTGTCTGATTT
AAAGCTTCCCCATTCCGTGTTGATTACCTAGAGAACATCAAACAATATTGTAGTTAGGCTAGTAGAGTACAGCAGTGAAGATACGATGCATGTGTACGCCCATACGATGTTTA
ATGGGACTTGTTCATTCCAAGCAGAAAGGGAATTTTAAACACCAACCATATTTCCGGATTTAGGCAGTAATCCAGTCCATTTTGGAATTGCATTAGCATGCTCAACATATGCT
GCATTTGTGCCAGTGCCCAAAATTACAGCAGCAACAACATCGTTATCCACATATCTCCCACCAGCCAATGTGCCGACTGTGTCATTAACCTACAAAATATCAACAATCAGCCT
ACATAAAGGAAAATCTCAAAAAAGAAGAAAGTTCTGCATGACGCAAAGCATGTTTCCATCCCCATCAACAGAAAAGATTAACGAGTAGCATCAGGACACACAAAAAATAAA
AATTGCATTATTTTGCGGGAGTAAAAAATAGATATGGTGTATTACCAATGCTGAAACTTTCATATCCAACCCCTGCCTCTCCATGGCCTTGCTCAACTCAGCCACAACATCCTC
ACCAACCTGATTAGGGAGTAACCAAATAACAAATTAAAATCAAAGTAATTGTGGTTATTATGGGATGATTGAAAAATTGGCTCAGAGACCATCAACCGAGAAACTTAATATT
CTTAAGTATCACAAATATATTCTTCCTTACCAAGCTCTACATTTCATCGTTGTGATAAACTCAGATAACATCAACCCAGAAACATACTACTCTTACTGCATCATAAAAATGTGCT
TCCTAACCAAGCTTTACATTTCATCATTGAGATAAAGAAATCAGTTTGACCATAATGTAATGAGTGATAACGAAGTATCAGTGAACTAAATGAATCAACATGAAATAATATTTA
ATAACATTTCATCATTGAGATAAAGAAATCAGTTTGACCATAATGCAATGACTGATTGAGAAGTATCAGAGAACTAACTGAGTCAAAATGAAATAATATTTAATAGCCACACG
CCCACACCACTATTAAAGTTCAACACTTGTGGCATATACAAATATACAATGGTTCAGAAGGCAAACCTCTTTTAGCCCTCCATGCCATCTAGAAATGACTAAAACATCTCTATT
ATATAGTTAACTAAATTAAAATTGAAAGTGATTGACGAAGAGATGACAACTAGTACAATAGAAAAAATGTCATAAGTGCTAAGAGGCAAATCCTTTTTTCATTTCTCTCCTTC
AAGGTCAGACATTATCTTAGATGGTTTGGCTAAGAAAAGGGTGTGTTAAATCTGCTCTACTATACCAAAAGAGTGATCAAGGACAAGAACCGAGTCCCTTGCCATATCAAAT
AAAGGTTAACCACCAGAAACTTGCCATATCAATAAAAATACACTTGAAGCCATGTCAACTAATGTCCTAGGAGTTTGCTGGACATAATCTGCATAGTGTTCTGAATTTACCGT
GCCGTTGATCGAAAAACCCTTTGTCCACTTGATGAGCGTTCCAGATGATATCGACAATTGGTTCACTGGAAAGGAAAAGGTGAAACCCAGCTCTCTCTGCCTCCCCTCTGG
GAGGTGGAAATCGTTGCCTTCGGTATCCACAAATTTTGCCAATGCAGCAGCAATGAAATCAAACAGTTCCTGTATCAATAGTTATTAAAGATACATTAGGAACCACGCGCGT
AAATTTGGCAACTCCAGGAAACTTGAGAATCAAACAGCAGGAAACTCACCATGGAAGTCCCAACCATGAGGTGAGGTGGGATAGAGACTTCCTCGGACTGTTGCTTGGCA
