

# Detection and identification of transgenic events by next generation sequencing combined with enrichment technologies

Frédéric Debode, Julie Hulin, Benoît Charlotheaux, Wouter Coppieters, Marc Hanikenne, Latifa Karim and Gilbert Berben

**Supplementary Table S1. Sequences of the probes used for enrichment**

| Structural element | Probe sequences  |
|--------------------|--|
| p35S               | CGATGGCCTTTCCTTTATCGCAATGATGGCATTGTAGGAGCCACCTTCCCTTTTCCACTATCTTCACA<br>ATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTCCGGATATTA  |
|                    | CCCTTTGTTGAAAAGTCTCAATTGCCCTTGGTCTTCTGAGACTGTATCTTTGATATTTTTGGAGTAG<br>ACAAGTGTGTCGTGCTCCACCATGTTGACGAAGATTTCTTCTTGTCATTGA   |
|                    | TTTTGACTGCATCTTTAACCTTCTTGGGAAGGATTTGATTTCTGGAGATTATTGCTCGGGTAGATC<br>GTCTTGATGAGACCTGCTGCGTAAGCCTCTCTAACCATCTGTGGGTTAGCAT   |
|                    | TGGAGATATCACATCAATCCACTTGCTTTGAAGACGTGGTTGGAACGTCTTCTTTTTCCACGATGCTC<br>CTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGCAGAGGCATCTTCAA  |
|                    | TCTTGAAAGATCTGCTAGAGTCAGCTTGTGACGCTGTTCTCTCAAATGAAATGAACTTCTTATATA<br>GAGGAAGGGTCTTGCGAAGGATAGTGGGATTGTGCGTCATCCCTTACGTCAG   |
|                    | GTCGTAAGAGACTCTGTATGAACTGTTCCGAGTCTTTACGGCGAGTTCTGTTAGGTCTCTATTTG<br>AATCTTTGACTCCATGGCCTTTGATTCAGTGGGAACTACCTTTTTAGAGACTC   |
|                    | CAATCTCTATTACTTGCCTTGGTTTGTGAAGCAAGCCTTGAATCGTCCATACTGGAATAGTACTTCTG<br>ATCTTGAGAAATATATCTTTCTCTGTGTTCTTGATGCAGTTAGTCCTGAATC |
|                    | GAAGGTATTTGATTTCTGGAGATTATTGCTCGGGTAGATCGTCTTGATGAGACCTGCTGCGTAAGC<br>CTCTCTAACCATCTGTGGGTTAGCATTCTTTCTGAAATTGAAAAGGCTAATCT  |
| pFMV               | AATTCTCAGTCCAAAGCCTCAACAAGGTCAGGGTACAGAGTCTCCAAACCATTAGCCAAAAGCTACA<br>GGAGATCAATGAAGAATCTTCAATCAAAGTAACTACTGTTCCAGCACATGCA  |
|                    | AACAAAATAACGTGAAAAGAGCTGTCCTGACAGCCCACTCACTAATGCGTATGACGAACGCAGTG<br>ACGACCACAAAAGAATCCCTCTATATAAGAAGGCATTCCATCCATTTGAAGG    |
|                    | AAAGTAATCTTGCAACATCGAGCAGCTGGCTTGTGGGGACCAGACAAAAAGGAATGGTGCAGAA<br>TTGTTAGGCGCACCTACCAAAGCATCTTTGCCCTTATTGCAAAGATAAAGCAG    |
|                    | ATTCCTCTAGTACAAGTGGGGAACAAAATAACGTGGAAAAGAGCTGTCCTGACAGCCCACTCACTAA<br>TGCGTATGACGAACGCAGTGACGACCACAAAAGAATTCCCTCTATATAAGAAG |
|                    | TACAGGAGATCAATGAAGAATCTTCAATCAAAGTAACTACTGTTCCAGCACATGCATCATGGTCAGT<br>AAGTTTCAGAAAAAGACATCCACCGAAGACTTAAAGTTAGTGGGCATCTTTG  |
|                    | AAATTGGTATCGCCAAAACCAAGAAGGAACTCCATCCTCAAAGGTTTGTAAAGGAAGAATTCTCAGT<br>CCAAAGCCTCAACAAGGTCAGGGTACAGAGTCTCCAAACCATTAGCCAAAAGC |

|      |  |
|------|--|
|      | GGTGCAGAATTGTTAGGCGCACCTACCAAAGCATCTTTGCCTTTATTGCAAAGATAAAGCAGATTC<br>CTCTAGTACAAGTGGGGAACAAAATAACGTGGAAAAGAGCTGTCCTGACAGCC    |
|      | TCATGGTCAGTAAGTTTCAGAAAAAGACATCCACCGAAGACTTAAAGTTAGTGGGCATCTTTGAAAG<br>TAATCTTGCAACATCGAGCAGCTGGCTTGTGGGGACCAGACAAAAAAGGAAT    |
|      | CACTCACTAATGCGTATGACGAACGCAGTGACGACCACAAAAGAATTAGCTTGAGCTCAGGATTTAG<br>CAGCATTCCAGATTGGGTTCAATCAACAAGGTACGAGCCATATCACTTTATTC   |
| pUbi | GCCCGGTAGTTCTACTTCTGTTTCATGTTTGTGTTAGATCCGTGTTTGTGTTAGATCCGTGCTGCTAGC<br>GTTTCGTACACGGATGCGACCTGTACGTCAGACACGTTCTGATTGCTAACTTG |
|      | TGGCATATGCAGCATCTATTCATATGCTCTAACCTTGAGTACCTATCTATTATAATAACAAGTATGTT<br>TTATAATTATTTTGATCTTGATATACTTGGATGATGGCATATGCAGCAGCT    |
|      | CCAGTGTTTCTCTTTGGGGAATCCTGGGATGGCTCTAGCCGTTCCGCAGACGGGATCGATTTTCATG<br>ATTTTTTTGTTTCGTTGCATAGGGTTTGGTTGCCCTTTTCTTTATTTCAAT     |
|      | AGTTTTTTTATTTAATAATTTAGATATAAAATAGAATAAAATAAAGTACTAAAATTAACAATACCC<br>TTAAGAAATTAAAAACTAAGGAAACATTTTTCTTGTTTCGAGTAGAT          |
|      | AATTTTGGAACTGTATGTGTGTGCATACATCTTCATAGTTACGAGTTTAAAGATGGATGGAAATATCG<br>ATCTAGGATAGGTATACATGTTGATGTGGGTTTTACTGATGCATATACATGA   |
|      | AATGCCAGCCTGTTAAACGCCGTCGACGAGTCTAACGGACACCAACCAGCGAACCAGCAGCGTCG<br>CGTCGGGCCAAGCGAAGCAGACGGCAGGCATCTCTGTCGCTGCCTCTGGACCCC    |
|      | CCATTTTATTAGTACATCCATTTAGGGTTTAGGGTTAATGGTTTTTATAGACTAATTTTTTTAGTACATC<br>TATTTTATTCTATTTTAGCCTCTAAATTAAGAAACTAAACTCTATTTT     |
|      | GATGATGGCATATGCAGCAGCTATATGTGGATTTTTTTAGCCCTGCCTTCATACGCTATTTATTTGCT<br>TGGTACTGTTTCTTTTGTGATGCTCACCTGTTGTTTGGTGTACTTCTGC      |
|      | CAGCTACGGGGGATTCCCTTTCCACCGCTCCTTCGCTTTCCCTTCCTCGCCCGCGTAATAAATAGA<br>CACCCCTCCACACCCTCTTTCCCAACCTCGTGTGTTTCGGAGCGCACACAC      |
|      | GGTGGATTTATTAATTTGGATCTGTATGTGTGTGCCATACATATTCATAGTTACGAATTGAAGATGAT<br>GGATGGAAATATCGATCTAGGATAGGTATACATGTTGATGCGGGTTTTACT    |
|      | ATTCACGTCCAAATGGGGGCCACTTGGCTGCAGTGCAGCGTGACCCGGTCGTGCCCTCTCTAGA<br>GATAATGAGCATTGCATGTCTAAGTTATAAAAAATTACCACATATTTTTTTTGTG    |
|      | GATGCATATACAGAGATGCTTTTTGTTGCTTGGTTGTGATGATGTGGTGTGGTTGGGCGGTCGTTT<br>ATTCGTTCTAGATCGGAGTAGAATACTGTTTCAAACCTACCTGGTGTATTTATT   |
|      | GTTAGACATGGTCTAAAGGACAATTGAGTATTTTGACAACAGGACTCTACAGTTTTATCTTTTTAGTG<br>TGCATGTGTTCTCCTTTTTTTTCAAATAGCTTACCTATATAATACTTCAT     |
|      | ACACTTGTGTTGAAGTGCAGTTTATCTATCTTTATACATATATTTAAACTTTACTCTACGAATAATATAAT<br>CTATAGTACTACAATAATATCAGTGTTTTAGAGAATCATATAAATGAACA  |
|      | ATATGCCGTGCACTTGTGTTGTCGGGTCATCTTTTCATGCTTTTTTTTGTCTTGGTTGTGATGATGTGG<br>TCTGGTTGGGCGGTCGTTCTAGATCGGAGTAGAATTCTGTTTCAAACCTACCT |
|      | TCTCGAGAGTTCGGCTCCACCGTTGGACTTGTCCGCTGTCCGCATCCAGAAATTGCGTGGCGGAG<br>CGGCAGACGTGAGCCGGCACGGCAGGCGGCCTCCTCCTCTCACGGCACCGG       |
|      | ACACAACCAGATCTCCCCAAATCCACCCGTCCGCACCTCCGCTTCAAGGTACGCCGCTCGTCCCTC<br>CCCCCCCCCTCTCTACCTTCTCTAGATCGGCGTTCGGTCCATGGTTAGG        |

|              |  |
|--------------|--|
| pNOS         | AGAATTAAGGGAGTCACGTTATGACCCCGCCGATGACGCGGGACAAGCCGTTTTACGTTTGGAAC<br>TGACAGAACCGCAACGATTGAAGGAGCCACTCAGCCGCGGGTTTCTGGAGTTA   |
|              | ATGAGCTAAGCACATACGTGAGAAACCATTATTGCGCGTTCAAAGTGCCTAAGGTCACTATCAGC<br>TAGCAAATATTTCTTGTCAAAAATGCTCCACTGACGTTCCATAAATCCCCTC    |
|              | GATCCCGTCAATTCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATAATGT<br>GTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGGATCATGAGCGG   |
|              | TTCAAAGTGCCTAAGGTCACTATCAGCTAGCAAATATTTCTTGTCAAAAATGCTCCACTGACGTTT<br>CATAAATCCCCTCGGTATCCAATTAGAGTCTCATATTTCACTCTCAATCAAA   |
| pmas         | AGTGTACCAAACAACGCTTTACAGCAAGAACGGAATGCGCGTGACGCTCGCGGTGACGCCATTTG<br>CCTTTTCAGAAATGGATAAATAGCCTTGCTTCTATTATATCTTCCCAAATTAC   |
|              | TTTTCTTGTGGCCGTCTTTGTTTATATTGAGATTTTTCAAATCAGTGCGCAAGACGTGACGTAAGTAT<br>CCGAGTCAGTTTTATTTTCTACTAATTTGGTCGTTATTTTCGGCGTGTAG   |
|              | GGCACCCGACAGGTCCGTCTTGACAAAAAGAACCGGGCGCCCTGCGCTGACAGCCGGAACAC<br>GGCGGCATCAGAGCAGCCGATTGTCTGTTGTGCCAGTCATAGCCGAATAGCCTCT    |
|              | ATTCGCCTTTTCAGAAATGGATAAATAGCCTTGCTTCTATTATATCTTCCCAAATTACCAATACATT<br>ACACTAGCATCTGAATTTATAACCAATCTCGATACACCAAATCGATAAGC    |
|              | CCACCCAAGCGGCCGGAGAACCTGCGTGCAATCCATCTTGTTCAATCATAGATCCCCGTGTGGAAG<br>ATATGAATTTTTTGTAGAACTAGATAAGATTAATGAATATCGGTGTTTTGGTT  |
|              | GACATGGCAACCGGGCCTGAATTTGCGGGTATTCTGTTTCTATTCCAATTTTTCTTGATCCGCAG<br>CCATTAACGACTTTTGAATAGATACGCTGACACGCCAAGCCTCGCTAGTCAA    |
|              |  |
| Ps7s7        | CTTTACAGCTATATCGATGTGACGTAAGATGCTTTACTTCGCTTCAAGTAAAGTAGGAAATTGCTCG<br>CTAAGTTATTTCTTTCTGAAAGAAATTAATTTAATTCTAATTAATTAATG    |
|              | ATTAGATGACACGTGGACGCTCAGGATCTGTGATGCTAGTGAAGCGCTTAAGCTGAACGAATCTGA<br>CGGAAGAGCGGACATACGCACATGGATTATGGCCACATGTCTAAAGTGTATCT  |
|              | AATAGCCCTTGGATTAGATGACACGTGGACGCTCAGGATCTGTGATGCTAGTGAAGCGCTTAAGCT<br>GAACGAATCTGACGGAAGAGCGGACATACGCACATGGATTATGGCCACATGTC  |
|              | GAGACATCACGTGACTCTCACGTGATGTCTCCGCGACAGGCTGGCACGGGGCTTAGTATTACCCC<br>GTGCCGGGATCAGAGACATTTGACTAAATATTGACTTGGAAATAATAGCCCTTGG |
|              | TAATTAATAGTAATTATGATTAATTATGAGATAAGAGTTGTTAATGCTTATGAGGAATAAAGAATG<br>ATTAATATTGTTAATTTTATTCCGCGAAGCGGTGTGTTATGTTTTGTTG      |
|              | TAAAGTGTATCTCTTTACAGCTATATCGATGTGACGTAAGATGCTTTACTTCGCTTCAAGTAAAGTA<br>GGAAATTGCTCGCTAAGTTATTTCTTTCTGAAAGAAATTAATTTAATTCTAA  |
|              | ATGTTTTGTTGGAGACATCACGTGACTCTCACGTGATGTCTCCGCGACAGGCTGGCACGGGGCTT<br>AGTATTACCCCGTGCCGGGATCAGAGACATTTGACTAAATATTGACTTGAAT    |
|              | TATTTCTTTCTGAAAGAAATTAATTTAATTCTAAATTAATTAATGAGTGGCTATAAATAGTGTGATG<br>CTACCTCACATCGTATTCTTCTTCGCATCGTCTGTTCTGGTTTTAAGCGA    |
|              | AGTGGCCTGCAGTAATTAATAGTAATTATGATTAATTATGAGATAAGAGTTGTTAATGCTTATGAG<br>GAATAAAGAATGATTAATATTGTTAATTTTATTCCGCGAAGCGGTGTGTT     |
|              |  |
| pRice_actin1 | ATTTTTGAGAAAAATATATATTCAGGCGAATTCACAATGAACAATAATAAGATTAATAATAGCTTGCCC<br>CCGTTGCAGCGATGGGTATTTTTCTAGTAAATAAAAGATAAACTTAGAC   |



