The original Sanger sequencing file in seq format is changed to fasta format with TextEdit. The forward Sanger sequencing files are stored as F000~003.fasta and the reverse Sanger sequencing files are stored as R004~006.fasta, which are shown below:
the two sequences match for 50 continuous nucleotides. The numbers in red files and the results are shown below. The blue highlight shows the first time that

EMBOSS needle is used to align the tandemly arranged Sanger sequencing files with EMBOSS revseq, which are stored as R004~006rev.fasta.

2. The reverse Sanger sequencing files are changing to the corresponding forward Sanger sequencing files with EMBoss revseq, and are stored as R004~006rev.fasta.

3. EMBoss needle is used to align the tandemly arranged Sanger sequencing files and the results are shown below. The blue highlight shows the first time that the two sequences match for 50 continuous nucleotides. The numbers in red files and the results are shown below. The blue highlight shows the first time that
F000                901 AGAAGAATGGCCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGCCTGAC
F001                223 AGAAGAATGGCCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGCCTGAC
F000                1000
F001                321
F000                1048
F001                371
F000                1090
F001                419
F000                1132
F001                469
F000                1177
F001                515
F000                1196
F001                665
F000                1196
F001                615
F000                1196
F001                1196
F000                1196
F001                715
F000                1196
F001                765
F000                1196
F001                815
F000                1196
F001                1196
F000                1196
F001                865
F000                1196
F001                915
F000                1196
F001                1196
F000                1196
F001                1015
F000                1196
F001                1049
F001.fasta-F002.fasta
F001                1 CTGTATTTCGATCAGTTCGGGGCCACTTCCTGATCGAGGGCGACCTGAAC
F002                0
F001                50
F002                0
F001                100
F002                0
F001                150
F001
1033 CC--AGCACAG--CTGCTGTA---------------------------------------- 1049
F002
393 CCAAGCACAGCTGCTGACGAGTACTTCACCAGCTGCTATGATGAGCTGAC 442
F001
1049 --------------------------------------------------------------- 1049
F002
443 CAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCTTCTGAGCAG 592
F001
1049 --------------------------------------------------------------- 1049
F002
493 GCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAA 542
F001
1049 --------------------------------------------------------------- 1049
F002
543 GTGACCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTT 592
F001
1049 --------------------------------------------------------------- 1049
F002
593 CGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCTGGG 642
F001
1049 --------------------------------------------------------------- 1049
F002
643 GCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGAC 692
F001
1049 --------------------------------------------------------------- 1049
F002
693 AATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACT 742
F001
1049 --------------------------------------------------------------- 1049
F002
743 GTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACC 792
F001
1049 --------------------------------------------------------------- 1049
F002
793 TGTTGCACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACACCGGC 842
F001
1049 --------------------------------------------------------------- 1049
F002
843 CAGCATCTGCGACTGAGTGTGACAGCTCTGAATGATGACCGCCAAGCCGAA 892
F001
1049 --------------------------------------------------------------- 1049
F002
893 GCAGCGAGAGAATTCTGCAAGCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCC 100
F001
1049 --------------------------------------------------------------- 1049
F002
943 TGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTC 1092
F001
1049 --------------------------------------------------------------- 1049
F002
993 TCATGCAGCTGATCACGACGACAGCTGACTTAAGAGACTGCAGATGCCAG 1106