ACACGTTTCTCCCTTCCTGCGAGTTGGACTCGGAGTACACGGAAGTTGGTTCCTCCAAGATCCAGCGCATAAAACAGCCCATGCTCGTCTCTGCAGACATAGCACCCAGAT
GATTAAACACAAAAGTTCAAGAGTCAAGAAAAAAAAATGCATGTTTATCAGTGAATTGCACATAATAGTCTGAACGCAACAAAATGTTAGATTTTAAGAACTTGAGTTTCAGT
GAATGATGCTCAACATGTTAAAGCATTTATATACTATCTGTTATAATTCGTGACCAATGTTTTCTCAACATGAGTTCGATCTGAAATTTCCTAGTGTATAAGAGGATGATCCTAG
AAAAACTAAGCAATTCACAGGTTATACGATTAGGTCGAGTAATTGATATAACTCAATAGAGAAATTGCTGACACAAACCATAAAGATGTATCAACGTTTCTAAAGAATTAAGA
GAATTGAGCTTATAGGAGTAAATCTGAGATAACTACGCATGTCATTAATCGAAAAACTAAGCACTTGACCGATTATATCTAGAGATGATTAGGTGGACTGGCAGATATAACTG
AACAAACAAATTATTGAAAGAAACGATAAAGACATATGAACTTTTCTACAGAATTCACAGTATAGGTGTAAACATGCGCTAACTATCAATGCCATTAATTGGACAGTACATGT
ACATCTCACTAACGGGCTATCAAATAGGCTCGCCCCACCAGGCACTCATCGCAATCTCATCATGTCACTCCAAGACTGGTGACATTCTCATTGACACCATATCCTAAATCATC
AAGGTTGGTTTGGAACTACCACGTGATCTGCAAGCATGGGATTTGGCAGGGCGGAATTGGTAGGTAAAAGTATCCTATCCTGGCCCATGATCCATCTAGAATCGGCAGAAGG
GTCGATCTTCCGCAATAAGATCCGGGCATTTCGCCGGCAGGGCGGTCATTCCTAATCAAGAGCGGGAGAGAAGAATTGCTGACGCTTCCGCGATCACGGACTCAGAATTCG
TCCAGGCGCGACGACAGTCCCAACAACAGACGGCGTCTTTGGGATCACAACAACAAGTCAACCGAACGCGCAAAAAGGAAAGGAAAGGGAAAAAAAAAGGAAAGACGG
GATCGGATCGGATCAGATCGGACAAGAATGGGATCTTTGTTACCCGGTGGGGAGGTTGTCGACGTATGAGATGAGCATCTTGAGCTGTGAGTGGATGTCCCCGCGCAGTCC
GCGCTCCATCTCGGCGACGAGCGCGTCGGAGATGCTCCTGAGCAGGGCGGTGGGCGTGGAGAGCGCGCGCTCGACCTCCTCGATCACGTCCGCCACCTTGCGCTTCCGCT
CGGCCTCGGCGGCGCCGTGCAGCTCCGCCGCCCGCTTCTGGCGGCGCCGGCCCAGCACGACCGCCGTGCCCACCGCCGCGCACACCACCACCGCCGTGCCCACCGCCGC
CGCCTTCCCCATCGCTGAGATCAGCCTCTCTCCGCCCAGCGAACGAGCGATCGACCGGCAAGAATCGCAGAGGAATTCCGCTCTGGATCCGAGGGAGTGAAGGAAGGGAT
GGTTCGCCCGGGCGCGTAGGAGAGGGACTGGCCCTCGGCGTTTTAACCGGTAACTTTAACTGGCCCTTGGCTTTGCGGCATTTATGGATCCACCTATAAATACACGAACGGC
CGGGATTTGTCCGGGGGACACGTGGACGGACGGGATCGTGGCCTTGGAGAAGTGGATCTGGATTTTTTGGATTATTAGGGAACGGATTTCTGCTGTGGAAATATTACTAG
ACGGATTTACTCCCAGATATACTTTACGGTAAGTACCTCCCGTGTAATTACCTACGGAGTATATGATTTGGCGAGTAGATCCTGAATTAGAGCGCTGACTTGTGGCGGCGGCC
CGCGTCCCAGTGGG