|       |   |
|-------|---|
|       | CTTATGGGGAATTTATATTTTACTTACTATAGAGCTTTCATACCTTTTTTTTACCTTGGATTAGTTAAT<br>ATATAATGGTATGATTCATGAATAAAAATGGGAAATTTTTGAATTTGTAC   |
|       | TTATCTCAGATGTTAAGATTTTTCTTAAAGTAAATTCITTCAAATTTTAGCTAAAAGTCTTGAATAACTA<br>AAGAATAATACACAATCTCGACCACGGAAAAAACACATAATAAATTTG    |
|       | AAATAAGAACTTTCAAAAAACAGAACTATGTTTAAATGTGTAAAGATTAGTCGCACATCAAGTCATCTGT<br>TACAATATGTTACAACAAGTCATAAGCCCAACAAAGTTAGCACGTCTAAAT |
| pTA29 | GAAGCTGTGCTAGAGAAGATGTTTATTCTAGTCCAGCCACCCACCTTATGCAAGTCTGCTTTTAGCTT<br>GATTCAAAACTGATTTAATTTACATTGCTAAATGTGCATACTTCGAGCCTA   |
|       | TGCAAATTATCCAGTTATACTTAGCTAGATATCCAATTTTGAATAAAAATAGCTCTTGATTAGTAAACC<br>GGATAGTGACAAAGTCACATATCCATCAAACCTTCTGGTGCTCGTGGCTAAG |
|       | GATAGATTGAAAGAAGTGCAGGGAAGCGGTTAACTGGAACATAACACAATGTCTAAATTAATTGCAT<br>TCGCTAACCAAAAAGTGATTACTCTCTCCGGTCCACAATAAGTTATTTTTTG   |
|       | GCAAAGTATCACAATTTTCTATCGAGAGATAGATTGAAAGAAGTGCAGGGAAGCGGTTAACTGGAAC<br>ATAACACAATGTCTAAATTAATTGCATTCGCTAACCAAAAAGTGATTACTCT   |
|       | TGTCGCTTTAATTCGAGTAGGATGTATATATTAGTACATAAAAAATCATGTTTGAATCATCTTTCATAA<br>AGTGACAAGTCAATTGTCCCTTCTGTTTGGCACTATATTCAATCTGTAA    |
|       | TTCTGATCGACATGGGGTTAAATTTAAATTGGGACACATAAATAGCCTATTTGTGCAAATCTCCCA<br>TCGAAAATGACAGATTGTTACATGGAAAACAAAAAGTCCTCTGATAGAAGTC    |
| pMTL  | ACTAACGGCCGTCGTAAGAAGGTCCATCATGCGTGGAGCGAGCCATTTGGTTGGTTGTCAGGCCG<br>CAGTTAAGGCCTCCATATATGATTGTGCTCGGGCCATAACAGCATCTCCTCCA    |
|       | CCAGTTTATTGTAAGAATAAATTAAGTAGAGATATTTGTCGTCGGGCAGAAGAACTTGGACAAGAA<br>GAAGAAGCAAGCTAGGCCAATTTCTTGCCGGCAAGAGGAAGATAGTGGCCTCT   |
|       | TCCATCAATTTAGAAATGAGCAAGCAAGCACCCGATCGCTCACCCCATATGCACCAATCTGACTCCC<br>AAGCTCTGTTTCGCATTAGTACCGCCAGCACTCCACCTATAGCTACCAATTGA  |
|       | ATGCAATAATTCAGGGAATTCGTCAGCCAATTGAGGTAGCTAGTCAACTTGACATTGGTGCGAGCA<br>ATTTTCCGCACTCAGGAGGGCTAGTTTGGAGGTCCAAAACTATAGGAGATTA    |
|       | TGCTGGATTACTAAATCGAAATGGACGTAATATTCCAGGCAAGAATAATTGTTTCGATCAGGAGAC<br>AAGTGGGGCATTGGACCGGTTCTTGCAAGCAAGAGCCTATGGCGTGGTGACA    |
|       | GGAGAGCGAGAGGCCAGAAATTTATAGCGACTGACGCTGTGGCAGGCACGCTATCGGAGGTTACG<br>ACGTGGCGGGTCACTCGACGCGGAGTTCACAGGTCTATCCTTGCATCGCTCGGC   |
|       | TCCATCCTCACTTGCTATAAAAAGAGGTGTCCATGGTGCTCAAGCTCAGCCAAGCAAATAAGACGAC<br>TTGTTTCATTGATTCTTCAAGAGATCGAGCTTCTTTTGCACCACAAGGTCGAG  |
|       | AACAACACCCTCGCCATGCCGCTCGACTACAGCCTCCTCCTCGTCGGCGCCACAACGAGGGAGCC<br>CGTGGTCGCAGCCACCGACCAGCATGTCTCTGTGCTCCTCGTCCGACCTCGACATG |
|       | CAAGCACGGCATGCATAGTATATGTATAGCTTGTTTTAGTGTGGCTTTGCTGAGACGAATGAAAGCA<br>ACGGCGGGCATATTTTTCAGTGGCTGTAGCTTTCAGGCTGAAAGAGACGTGGC  |
|       | GACCTTTCAGCCTAAGCAGATCGATTGATCGTTAGAGTCAAAGAGTTGGTGGTACGGGTACTTTAA<br>CTACCATGGAATGATGGGGCGTGATGTAGAGCGGAAAGCGCCTCCCTACGCGG   |
|       | AAGAGGCTAAAATCCTCTCCTTATTTAATTTTAAATAAGTAGTGTATTTGTATTTAACTCCTCCAACC<br>CTTCCGATTTTATGGCTCTCAAACCTAGCATTAGTCTAATGCATGCATGCT   |

|      |   |
|------|---|
|      | TAAAGAAAATCGACGCAGGACAACCGTCAAACGGGTGCTTTAATATAGTAGATATATATATATAGAG<br>AGAGAGAGAAAGTACAAAGGATGCATTTGTGTCTGCATATGATCGGAGTATT   |
|      | TGGCTAGAGGTCGTATGGGGTTGTTAATAGCATAGCTAGCTACAAGTTAACCGGGTCTTTTATATTTA<br>ATAAGGACAGGCAAAGTATTACTTACAAATAAAGAATAAAGCTAGGACGAAC  |
|      | CGCGGGAAAGCACGACGACTTACGAAGGTTAGTGGAGGAGGAGGACACACTAAAATCAGGACGCA<br>AGAAACTCTTCTATTATAGTAGTAGAGAAGAGATTATAGGAGTGTGGGTTGATTC  |
|      | GATGTTATTTAAGAAAATATGTTGTTAACAGAATAAGATTAATATCGAAATGGAAACATCTGTAAATTA<br>GAATCATCTTACAAGCTAAGAGATGTTACGCTTTGAGAACTTCTTCAGA    |
|      | TCATGACCGTAGAAGTAGCTCTCAAGACTCAACGAAGGCTGCTGCAATTCCACAAATGCATGACAT<br>GCATCCTTGTAACCGTCGTCGCCGCTATAAACACGGATAACTCAATTCCCTGC   |
|      | CGGCGCGTTGCCATACATCATGCCTCCATCGATGATCCATCCTCACTTGCTATAAAAAGAGGTGTC<br>CATGGTGCTCAAGCTCAGCCAAGCAAATAAGACGACTTGTTTCATTGATTCTT   |
|      | TCATGGCAAACAGTCGGACGCCAGCACCAGACTGACGACATGAGTCTCTGAAGAGCCCGCCACCTA<br>GAAAGATCCGAGCCCTGCTGCTGGTAGTGGTAACCATTTTCGTCGCGCTGACGC  |
|      | AGTTTATATATCGGCGTGATGATGATGCTCCTAGCTAGAAATGAGAGAAGAAAAACGGACGCGTGT<br>TGGTGTGTGCAATGGCGTCCATCCTTCCATCAGATCAGAACGATGAAAAAGT    |
|      | TATTCTATTTTTCATGATGTATCTTCTCTAACATGCCTTAATTTGCAAATTTGGTATAACTACTGATTG<br>AAAATATATGTATGTAAAAAATACTAAGCATATTTTTGAAGCTAAACAT    |
|      | GCGGAGTTTACGGGGACTTATCCTTACGACGTGCTCTAAGGTTGCGATAACGGGCGGAGGAAGGC<br>GTGTGGCGTGCGGAGACGTTTATACACGTAGTGTGCGGGAGTGTGTTTCGTAGA   |
|      | TGCACATGACAACAATTGTAAGAGGATGGAGACCACAACGATCCAACAATACTTCTGCGACGGGCT<br>GTGAAGTATAGAGAAGTTAAACGCCAAAAGCCATTGTGTTTGAATTTTTAGT    |
| tOCS | GGATGATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCACCCCCTGCTTTAATGAGATA<br>TGCGAGACGCCTATGATCGCATGATATTTGCTTTCAATTCTGTTGTGCACGTTG   |
|      | ATCCAATTTTAAAAAAGCGGCAGAACCGGTCAAACCTAAAAGACTGATTACATAAATCTTATTCAA<br>TTTCAAAGGCCCCAGGGGCTAGTATCTACGACACACCGAGCGGCGAACTAA     |
|      | GTACCCTACTACTTATATGTACAATATTAATAATGAAAACAATATATTGTGCTGAATAGGTTTATAGCGA<br>CATCTATGATAGAGCGCCACAATAACAAACAATTGCGTTTTATTATTACAA |
|      | TAACGTTCACTGAAGGGAACCTCCGGTCCCCGCCGGCGCATGGGTGAGATTCCTTGAAGTTGAG<br>TATTGGCCGTCCGCTCTACCGAAAGTTACGGGCACCATTCAACCCGGTCCAGCA    |
|      | TTCAACCCGGTCCAGCACGGCGCGGGTAACCGACTTGCTGCCCGAGAATTATGCAGCATTTTT<br>TTGGTGTATGTGGGCCCTGTACAGCGGCCGCGTTAACGCGTATACTCTAGAGCG     |
|      | TAAAAACCTGAGCATGTGTAGCTCAGATCCTTACCGCCGGTTTTCGGTTCAATTCTAATGAATATATCA<br>CCCGTTACTATCGTATTTTTATGAATAATATTCTCCGTTCAATTTACTGATT |
|      | GGACTCTGGGGTTCGAAATGACCGACCAAGCGACGCCAACCTGCCATCACGAGATTTGATTCCA<br>CCGCCGCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCT     |
| tE9  | TTGTACCATTTGTTGTGCTTGTAAATTTACTGTGTTTTTTATTCGGTTTTCGCTATCGAACTGTGAAATG<br>GAAATGGATGGAGAAGAGTTAATGAATGATATGGTCCTTTTGTTCATTCTC |
|      | AAATTAATATTATTTGTTTTTCTTATTTGTTGTGTGTTGAATTTGAAATTATAAGAGATATGCAAAC<br>ATTTTGTGTTGAGTAAAAATGTGTCAAATCGTGGCCTCTAATGACCGAAG     |

|        |   |
|--------|---|
|        | TTTAGACATTTATGGACTTTTCCTTTATGTAATTTTCCAGAATCCTTGTGTCAGATTCTAATCATTGCTTTAT<br>AATTATAGTTATACTCATGGATTTGTAGTTGAGTATGAAAATATTTTTTA   |
|        | TTAATATGAGGAGTAAACACTTGTAGTTGTACCATTATGCTTATTCACTAGGCAACAAATATATTTTC<br>AGACCTAGAAAAGCTGCAAATGTTACTGAATACAAGTATGTCCTCTTGTGT       |
|        | CAGATTCTAATCATTGCTTTATAATTATAGTTATACTCATGGATTTGTAGTTGAGTATGAAAATATTTTT<br>TAATGCATTTTATGACTTGCCAATTGATTGACAACATGCATCAATCGACC      |
|        | AGCTTTTCGTTTCGTATCATCGGTTTCGACAACGTTTCGTCAAGTTCAATGCATCAGTTTCATTGCGCACA<br>CACCAGAATCCTACTGAGTTTGTAGTATTATGGCATTGGGAAAACCTGTTTTTC |
| tNOS   | GTATTAATGTATAATTGCGGGACTCTAATCATAAAAACCCATCTCATAAATAACGTCATGCATTACAT<br>GTTAATTATTACATGCTTAACGTAATTCAACAGAAATTATATGATAATCAT       |
|        | TCGACCGCGGTACCCGGAATTCGAGCTCGGTACCCGGGGATCTCCCGATCTAGTAACATAGATGA<br>CACCGCGCGGATAATTTATCCTAGTTTGC GCGCTATATTTTGTCTATCGC          |
|        | GCATTACATGTTAATTATTACATGCTTAACGTAATTCAACAGAAATTATATGATAATCATCGCAAGACC<br>GGCAACAGGATTCAATTTAAGAACTTTATTGCCAATGTTTGAACGATC         |
| tg7    | TCTTGAAAGAAATATAGTTTAAATATTTATTGATAAAATAACAAGTCAGGTATTATAGTCCAAGCAAAA<br>ACATAAATTTATTGATGCAAGTTTAAATTCAGAAATATTTCAATAACTGAT      |
|        | ATGCAAGTTTAAATTCAGAAATATTTCAATAACTGATTATATCAGCTGGTACATTGCCGTAGATGAAA<br>GACTGAGTGCATATTATGTGTAATACATAAATTGATGATATAGCTAGCTTA       |
| tpinII | AGACTTGTCCATCTTCTGGATTGGCCAACCTTAATTAATGTATGAAATAAAAGGATGCACACATAGTGA<br>CATGCTAATCACTATAATGTGGGCATCAAAGTTGTGTGTTATGTGTAATTAC     |
|        | GATGAACCAGATGCATTTTCATTAACCAATCCATATACATATAAATTAATCATATATAATTAATATCA<br>ATTGGGTTAGCAAAACAAATCTAGTCTAGGTGTGTTTTGCGAATTGCGGC        |
|        | TAGTTATCTGAATAAAAGAGAAAGAGATCATCCATATTTCTTATCCTAAATGAATGTCACGTGTCTTTA<br>TAATTTCTTATGATGAACCAGATGCATTTTCATTAACCAATCCATATACATAT    |
| t35S   | TCACTCTCGGCATGGACGAGCTGTACAAGTGAAGATCCACCTAGTCTAGAGTCCGCAAAAATCACCA<br>GTCTCTCTCTACAAATCTATCTCTCTATTTTTCTCCAGAATAATGTGTGAG        |
|        | CTCTATTTTTCTCCAGAATAATGTGTGAGTAGTTCCAGATAAGGGAATTAGGGTTCTTATAGGGTTT<br>CGCTCATGTGTTGAGCATATAAGAAACCCCTAGTATGATTTTGTATTTGTAA       |
| gus    | AGCCAGACAGAGTGTGATATCTACCCGCTTCGCGTCCGCATCCGGTCAGTGCCAGTGAAGGGCGA<br>ACAGTTCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCATGAA       |
|        | GGAATGGTGATTACCGACGAAAACGGCAAGAAAAAGCAGTCTTACTTCCATGATTTCTTTAACTATG<br>CCGGGATCCATCGCAGCGTAATGCTCTACACCACGCCGAACACCTGGGTGGAC      |
|        | ACTGGACAAGGCACTAGCGGGACTTTGCAAGTGGTGAATCCGCACCTCTGGCAACCGGGTGAAGG<br>TTATCTCTATGAACTGTGCGTCACAGCCAAAAGCCAGACAGAGTGTGATATCTAC      |
|        | GAGCATCAGGGCGGCTATACGCCATTTGAAGCCGATGTCACGCCGATGTTATTGCCGGGAAAAGT<br>GTACGTATCACCGTTTGTGTGAACAACGAACTGAACTGGCAGACTATCCCGCCG       |
|        | GATATCACCGTGGTGACGCATGTCGCGCAAGACTGTAACCACGCGTCTGTTGACTGGCAGGTGGTG<br>GCCAATGGTGTGTCAGCGTTGAACTGCGTGATGCGGATCAACAGGTGGTTGCA       |
| gox    | ATGGCTGAGAACCACAAGAAGGTTGGTATCGCTGGAGCTGGAATCGTTGGTGTGTTGCACTGCTTTG<br>ATGCTTCAACGTCGTGGATTCAAGGTTACCTTGATTGATCCAAACCCACCAGGT     |