F002.fasta-F003.fasta
F002
1 GGGATCATCTGGAAGTGGACGGCACCGAGGAACTGCTCGTGAAGCTGAAC 50
F003
0 --------------------------------------------------------------- 0
F002
51 AGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGCATCCC 100
F003
0 --------------------------------------------------------------- 0
F002
101 CCACCATGCCACCTGCGGAGAGCTGCAACGACCATTCTGCGCCGCGCAAGGAAG 150
F003
0 --------------------------------------------------------------- 0
F002
151 ATTTTTAACCATTCTGAAAGGACAAACCGGGAAAGATCGAGAAGATCCTG 200
F003
  455 GGTG-GACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATC 503
F002
  1098 TGGTATTCG----------------------------------------- 1106
  504 --GTGATCGAATGGCCAGAGAGAACCAGACCACCCAGAAGGGGACAAG 551
F002
  1106 -------------------------------------------------- 1106
F003
  552 AACAGGCCGGAGAAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGG 601
F002
  1106 -------------------------------------------------- 1106
F003
  602 CAGCCAGATCTGAAAGACCAACCCCGTGGAACACCCAGCTGCAAGACG 651
F002
  1106 -------------------------------------------------- 1106
F003
  652 CAGCCAGATCCTGAAAGAACCCCGTGGAAAACACCCAGCTGCAGAACG 701
F002
  1106 -------------------------------------------------- 1106
F003
  702 CAGGAACTGGAATCAACCAGGCTGCTCGACTAGTGTGGACCATATCGT 751
F002
  1106 -------------------------------------------------- 1106
F003
  752 GCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAAGTGCTGACCA 801
F002
  1106 -------------------------------------------------- 1106
F003
  802 GAAGGGCAAAACCCGGGCAAGACGCAAGAGAGTGGCCCTCAGAGAGGT 851
F002
  1106 -------------------------------------------------- 1106
F003
  852 AGAAGCTGTACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGAC 901
F002
  1106 -------------------------------------------------- 1106
F003
  902 CAGGAACTGGACATCAACCGGCTGTCCGACTACGATGTGGACCATATCGT 951
F002
  1106 -------------------------------------------------- 1106
F003
  952 GCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAAGTGCTGACCA 1001
F002
  1106 -------------------------------------------------- 1106
F003
  1002 TCACAAAGCACTGGCACAGATCTGACTCCCGGATGACACTTAGTACGAC 1051
F002
  1106 -------------------------------------------------- 1106
F003
  1052 GAGATGACAGCCTGATCGGCAGTGAAGTGATCACCCTTGAGTTCAGCTTG 1101
F002
  1106 -------------------------------------------------- 1106
F003
  1102 GGTTCCGATTCGCAAAGAATTCCTAGCTTACCAAATTGC 1140

F003.fasta-R004rev.fasta
  1 -----------------------GGCCAGGCTACACGA-TCTGCTGAA-T 25
R004rev
  1 GATGGTCTGATTCAACGCCTCCTGGGCACACATCACGATTCTGCTGAAGC 50
F003
  26 TATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGG 75
R004rev
  51 AATC-AGGAC-ATGACTGCTTGACCAATGA-GAAAACGA-GACA-TCTGA 95
F003
  76 AAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAG 125
R004rev
  96 AAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAG 143
F003
  126 GAACGGCTGAAAACCTATGCCCCACCTGTTCGACGACAAAGTGATGAGC 174
R004rev
  144 GAACGGCTGAAAACCTATGCCCCACCTGTTCGACGACAAAGTGATGAGC 193
F003
  175 AGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTG 224
R004rev
  194 AGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTG 243
4. The numbers with red highlight are F000 (1-700), F001 (23-700), F002 (44-697), F003 (56-174), R004rev (194-900), R005rev (116-742), R006rev (274-1199), which are shown in red highlight in the corresponding Sanger sequencing files as below.