Firstly, we are going to have a look at how we design primers for specific regions of a DNA sequence. Primer design can be done manually or by using an appropriate software tool.

1. **Click on the Primer design link in blackboard to open the Primer 3 web page.**
2. **Copy the sequence above and paste it into the first box on the web page.**
3. **Change the product size to min 1800, opt 2000 and max 2200.**
4. **Change the primer size max to 24**
5. **In Included region put 1, 2200 (this is a specific region with which we want to design our primers within our sequence, 1 being the first codon and 2200 being the base region upto where we want our primers)**

Your output should look like this:

Primer3 Output

```
No mispriming library specified
Using 1-based sequence positions
OLIGO            start  len      tm     gc%   any    3' seq
LEFT PRIMER         25   19   60.73   57.89  4.00  2.00 GCAGCTTCGACATCGTCTG
RIGHT PRIMER      2024   20   59.84   45.00  7.00  3.00 GCTCATGATGCTGCATTGTT
SEQUENCE SIZE: 5960
INCLUDED REGION SIZE: 2200

PRODUCT SIZE: 2000, PAIR ANY COMPL: 6.00, PAIR 3' COMPL: 0.00

    1 TTTCCGACGACGACGACAGCAAGGGCAGCTTCGACATCGTCTGACGAAATACGGTAGTGA
                                >>>>>>>>>>>>>>>>>>>

   61 CCTCAGAAAGTCAGAATACAAGTAGTAGCAATAAGGTACAGTAGCAAGAGTTGTGCAGTG


  121 CAGTTTCTTAGGGAGAGCAGAGCAGATAAAAATGGTGGAATCGACACTGTGAGCCATCAT


  181 GCAGGTGTATAGGCAACTAAGTTTGATCAGATAGTATCATTTTGAAAACAAGATTGTGCA


  241 TTCAGACATATATATAGTGGACAAATGAGGTAAACACAGCCAAAAGACGGTTAAATTGGA


  301 CTATGTCGATCCCCATCTTTATTGGGAAGAGAAAAATACAAAAGGATTAACAACAAAATG


  361 CTTGGATTACAACACCAGACAAGGGAAAAAGAATCTGGCAGAACAGGAATGGCATTGTTC


  421 AAACCTCCATCTCCATTGTTCCTGCCATCCAAACGTTTATGCATGCCATTAGATAAACCT


  481 CCAAACTCTTTGGCGCTATACTAGCTATTATCCCGCTGCAAAGATGTCTGCAGAATAGCC


  541 GAAGGAAGGTAAAATTCCACCACCAGGGGAGCCCAGAATATTAGGAAGATTTGAAGAGAA


  601 AACGAAGAGGAGAGGCTACACTCAAGAGTTCAGTCCATCACGGCAGTTATCAAAAGACTA


  661 GTCGACCTCCGCATACTGGGAGTGCGAGGCCGCAAGGAGAGCGGCTCCGATGCCAGATCC


  721 GTCGTTGGCCAGCTTAGCAACCAGCGACGAAGAGGCTTCCTCCCCGAGGAGATCTGTAAG


  781 AGTTGCTTCTAAGCAACTGCTGAACTTCTTGTAATGTTCATAGAGCCCACCATCCAAGGC


  841 AACGACGGTCCTTGGCATCGTGCTGCCATCAGTTGGCACTTTGTCACGGCCTAGCTTCTT


  901 CAGGATGCCATATATACCAGCAGCGGCCAAGCGTGCTCCGCGCTCAGCAACAATGTCACA


  961 TATGTGACGAGTAATGTATCTTACTTCCAGGGAAGTATCTGCAACCTGAACATGAAATCG


 1021 AGATGTCATACAAGTAATCAATGGTCATTGAACATTTCTGGTATCCAGTTAGGTTCACGC


 1081 AAGTCATACCCCAACGATATCCTTTAGCTTAGTTCCCAGAACTTTAAGGTCATGTGATGC