|          |  |
|----------|--|
|          | ATGACCGCAACCCTCGTTTCTGAGCTCCTCGCAGGTGAGAAGACCTCTATCGACATCTCTCCATTCCG<br>CACCAAACCGTTTCGGTATTGGTAAGTCCAAGCAAACCTGGTCCTGCATCCTAA |
|          | ATCCCAGATTCCCTTCCAGTGATTGGTCGTGCTACCCGTA CTCCAGACGTTATCTACGCTTTCGGTC<br>ACGGTCACCTCGGTATGACTGGTGCTCCAATGACCGCAACCCTCGTTTCTGAG  |
|          | ATCGCCAACCCAGAAGCTGCTCCACGTATTCCAAC TACCGATGCTTCTGGAAAGTTCATCGCTACTC<br>CTATGGAGATGGGTCTTCGTGTTGCTGGAACCGTTGAGTTCGCTGGTCTCACT  |
|          | CCTTTGATCAAGTCCTTGGCTGAGGAGGCTGATGCTAGCCACCTTATCCGTCACGAAGGTCACCTTA<br>CCGTGTACCGTGGAGAAGCAGACTTCGCCAAGGACCGTGGAGGTTGGGAACCTT  |
|          | ACCAACGGTGTTCTTGCTGTTGATGCAGCTGTTGTTGCAGCTGGTGCACTCCAAGTCTCTTGCTA<br>ACTCCCTTGGTGATGACATCCATTGGATAACCGAACGTGGATAACCATCGTG      |
|          | CGTCGTCTCAACGGTGTTTCGTACTCAAATCCTCAGCGCTGATGCATTGCGTGATTTTCGATCCTAACT<br>TGTCTACGCCTTACCAAGGGAATCCTTATCGAAGAGAACGGTCACACCATC   |
|          | ATCCGTTTCAGCTACTTTCCAACCATCATGCCTTGTTGATTTCGTTTCTTGCTTGCTGGAAGACCAA<br>CAAGGTGAAGGAGCAAGCTAAGGCACTCCGTAACCTCATCAAGTCCACTGTG    |
|          | GAAGGTGCTTCTTTCGGTAACGCTGGTTGCTTCAACGGTTCCTCCGTTGTTCCAATGTCCATGCCAG<br>GAAACTTGACTAGCGTTCCAAAGTGGCTTCTTGACCCAATGGGTCCATTGTCC   |
|          | AACCCACAAGGTCTCGTACTCTCTTGTTTCGTGCTTTCATCGCTAACGGTGGAGAGTTCGTGTCTG<br>CTCGTGTATCGGATTCGAGACTGAAGTCTGCTCTCAAGGTATCACCACC        |
|          | GCTGCTCCTAACTGGAAGCGTGCTCACGTTCTCTACACTCACGCTCGTAAGTTGCTTCCAGCTCTCG<br>CTCCTGCCAGTTCTGAAGAAGTTACTCCAAGTGGATGGGTTTCCGTCCAAGC    |
| cryIAb_1 | GACCCACCAACCCAGCCCTGCGCGAGGAGATGCGCATCCAGTTCAACGACATGAACTCTGCCCT<br>GACCACCGCCATCCCCTCTTCGCTGTCCAGA ACTACCAGGTCCCTCTCCTGTCT    |
|          | ATCGAACAGTTAATTAACCAAAGAATAGAAGAATTCGCTAGGAACCAAGCCATCTCTAGACTGGAGG<br>GCCTGAGCAACCTGTACCAGATCTACGCCGAGAGCTTCCGCGAGTGGGAGGCT   |
|          | AGCCAACGCTACAGGGTCAGGATCAGGTACGCTTCCACCACCAACCTGCAGTTCCACACCAGCATC<br>GACGGCAGGCCCATCAACCAGGGCAACTTCAGCGCCACCATGAGCAGCGGCAGC   |
|          | TTCAGCAACAGCAGCGTGAGCATCATCAGGGCTCCCATGTTTCAGCTGGATCCACCGCAGCGCTGAG<br>TTCAACAACATCATTCCAAGTAGCCAGATCACTCAGATCCCCTCACCAGAGC    |
|          | ATGGCTGACAACAACCCCAACATCAACGAGTGCATCCCCTACA ACTGCCTGAGCAACCCAGAGGTG<br>GAGGTGCTGGGTGGTGAGAGGATCGAGACCGTTACACCCCATCGACATCAGC    |
|          | GTATACCGCAAGAGCGGCACTGTGGACAGCCTGGACGAGATCCCACCCAGAACAACAACGTGCC<br>CCCTAGGCAGGGGTTCTCTCATCGCCTCTCACACGTGAGCATGTTCCGCGAGCGGC   |
|          | TCCAACGGCAGCAGCGTGTTCCACCTGTCTGCCACGTGTTCAACAGCGGCAACGAGGTGTACATC<br>GACAGGATCGAGTTTGTCCAGCTGAGGTGACCTTCGAAGCTGAGTACGACTGA     |
|          | CTGAGCCTGACCCAGTTCCTGCTGAGCGAGTTCGTGCCTGGTGCTGGCTTCGTGCTGGGACTAGTG<br>GACATCATCTGGGGCATCTTCGGTCCCAGCCAGTGGGATGCCTTCTGGTGCG     |
|          | TTCAGGGGCTCTGCCAGGGCATCGAGGGCAGCATCCGAGCCCCACCTGATGGACATCCTGAA<br>CAGCATCACCATCTACACTGACGCCACAGGGGTGAGTACTACTGGTCTGGCCAC       |
|          | CGCACCTGAGCAGCACCTGTACAGGAGGCCCTTCAACATCGGCATCAACAACCGAGCAGCTGAG<br>CGTGCTGGATGGCACCGAGTTCGCCTACGGCACCAGCAGCAACCTGCCAGCGCC     |



|          |   |
|----------|---|
|          | CAGATCATGGCTTCTCCCGTGGGCTTCAGCGGTCCCGAGTTCACCTTCCCCCTGTACGGCACAATG<br>GGCAACGCTGCCCCACAGCAGAGGATCGTGGCCAGCTGGGCCAGGGCGTGTAC   |
|          | GTCTTGGATATCGTCAGTCTCTTCCCAACTACGACAGCAGGACCTACCCTATCCGGACTGTGAGC<br>CAGCTGACCCGCGAGATCTACACCAACCCCGTGTGGAGAACTTCGACGGCAGC    |
|          | AACCTGCAGAGCGGAAGCTTCCGCACTGTGGGCTTCACTACCCATTCAACTTCTCCAACGGCAGC<br>AGCGTGTTACCCTGTCTGCCACGTGTTCAACAGCGGCAACGAGGTGTACATC     |
|          | ACCAACCTGGGCTCCGGGACTAGCGTTGTCAAGGGACCAGGGTTCCTGGAGGGCAGATCCTGAG<br>GAGGACCAGCCAGGCCAGATCAGCACCTTAAGGGTGAACATCACCGCTCCCCTC    |
|          | ATCGGCAACTACACCGACCACGCAGTGAGATGGTACAACACTGGGCTTGAGAGGGTCTGGGGTCC<br>CGACAGCCGCGACTGGATCAGGTACAACCAGTTCAGGCGTGAACACTACTCTCACC |
|          | GTCTATGTGCAAGCTGCCAACCTCCATCTCAGCGTCCTTCGCGACGTGAGCGTCTTTGGGCAGAGG<br>TGGGGGTTGACGCTGCCACCATCAACAGCCGCTACAACGACCTGACGCGTCTG   |
| crylab_2 | TGGATTCATCGCAGCGCGGAGTTCAACAATATCATTCCGTCTCCCAAATCACCAAATCCCCCTCA<br>CCAAGTCCACCAACCTGGGCAGCGCACCTCCGTGGTGAAGGGCCAGGCTTC      |
|          | CACCTGATGGACATCCTCAACAGCATCACTATCTACACCGATGCCACCGCGGCGAGTACTACTGG<br>TCCGGCCACCAGATCATGGCCTCCCCGGTCCGCTTCAGCGGCCCCGAGTTTACC   |
|          | CGGTACGCGCTCACTCCGCCCTCTGCCTTTGTTACTGCCACGTTTCTCTGAATGCTCTCTTGTGTGG<br>TGATTGCTGAGAGTGGTTTAGCTGGATCTAGAATTACACTCTGAAATCGTGTT  |
|          | TCAACTTCTCCAACGGCTCCAGCGTTTTACCCTGAGCGCCACGTGTTCAATTCCGGCAATGAGG<br>TGACATTGACCGCATTGAGTTCGTGCCAGCCGAGGTCACCTTCGAAGCCGAG      |
|          | CAGGCCATCAGCCGCCTGGAGGGCCTCAGCAACCTCTACCAAATCTACGCTGAGAGCTTCCGCGA<br>GTGGGAGGCCGACCCACTAACCAGCTCTCCGCGAGGAGATGCGCATCCAGTTC    |
|          | AACGACATGAACAGCGCCCTGACCACCGCCATCCACTCTTCGCCGTCCAGAACTACCAAGTCCCG<br>CTCCTGTCCGTGTACGTCCAGGCCGCAACCTGCACCTCAGCGTGCTGAGGGAC    |
|          | TCATCATGAGAATCCTTCTTAGATGTAAGACCACCTGCAGCAGATGACTTTGATCTTGTTGTTGGGC<br>GCCGACCAGATTGAGCCATTGCAGCTGTTAATGATGCACCAGCCGTGGTGCCA  |
|          | TTGATGTTTATCTCTGCTCCTTTATTGTGACCATAAGTCAAGATCAGATGCACTTGTTTAAATATTGT<br>TGTCTGAAGAAATAAGTACTGACAGTATTTTATGATGCATTGATCTGCTTGTT |
|          | ATCCTACTAGACTGGCTGAGCACAACAGGGAAACGTTCCAGCGAAGACAATGCACTTTGCAGTTTCT<br>GAACCAACAGTGCCATGGGAGTTTCATCTCTCTCGACCAATCGAATTTCCGCG  |
|          | ATAATAATTTGTTTCGTCTGCAGAGCTTATTATTTGCCAAAATTAGATATTCCTATTCTGTTTTTGT<br>GTGCTGTAAATTGTTAACGCCTGAAGGAATAAATAAATGACGAAAT         |
|          | GAGGACTTTCGGTAGCCTTCTTTCAATTCGAATTTGCTTGCAGCAGTCAGGTCTTTTGATTCATC<br>TGAGTTTGGCTTTACAATAGCTTTTCTTTTCTTTGGCAGTACTAGTGCTT       |
|          | ACGGGCGGCGACATCCTGCGCAGGACCTCCCCGGGCCAGATCAGCACCTCCGCGTCAACATCAC<br>CGCTCCCCTGTCCAGAGGTACCGCGTCAGGATTCGCTACGCTAGCACCACCAAC    |
|          | TACGACCTGGAGAGAGCCCAGAAGGCTGTCAATGAGCTTTCACGTCCAGCAATCAGATCGGCCTG<br>AAGACCGACGTCACTGACTACCACATCGACCAAGTCTCCAACCTCGTGGAGTGC   |
|          | TTTCCTCTACGGCACGATGGGCAACGCCGCTCCACAACAACGCATCGTCGCTCAGCTGGGCCAG<br>GGCGTCTACCGCACCTGAGCTCCACCCTGTACCGCAGGCCCTTCAACATCGGT     |

|          |   |
|----------|---|
|          | CTGGAGCGCGTCTGGGGCCCTGATTCTAGAGACTGGATTCGCTACAACCAGTTCAGGCGCGAGCT<br>GACCCTCACCGTCTGGACATTGTGCCCTCTCCCGAACTACGACTCCCGCACC     |
|          | TGCCTCAGCAACCCTGAGGTGAGGTGCTCGGCGGTGAGCGCATCGAGACCGGTTACACCCCAT<br>CGACATCTCCCTCTCCCTCACGCAGTTCCTGCTCAGCGAGTTCGTGCCAGGCGCT    |
|          | CAGCTCGACCGCGGCTGGCGCGGCAGCACCGACATCACGATCCAGGGCGGCGACGATGTGTTCAA<br>GGAGAACTACGTGACTCTCCTGGGCACTTTCGACGAGTGCTACCCTACCTACTTG  |
|          | TACCAGAAGATCGATGAGTCCAAGCTCAAGGCTTACACTCGCTACCAGCTCCGCGGCTACATCGAA<br>GACAGCCAAGACCTCGAGATTTACCTGATCCGCTACAACGCCAAGCACGAGACC  |
|          | GTCAGCGTGTGGCCAGAGGTGGGGCTTCGACGCCGCCACCATCAACAGCCGCTACAACGACCT<br>CACCAGGCTGATCGGCAACTACCCGACCACGCTGTCCGCTGGTACAACACTGGC     |
|          | TGTTGTAACAAAATTTAAAAATAAAGAGTTTCCTTTTTGTTGCTCTCCTTACCTCCTGATGGTATCTAG<br>TATCTACCAACTGACACTATATTGCTTCTTTACATACGTATCTTGCTCG    |
|          | ATCAACAACCAGCAGCTGTCCGTCCTGGATGGCACTGAGTTCGCCTACGGCACCTCCTCCAACCTG<br>CCCTCCGCTGTCTACCGCAAGAGCGGCACGGTGGATTCCCTGGACGAGATCCCA  |
|          | GTCAACGTGCCCGGTAAGTTCCTCTGGCCGCTGAGCGCCCCAGCCCGATCGGCAAGTGTGC<br>CCACCACAGCCACCACTTCTCCTTGGACATCGATGTGGGCTGCACCGACCTGAAC      |
|          | TACCCGATCCGCACCGTGTCCCAACTGACCCGCGAAATCTACACCAACCCCGTCTGGAGAACTTC<br>GACGGTAGCTTCAGGGGCAGCGCCAGGGCATCGAGGGCTCCATCAGGAGCCCA    |
|          | CTGCAATTCCACACCTCCATCGACGGCAGGCCGATCAATCAGGGTAACCTTCTCCGCCACCATGTCC<br>AGCGGCAGCAACCTCCAATCCGGCAGCTTCCGCACCGTGGGTTTCACCACCCCC |
|          | CTGCCTGTGCTGATTACTTGCCGTCCTTTGTAGCAGCAAATATAGGGACATGGTAGTACGAAACGA<br>AGATAGAACCTACACAGCAATACGAGAAATGTGTAATTTGGTGTAGCGGTA     |
|          | TACTAGACTGGCTGAGCACAACAGGGAAACGTTCCAGCGAAGACAATGCACTTTGCAGTTTCTGAAC<br>CAACAGTGCCATGGGAGTTTCATCTCTCTCGACCAATCGAATTTCCGCGGGCC  |
|          | TTTATTTAAGCACATGTTGGTGTATAGGGCACTTGGATTGAGAAGTTTGCTGTTAATTTAGGCACAG<br>GCTTCATACTACATGGGTCAATAGTATAGGGATTTCATATTATAGGCGATACT  |
|          | ATGCCTTCTCCCTAGTGTGACCAGTGTTACTCACATAGTCTTTGCTCATTTCATTGTAATGCAGATA<br>CCAAGCGGCCATGGACAACAACCCAAACATCAACGAGTGCATCCCGTACAAC   |
|          | CTCTCCGATGAGTTCTGCCTCGACGAGAAGAAGGAGCTGTCCGAGAAGGTGAAGCATGCCAAGCGT<br>CTCAGCGACGAGAGGAATCTCCTCCAGGACCCCAATTTCCGCGGCATCAACAGG  |
|          | CCACAGAACAACAATGTGCCCCCAGGCAGGGTTTTTCCACAGGCTCAGCCACGTGTCCATGTTT<br>CGCTCCGGCTTCAGCAACTCGTCCGTGAGCATCATCAGAGCTCCTATGTTCTCC    |
|          | GGCTTCGTCTGGGCTCGTGGACATCATCTGGGGCATCTTTGGCCCTCCAGTGGGACGCCTTC<br>CTGGTGCAAATCGAGCAGCTCATCAACCAGAGGATCGAGGAGTTCGCCAGGAAC      |
|          | AGCACAATATTTGATGAATAATCCCGAAGTATTTTTCGCCATGAGCTCGGGAAAGACGAAGCTTGA<br>AGGGTTGAGCCAGAGCACTAAGACCTGAAGTCAGACGAGACCCTCCAATACCA   |
|          | GGAACCCAGATTGAGAATTATTACCAGATGGAATTATAGGCTTCGATGCAACCTCACTGCGTTGAA<br>CTCTAGGCCAAAGGAATTCTTCAACAGATGCAAGACTAGCAAATGGGTGATA    |
| cryIAb_3 | GGCAACTACCCGACCACGCCGTGCGCTGGTACAACACCGGCCTGGAGCGCGTGTGGGGTCCCG<br>ACAGCCGCGACTGGATCAGGTACAACCAGTTCGCCGCGAGCTGACCCTGACCGTG    |