> F000

```
GCTGACTCTTCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGG
AAGTTCTGTTCCAGGGGCCCCATATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACG
GAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCA
GCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGG
CCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGG
TGGACGACAGCTTCTCAGCTGGGAGGCGACCCAAACTGAGGAAGGCTGGTTCAAGGAGTGACGAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGG
TGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACG
```
5. The nucleotides in red highlight in step 4 are joined via copy and paste to obtain the merged Sanger sequence, as shown below.

```plaintext
>gene
TTCCCTTGGACACCAGTCATCTGTAGCTGACCAATCATCGCTGCTGAGAACCTGGCTGACAGGTGGCCACCCCGCATCCTGGGCTGATCTATCTGGTGGACGACAGCTTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGGAGTCCCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCCAAGTTCTGTTCCAGGGGCCCCATATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGCGGATCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGCTGACTCTTCCCTCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGGAAGGTCTGAAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGAGCGATTTCCGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACCCAAAGAAGAAGATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAATACTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAATACTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAATACTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTTGAACACGACGACAGATCCGGCCGAAGCAAGGAGAAGATGATCCACACTCATATACATACACCTGACATCTCATATCATACACACCGGCTGGTGGAAGGGCAGAAGATACAGCTGACCAATACATAGCTGAGACAGATCCGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTTGAACACGACGACAGATCCGGCCGAAGCAAGGAGAAGATGATCCACACTCATATACATACACCTGACATCTCATATCATATCATACACACCGGCTGGTGGAAGGGCAGAAGATACAGCTGACCAATACATAGCTGAGACAGATCCGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTTGAACACGACGACAGATCCGGCCGAAGCAAGGAGAAGATGATCCACACTCATATACATACACCTGACATCTCATATCATATCATACACACCGGCTGGTGGAAGGGCAGAAGATACAGCTGACCAATACATAGCTGAGACAGATCCGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTTGAACACGACGACAGATCCGGCCGAAGCAAGGAGAAGATGATCCACACTCATATACATACACCTGACATCTCATATCATATCATACACACCGGCTGGTGGAAGGGCAGAAGATACAGCTGACCAATACATAGCTGAGACAGATCCGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTTGAACACGACGACAGATCCGGCCGAAGCAAGGAGAAGATGATCCACACTCATATACATACACCTGACATCTCATATCATATCATACACACCGGCTGGTGGAAGGGCAGAAGATACAGCTGACCAATACATAGCTGAGACAGATCCGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTTGAACACGACGACAGATCCGGCCGAAGCAAGGAGAAGATGATCCACACTCATATACATACACCTGACATCTCATATCATATCATACACACCGGCTGGTGGAAGGGCAGAAGATACAGCTGACCAATACATAGCTGAGACAGATCCGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTTGAACACGACGACAGATCCGGCCGAAGCAAGGAGAAGATGATCCACACTCATATACATACACCTGACATCTCATATCATATCATACACACCGGCTGGTGGAAGGGCAGAAGATACAGCTGACCAATACATAGCTGAGACAGATCCGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTTGAACACGACGACAGATCCGGCCGAAGCAAGGAGAAGATGATCCACACTCATATACATACACCTGACATCTCATATCATATCATACACACCGGCTGGTGGAAGGGCAGAAGATACAGCTGACCAATACATAGCTGAGACAGATCCGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTTGAACACGACGACAGATCCGGCCGAAGCAAGGAGAAGATGATCCACACTCATATACATACACCTGACATCTCATATCATATCATACACACCGGCTGGTGGAAGGGCAGAAGATACAGCTGACCAATACATAGCTGAGACAGATCCGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCTTCAAGTACTTTTGAACACGACGACAGATCCGGCCGAAGCAAGGAGAAGATGATCCACACTCATATACATACACCTGACAT11
```
The manual joined sequence is aligned to the script merged sequence with EMBOSS needle and the results is shown below. As can be seen, the two sequence are identical.

```plaintext
7. The manual joined sequence is aligned to the script merged sequence with EMBOSS needle and the results is shown below. As can be seen, the two sequence are identical.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Merged Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GCTGACTCTCTTCCCTCTAGATAAATTGTTTTAATTCTTTAAGAGGAGATAT 50</td>
</tr>
<tr>
<td>51</td>
<td>ACCATGGGCAAGAGAGCCATCATCTACATCATCAGAACGCGGCTGGAATGT 100</td>
</tr>
<tr>
<td>101</td>
<td>TGCTGTTTCAGGGCCCCTGATATGCTACATGACTGGTGGGACACAAATTG 150</td>
</tr>
<tr>
<td>151</td>
<td>TGCAGGGCCATCAGAACAGGGTCCGCTCTCATCAGAGCTGACCTGCCCAAACAAT 200</td>
</tr>
<tr>
<td>251</td>
<td>GGCTGCGCGCTGACAGGAGAAGTGCTGCTTCAAGAAGAGGATGATTCTCCGCGGTGGAATG 300</td>
</tr>
<tr>
<td>301</td>
<td>AGTGGCTGGAACCCGGCAGCATCAGCCATGACGTCCGCTGGAACCCCAATGG 350</td>
</tr>
<tr>
<td>351</td>
<td>GCCCTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACCCAAAGAAGAAG 400</td>
</tr>
</tbody>
</table>
```