 1141 ATCATGGTGCATGGCTGACATATCCGGGGTCCTGCAATGCATAAGATTCAGTAGTTGAAT
```

```
1201 TTATACGGAGTAAGAATACACTACAAAACTTCATACAGATAAACGCTTTAATCGTAATTG

1261 AGCTCATGTTCAGAAAACATTTGTCAAGATTTAACACAGATTTCTGCTACTAATGGCATG

1321 GGCTCTTCATTAGTTCTCAAGAATTCTATGCTATTCAATACACATATCAATCAACAAATA

1381 GTATATCTACTTAGCCTCAAGTAAATATTCGATTAGTTCCAAGAAGAGGAGCATAACACT

1441 ATTCACCATCACTCTTTTGAGCCACTTGAAAAAACAAAAGAAAAAGGAAAACAAACATTT

1501 GGTTTTGAATTAGAATATTAAGGTGTGCACGAAACAAACTCTATCCAAGTAAGCAATTGC

1561 ACACGCCATCCACCAAATTATTCTAATACACCAGAGAAAAAAGATACCAAGCTTTGAATT

1621 GATCGCAGTTGAGGACAAAACAGAGCATTGACCAGATGTAAAGTATTTCATCACCAGTTA

1681 TGATGAATCGGCATTAAATTTTGATTTTGCCACCTCACATAGATTACTTGACTGAAGTAT

1741 ATGTTAACTTCAGAAGGCCCGAAATTTGCCAATTCTTTTATAGCTTTGGGCTTTGGCCTT

1801 ACACGGACTAGTCAACTGATGTAACATGATAAATTCGGACGAGAATCTGTTTCCTTGCGC

1861 AGTTTTGAAAGCTCAATATTAGTGTGTAAAAATTCTTGATCACCGTTTAGCTACAGCAAA

1921 GATGAAGAGTTCAATGACTAATGGGTCAAGAAAGTATACCTCAGTACAAACGGCTGTTCC

1981 AACTTAGGTGGAACAACATCTCCAAACAATGCAGCATCATGAGCCAGTTTCAGTAAGATT
                             <<<<<<<<<<<<<<<<<<<<

2041 CTTCGCACAATCTCTCCAAGATACATGCCAGAAATCATCTTTTCATATATCTGTTACCCA

2101 GAAGATATTCGCATATATTAGTGCAGTTTAGCAACAAAACATAAATGGAAGTGTAAATAC

2161 CACCGAGATAAACACCTGTTCTCCAGGATTCAGACTTTCAAAATCCATAGCTTTATCGTA
                                             ....................

2221 TTCTGAAAGTGGAAGCTTGTCTGATTTAAAGCTTCCCCATTCCGTGTTGATTACCTAGAG
     ............................................................

2281 AACATCAAACAATATTGTAGTTAGGCTAGTAGAGTACAGCAGTGAAGATACGATGCATGT
     ............................................................

2341 GTACGCCCATACGATGTTTAATGGGACTTGTTCATTCCAAGCAGAAAGGGAATTTTAAAC
     ............................................................

2401 ACCAACCATATTTCCGGATTTAGGCAGTAATCCAGTCCATTTTGGAATTGCATTAGCATG
     ............................................................

2461 CTCAACATATGCTGCATTTGTGCCAGTGCCCAAAATTACAGCAGCAACAACATCGTTATC
     ............................................................

2521 CACATATCTCCCACCAGCCAATGTGCCGACTGTGTCATTAACCTACAAAATATCAACAAT
     ............................................................

2581 CAGCCTACATAAAGGAAAATCTCAAAAAAGAAGAAAGTTCTGCATGACGCAAAGCATGTT
     ............................................................

2641 TCCATCCCCATCAACAGAAAAGATTAACGAGTAGCATCAGGACACACAAAAATAAAAATT
     ............................................................

2701 GCATTATTTTGCGGGAGTAAAAAATAGATATGGTGTATTACCAATGCTGAAACTTTCATA
```

....