|       |  |
|-------|--|
|       | CGCGGCAGCGCCCAGGGCATCGAGGGCAGCATCCGCAGCCCCACCTGATGGACATCCTGAACA<br>GCATCACCATCTACACCGACGCCACCGCGGGAGTACTACTGGAGCGGCCACCAG    |
|       | CTGGACATCGTGAGCCTGTTCCCAACTACGACAGCCGCACCTACCCCATCCGCACCGTGAGCCAG<br>CTGACCCGCGAGATTTACACCAACCCCGTGCTGGAGAACTTCGACGGCAGCTT   |
|       | AACCTGGGCAGCGGCACCAGCGTGGTGAAGGGCCCCGGCTTACCGGCGGGCAGATCCTGCGCC<br>GCACCAGCCCCGGCCAGATCAGCACCCCTGCGCGTGAACATCACCGCCCCCTGAGC  |
|       | AGCAACAGCAGCGTGAGCATCATCCGTGCACCTATGTTACAGCTGGATTACCGCAGTGCCGAGTTC<br>AACAACATCATCCCCAGCAGCCAAATCACCCAGATCCCCCTGACCAAGAGCACC |
|       | GAGCAGCTGATCAACCAGCGCATCGAGGAGTTCGCCCCAACCAGGCCATCAGCCGCCTGGAGGG<br>CCTGAGCAACCTGTACCAAATCTACGCCGAGAGCTTCCGCGAGTGGGAGGCCGAC  |
|       | CTGCAGAGCGGCAGCTTCCGCACCGTGGGCTTACCACCCCCTTCAACTTCAGCAACGGCAGCAG<br>CGTGTTACCCTGAGCGCCACGTGTTAACAGCGGCAACGAGGTGTACATCGAC     |
|       | TACCGCAAGAGCGGCACCCTGGACAGCCTGGACGAGATCCCCCTCAGAACAACAACGTGCCACC<br>TCGACAGGGCTTACGCCACCGTCTGAGCCACGTGAGCATGTTCCGCAGTGGCTTC  |
|       | TACGTGCAGGCCGCCAACCTGCACCTGAGCGTGCTGCGCGACGTACGCGTGTTCGGCCAGCGCTG<br>GGGCTTCGACGCCGCCACCATCAACAGCCGCTACAACGACCTGACCCGCCTGATC |
|       | CGCATCGAGTTCGTGCCGCGGAGGTGACCTTCGAGGGCCGAGTACGACCTGGAGAGGGCTCAGAA<br>GGCCGTGAACGAGCTGTTACCAGCAGCAACCAGATCGGCCTGAAGACCGACGTG  |
|       | CAGCGTACCGCGTCCGCATCCGCTACGCCAGCACCACCAACCTGCAGTTCACACCAGCATCGA<br>CGGCCGCCCATCAACCAGGGCAACTTCAGCGCCACCATGAGCAGCGGCAGCAAC    |
|       | ACCCTGAGCAGCACCCCTGTACCGTCGACCTTTCAACATCGGCATCAACAACCAGCAGCTGAGCGTG<br>CTGGACGGCACCGAGTTCGCCTACGGCACCAGCAGCAACCTGCCAGCGCCGTG |
|       | ACCTTCGAGGCCGAGTACGACCTGGAGAGGGCTCAGAAGGCCGTGAACGAGCTGTTACCAGCAG<br>CAACCAGATCGGCCTGAAGACCGACGTGACCGACTACCACATCGATCAGGTGTAG  |
|       | ATGGACAACAACCCCAACATCAACGAGTGCATCCCCTACAACCTGCCTGAGCAACCCCGAGGTGGAG<br>GTGCTGGGCGGGCAGCGCATCGAGACCGGCTACACCCCATCGACATCAGCCTG |
|       | CCCACCAACCCCGCCCTGCGCGAGGAGATGCGCATCCAGTTC AACGACATGAACAGCGCCCTGAC<br>CACCGCCATCCCCCTGTTCCGGTGCAGAACTACCAGGTGCCCTGCTGAGCGTG  |
|       | AGCCTGACCCAGTTCCTGCTGAGCGAGTTCGTGCCGCGCGCCGGCTTCGTGCTGGGCCTGGTGGA<br>CATCATCTGGGGCATCTTCGGCCCCAGCCAGTGGGACGCCTTCTGGTGCAGATC  |
|       | ATCATGGCCAGCCCCGTGCGCTTACGCGGCCCGAGTTCACCTTCCCCCTGTACGGCACCATGGG<br>CAACGCTGCACCTCAGCAGCGCATCGTGGCACAGCTGGGCCAGGGAGTGTACCGC  |
| cry1B | GTGTA CTCTTGACTCACAGGTCTGCTGACCGCACTAACACCATCGGCCAAACAGGATCACCCAA<br>ATCCCAATGGTGAAGGCCTCCGAGCTGCCACAAGGCACCACCGTGGTGAGGGGA  |
|       | GCCCAGTTGACCAGAGAGGTTTACACCGATGCTATCGGTGCTACTGGCGTGAACATGGCCAGCATG<br>AACTGGTACAACAACAACGCCCATCCTTCTCCGCCATCGAAGCTGCCGCTATC  |
|       | ATGGTTACCTCCAACCGTAAGAACGAGAACGAGATCATCAACGCTGTGTCCAACCACTCCGCTCAG<br>ATGGACCTCCTCCAGACGCTCGCATCGAGGACTCCTTGTGCATCGCTGAGGGC  |
|       | TTCCTCAACGCTATGCCACTCTTCGCCATCAGGAACCAAGAGGTCCCATTGTTGATGGTGTACGCTC<br>AGGCTGCCAACTTGACCTCTTGTTGTTGAGGGACGCCTCCCTCTTCGGCAGC  |

|        |  |
|--------|--|
|        | AGGTCCCCACACCTCCTGGACTTCCTGGAGCAACTCACCATCTTCAGCGTTTCTCCCGCTGGAGC<br>AACACCAGGCACATGACCTACTGGCGTGGTCACACCATCCAGTCCCGTCCCATC  |
|        | CCAGGCTTCACTGGCGGCGACATCCTCAGGAGGACCAACACTGGCGGCTTCGGTCCCATCAGGGT<br>CACCGTCAACGGCCCATTGACCCAAAGATACCGCATCGGCTTCGGCTACGCTTCC |
|        | GGAGGTGGCTTGAACACCTCCACCCACGGAGCTACTAACACCTCCATCAACCCAGTGACCTTGCGT<br>TTCGCCTCTCGTGACGTGTACAGGACCGAGAGCTACGCCGGTGTCTCCTCTGG  |
|        | AACAACATCGACCCATTTCGTGACGCGCTCCACCGTCCAAACCGGCATCAACATCGCTGGCCGCATC<br>CTTGGCGTGTGGGGCTCCATTTCGCAGGACAGCTTGCTAGTTTCTACTCCTTC |
|        | GGCATCTACCTCGAGCCCATTACGGCGTGCCAACCGTGAGGTTCAACTTCACCAACCCACAGAAC<br>ATCTCCGATCGCGGCACTGCGAACTACTCCAACCATAACGAGTCCCCAGGCCTC  |
|        | ACCAACGCTGCTTCTGGGTGAGGTACAACCAGTTCGCGAGGGACCTCACCTTGGGTGTCTTGGAC<br>CTCGTGGCCCTCTTCCCAAGCTACGACACCCGCACCTACCCCATCAACACCTCT  |
|        | ACCGTGGACTTCGACTTCTTCGTGTCCCGTGGTGGCACCACCGTAAACAACCTTCAGGTTCTCAGG<br>ACCATGAACAGCGGTGACGAGCTCAAGTACGGCAACTTCGTGAGGCGTGCCTTC |
|        | CAGGGTTTGGGTGACTCCTTCAGGGCCTATCAACAGTCTCTGGAGGACTGGCTCGAGAACCGCGAC<br>GACGCTAGGACCAGGAGCGTCTCTACACCCAGTACATCGCTTTGGAATTGGAC  |
|        | CAGTTGAAGGACTCCGAGACTGAGTTGCCACCAGAGACTACTGAGCGTCCAAACTACGAGTCTTAC<br>AGCCACAGGTTGTCCACATCGGCATCATCCTGCAGTCTAGGGTGAACGTGCCA  |
|        | GACATCATCCGCACCTCCATCCAAGGCCTCAGCGGCAACGGCGAGGTGTACATCGACAAGATCGA<br>GATCATCCAGTGACCGCCACCTTCGAGGCCGAGTACGACTTGAGAGGTAATGA   |
|        | GAATTCGGCCTCACTAGCCAGGAGATCCAGAGGTAACGAGCGCCAAGTCGAGAGGACTAGGGA<br>CTACTCCGACTACTGCGTGGAGTGGTACAACACTGGTCTCAACAGCTTGAGGGGC   |
|        | CTCGTCGGCGAGTTGTGGCCAGGGGCAGGGACCAATGGGAGATCTTCTTGGAACACGTCGAGCA<br>ACTCATCAACCAGCAGATCACCGAGAATGCCAGGAACACCGCTCTCGCTCGTCTC  |
|        | ACCACCCCATTCACCTTCACCCAGATCCAAGACATCATCCGCACCTCCATCCAAGGCCTCAGCGGC<br>AACGGCGAGGTGTACATCGACAAGATCGAGATCATCCAGTGACCGCCACCTTC  |
| cry1Aa | TTGGACATTGTGTCTCTTCCCGAACTATGACTCCAGAACCTACCCTATCCGTACAGTGTCCCAAC<br>TTACCAGAGAAATCTATACTAACCCAGTTCCTTGAGAACTTCGACGGTAGCTTC  |
|        | GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGAACCAGGCCATCTCTAGGTTGGAAGGA<br>TTGAGCAATCTCTACCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT |
|        | AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCTCATGGATTCATCGTAGTGCTGAGTTCA<br>ACAATATCATTCTTCTCTCAAATCACCCAAATCCCATTGACCAAGTCTACT   |
|        | ACCCTGTCTTCCACCTTGTACAGAAGACCCTTCAATATCGGTATCAACAACCAGCAACTTTCGGTTCT<br>TGACGGAACAGAGTTCGCCTATGGAACCTCTTCTAACTTGCCATCCGCTGTT |
|        | TACGTTCAAGCAGCTAATCTTACCTCAGCGTGCTTCGAGACGTTAGCGTGTGGGGCAAAGGTGG<br>GGATTGATGCTGCAACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT    |
|        | ATGGACAACAACCCAAACATCAACGAATGCATTCCATACAACTGCTTGAGTAACCCAGAAGTTGAAG<br>TACTTGGTGGAGAACGCATTGAAACCGGTTACACTCCCATCGACATCTCCTTG |
|        | CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAATTCAACGACATGAACAGCGCCTTGACCA<br>CAGCTATCCCATTGTTGCGAGTCCAGAATAACCAAGTTCCTCTCTTGTCCGTG |

|          |  |
|----------|--|
|          | CAAAGATATCGTGTGTCAGGATTCGTTACGCATCTACCACTAACTTGAATTCCACACCTCCATCGACG<br>GAAGGCCTATCAATCAGGGTAACTTCTCCGCAACCATGTCAAGCGGCAGCAAC  |
|          | ATCATGGCCTCTCCAGTTGGATTGAGCGGGCCCGAGTTTACCTTTCCTCTCTATGGAACATGGGAA<br>ACGCCGCTCCACAACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA    |
|          | TTGCAATCCGGCAGCTTCAGAACCGTCGGTTTCACTACTCCTTTCAACTTCTCTAACGGATCAAGCG<br>TTTTACCCTTAGCGCTCATGTGTTCAATTCTGGCAATGAAGTGTACATTGAC    |
|          | TACAGAAAGAGCGGAACCGTTGATTCCCTGGACGAAATCCCACCACAGAACAACAATGTGCCACCC<br>AGGCAAGGATTCTCCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATT     |
|          | CGTGGTTCTGCCAAGGTATCGAAGGCTCCATCAGGAGCCCACACTTGATGGACACCTTGAACAGC<br>ATAACTATCTACACCGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG    |
|          | GGAAACTACACCGACCACGCTGTTTCGTTGGTACAACACTGGCTTGGAGCGTGTCTGGGGTCTGAT<br>TCTAGAGATTGGATTAGATAACAACAGTTCAGGAGAGAATTGACCCTCACAGTT   |
|          | TTCTCTAACGGATCAAGCGTTTTACCCTTAGCGCTCATGTGTTCAATTCTGGCAATGAAGTGTACAT<br>TGACCGTATTGAGTTTGTGCCTGCCGAAGTTACCTTCGAGGCTGAGTACTGA    |
|          | TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTGCTGGGTTCTCGGACTAGTTGACA<br>TCATCTGGGGTATCTTTGGTCCATCTCAATGGGATGCATTCCCTGGTGCAAATT      |
|          | AACCTTGGATCTGGAACCTTCTGTCGTGAAAGGACCAGGCTTCACAGGAGGTGATATTCTTAGAAGAA<br>CTTCTCCTGGCCAGATTAGCACCCCTCAGAGTTAACATCACTGCACCACTTTCT |
| cry1Ab/c | TACGTTCAAGCAGCTAATCTTCACCTCAGCGTGCTTCGAGACGTTAGCGTGTGGGGCAAAGGTGG<br>GGATTGATGCTGCAACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT     |
|          | TCCTTGGATAATCTCCAATCCAGCGATTTTCGGTTACTTTGAAAGTGCCAATGCTTTTACATCTTCACT<br>CGGTAACATCGTGGGTGTTAGAACTTTAGTGGGACTGCAGGAGTGATTATC   |
|          | CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAATTCAACGACATGAACAGCGCCTTGACCA<br>CAGCTATCCCATTGTTTCGAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTG   |
|          | TTGGACATTGTGTCTCTTCCCGAACTATGACTCCAGAACCTACCCTATCCGTACAGTGTCCCAAC<br>TTACCAGAGAAATCTATACTAACCCAGTTCCTGAGAACTTCGACGGTAGCTTC     |
|          | GGAAACTACACCGACCACGCTGTTTCGTTGGTACAACACTGGCTTGGAGCGTGTCTGGGGTCTGAT<br>TCTAGAGATTGGATTAGATAACAACAGTTCAGGAGAGAATTGACCCTCACAGTT   |
|          | ACCTTGTCTTCCACCTGTACAGAAGACCCTTCAATATCGGTATCAACAACCAGCAACTTTCCGTTCT<br>TGACGGAACAGAGTTCGCCTATGGAACCTCTTCTAATTGCCATCCGCTGTT     |
|          | AGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCTCTTGGATACATCGTAGTGCTGAGTTCA<br>ACAACATCATCGCATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC   |
|          | CGTGGTTCTGCCAAGGTATCGAAGGCTCCATCAGGAGCCCACACTTGATGGACATCTTGAACAGC<br>ATAACTATCTACACCGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG    |
|          | GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGAACCAGGCCATCTCTAGGTTGGAAGGA<br>TTGAGCAATCTCTACCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT   |
|          | TTTCTTCAACGGTTCTGTCAATTCAGGACCAGGATTCAGTGGTGGAGACCTCGTTAGACTCAACA<br>GCAGTGGAATAACATTCAGAATAGAGGGTATATTGAAGTTCCAATTCATTC       |
|          | CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATGCTTCTGTGACCCCTATTCACCTCAACG<br>TTAATTGGGGTAATTCATCCATCTTCTCCAATACAGTTCCAGCTACAGCTACC   |

