(A)

CTGCCAGGGAATCATCCTGAGGTACGTGCGGTCTGTGTGCTTAGGACAGACAAGTTGTACAACACTACCCGTTTTGTATAGCTAAGCT

CATAAAAAGCAAAGCATCAGGCAACGGCGCCTATCAGTCGTTCCTTTTTCATCTACGGTCAATGGGCCGCCCACATCGAGAGGCAAGG

ATGACGACCATCGCCCCCCCGCACAGCGCTGGCGCTGATCACGATCAGCCACTGGCCGTCGCACATGCGGCCGTAGAAGCCCGGCTG

MetThrThrIleAlaProProHisSerAlaGlyAlaAspHisAspGlnProLeuAlaValAlaHisAlaAlaValGluAlaArgLeu

*crtE* →

GCGGCATTGCTGCCATCGGCAGGTCACGAGCGTGACCTAGTGGTTGCGGCCATGCGCGACTGCACCCTGGCGCCAGGCAAGCGCTTG

AlaAlaLeuLeuProSerAlaGlyHisGluArgAspLeuValValAlaAlaMetArgAspCysThrLeuAlaProGlyLysArgLeu

CGGCCCTTGCTGTTACTCTTCACGGCACAAGGCCTGGGCGCGTGCCCACACGCCGCGCTCGACCTTGGCTGCGCCGTGGAGATGGTT

ArgProLeuLeuLeuLeuPheThrAlaGlnGlyLeuGlyAlaCysProHisAlaAlaLeuAspLeuGlyCysAlaValGluMetVal

CATGCAGCCTCCCTTGTACTGGACGACCTCCCCTGTATGGACAACGCCGCCTTGCGCCGCGGTCGCCCTACCTTGCACCTGGCGTTT

HisAlaAlaSerLeuValLeuAspAspLeuProCysMetAspAsnAlaAlaLeuArgArgGlyArgProThrLeuHisLeuAlaPhe

GGGGAAGATGTGGCAGTGTTGACGGCCATTGCCTTGCTCTCCCGTGCATTCGGGGTTGTGGCGGGGTTGCAGACGGTCGGGCCGGAT

GlyGluAspValAlaValLeuThrAlaIleAlaLeuLeuSerArgAlaPheGlyValValAlaGlyLeuGlnThrValGlyProAsp

GTGCGCAGCGACCTGGTCGTGACCCTCACCTCGGCAATCGGTGTCGATGGGCTGGTCAAGGGGCAGTTGCAGGATCTGCGCGACGGT

ValArgSerAspLeuValValThrLeuThrSerAlaIleGlyValAspGlyLeuValLysGlyGlnLeuGlnAspLeuArgAspGly

GCACGGCCGCGCACGGCGGCTGAAATTGCCGAGACCAATCAACTCAAGACCGGGGCCTTGTTCTCGGCCTTGCTGGGTATGGCCGGC

AlaArgProArgThrAlaAlaGluIleAlaGluThrAsnGlnLeuLysThrGlyAlaLeuPheSerAlaLeuLeuGlyMetAlaGly

CGGTTGGCGGGGGCTGGCGAAACCCAGCAGGGCACCTTGCAGCAGTTGGCGGGGGAATTGGGGCAAGCGTTTCAATTGCACGATGAT

ArgLeuAlaGlyAlaGlyGluThrGlnGlnGlyThrLeuGlnGlnLeuAlaGlyGluLeuGlyGlnAlaPheGlnLeuHisAspAsp

TTGCATGACCGTGACCCTGAGAGCGGCAAGAGCATGGGGTTGGACCAAGGCAAGTCGACCCTGCTGGCCATGTGTGGCGAGATGCAG

LeuHisAspArgAspProGluSerGlyLysSerMetGlyLeuAspGlnGlyLysSerThrLeuLeuAlaMetCysGlyGluMetGln