2761 TCCAACCCCTGCCTCTCCATGGCCTTGCTCAACTCAGCCACAACATCCTCACCAACCTGA
....

2821 TTAGGGAGTAACCAAATAACAAATTAAAATCAAAGTAATTGTGGTTATTATGGGATGATT
....

2881 GAAAAATTGGCTCAGAGACCATCAACCGAGAAACTTAATATTCTTAAGTATCACAAATAT
....

2941 ATTCTTCCTTACCAAGCTCTACATTTCATCGTTGTGATAAACTCAGATAACATCAACCCA
....

3001 GAAACATACTACTCTTACTGCATCATAAAAATGTGCTTCCTAACCAAGCTTTACATTTCA
....

3061 TCATTGAGATAAAGAAATCAGTTTGACCATAATGTAATGAGTGATAACGAAGTATCAGTG
....

3121 AACTAAATGAATCAACATGAAATAATATTTAATAACATTTCATCATTGAGATAAAGAAAT
....

3181 CAGTTTGACCATAATGCAATGACTGATTGAGAAGTATCAGAGAACTAACTGAGTCAAAAT
....

3241 GAAATAATATTTAATAGCCACACGCCCACACCACTATTAAAGTTCAACACTTGTGGCATA
....

3301 TACAAATATACAATGGTTCAGAAGGCAAACCTCTTTTAGCCCTCCATGCCATCTAGAAAT
....

3361 GACTAAAACATCTCTATTATATAGTTAACTAAATTAAAATTGAAAGTGATTGACGAAGAG
....

3421 ATGACAACTAGTACAATAGAAAAAATGTCATAAGTGCTAAGAGGCAAATCCTTTTTTCAT
....

3481 TTCTCTCCTTCAAGGTCAGACATTATCTTAGATGGTTTGGCTAAGAAAAGGGTGTGTTAA
....

3541 ATCTGCTCTACTATACCAAAAGAGTGATCAAGGACAAGAACCGAGTCCCTTGCCATATCA
....

3601 AATAAAGGTTAACCACCAGAAACTTGCCATATCAATAAAAATACACTTGAAGCCATGTCA
....

3661 ACTAATGTCCTAGGAGTTTGCTGGACATAATCTGCATAGTGTTCTGAATTTACCGTGCCG
....

3721 TTGATCGAAAAACCCTTTGTCCACTTGATGAGCGTTCCAGATGATATCGACAATTGGTTC
....

3781 ACTGGAAAGGAAAAGGTGAAACCCAGCTCTCTCTGCCTCCCCTCTGGGAGGTGGAAATCG
....

3841 TTGCCTTCGGTATCCACAAATTTTGCCAATGCAGCAGCAATGAAATCAAACAGTTCCTGT
....

3901 ATCAATAGTTATTAAAGATACATTAGGAACCACGCGCGTAAATTTGGCAACTCCAGGAAA
....

3961 CTTGAGAATCAAACAGCAGGAAACTCACCATGGAAGTCCCAACCATGAGGTGAGGTGGGA
....

4021 TAGAGACTTCCTCGGACTGTTGCTTGGCAACACGTTTCTCCCTTCCTGCGAGTTGGACTC
....

4081 GGAGTACACGGAAGTTGGTTCCTCCAAGATCCAGCGCATAAAACAGCCCATGCTCGTCTC
....

4141 TGCAGACATAGCACCCAGATGATTAAACACAAAAGTTCAAGAGTCAAGAAAAAAAATGCA
....

4201 TGTTTATCAGTGAATTGCACATAATAGTCTGAACGCAACAAAATGTTAGATTTTAAGAAC
....

```
4261 TTGAGTTTCAGTGAATGATGCTCAACATGTTAAAGCATTTATATACTATCTGTTATAATT
     ............................................................

4321 CGTGACCAATGTTTTCTCAACATGAGTTCGATCTGAAATTTCCTAGTGTATAAGAGGATG
     ............................................................

4381 ATCCTAGAAAAACTAAGCAATTCACAGGTTATACGATTAGGTCGAGTAATTGATATAACT
     ............................................................

4441 CAATAGAGAAATTGCTGACACAAACCATAAAGATGTATCAACGTTTCTAAAGAATTAAGA
     ............................................................

4501 GAATTGAGCTTATAGGAGTAAATCTGAGATAACTACGCATGTCATTAATCGAAAAACTAA
     ............................................................

4561 GCACTTGACCGATTATATCTAGAGATGATTAGGTGGACTGGCAGATATAACTGAACAAAC
     ............................................................

4621 AAATTATTGAAAGAAACGATAAAGACATATGAACTTTTCTACAGAATTCACAGTATAGGT
     ............................................................

4681 GTAAACATGCGCTAACTATCAATGCCATTAATTGGACAGTACATGTACATCTCACTAACG
     ............................................................

4741 GGCTATCAAATAGGCTCGCCCCACCAGGCACTCATCGCAATCTCATCATGTCACTCCAAG
     ............................................................

4801 ACTGGTGACATTCTCATTGACACCATATCCTAAATCATCAAGGTTGGTTTGGAACTACCA
     ............................................................

4861 CGTGATCTGCAAGCATGGGATTTGGCAGGGCGGAATTGGTAGGTAAAAGTATCCTATCCT
     ............................................................

4921 GGCCCATGATCCATCTAGAATCGGCAGAAGGGTCGATCTTCCGCAATAAGATCCGGGCAT
     ............................................................

4981 TTCGCCGGCAGGGCGGTCATTCCTAATCAAGAGCGGGAGAGAAGAATTGCTGACGCTTCC
     ............................................................

5041 GCGATCACGGACTCAGAATTCGTCCAGGCGCGACGACAGTCCCAACAACAGACGGCGTCT