|        |   |
|--------|---|
|        | TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTGCTGGGTTTCGTTCTCGGACTAGTTGACA<br>TCATCTGGGGTATCTTTGGTCCATCTCAATGGGATGCATTCCTGGTGCAAATT |
|        | ATCATGGCCTCTCCAGTTGGATTGAGCGGGCCCGAGTTTACCTTTCCTCTCTATGGAACATGGGAA<br>ACGCCGCTCCACAACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA   |
|        | AGATTCGAGTTCATTCCAGTACTGCAAACTCGAGGCTGAGTACAACCTCGAAAGAGCCCAGAAG<br>GCTGTAATGCCCTCTCACCTCTACAAACCAGCTTGGACTCAAGACAAATGTGA     |
|        | TACAGAAAGAGCGGAACCGTTGATTCCCTTGGACGAAATCCCACCACAGAACAACAATGTGCCACCC<br>AGGCAAGGATTCTCCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATT   |
|        | ATGGACAACAACCCAAACATCAACGAATGCATTCCATACTGCTTGAGTAACCCAGAAGTTGAAG<br>TACTTGGTGGAGAACGCATTGAAACCGGTTAACTCCCATCGACATCTCCTTG      |
|        | GACAGATTCGAGTTCATTCCAGTACTGCAAACTCGAGGCTGAGTACAACCTCGAAAGAGCCCAG<br>AAGGCTGTAATGCCCTCTCACCTCTACAAACCAGCTTGGACTCAAGACAAATG     |
| cry1Ac | CCTACTAACCCAGCTCTCCGCGAGGAAATGCGTATTCAATTCAACGACATGAACAGCGCCTTGACCA<br>CAGCTATCCCATTGTTTCGCAGTCCAGAACTACCAAGTTCCTCTCTTGTCCGTG |
|        | TACGTTCAAGCAGTAATCTTCACCTCAGCGTGCTTCGAGACGTTAGCGTGTGGGGCAAAGGTGG<br>GGATTGATGCTGCAACCATCAATAGCCGTTACAACGACCTTACTAGGCTGATT     |
|        | ATGGACAACAACCCAAACATCAACGAATGCATTCCATACTGCTTGAGTAACCCAGAAGTTGAAG<br>TACTTGGTGGAGAACGCATTGAAACCGGTTAACTCCCATCGACATCTCCTTG      |
|        | ACCTTGTCTTCCACCTGTACAGAAGACCCTTCAATATCGGTATCAACAACCAGCAACTTTCGGTCT<br>TGACGGAACAGAGTTCGCCTATGGAACCTCTTCTAAGTCCATCCGCTGTT      |
|        | TTTCTCTTCAACGGTCTGTCAATTCAGGACCAGGATTCAGTGGTGGAGACCTCGTTAGACTCAACA<br>GCAGTGGAATAACATTCAGAATAGAGGGTATATTGAAGTTCCAATTCATTC     |
|        | GAGCAGTTGATCAACCAGAGGATCGAAGAGTTCGCCAGGAACCAGGCCATCTCTAGGTTGGAAGGA<br>TTGAGCAATCTCTACCAAATCTATGCAGAGAGCTTCAGAGAGTGGGAAGCCGAT  |
|        | TCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTGCTGGGTTTCGTTCTCGGACTAGTTGACA<br>TCATCTGGGGTATCTTTGGTCCATCTCAATGGGATGCATTCCTGGTGCAAATT |
|        | AGATTCGAGTTCATTCCAGTACTGCAAACTCGAGGCTGAGTACAACCTTGGAGAGAGCCCAGAAG<br>GCTGTGAACGCCCTCTTACCTCCACCAATCAGCTTGGCTTGAAAATAACGTT     |
|        | GACAGATTCGAGTTCATTCCAGTACTGCAAACTCGAGGCTGAGTACAACCTTGGAGAGAGCCCAG<br>AAGGCTGTGAACGCCCTCTTACCTCCACCAATCAGCTTGGCTTGAAAATAAC     |
|        | CCATCCACATCTACCAGATATAGAGTTCGTGTGAGGTATGCTTCTGTGACCCCTATTCACCTCAACG<br>TTAATTGGGGTAATTCATCCATCTTCTCCAATACAGTTCCAGCTACAGCTACC  |
|        | GGAAACTACACCGACCACGCTGTTTCGTTGGTACAACACTGGCTTGGAGCGTGTCTGGGGTCTGAT<br>TCTAGAGATTGATTAGATACAACCAGTTCAGGAGAGAATTGACCCTCACAGTT   |
|        | AGCAACAGTTCGTTGAGCATCATCAGAGCTCCTATGTTCTCTTGGATACATCGTAGTGCTGAGTTCA<br>ACAACATCATCGCATCCGATAGTATTACTCAAATCCCTGCAGTGAAGGGAAAC  |
|        | TACAGAAAGAGCGGAACCGTTGATTCCCTTGGACGAAATCCCACCACAGAACAACAATGTGCCACCC<br>AGGCAAGGATTCTCCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGATT   |
|        | CGTGGTTCGCCCAGGTATCGAAGGCTCCATCAGGAGCCCACACTTGTATGGACATCTTGAACAGC<br>ATAACTATCTACACCGATGCTCACAGAGGAGAGTATTACTGGTCTGGACACCAG   |

|            |  |
|------------|--|
|            | TCCTTGGATAATCTCCAATCCAGCGATTTTCGGTTACTTTGAAAGTGCCAATGCTTTTACATCTTCACT<br>CGGTAACATCGTGGGTGTTAGAAACTTTAGTGGGACTGCAGGAGTGATTATC  |
|            | TTGGACATTGTGTCTCTCTTCCCGAACCTATGACTCCAGAACCTACCCTATCCGTACAGTGTCCTAAC<br>TTACCAGAGAAATCTATACTAACCCAGTTCTTGAGAACTTCGACGGTAGCTTC  |
|            | ATCATGGCCTCTCCAGTTGGATTGAGCGGGCCCGAGTTTACCTTTCCTCTCTATGGAACCTATGGGAA<br>ACGCCGCTCCACAACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTACAGA  |
| Cry1A105_1 | GGCTATGTTACCAAGGAGCTTGAGTACTTTCTGAGACCGACAAAGTGTGGATCGAGATCGGTGAA<br>ACCGAGGGAACCTTCATCGTGGACAGCGTGGAGCTTCTTGTATGGAGGAATAA     |
|            | GTTTACAGAAAGAGCGGAACCGTTGATTCTTGGACGAAATCCCACCACAGAACAACAATGTGCCAC<br>CCAGGCAAGGATTCTCCACAGGTTGAGCCACGTGTCCATGTTCCGTTCCGGA     |
|            | TTGTCCTTGACACAGTTTCTGCTCAGCGAGTTCGTGCCAGGTGCTGGGTTCTCGGACTAGTTG<br>ACATCATCTGGGGTATCTTTGGTCCATCTCAATGGGATGCATTCTGGTGAA         |
|            | AATGAGGACCTCGGAGTCTGGGTCATCTTCAAGATCAAGACCCAAGACGGACACGCAAGACTTGGC<br>AACCTTGAGTTTCTCGAAGAGAAACCATTGGTCCGTGAAGCTCTCGCTCGTGTG   |
|            | GTCACCCTCTCCGGAACCTTCGACGAGTGTACCCTACCTACTTGTACCAGAAGATCGATGAGTCCA<br>AACTCAAAGCCTTACCAGGTATCAACTTAGAGGCTACATCGAAGACAGCCAA     |
|            | AGATGCGCTCCACACCTTGAGTGAATCCTGACTTGGACTGCTCCTGCAGGGATGGCGAGAAGTGT<br>GCCACCATTCTCATCACTTCTCCTTGGACATCGATGTGGGATGTAAGTACTGACCTG |
|            | GTGTACGTTCAAGCAGCTAATCTTACCTCAGCGTGCTTCGAGACGTTAGCGTGTGGGCAAAGGT<br>GGGATTGATGCTGCAACCATCAATAGCCGTTACAACGACCTTACTAGGCTG        |
|            | GACCTTGAAATCTACTCGATCAGGTACAATGCCAAGCAGACCGTGAATGTCCAGGTAAGTGTG<br>CCCTCTGGCCACTTCTGCCAATCTCCATTGGGAAGTGTGGAGAGCCTAAC          |
|            | GCCATGGACAACAACCCAAACATCAACGAATGCATTCCATACAACCTGCTTGAGTAACCCAGAAGTTG<br>AAGTACTTGGTGGAGAACGCATTGAAACCGTTACTACTCCATCGACATCTCC   |
|            | GTTACTGACTATCACATTGACCAAGTGTCCAACCTTGGTCCACTACCTTAGCGATGAGTTCTGCCTCG<br>ACGAGAAGCGTGAACCTCTCCGAGAAAGTTAAACACGCCAAGCGTCTCAGCGAC |
|            | TTCAGCAACAGTTCCGTGAGCATCATCAGAGCTCCTATGTTCTTGGATACACCGTAGTGCTGAGT<br>TCAACAACATCATTGCATCCGACAGCATTACTCAAATACCTTGGTGAAGCA       |
|            | CGTGAGAACCCTTGCAGTTCAACAGAGGTTACAGGGACTACACACCACTTCCAGTTGGCTATGTTA<br>CCAAGGAGCTTGAAGTCTTCTGAGACCGACAAAGTGTGGATCGAGATCGGT      |
|            | CAGATCATGGCCTCTCCAGTTGGATTGAGCGGGCCCGAGTTTACCTTTCCTCTCTATGGAACCTATGG<br>GAAACGCCGCTCCACAACAACGTATCGTTGCTCAACTAGGTCAGGGTGTCTAC  |
|            | CCCCAAAGGTATCGTGCAAGAATCCGCTATGCCTCTACTACAAATCTCAGGATCTACGTGACTGTTG<br>CAGGTGAAAGGATCTTGTGTTGTCAGTTCAACAAGACTATGGATACCGGTGAC   |
|            | ATTGGAAACTACACCGACCAGCTGTTCTGTTGGTACAACACTGGCTTGGAGCGTGTCTGGGGTCT<br>GATTCTAGAGATTGGATTAGATAACAACAGTTTCAAGGAGAGAATTGACCCTCACA  |
|            | AGAACCTTGTCTTCCACCTTGTACAGAAGACCTTCAATATCGGTATCAACAACAGCAACTTCCGT<br>TCTTGACGGAACAGAGTTGCCTATGGAACCTTCTTAACCTTGCATCCGCT        |
|            | GAGAGGAATCTCTTGAAGACTCCAACCTTCAAAGACATCAACAGGCAGCCAGAACGTGGTTGGGGT<br>GGAAGCACCGGGATCACCATCCAAGGAGGCGACGATGTGTTCAAGGAGAACTAC   |

|            |   |
|------------|---|
|            | CATACACTTCAGTCAGGTA<br>CTACTGTTGTCAGAGGTC<br>CAGGGTTTACAGGAGGAG<br>ACATTCTTCGTCGCACA<br>AGTGGAGGACCCTTTGCT<br>TACACTATTGTTAACATCA<br>ATGGCCAATTG  |
|            | GACAGGTTTGAATTGATT<br>CCAGTACTGCAACCCTCG<br>AGGCTGAGTACAACCTTG<br>AGAGAGCCCAGAAGGCT<br>GTGAACGCCCTCTTACCT<br>CCACCAATCAGCTTGGCT<br>TGAAAATAAC     |
|            | GTTTTGGACATTGTGTCT<br>CTCTCCCGAACTATGACT<br>CCAGAACCTACCCTATCC<br>GTACAGTGTCCCAACTT<br>ACCAGAGAAATCTATACT<br>AACCCAGTTCTTGAGAA<br>CTTCGACGGTAGC   |
|            | GAACAGAACAAATCAGCG<br>TTCGGTCTGGTTGTGCCT<br>GAGTGGGAAGCTGAAGT<br>GTCCCAAGAGGTTAGAG<br>TCTGCCAGGTAGAGGCT<br>ACATTCTCCGTGTGACCG<br>CTTACAAGGAGGGA   |
|            | GAACTTGAGGGACGTATC<br>TTTACCGCATTCTCCTTG<br>TACGATGCCAGAAACGTC<br>ATCAAGAACGGTGAAC<br>TCAACAATGGCCTCAGCT<br>GCTGGAATGTGAAAGGTC<br>ATGTGGACGTGGAG  |
|            | TTGCAAGCCGACACCAAC<br>ATCGCCATGATCCACGCC<br>GCAGACAAACGTGTGCAC<br>AGCATTCTGTGAGGCTT<br>ACTTGCCGTGATCCCTGG<br>TGTGAACGCTGCCATCT<br>TCGAG           |
|            | TACGGTGAGGGTTGCGTG<br>ACCATCCACGAGATCGAG<br>AACAACACCGACGAGCT<br>TAAAGTTCTCCAACGCG<br>TCGAGGAAGAAATCTAT<br>CCCAACAACACCGTTACT<br>TGCAACGACTACACT  |
|            | CCTTTGACATTCCAATCT<br>TTTAGCTACGCAACTATCA<br>ACACAGCTTTTACATTCC<br>CAATGAGCCAGAGTAGCT<br>TTCACAGTAGGTGCTGAC<br>ACTTTCAGCTCAGGGAAT<br>GAAGTTTACATC |
|            | GATCCTACTAACCCAGCT<br>CTCCGCGAGGAAATGCGT<br>ATTCAATTCAACGACATGA<br>ACAGCGCCTTGACCACAG<br>CTATCCCATTGTTCCGAG<br>TCCAGAACTACCAAGTT<br>CCTCTCTTGTC   |
|            | ATTGAGCAGTTGATCAAC<br>CAGAGGATCGAAGAGTT<br>CGCCAGGAACCAGGCCAT<br>CTCTAGGTTGGAAGGAT<br>TGAGCAATCTCTACCAAT<br>CTATGCAGAGAGCTTCAG<br>AGAGTGGGAAGCC   |
|            | AAGAGAGCAGAGAAGAAG<br>TGGAGGGACAAACGTGAG<br>AAACTCGAATGGGAAACT<br>AACATCGTTTACAAGGAG<br>GCCAAAGAGTCCGTGGAT<br>GCTTTGTTGTTGAACTCC<br>CAATATGATCAG  |
|            | GTGAATCAGGAAGAGTAC<br>GGAGGTGCCTACACTAGC<br>CGTAACAGAGGTTACAAC<br>GAAGCTCCTTCCGTTCT<br>GCTGACTATGCCTCCGT<br>GTACGAGGAGAAATCCT<br>ACACAGATGGCAGA   |
|            | TTCCGTGGTTCTGCCCA<br>AGGTATCGAAGGCTCCAT<br>CAGGAGCCCACACTTGAT<br>GGACATCTTGAACAGCAT<br>AATCTACACCGATGCTCA<br>CAGAGGAGAGTATTACT<br>GGTCTGGACAC     |
| Cry1A105_2 | GTTCAAGGAGAACTACGT<br>CACCTCTCCGGAACTTTC<br>GACGAGTGCTACCCTACT<br>ACTTGTACCAGAAGATCG<br>ATGAGTCCAAACTCAAAG<br>CCTTACCAGGTATCAACT<br>TAGAGGCTA     |
|            | GAATGAAGTTTACATCGA<br>CAGGTTTGAATTGATTCC<br>AGTACTGCAACCCTCGAG<br>GCTGAGTACAACCTTGAG<br>AGAGCCCAGAAGGCTGT<br>GAACGCCCTCTTACCTCC<br>ACCAATCAGCT    |
|            | CATCGAAGACAGCCAAGAC<br>CTTGAAATCTACTCGATC<br>AGGTACAATGCCAAGCAC<br>GAGACCGTGAAATGTC<br>CCAGTACTGGTCCCTCT<br>GGCCACTTCTGCCAATCT<br>CCCATTGGGAA     |
|            | TTCGTGCCAGGCGCTGGC<br>TTGCTGCTGGCCTCGTGG<br>ACATCATCTGGGGCATCT<br>TTGGCCCCTCCAGTGGG<br>ACGCCTTCTGGTGCAAAT<br>CGAGCAGCTCATCAACC<br>AGAGGATCGAG     |
|            | AGCTCTCGCTCGTGTGA<br>AGAGAGCAGAGAAGAAG<br>TGGAGGGACAAACGTGAG<br>AAACTCGAATGGGAACT<br>AACATCGTTTACAAGGAG<br>GCCAAAGAGTCCGTGGAT<br>GCTTTGTTGTTGAA   |
|            | CAAGCGTCTCAGCGACGAG<br>AGGAATCTCTTGCAAGACT<br>CCAACCTCAAAGACATCA<br>ACAGGCAGCCAGAACGT<br>GGTTGGGGTGAAGCACCG<br>GGATCACCATCCAAGGAG<br>GCGACGATGT   |
|            | TGGCTTGAAAATAACGTT<br>ACTGACTATCACATTGACC<br>AAGTGTCCAACCTGGTCA<br>CCTACCTTAGCGATGAG<br>TTCTGCCTCGACGAGAAG<br>CGTGAACCTCTCCGAGAA<br>AGTTAAACACGC  |