GTGCGCCAGCGGTTGGAAGCCCATTTGCAGAGCGTCGAACAGTGCCTGACCAGCCTGGGCCTTCAGGCCTCACCGATTGGCGGGGTC

ValArgGlnArgLeuGluAlaHisLeuGlnSerValGluGlnCysLeuThrSerLeuGlyLeuGlnAlaSerProIleGlyGlyVal

TTGCGAAAGGTGTTCGCTGCGAAGGGTGTGCCGGCTGGCGGCAGCACCGTATGGCGTCAAGGCCTGAGCGGGCTTGCGGTGGGTGGG

LeuArgLysValPheAlaAlaLysGlyValProAlaGlyGlySerThrValTrpArgGlnGlyLeuSerGlyLeuAlaValGlyGly

GCGGTCATGCATCACGTCTAA

AlaValMetHisHisVal\*\*\*

(B)

ATGTGGGCGGCCCATTGACCGTAGATGAAAAAGGAACGACTGATAGGCGCCGTTGCCTGATGCTTTGCTTTTTATGAGCTTAGCTAT

ACAAAACGGGTAGTGTTGTACAACTTGTCTGTCCTAAGCACACAGACCGCACGTACCTCAGGATGATTCCCTGGCAGTGGAGCCGGC

ATGAACAAAAGCCCGTTGAGTCAACGCAAGGATGATCACCTCGATATCGTTCTGGCCGCCGGCACCATCTCGCATGGGGTCGATGCT

MetAsnLysSerProLeuSerGlnArgLysAspAspHisLeuAspIleValLeuAlaAlaGlyThrIleSerHisGlyValAspAla

*idi* →

GGCTGGTCGGCGCTGCGTTTCGAACATTGTGCACTGCCTGAGATGGCCCTGGAAGACGTCAACCTGGGTACCGAGCTGTTCGGCATG

GlyTrpSerAlaLeuArgPheGluHisCysAlaLeuProGluMetAlaLeuGluAspValAsnLeuGlyThrGluLeuPheGlyMet

GCACTGCGCGCACCACTGCTGATCAGTTCGATGACCGGCGGTGCTGAGCGCGCACGGTTCATCAACCAGCACCTGGCCGAAGCTGCG

AlaLeuArgAlaProLeuLeuIleSerSerMetThrGlyGlyAlaGluArgAlaArgPheIleAsnGlnHisLeuAlaGluAlaAla

CAGGCTTTGGGCCTGGCCATGGGCGTGGGTTCGCAGCGAGTTAGCTTGCGCAGCGGCCAGGACCAGGGGCTGACCACGGAGCTGAGA

GlnAlaLeuGlyLeuAlaMetGlyValGlySerGlnArgValSerLeuArgSerGlyGlnAspGlnGlyLeuThrThrGluLeuArg

AGGCTCGCACCCGACGTCCCGCTCCTGGCAAACCTGGGCGCCGCCCAGTTACGCGAGGCTGATGGTGTAGACCTTGCCCGCCGTGCG

ArgLeuAlaProAspValProLeuLeuAlaAsnLeuGlyAlaAlaGlnLeuArgGluAlaAspGlyValAspLeuAlaArgArgAla

GTCGATACACTCGATGCCAGCGCCCTGATCATTCATCTCAACCCCTTGCAGGAAGCCGTTCAGCCCGAGGGGGACCGCAACTGGAAC

ValAspThrLeuAspAlaSerAlaLeuIleIleHisLeuAsnProLeuGlnGluAlaValGlnProGluGlyAspArgAsnTrpAsn

GGGGTGCTCGCCTGCATCGAGCACGTCGCCGGGCAGCTTCACGCGCCCATCATCGTCAAGGAAGTAGGCGCCGGTATCTCCCCAGGC

GlyValLeuAlaCysIleGluHisValAlaGlyGlnLeuHisAlaProIleIleValLysGluValGlyAlaGlyIleSerProGly