     ............................................................

5101 TTGGGATCACAACAACAAGTCAACCGAACGCGCAAAAAGGAAAGGAAAGGGAAAAAAAAG
     ............................................................

5161 GAAAGACGGGATCGGATCGGATCAGATCGGACAAGAATGGGATCTTTGTTACCCGGTGGG
     ............................................................

5221 GAGGTTGTCGACGTATGAGATGAGCATCTTGAGCTGTGAGTGGATGTCCCCGCGCAGTCC
     ............................................................

5281 GCGCTCCATCTCGGCGACGAGCGCGTCGGAGATGCTCCTGAGCAGGGCGGTGGGCGTGGA
     ............................................................

5341 GAGCGCGCGCTCGACCTCCTCGATCACGTCCGCCACCTTGCGCTTCCGCTCGGCCTCGGC
     ............................................................

5401 GGCGCCGTGCAGCTCCGCCGCCCGCTTCTGGCGGCGCCGGCCCAGCACGACCGCCGTGCC
     ............................................................

5461 CACCGCCGCGCACACCACCACCGCCGTGCCCACCGCCGCCGCCTTCCCCATCGCTGAGAT
     ............................................................

5521 CAGCCTCTCTCCGCCCAGCGAACGAGCGATCGACCGGCAAGAATCGCAGAGGAATTCCGC
     ............................................................

5581 TCTGGATCCGAGGGAGTGAAGGAAGGGATGGTTCGCCCGGGCGCGTAGGAGAGGGACTGG
     ............................................................

5641 CCCTCGGCGTTTTAACCGGTAACTTTAACTGGCCCTTGGCTTTGCGGCATTTATGGATCC
     ............................................................

5701 ACCTATAAATACACGAACGGCCGGGATTTGTCCGGGGGACACGTGGACGGACGGGATCGT
     ............................................................

5761 GGCCTTGGAGAAGTGGATCTGGATTTTTTTGGATTATTAGGGAACGGATTTTCTGCTGTG
     ............................................................
```

```
 5821 GAAATATTACTAGACGGATTTACTCCCAGATATACTTTACGGTAAGTACCTCCCGTGTAA
      ............................................................

 5881 TTACCTACGGAGTATATGATTTGGCGAGTAGATCCTGAATTAGAGCGCTGACTTGTGGCG
      ............................................................

 5941 GCGGCCCGCGTCCCAGTGGG
      ....................

KEYS (in order of precedence):
...... vector sequence
>>>>>> left primer
<<<<<< right primer

ADDITIONAL OLIGOS
                    start  len      tm     gc%   any    3' seq

 1 LEFT PRIMER         25   19   60.73   57.89  4.00  2.00 GCAGCTTCGACATCGTCTG
   RIGHT PRIMER      2020   20   60.08   40.00  7.00  0.00 ATGATGCTGCATTGTTTGGA
   PRODUCT SIZE: 1996, PAIR ANY COMPL: 6.00, PAIR 3' COMPL: 0.00

 2 LEFT PRIMER         20   18   59.94   55.56  4.00  2.00 CAAGGGCAGCTTCGACAT
   RIGHT PRIMER      2020   20   60.08   40.00  7.00  0.00 ATGATGCTGCATTGTTTGGA
   PRODUCT SIZE: 2001, PAIR ANY COMPL: 6.00, PAIR 3' COMPL: 1.00

 3 LEFT PRIMER         25   19   60.73   57.89  4.00  2.00 GCAGCTTCGACATCGTCTG
   RIGHT PRIMER      2018   20   59.81   45.00  7.00  0.00 GATGCTGCATTGTTTGGAGA
   PRODUCT SIZE: 1994, PAIR ANY COMPL: 6.00, PAIR 3' COMPL: 2.00

 4 LEFT PRIMER         20   18   59.94   55.56  4.00  2.00 CAAGGGCAGCTTCGACAT
   RIGHT PRIMER      2018   20   59.81   45.00  7.00  0.00 GATGCTGCATTGTTTGGAGA
   PRODUCT SIZE: 1999, PAIR ANY COMPL: 6.00, PAIR 3' COMPL: 1.00

Statistics
         con   too    in    in          no    tm    tm  high  high        high
         sid  many   tar  excl   bad    GC   too   too   any    3'  poly   end
        ered    Ns   get   reg   GC% clamp   low  high compl compl     X  stab    ok
Left    2669     0     0     0    10     0  1625   322    39    50     0    32   591
Right   2663     0     0     0    20     0  1786   185     1    16     0    45   610
Pair Stats:
considered 2225, unacceptable product size 1767, high end compl 54, ok 404
primer3 release 1.1.4
```