|  |
|--|
| CTACTGGTCCGGCCACCAGATCATGGCCTCCCCGGTCCGGCTTCAGCGGCCCCGAGTTTACCTTTCC<br>TCTCTACGGCACGATGGGCAACGCCGCTCCACAACAACGCATCGTCGCTCAGCT  |
| ATGCGCATCCAGTTCAACGACATGAACAGCGCCCTGACCACCGCCATCCCACTCTTCGCCGTCCAG<br>AACTACCAAGTCCCGCTCCTGTCCGTGTACGTCCAGGCCGCCAACCTGCACCTC   |
| GAGTTCGCCAGGAACCAGGCCATCAGCCGCCTGGAGGGCCTCAGCAACCTCTACCAAATCTACGCT<br>GAGAGCTTCCGCGAGTGGGAGGCCGACCCCACTAACCCAGCTCTCCGCGAGGAG   |
| TTGCAACGACTACACTGTGAATCAGGAAGAGTACGGAGGTGCCTACACTAGCCGTAACAGAGGTTA<br>CAACGAAGCTCCTTCCGTTCTGTGACTATGCCTCCGTGTACGAGGAGAAATC     |
| CTCCCAATATGATCAGTTGCAAGCCGACACCAACATCGCCATGATCCACGCCGACAGACAAACGTGT<br>GCACAGCATTCTGTAGGCTTACTTGCCCTGAGTTGTCCGTGATCCCTGGTGTGAA |
| TATGGATACCGGTGACCCTTTGACATTCCAATCTTTTAGCTACGCAACTATCAACACAGCTTTTACAT<br>TCCCAATGAGCCAGAGTAGCTTACAGTAGGTGCTGACACTTTCAGCTCAGG    |
| CTACACAGATGGCAGACGTGAGAACCCTTGCGAGTTCAACAGAGGTTACAGGGACTACACACCACT<br>TCCAGTTGGCTATGTTACCAAGGAGCTTGAGTACTTTCCTGAGACCGACAAAAGT  |
| GGGATGTACTGACCTGAATGAGGACCTCGGAGTCTGGGTATCTTCAAGATCAAGACCCAAGACGG<br>ACACGCAAGACTTGGCAACCTTGAGTTTCTCGAAGAGAAACCATTGGTCCGTGA    |
| CATCAATGGCCAATTGCCCAAAGGTATCGTGCAAGAATCCGCTATGCCTCTACTACAAATCTCAGG<br>ATCTACGTGACTGTTGCAGGTGAAAGGATCTTTGCTGGTCAGTTCAACAAGAC    |
| CGCTTACAAGGAGGGATACGGTGAGGGTTGCGTGACCATCCACGAGATCGAGAACAACACCGACG<br>AGCTTAAGTTCTCCAACCTGCGTCGAGGAAGAAATCTATCCAACAACACCGTTAC   |
| TCATGTGGACGTGGAGGAACAGAACAATCAGCGTTCCGTCCTGGTTGTGCCTGAGTGGGAAGCTGA<br>AGTGTCCCAAGAGGTTAGAGTCTGTCCAGGTAGAGGCTACATTCTCCGTGTGAC   |
| GAACCTCGACGGTAGCTTCAGGGGCAGCGCCAGGGCATCGAGGGCTCCATCAGGAGCCACACC<br>TGATGGACATCCTCAACAGCATCACTATCTACACCGATGCCACCGCGGGCAGTA      |
| GGCTATGTTACCAAGGAGCTTGAGTACTTTCTGAGACCGACAAAAGTGTGGATCGAGATCGGTGAA<br>ACCGAGGGAACCTTCATCGTGACAGCGTGGAGCTTCTTTGATGGAGGAATAA     |
| CGCTGCCATCTTCGAGGAACTTGAGGGACGTATCTTTACCGCATTCTCCTTGTACGATGCCAGAAAC<br>GTCATCAAGAACGGTGACTTCAACAATGGCCTCAGCTGCTGGAATGTGAAAGG   |
| CAACCTGCCCTCCGCTGTCTACCGCAAGAGCGGCACGGTGGATTCCCTGGACGAGATCCCACCAC<br>AGAACAACAATGTGCCCCCAGGCAGGGTTTTTCCCACAGGCTCAGCCACGTGTC    |
| GTGTGGAGAGCCTAACAGATGCGCTCCACACCTTGAGTGGAATCCTGACTTGGACTGCTCCTGCAG<br>GGATGGCGAGAAGTGTGCCACCATTCTCATCACTTCTCCTTGGACATCGATGT    |
| AGCGTGCTGAGGGACGTGAGCGTGTGTTGGCCAGAGGTGGGGCTTCGACGCCGCCACCATCAACAG<br>CCGCTACAACGACCTCACCAGGCTGATCGGCAACTACACCGACCACGCTGTCCGC  |
| GGGCCAGGGCGTCTACCGCACCTGAGCTCCACCCTGTACCGCAGGCCCTTCAACATCGGTATCA<br>ACAACCAGCAGCTGTCCGTCTGGATGGCACTGAGTTCGCTACGGCACCTCCTC      |
| ACCCTTGGTGAAAGCACATACACTTCAGTCAGGTAATACTGTTGTCAGAGGTCCAGGGTTTACAGGA<br>GGAGACATTCTTCGTGCAACAAGTGGAGGACCCCTTTGCTTACACTATTGTTAA  |
| CATGTTCCGCTCCGGCTTCAGCAACTCGTCCGTGAGCATCATCAGAGCTCCTATGTTCTCTTGGATA<br>CACCGTAGTGCTGAGTTCAACAACATCATTGCATCCGACAGCATTACTCAAAAT  |

|          |  |
|----------|--|
|          | TGCATCCCGTACAACCTGCCTCAGCAACCCTGAGGTGAGGTGCTCGGCGGTGAGCGCATCGAGAC<br>CGGTTACACCCCATCGACATCTCCCTCTCCCTCACGCAGTTCCTGCTCAGCGAG    |
|          | TGGTACAACACTGGCCGTCTGGACATTGTGTCCCTTTCCCGAACTACGACTCCCGCACCTACCC<br>GATCCGCACCGTGTCCCAACTGACCCGCGAAATCTACACCAACCCCGTCTGGA      |
| Cry2Ab2  | AACTCCGTCTGAACTCTGGTCGCACCACCATCTGCGACGCCTACAACGTGCGGGCGCATGATCCA<br>TTCAGCTTCCAGCACAAAGAGCCTCGACACTGTTTTCAGAAGGAGTGGACGGAGTGG |
|          | TCCATGATCCACCTGGCGCCCAATGATTACACCGGCTTACCATCTCTCCAATCCACGCCACCCAAG<br>TGAACAACCAGACACGCACCTTCATCTCCGAGAAGTTTCGGCAACCAGGGCGAC   |
|          | GCAAACGTGGAGGAGTTCAACCGCCAAGTGGACAACCTCCTCAACCCCAACCGCAATGCGGTGCCT<br>CTGTCCATCACTTCTCCGTGAACACCATGCAACAACCTGTTCTCAACCGCTTG    |
|          | AAGAAGAACAACCACAGCCTGTACCTGGACCCCATCGTCGGCACGGTGGCCAGCTTCTTCTCAAG<br>AAGGTGCGCTCTCTCGTCGGGAAGCGCATCCTCTCGGAACTCCGCAACCTGATC    |
|          | CTTGGCCTCCGGAGCGGCGCCTTACGGCGCGTGGAAATTCTAACTACTTCCCGACTACTTCATC<br>AGGAACATCTCTGGTGTCTCTCGTCGTCGCAACGAGGACCTCCGCCGTCCA        |
|          | TCCCTGAGGTTGAGCAGAACAACACCACCGCCAGGTACACCCTGCGCGGCAACGGCAACAGCTA<br>CAACCTGTACCTGCGCGTCAGCTCCATTGGCAACTCCACCATCAGGGTCACCATC    |
|          | TGCTCCACCTTCTGCCGCGCTGCTCACCCCGTTCGTGAGGTCTGGCTCGACAGCGGCTCCGA<br>CCGCGAGGGCGTGGCCACCGTCACCAACTGGCAAACCGAGTCTTCGAGACCACC       |
|          | CCTCAGTTCAGATGCAAGGCTACCAGCTGCTCCTGCTGCCACTCTTGTCTCAGGCTGCCAACCTG<br>CACCTCTCCTTCATTTCGTGACGTGATCCTCAACGCTGACGAGTGGGGCATCTCT   |
|          | AACATTGTTGGCCTCCCCGGCTCCACCACAACCTCATGCTCTGCTTGCTGCCAGAGTGAACACTCC<br>GGCGGCATCTCGAGCGGCGACATTGGTGCATCGCCGTTCAACCAGAACTTCAAC   |
|          | CTGGAGTTCAGGACCTACATGTTCTGAACGTGTTTCGAGTACGTCAGCATCTGGTCGCTCTTCAAGT<br>ACCAGAGCCTGCTGGTGTCCAGCGGCGCAACCTCTACGCCAGCGGCTCTGGT    |
|          | CTGCACTACAACGAGATCAGGAACATCGCCTCTCCGTCCGGGACGCCCGGAGGTGCAAGGGCGTA<br>CATGGTGAGCGTCCATAACAGGAAGAACAACATCCACGCTGTGCATGAGAACGGC   |
|          | TGGTGGCCAGCAGCAACTCCGACGTCCCGCTGGACATCAACGTGACCCTGAACTCTGGCACCCAG<br>TTCGACCTCATGAACATCATGCTGGTGCCAACTAACATCTCGCCGCTGTACTGAT   |
|          | CCCCAACAACTCAGAGCTTACCAGCCAGGACTGGCCATTCTGTATTGTTGTTCCAAGTCAACT<br>CCAACACTAGTCTCAACGGCTTCTCTGGTGTCTCGCCTCTCAACACCTTCCCC       |
|          | GCAGCCACGCTGAGGACCTACCGCGACTACCTGAAGAATAACACCAGGACTACTCCAACCTATTGC<br>ATCAACACCTACCAGTCGGCCTTCAAGGGCCTCAATACGAGGCTTACGACATG    |
|          | TTTCCATCTGGCTCCACCAACCTCATGCAAGACATCCTCAGGGAGACCGAGAAGTTTCTCAACCAGC<br>GCCTCAACACTGATACCCTTGTCTCGCGTCAACGCTGAGCTGACGGGTCTGCAA  |
|          | AACGGGAGGGTGTACACAGCCACCAATGTGAACACGACGACCAACAATGATGGCGTCAACGACAAC<br>GGCGCCCGCTTCAGCGACATCAACATTGGCAACGTGGTGGCCAGCAGCAACTCC   |
| Cry3A055 | TACAACGTGGGTCTAGACAAGCTCCGCGGCAGCAGCTACGAGAGCTGGGTGAACTTCAACCGCTAC<br>CGCCGCGAGATGACCCTGACCGTGCTGGACCTGATCGCCCTGTTCCCCCTGTAC   |
|          | GTGGCAGGCCCGCTTACCAGCGGCGGACATCATCCAGTGCACCGAGAACGGCAGCGCCGCCA<br>CCATCTACGTGACCCCGACGTGAGCTACAGCCAGAAGTACCAGCGCCCGCATCCAC     |