7. The manual joined sequence is aligned to the script merged sequence with EMBOSS needle and the results is shown below. As can be seen, the two sequence are identical.
merged 1301 AGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGA 1350
gene 1351 TGGACGCCACGGAGAAGGCTGTCAGTTGGAAGCTGAAAGAGGACCTGCTG 1400
merged 1351 TGGACGCCACGGAGAAGGCTGTCAGTTGGAAGCTGAAAGAGGACCTGCTG 1400
gene 1401 CGGAAGCGGACCTTGGACAAAGGGAAGCCTTGGACAAAGGGAAGCCTTGGACAA 1450
merged 1401 CGGAAGCGGACCTTGGACAAAGGGAAGCCTTGGACAAAGGGAAGCCTTGGACAA 1450
gene 1451 GGGAAGACTGGCAGCCTCATTGGGACAGGAAGAGATTTTTACCCATTCC 1500
merged 1451 GGGAAGACTGGCAGCCTCATTGGGACAGGAAGAGATTTTTACCCATTCC 1500
gene 1501 TGAAAGCAGAACCCGAGAACAGCAGATCTGACCTTGGCAGGATCCC 1550
merged 1501 TGAAAGCAGAACCCGAGAACAGCAGATCTGACCTTGGCAGGATCCC 1550
gene 1551 TACTACGTGAGCCCTCTGGCCAGGGGAAACAGCAGATTCGCTGGATGAC 1600
merged 1551 TACTACGTGAGCCCTCTGGCCAGGGGAAACAGCAGATTCGCTGGATGAC 1600
gene 1601 CAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGG 1650
merged 1601 CAGAAAGAGCGAGGAAACCATCACCCCCTGGAACTTCGAGGAAGTGGTGG 1650
gene 1651 TACTACGTGAGCCCTCTGGCCAGGGGAAACAGCAGATTCGCTGGATGAC 1700
merged 1700 AAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGA 1750
gene 1701 AAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGA 1750
merged 1701 AAGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGA 1750
gene 1751 GTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGG 1800
merged 1751 GTACTTCACCGTGTATAACGAGCTGACCAAAGTGAAATACGTGACCGAGG 1800
gene 1801 GAATGAGAAAGCCCGCCTTCCTGAGCCGAGGAGAAAAGGACATTCTG 1850
merged 1801 GAATGAGAAAGCCCGCCTTCCTGAGCCGAGGAGAAAAGGACATTCTG 1850
gene 1851 GACCTGCTGTTCAAGACCAACCCGAAAGTGACCAGCTGAAACAAGGAGAAGTGGGAGG 1900
merged 1900 GGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCG 1950
merged 1901 GGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCG 1950
gene 1951 TGGAAGATCCTGGCTACCGCCTGACAGTGAAAGATAGTGGAAACAGTGAGCAGG 2000
merged 1951 TGGAAGATCCTGGCTACCGCCTGACAGTGAAAGATAGTGGAAACAGTGAGCAGG 2000
gene 2001 ATTATCAAGGACAAGGACTTCCTTGGACAAATGAGGAAAAAAGGAGAATCTGCTTCCTCCGAG 2050
merged 2001 ATTATCAAGGACAAGGACTTCCTTGGACAAATGAGGAAAAAAGGAGAATCTGCTTCCTCCGAG 2050
gene 2051 GAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCGACGAGG 2100
merged 2051 GAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCGACGAGG 2100