The arrows below the lines of sequence show the location of the primers. From this output we can see the top primers picked by Primer 3 and some alternatives below the sequence output. This way we can choose which primers we deem the most suitable depending on the criteria.

VALIDATE OUR PRIMERS

Next we need to validate our primers to check that they will amplify a region of sequence specific to what we need.

1. **Click on the NCBI Primer BLAST link on Blackboard**
2. **Copy and paste the left primer and right primer output from primer 3 into the forward primer and reverse primer parameter boxes respectively.**
3. **Change the PCR product size to 1000 min and 2200 max**
4. **Change the Organism box in primer pair specificity checking parameters from homo sapiens to plants.**
5. **Click show results in a new window and select GET PRIMERS**

We should now see a list of products that the primers are hitting within the database and we can check that these are hitting sensible results. This enables us to check the selectivity of the

primers we are designing and whether we need to go back and redesign some more appropriate primers.

2. DNA Sequence Mutations

In this module we have determined the presence of a nematode drug resistant strain and a drug susceptible strain. From the drug screening assay we looked at in workshop 1 we are able to identify and observe the phenotype using microscopy. To further analyse the presence of any drug resistant strains and the drug susceptible strains, we can extract DNA from the worms and used PCR to amplify a region of DNA in order to determine whether a potential sequence mutation may be present in our nematode strains.

In this section of the workshop we will align the DNA sequences from the nematode stains and identify any sequence mutations. We will then BLAST the sequences using the NCBI database

The DNA sequences below are from our laboratory samples:

Susceptible strain

tagtgaagttagtttagataatgaaaatgtataagaaaaccctataaaatgttttcgactttttgttcaatgtttatta
aaaagtgtctatttctgagtttctgaactttcacaattttgaaagtcgacaaaaaaatgatttcttatattttgatcaga
ctaagggcataattatttaagcattatagatcttcaaacacgcaataatactatatttctcaaaaagacaaaatatttcg
agaagttggcaagtcgaaatttcaaagttccctaaactaatggaatgtagtacattttcatattcttcgttttgttcgat
ttcttttaatttcccacaacatattttttaaagtaaatacctcagtatctccagattcttcatatctttcataagccaa
tctttgatctgtccattcttctctgaatgtgaattgagcactgtattccatatttacgtcatcaatttttgatattgatc
gaagataaatgtttactgtcactagaacgggtccaccggtatctgatagtttagtttttgaattaattctgtcataatag
aagtttaaatttgtatttttgccaaggtttgcaaattttcaagggtcaaaaatcttttgaaatactttttgggattatg
caataagtggtttttcaagtgcaaaactcaatttcttacagaaaaaactcctgaacctgccgattcagcaaagtaaattt
tcaagaaccgtattttcacagttttcaaaatctagattaattcatataggttcggtgcttagtaattctcgtggtgagac
ccatctgaatcgagttgcttggctataggtcagcgtttcagaaaaatgcaaaacaactataactatttgtttttcattag
ttaacttggtttctgaattcaattccaataatttcttgaatttaaaaactaaaaatactcacctggccacgtggcattca
ttccacgtggtctaactctccaatcatagtctttcaatatt

Resistant strain

tagtgaagttagtttagataatgaaaatgtataagaaaaccctataaaatgttttcgactttttgttcaatgtttatta
aaaagtgtctatttctgagtttctgaactttcacaattttgaaagtcgacaaaaaaatgatttcttatattttgatcaga
ctaagggcataattatttaagcattatagatcttcaaacacgcaataatactatatttctcaaaaagacaaaatatttcg
agaagttggcaagtcgaaatttcaaagttccctaaactaatggaatgtagtacattttcatattcttcgttttgttcgat
ttcttttaatttcccacaacatattttttaaagtaaatacctcagtatctccagattcttcatatctttcataagccaa
tctttgatctgtccattcttctctgaatgtgaattgagcactgtattccatatttacgtcatcaatttttgatattgatc
gaagataaatgtttactgtctctagaacgggtccaccggtatctgatagtttagtttttgaattaattctgtcataatag
aagtttaaatttgtatttttgccaaggtttgcaaattttcaagggtcaaaaatcttttgaaatactttttgggattatg
caataagtggtttttcaagtgcaaaactcaatttcttacagaaaaaactcctgaacctgccgattcagcaaagtaaattt
tcaagaaccgtattttcacagttttcaaaatctagattaattcatataggttcggtgcttagtaattctcgtggtgagac
ccatctgaatcgagttgcttggctataggtcagcgtttcagaaaaatgcaaaacaactataactatttgtttttcattag
ttaacttggtttctgaattcaattccaataatttcttgaatttaaaaactaaaaatactcacctggccacgtggcattca
ttccacgtggtctaactctccaatcatagtctttcaatatt

We can conduct a sequence alignment on our sequences to see if we can observe any mutations occurring between our susceptible and resistant strains.

1. **Click on the NCBI alignment link in Blackboard.**
2. **Copy and paste the first sequence into the box at the top of the page and the second sequence in the box below it on the NCBI alignment webpage.**
3. **Make sure highly similar sequences (megablast) is selected**
4. **Click show results in a new window**
5. **Click BLAST**

The results showing the alignment of the two sequences will be shown in a new window and we can now look to see if there are any mutations within our sequence. The sequences should be very similar as they are from the same species of nematode.

Click on Graphics this will bring up a new window (here we will be able to see if any sequence changes are highlighted). Here we can see exactly what and where the mutation is occurring.

Click on the button with the highlighted mutation in the middle of the page (zoom to sequence). We can then see the two sequences aligned and the highlighted mutation in red.

You can use the pan left and right arrows to look along the sequence for any further changes within the two alignments.

Sequence BLAST

We can also BLAST our sequence to see which organism it hits within the NCBI database.

1. **Click on NCBI BLAST in Blackboard**
2. **Click on nucleotide BLAST**
3. **Make sure align sequences is unchecked**
4. **Copy and paste the top sequence in the box at the top of the webpage.**
5. **In the Organism box type Metazoa**
6. **Click results in new window**
7. **Click BLAST**