|          |  |
|----------|--|
|          | TACGCCAGCACCAGCCAGATCACCTTCACCCTGAGCCTGGACGGGGCCCCCTTCAACCAATACTAC<br>TTCGACAAGACCATCAACAAGGGCGACACCCTGACCTACAACAGCTTCAACCTG   |
|          | ATGACGGCCGACAACAACACCGAGGCCCTGGACAGCAGCACCACCAAGGACGTGATCCAGAAGGG<br>CATCAGCGTGGTGGGCGACCTGCTGGGCGTGGTGGGCTTCCCCTTCGGCGGCGCC   |
|          | CTGGCCAGCTTCAGCACCCCTTTTCGAGCTGAGCGGCAACAACCTCCAGATCGGCGTGACCGGCCT<br>GAGCGCCGGCGACAAGGTGTACATCGACAAGATCGAGTTCATCCCCGTGAACCTAG |
|          | GACGTGCGCCTGTACCCCAAGGAGGTGAAGACCGAGCTGACCCGCGACGTGCTGACCGACCCCAT<br>CGTGGGCGTGAACAACCTGCGCGGCTACGGCACCACTTCAGCAACATCGAGAAC    |
|          | TACATCCGCAAGCCCCACCTGTTGACTACCTGCACCGCATCCAGTTCACACGCGTTTCCAGCCC<br>GGCTACTACGGCAACGACAGCTTCACTACTGGAGCGGCAACTACGTGAGCACC      |
|          | AACCTGGCCGTGTGGCCCTCTGCAGTGTACAGCGGCGTGACCAAGGTGGAGTTCAGCCAGTACAA<br>CGACCAGACCGACGAGGCCAGCACCCAGACCTACGACAGCAAGCGCAACGTGGGC   |
|          | AAGGCACTGGCCGAGCTACAGGGCCTCCAGAACAACGTGGAGGACTATGTGAGCGCCCTGAGCAG<br>CTGGCAGAAGAACCCCGCTGCACCGTTCGCAACCCCCACAGCCAGGGCCGCATC    |
|          | ACCATCCCCGTGCTGACCTGGACCCACAAGAGCGTCGACTTCTTCAACATGATCGACAGCAAGAAG<br>ATCACCCAGCTGCCCTGGTGAAGGCCTACAAGCTCCAGAGCGGCGCCAGCGTG    |
|          | CTGGTGAGCTTCTACACCAACTTCTGAACACCATCTGGCCCAGCGAGGACCCCTGGAAGGCCTTC<br>ATGGAGCAGGTGGAGGCCCTGATGGACCAGAAGATCGCCGACTACGCCAAGAAC    |
|          | CTGCTGAAGGACGCCAAATCTACGGAGAGGAGTGGGGCTACGAGAAGGAGGACATCGCCGAGTT<br>CTACAAGCGCCAGCTGAAGCTGACCCAGGAGTACACCGACCACTGCGTGAAGTGG    |
|          | CGCCCCAGCATCGGCAGCAACGACATCATCACCAGCCCCTTCTACGGCAACAAGAGCAGCGAGCC<br>CGTGCAGAACCTTGAGTTCAACGGCGAGAAGGTGTACCGCGCCGTGGCTAACACC   |
|          | GCCGTGAGCTGGGACAGCATCGACCAGCTGCCCCCGAGACCACCGACGAGCCCCTGGAGAAGG<br>GCTACAGCCACCAGCTGAACTACGTGATGTGCTTCTGATGCAGGGCAGCCGCGGC     |
|          | CGCGAGCTGTTTCAGCCAGGCCGAGAGCCACTTCCGCAACAGCATGCCAGCTTCGCCATCAGCGG<br>CTACGAGGTGCTGTTCTGACCACCTACGCCAGGCCGCAACACCCACCTGTTCT     |
| Cry34Ab1 | CAGCCTCTACTTCGACAACCCGTTTCGCCGGCTCCAACAATACGACGGCCACTCCAACAAGTCCA<br>GTACGAGATCATACCCAGGGCGGCTCCGGCAACCAGTCCCACGTGACCTACAC     |
|          | CAGTACGAGATCATACCCAGGGCGGCTCCGGCAACCAGTCCCACGTGACCTACACCATCCAGAC<br>CACCTCCTCCCGCTACGGCCACAAGTCTGAGTCATGAGTCATGAGTCAGTTAAC     |
|          | GTCGACTCTAGAGGATCCACACGACACCATGTCCGCCCCGAGGTGCACATCGACGTGAACAACAA<br>GACCGGCCACACCCTCCAGCTGGAGGACAAGACCAAGCTCGACGGCGGCAGGTG    |
|          | GCGCACCTCCCCGACCAACGTGGCCAACGACCAGATCAAGACCTTCGTGGCCGAATCCAACGGCT<br>TCATGACCGGCACCGAGGGCACCATCTACTACTCAATTAATGGCGAGGCCGAGAT   |
| Cry35Ab1 | TGGGGCACCCGAGATCGACCAGAAGACCACCATCATCAACACCCTCGGCTTCCAGATCAACATCGAC<br>AGCGGCATGAAGTTCGACATCCCGGAGGTGGGCGGCGGTACCGACGAGATCAAG  |
|          | AAGAACACCCAGATCAAGACCACCCCGTACTACATCCTCAAGAAGTACCAGTACTGGCAGAGGGCC<br>GTGGGCTCCAACGTGCGGCTCCGCCCGCACGAGAAGAAGTCTACACCTACGAG    |
|          | AACAACGACAAGATTAATGTGTCAACCTACTCCTCCACCAACTCCATCCAGAAGTGGCAGATCAAGG<br>CCAACGGCTCCTCCTACGTGATCCAGTCCGACAACGGCAAGGTGCTCACCGCC   |

|         |   |
|---------|---|
|         | GGCACCGGCCAGGCCCTCGGCCTCATCCGCTCACCGACGAGTCCTCCAACAACCCGAACCAGCA<br>ATGGAACCTGACGTCCGTGCAGACCATCCAGCTCCCGCAGAAGCCGATCATCGAC   |
|         | CTCACCATCACCTCCCTGGAGCTCTACCGCTACAACGGCTCCGAGATCCGCATCATGCAGATCCAG<br>ACCTCCGACAACGACACCTACAACGTGACCTCCTACCCGAACCACCAGCAGGCC  |
|         | ATGAACAAGAACGACGACGACATCGACGACTACAACCTCAAGTGGTTCCTCTTCCCGATCGACGAC<br>GACCAGTACATCATCACCTCCTACGCCGCAACAACCTGCAAGGTGTGGAACGTG  |
|         | CTCTACCGCTACAACGGCTCCGAGATCCGCATCATGCAGATCCAGACCTCCGACAACGACACCTAC<br>AACGTGACCTCCTACCCGAACCACCAGCAGGCCCTGCTGCTGCTGACCAACCAC  |
|         | ACCCAGCTCAACGAGGAGCTCAAGATCGAGTATTCACATGAGACGAAGATCATGGAGAAGTACCAG<br>GAGCAGTCCGAGATCGACAACCCGACCGACAGTCCATGAACTCCATCGGCTTC   |
|         | ACCAAGCTCAAGGACTACCCGAAGTACTCCCCGACCGGCAACATCGACAACGGCACCTCCCCGCA<br>GCTCATGGGCTGGACCCTCGTGCCGTGCATCATGGTGAACGACCCGAACATCGAC  |
| Epsps_1 | GAAGGCGGCGGCAGAGCGAGAATCGGATATTCGTCGATCATCGAAGGCGCGCGGTCTTCCGGCA<br>CCGTGACGCCCTTCAGCGTGGAGGAGCGAACGCGCAGGTCCGCCACGTCTTCGCC   |
|         | AAGCAGGGCCGCAACCAGCGGGAAGGCCGTGAGGACGGGTCGCCCGGCACGTGATGACTTGG<br>CCGGTGAGCTTGCCGCGGCCTTCAGGCGGATGGTGCGCACGCCGTCCGCATCCGT     |
|         | GCCGGCAAGGCGCGGGTTGATGACTTCGATGTGCGGCCCATTTCTGCAGCGTCAGGATGAGGC<br>CGGTGCGGGTGGGGTTCATCAGCACGTTGAGGATGGTGACGTCGGAGCCCGGAAC    |
|         | CTCGACGGTAAGGTTGGCGCCAAAGCCCTGCAGCATTTTTCCGTATGATCGCGCGTCATGATCGG<br>CTCGATGACCGTCGTGATGCCGGGCGTGTGAGGCCGGCGAGCAGCACGGCGGA    |
|         | CTTCACCTGTGCGGAGGCCATCGGCACGCGGTAGGTGATCGGCGTCGGCGTCTTCGGCCCGCGCA<br>AGGTAACGGGAAGACGGTCACCGTCTTCCGATTTCACTGCACGCCCATTTCCGC   |
|         | GCCGAAATCGAGCGGCGCCTCAGGCGCCAGGAGCCGCCATTGCCGACGCCATCGATGATCCAG<br>GTGTGCGCTTCTTACGGATCCTGGCGCCCATGGCCTGCATGGCCTTGCCCGTATT    |
|         | GATGACGTCCTCGCCTTCCAGAAGGCCGGTGATGCGCGTTTTACCGCTCGCGAGACCGCCGAACA<br>TGAAGGACCGGTGGGAGATCGACTTGTGCGCGGGAATGCGGACGGTTCCGGAAAAG |
|         | CGACACGAGGCCCATGACGAGGAAGCTCATGGCGATGCGGTGATCGAGATGGGTGGCGACGGCG<br>GCGCCCGAGGCGTTGCCGAGCCCTTGCCGTGAGGCGGCCACGCACGACGAGCGA    |
|         | TCATCAGGCAGCCTTCGTATCGGAGAGTTCGATCTTCGCGCCAGCCCGGCCATCAGGTCCATGAA<br>CTCCGGGAAGCTCGTGGCGATCATCGTGGCATCGTCCACCGTGACAGGGTTTTTC  |
|         | CGGGAATGCGGACGGTTCCGGAAGGCCAGAGGATTTGCGGGCGGTTGCGGGCCGGCTGCTTGC<br>ACCGTGAAGCATGCAGGCTGTAGCCACTGATGCTGAAATCCTAAAGGAACAAAAC    |
|         | CGTCTCGCCCTCATCGCAATCCACGCCATTGAGCTTGAGGCCATTGGCGACGGCCGAGAGGCGGT<br>CGCTTTCTTGACGCGGAGTTCTTCCAGACCGTTCATCACGGTCGCCCTTCCGC    |
|         | CAGCGGGTTCAACACGCGGCCCATCGGGCGCTTTGTGAGCGAGGCGTCGCCGATGAAGGTGCTGT<br>CGAAATCGTAGACCCCGACGAGGCCCATGGTCAGGCGGCAGCCCGTGCGGGCATT  |
| EPSPS_2 | TCTTCACCTTCCAAAAGACCGGTGATACGGGTTTTACCGCTAGCGAGACCTCCAAACATGAAGGAC<br>CTGTGGGAGATAGACTTGTACCTGGAATACGGACGGTTCAGAAAAGACCAGAG   |
|         | CCTTCATCGCAATCAACACCGTTGAGCTTGAGACCGTTTTGCGACAGCAGAAAGACGGTCGCTTTCCT<br>TAACACGGAGTTCTTCCAAACCGTTCATAACGGTAGCACCTTCAGCGAATGCA |

|         |   |
|---------|---|
|         | GCAGCAACCAATGGGAAAGCAGTAGAGGATGGATCACCTGGAACATCAATCACTTGACCGGTGAGC<br>TTACCACGACCTTCAAGACGGATGGTACGCACACCGTCAGCATCAGTCTCAACG  |
|         | TTCAACACACGACCCATTGGACGCTTAGTGAGAGAAGCGTCACCAATGAAAGTGCTATCGAAATCGT<br>AAACACCAACAAGACCCATAGTCAAACGGCAACCAGTTGCAGCGTTACCGAAA  |
|         | ACCTCCAAACATGAAGGACCTGTGGGAGATAGACTTGTACCTGGAATACGGACGGTTCAGAAAAG<br>ACCAGAGGACTTACGAGCAGTTGCTGGACGGCTGCTTGCACCGTGAAGCATGCA   |
|         | TGAGCGGAAGCCATAGGTACCCTGTAGGTGATTGGCGTTGGAGTCTTTGGTCCACGCAAGGTA<br>GGAAGACGATCACCGTCTTCAGACTTACCTGCACACCCATTTTCGCGAAGTGGG     |
|         | GCTGCAACAGCGAGAATTGGATACTCGTCGATCATAGAAGGAGCACGGTCTTCTGGAACAGTAACA<br>CCCTTCAAAGTAGAAGAACGAACACGCAAGTCAGCCACGTCTTCTCCACCAGCA  |
|         | GTAAGGTTAGCACAAAACCTTGAAGCATCTTTTTCAGTGTGGTCACGAGTCATGATTGGCTCGATAA<br>CAGTGGTGATACCTGGGGTGTGAGACCAGCAAGCAGAACAGCGGACTTCACT   |
|         | TCGAGAGGAGCCTCAGGAGCAAGGAGTCCACCGTTACCAACACCATCAATGATCCAAGTATCACCT<br>TCCTTACGGATTCTGGCACCCATAGCTTGCATAGCCTTACCAGTGTGATAACA   |
|         | GCAGCCTTAGTGTGCGAGAGTTCGATCTTAGCTCCAAGACCAGCCATCAAATCCATGAACTCTGGG<br>AAGCTAGTAGCGATCATAGTAGCATCATCAACAGTAACAGGGTTTTTCAGAAACG |
|         | AGACCATAACGAGGAAGCTCATAGCGATACGGTGATCGAGGTGGGTAGCGACAGCTGCTCCAGA<br>AGCGTTACCGAGACCCTTACCGTCAGGACGACCACGCACGACGAGAGAAGTCTCA   |
|         | AGACGTGGGTTGATCACTTCGATGTCGGCACCCATTTCTGCAGAGTCAAGATGAGACCAGTACGG<br>GTTGGGTTTCATCAAACGTTAAGGATGGTGACGTCGGAACCTGGAACAAGCAAG   |
| 2mEPSPS | CGACGACCACAGGATGGCGATGGCTTTCTCCCTTGCCGCCTGTGCCGAGGTCCCCGTCACCATCC<br>GGGACCCTGGGTGCACCCGGAAGACCTTCCCCGACTACTTCGATGTGCTGAGCAC  |
|         | GGCCGGCGCCGAGGAGATCGTGCTGCAGCCCATCAAGGAGATCTCCGGCACCGTCAAGCTGCCG<br>GGGTCAAAGTCGCTTTCAAACCGGATCCTCCTACTCGCCGCCCTGTCCGAGGGGAC  |
|         | GCAGGGTGATGTGAAGTTTGCTGAGGTAAGTGGAGATGATGGGAGCGAAGGTTACATGGACCGAGA<br>CTAGCGTAACTGTTACTGGCCACCGCGGGAGCCATTTGGGAGGAAACACCTCAA  |
|         | CGTGCTTGATGGAGTACCAAGAATGAGGGAGAGACCCATTGGCGACTTGGTTGTCGGATTGAAGCA<br>GCTTGGTGACAGATGTTGATTGTTTCTTGGCACTGACTGCCACCTGTTTCGTGT  |
|         | ATACAAGTCCCCTAAAAATGCCTATGTTGAAGGTGATGCCTCAAGCGCAAGCTATTTCTTGGCTGGT<br>GCTGCAATTACTGGAGGGACTGTGACTGTGGAAGTTGTGGCACCACCAAGTTT  |
|         | GATGGCTTTCTCCCTTGCCGCCTGTGCCGAGGTCCCCGTCACCATCCGGGACCCTGGGTGCACCC<br>GGAAGACCTTCCCCGACTACTTCGATGTGCTGAGCACTTTCGTCAGAATTAAGC   |
|         | TGGTGGAAGTTCCAGTTGAGGATGCTAAAGAGGAAGTGCAGCTCTTCTTGGGGAATGCTGGAAT<br>CGCAATGCGGTCTTGACAGCAGCTGTTACTGCTGCTGGTGGAAATGCAACTTA     |
|         | CAATGGAATCGGAGGGCTACCTGGTGGCAAGTCAAGCTGTCTGGCTCCATCAGCAGTCAGTACTT<br>GAGTGCCTTGCTGATGGCTGCTCCTTGGCTCTTGGGGATGTGGAGATTGAAAT    |
|         | CATTGATAAATTAATCTCCATTCCGTACGTCGAAATGACATTGAGATTGATGGAGCGTTTTGGTGTGA<br>AAGCAGAGCATTCTGATAGCTGGGACAGATTCTACATTAAGGGAGGTCAAAA  |
|         | AACAGTGGTTGATAACCTGCTGAACAGTGAGGATGTCCACTACATGCTCGGGGCCTTGAGGACTCT<br>TGGTCTCTCTGTGCAAGCGGACAAAAGCTGCCAAAAGAGCTGTAGTTGTTGGCTG |

|           |  |
|-----------|--|
|           | CGAGAGGATGGTTGCGATCCGGACGGAGCTAACCAAGCTGGGAGCATCTGTTGAGGAAGGGCCGG<br>ACTACTGCATCATCACGCCGCCGAGAAGCTGAACGTGACGGCGATCGACACGTA    |
|           | GGCGATTGATGTCAACATGAACAAGATGCCTGATGTCGCCATGACTCTTGCTGTGGTTGCCCTCTTT<br>GCCGATGGCCCCGACAGCCATCAGAGACGTGGCTTCCTGGAGAGTAAAGGAGAC  |
| bar       | CTGCGGGCGGCCGGCTTCAAGCACGGGAACTGGCATGACGTGGGTTTCTGGCAGCTGGACTTCAG<br>CCTGCCGGTACCGCCCCGTCCGGTCTGCCGTACCCGAGATCTGAACGGAGTGC     |
|           | TACATCGAGACAAGCACGGTCAACTTCCGTACCGAGCCGCAGGAACCGCAGGAGTGGACGGACGA<br>CCTCGTCCGTCTGCGGGAGCGCTATCCCTGGCTCGTCGCCGAGGTGGACGGCGAG   |
|           | AGTGCGCGTGGGCATCGCCCCGAGTTGGAGCTGGTACGGGAACTCATCGAACTCAACTGGCATAACC<br>CGCAATGGTGAAGTGAACCGCGGGCGGATCGCGTACGACCGTGCCCAGGAGGCCT |
|           | ACCCACCTGCTGAAGTCCCTGGAGGCACAGGGCTTCAAGAGCGTGGTTCGCTGTCATCGGGCTGCC<br>CAACGACCCGAGCGTGCATGCACGAGGCGCTCGGATATGCCCCCGCGGCATG     |
|           | GCTCGCTGTCATTTTCGAGACGCCATCTTTGGAAGCGGTGGCCGAATCCGTACTGCGCGGACTCGA<br>CGACGCGTAAAACGATCGACCACGTACACGAGTCCGGACACGGGGCGAGGAGGC   |
|           | GTCGCCGGCATCGCTACGCGGGCCCCTGGAAGGCACGCAACGCCTACGACTGGACGGCCGAGT<br>CGACCGTGTACGTCTCCCCCGCCACCAGCGGACGGGACTGGGCTCCACGCTCTAC     |
|           | CCGGTTCCGGCACCGAGGAAGACCGAAGGAAGACCACACGTGAGCCCAGAACGACGCCCGGCCG<br>ACATCCGCCGTGCCACCGAGGCGGACATGCCGGCGGTCTGCACCATCGTCAACCAC   |
| Pat       | CATGTCTCCGGAGAGGAGACCAGTTGAGATTAGGCCAGCTACAGCAGCTGATATGGCCGCGTTTG<br>TGATATCGTTAACCATTACATTGAGACGTCTACAGTGAACCTTAGGACAGAGCC    |
|           | GTCTGTGGTTGCTGTTATAGGCCTTCCAAACGATCCATCTGTTAGGTTGCATGAGGCTTTGGGATAC<br>ACAGCCCGGGGTACATTGCGCGCAGCTGGATAACAAGCATGGTGGATGGCATGA  |
|           | GAACGCTTACGATTGGACAGTTGAGAGTACTGTTTACGTGTACATAGGCATCAAAGGTTGGGCCTA<br>GGATCCACATTGTACACACATTTGCTTAAGTCTATGGAGGCGCAAGTTTTAA     |
|           | ACAAACACCACAAGAGTGGATTGATGATCTAGAGAGGTTGCAAGATAGATACCCTTGGTTGGTTGCT<br>GAGGTTGAGGGTGTGTGGCTGGTATTGCTTACGCTGGGCCCTGGAAGGCTAG    |
|           | CAGCTGGATAACAAGCATGGTGGATGGCATGATGTTGGTTTTTGGCAAAGGGATTTGAGTTGCCAG<br>CTCCTCCAAGGCCAGTTAGGCCAGTTACCCAGATCTGAGTCGACCTGCAGGCAT   |
| Prsv_cp_1 | GCCAAGACTTTCTTTAGTCCGCTGATGGGACACTATATGAAAAGCGTTCTAAGCAAGGAAGCATACG<br>TTAAGGACTTATTGAAATACTCAAGTGACATAATTGTTGGTGAAGTTGACCAC   |
|           | TCATTCTCTGGTTCATTTGCTGAGCTTGGACTTAAGTACGATTTCACTCAAAGGCCACCGAAATAACA<br>AGATTTGTGGTTCATGTCACATCGAGGTATTCTGATCGATGACATCTATATC   |
|           | ACGCACCAGATCCGCAGTTCTATCAATGGGTGCTTGAGCAGGCTCCATTCAATGAATTAGCGAAG<br>CAGGGCAGGGCTCCATATGTGTCTGAGGTTGGATTGAGGCGCTTATACTAGT      |
|           | ATATTTTGGAGCGAGAAAGGGGAGATTGCGCTGAGCTACTGGTGTATTATGAATCGAGAAGCACTGAT<br>GATCATCAGTTAACTTGCGGCAGTAAAACACATGTGTTTCACCAGTCTAAGTA  |
|           | AATGGATCATTAAAGGCCGAGCTGCGACCCGCAGAAAAAGTTCTAGCCAAAAAGACAAGATCTTTCA<br>CTGCGGCGCCTCTTGACACTTTGCTGGGAGCCAAAGTTTGTGTTGATGATTTT   |
|           | GTTGGGGCTTTGTACAAAGGAAAGAAAAGAAAATATTTGAAAGGATCAACGTTGAGCACAGACATA<br>CTTTGGTGCGGAAAAGTTGTGAGCGTTTGTATGAAGGGCGAATGGGTGTGTGG    |

|           |   |
|-----------|---|
|           | AATAATTGGTTCTACAGCAAAAATATGGAGTGTCTTGGACTGTCGGCATGACAAAAGTTTATAAAGG<br>CTGGGACGAATTCTTGAGAAAATTTCTGATGGCTGGGTTTATTGTGATGCA    |
|           | GAAATTACATACACACCTATTTTGACACCTGACGGGACGATTGTCAAGAAATTCAAAGGAATAATAG<br>TGGTCAACCTTCAACAGTTGTCGATAACACACTGATGGTTTTAATCACAATG   |
|           | TATTACGCGCTACGCAAGGCTGGTTACGACACGAAGGCTCAGGAGGACATGTGTGTGTTTTATATCA<br>ATGGTGATGATCTCTGTATTGCCATCCACCCGATCATGAACATGTTCTTGAC   |
|           | GATGGATCTCAGTTTGATAGCTCTTTGACACCATATTTGCTTAATGCTGTGTTATCAATTCGACTGTG<br>GGCTATGGAAGATTGGGATATTGGAGCTCAGATGCTTAAGAATCTGTATGGA  |
|           | GATGTTTTTGAGGAAAGTGTGCACAAGTTGTCGAATTGCTAAACGACCACGAGTGCCTGAACTCG<br>AATACATCACAGATAGCGAGGTTATCATACAAGCTCTAAATATGGATGCTGCT    |
|           | ATGGATAAGTTACACGGCAATCTTAAGGGAGTGAGCTCAGCTTCTAGCAATTTAGTCACGAAGCATG<br>TTGTGAAAGGCATTTGTCCCCTCTCCGTAGTTACCTGGAGTGTAATGAGGAA   |
|           | AAGCGTGGGTCAATGGATGAATTGGAGGCCTACATAGATAAAATATTTTGAGCGAGAAAGGGGAGAT<br>TCGCCTGAGCTACTGGTGTATTATGAATCGAGAAGCACTGATGATCATCAGTTA |
|           | CCCAAACTTGAACCTGAGAGAATTGTGGCCATCCTTGAGTGGGATAAATCCAAGCTTCCAGAACATC<br>GATTGGAAGCTATCACAGCAGCAATGATAGAGTCATGGGGATATGAGGAGCTA  |
| Prsv_cp_2 | GGCGTCTAAAAATGAAGCTGTGGATACCGGTCTGAATGAGAAGCTCAAAGAAAAAGAAAAGCAGAA<br>AGAAAAAGAAAAAGATAAACAACAAGATAAAGACAATGATGGAGCTAGTGACGG  |
|           | GCACTCTCTCCTGGGTATGCGCAATTGAATACTCGCGCTAGTGTGTTTGTGGGCTGGCTCGACC<br>CTGTTTCACCTTATAATACTATGTAAGCATTAGAATATAGTGTGGCTGCGCCAC    |
|           | AAACGATGTGTCAACTAGCACAAAACTGGAGAGAGAGATAGGGATGTCAATGCCGGAAGTGTGG<br>AACCTTCACTGTTCCGAGGATAAAGTCATTTACTGATAAGATGATCTTACCAAG    |
|           | GAAGTACATGCCGCGGTATGGAATCAAGAGAAATTTGACTGACATTAGTCTCGCTAGATATGCTTTC<br>GATTTCTATGAGGTGAATTCGAAAACACCTGATAGGGCTCGTGAAGCTCATAT  |
|           | AAATGATTATGGCCTTAATGATAACGAAATGCAAGTAATGTTAAATGGTTTGATGGTTTGGTGTATCG<br>AAAATGGTACATCTCCAGATATATCTGGTGTCTGGGTTATGATGGATGGGGA  |
|           | GCTGCGCCACCGCTTCTATTTTACAGTGAGGGTAGCCCTCCGTGCTTTTGTGTTATTTCGAGTTCTC<br>TGAGTCTCCATACAGTGTGGGTGGCCACGTGCTATTTCGAGCCTCTTGGAAATG |
|           | GCAGATGAAGGCTGCAGCGCTACGCAATACTAATCGCAAAATGTTTGGAAATGGACGGCAGTGTGAG<br>TAACAAGGAAGAAAACACGGAGAGACACACAGTGAAGATGTCAACAGAGACAT  |
|           | AATTAAGGGAAAAACTGTCCTTAATTTAAATCATCTTCTTCAGTATAATCCGAAACAAGTTGACATCTC<br>AAACTTCGCGCCACTCAATCTCAATTTGAGAAGTGGTATGAGGGAGTGAG   |
|           | AACCCAAGTCGATTATCCCATTAACCTTTGATTGAACACGCAACTCCTTCATTTAGGCAAATCATGG<br>CTCACTTCAGTAACGCGGCAGAGGCATACATCGCGAAGAGGAATGCAACTGA   |
| Hsp70     | GAAATTTTGATGTTTATCTCTGCTCCTTTATTGTGACCATAAGTCAAGATCAGATGCACTTGTTTTAAA<br>TATTGTTGTCTGAAGAAATAAGTACTGACAGTATTTTGATGCATTGATCTG  |
|           | ACTGACACTATATTGCTTCTCTTTACATACGTATCTTGCTCGATGCCTTCTCCCTAGTGTGACCAGT<br>GTTACTCACATAGTCTTTGCTCATTTCATTGTAATGCAGATACCAAGCGGCC   |
|           | GCGGTATTTATTTAAGCACATGTTGGTGTATAGGGCACTTGGATTGAGAAGTTTGTGTTAATTTAG<br>GCACAGGCTTCATACTACATGGGTCAATAGTATAGGGATTATATTATAGGC     |

CGTCTTCGGTACGCGCTCACTCCGCCCTCTGCCTTTGTTACTGCCACGTTTCTCTGAATGCTCTCTT  
GTGTGGTGATTGCTGAGAGTGGTTTAGCTGGATCTAGAATTACACTCTGAAAT

GATACTATAATAATTTGTTTCGTCTGCAGAGCTTATTATTTGCCAAAATTAGATATTCCTATTCTGTTTT  
GTTTGTGTGCTGTAAATTGTTAACGCCTGAAGGAATAAATATAAATGAC

CTTGTTTGTTGTAACAAAATTTAAAAATAAAGAGTTTCCTTTTTGTTGCTCTCCTTACCTCCTGATGGT  
ATCTAGTATCTACCAACTGACACTATATTGCTTCTCTTACATACGTATCT

CGTGTTCTGCCTGTGCTGATTACTTGCCGTCCTTTGTAGCAGCAAAATATAGGGACATGGTAGTACG  
AAACGAAGATAGAACCTACACAGCAATACGAGAAATGTGTAATTTGGTGCTTA



# Detection and identification of transgenic events by next generation sequencing combined with enrichment technologies

Frédéric Debode, Julie Hulin, Benoît Charlotheaux, Wouter Coppieters, Marc Hanikenne, Latifa Karim and Gilbert Berben

## Supplementary Data S2. Commands and parameters used for the bioinformatic workflow

```
# RMQ : REF_GENOME = database used for enrichment, preformatted with the command "bwa index REF_GENOME"

#### First part of the analysis : Analysis by counts
# Perform alignments
bwa mem -M -t 8 REF_GENOME File_sample01_R1.fastq.gz File_sample01_R2.fastq.gz | \
samtools view -bS - | samtools sort -o RESULT_BWA_sample01.sorted.bam
# Index the sequences
samtools index RESULT_BWA_sample01.sorted.bam
# Make count file
samtools idxstats RESULT_BWA_sample01.sorted.bam > RESULT_BWA_sample01.sorted.bam.counts
# The count file is processed with R, to convert raw counts in RPKM values and calculate the LOD

#### Second part of the analysis : make contigs
# Extract only sequences aligned against sequences present in the database
# (all the header of these sequences must be provided in the file "LIST_SEQ_DB")
samtools view RESULT_BWA_sample01.sorted.bam LIST_SEQ_DB | cut -f1 | sort | uniq >SAMPLE_01.id
seqtk subseq File_sample01_R1.fastq.gz SAMPLE_01.id >SAMPLE_01.R1.unic.fastq
seqtk subseq File_sample01_R2.fastq.gz SAMPLE_01.id >SAMPLE_01.R2.unic.fastq

# In fastq files, sequences identifiers are modified with a perl oneliner to distinguish reads R1 and R2
perl -i.bak -nle '~ s/(\^[M].+)/$1\1/; print' SAMPLE_01.R1.unic.fastq
perl -i.bak -nle '~ s/(\^[M].+)/$1\2/; print' SAMPLE_01.R2.unic.fastq

# Contigs are assembled with two programs in parallel
# First, assembly with SPADES
spades.py --pe1-1 SAMPLE_01.R1.unic.fastq --pe1-2 SAMPLE_01.R2.unic.fastq -o spades/

# Second, assembly with MIRA
# The program needs to write parameters in a "manifest" file.
##### EXAMPLE OF MANIFEST FILE #####
project = Enrichment
job = est,denovo,accurate
readgroup = SAMPLE_NAME
data = SAMPLE_01.R1.unic.fastq SAMPLE_01.R2.unic.fastq
technology = solexa
template_size = 0 350 autorefine
segment_placement = ---> <---
segment_naming = solexa
parameters = -GENERAL :number_of_threads= 8 \
              SOLEXA_SETTINGS \
##### END OF EXAMPLE MANIFEST FILE #####
# run mira
mira manifest

# Third, integrate all contigs obtained with CISA
# Firstly use merge.py ; to merge all contigs
# Need to create "merge.config" file, with parameters
##### EXAMPLE OF MERGE.CONFIG FILE #####
count=2
data=mira_contigs.fa,title=mira
data=spades_contigs.fa,title=spades
min_length=100
Master_file=SAMPLE01_CISA.fa
##### END EXAMPLE OF MERGE.CONFIG FILE #####

# Run first step of cisa
python Merge.py merge.config

# The terminal shows important information for the following step.
# Note the largest "Whole" value. For our example, we use 123456 as being this largest "Whole" value
# Secondly for CISA
# Need to create "cisa.config" file
##### EXAMPLE OF CISA.CONFIG FILE #####
genome=123456
infile=SAMPLE01_CISA.fa
outfile=SAMPLE01_CONTIG_CISA.fa
R2_Gap=0.95
CISA=<PATH_TO_CISA_DIRECTORY>
makeblastdb=/usr/bin/makeblastdb
blastn=/usr/bin/blastn
nucmer=/usr/bin/nucmer
##### END EXAMPLE OF CISA.CONFIG FILE #####
# Run second step of cisa
python cisa.py cisa.config
```