GTGGCCCGGCAATTGTACGACGCTGGCGTGTCCATGATCGATGTGGCCGGCGCCGGCGGCACGAGCTGGGCCGCTGTCGAGGCGGAG

ValAlaArgGlnLeuTyrAspAlaGlyValSerMetIleAspValAlaGlyAlaGlyGlyThrSerTrpAlaAlaValGluAlaGlu

CGCGCCAATGACCCTGCTGAGCGTGCGGTGGCAATGGCCTTCAGCGGCTGGGGCATTCCGACCGCCCAGAGCCTGCTTGCCGTGCGC

ArgAlaAsnAspProAlaGluArgAlaValAlaMetAlaPheSerGlyTrpGlyIleProThrAlaGlnSerLeuLeuAlaValArg

CGAGCCGTGCCCACGGTTGGCCTGATCGCCTCGGGTGGGGTGCGCGATGGGGTGGACGCGGCCAAGGCCATCCGCCTGGGGGCAGAC

ArgAlaValProThrValGlyLeuIleAlaSerGlyGlyValArgAspGlyValAspAlaAlaLysAlaIleArgLeuGlyAlaAsp

TTGGTGGGCCAGGCTGCCGGCGTGCTGACCTCTGCGATGACGTCTACCGAAGCGGTGGTCTCACATTTCGACGTGATCATCCGGCAA

LeuValGlyGlnAlaAlaGlyValLeuThrSerAlaMetThrSerThrGluAlaValValSerHisPheAspValIleIleArgGln

CTGCGTATCGCCTGCTTCTGCACCGGCTCCGCCACGCTGGCGCACTTGCGCCGGGCCCGACTGGTCAATGCCCAGCTCGACCCC

LeuArgIleAlaCysPheCysThrGlySerAlaThrLeuAlaHisLeuArgArgAlaArgLeuValAsnAlaGlnLeuAspPro

AAATGAGCCACTTCGCCCTGATTGCCCCACCCTACCCCAGCCATTTTGCAGCCATGCAGGCGCTGGGCGGCGAGTTGGTCGAACGC

Lys\*\*\*

--MetSerHisPheAlaLeuIleAlaProProTyrProSerHisPheAlaAlaMetGlnAlaLeuGlyGlyGluLeuValGluArg

*crtX* →

GGCCATCAGGTCAGTTTCTTTCACCAGCCCGAAGCCGAACGCTGGCTGAGTGACTCGCGGCTGGGCTTCGTCGCCATTGGCCATGCC

GlyHisGlnValSerPhePheHisGlnProGluAlaGluArgTrpLeuSerAspSerArgLeuGlyPheValAlaIleGlyHisAla

GAGTGCCCGCCGGGCAGCCTTGACCTCGCCATGCGTCGCGTGGCCGACCCCGCCGGGCCGCTGCGTCTGCGTCGCTTGATCAAGCAG

GluCysProProGlySerLeuAspLeuAlaMetArgArgValAlaAspProAlaGlyProLeuArgLeuArgArgLeuIleLysGln

TTGGCGCACACCACCGGCATGCTGATCAGCAACCTGCCGCAGGCCCTTAGGGACCATCGGGTCGACGCGATGATCTGCGACCAGATG

LeuAlaHisThrThrGlyMetLeuIleSerAsnLeuProGlnAlaLeuArgAspHisArgValAspAlaMetIleCysAspGlnMet

GAACCCGCAGGCGGCTTGGTGGCCGAGGCTTGCAACCTGCCGTTCTTTTCGGTGGCCTGCGCGTTGCCGATCAATCGCGAACCGGGC

GluProAlaGlyGlyLeuValAlaGluAlaCysAsnLeuProPhePheSerValAlaCysAlaLeuProIleAsnArgGluProGly

TTGCCATTGCCCGTGATGCCCTTCGGGCTCGGCTTGGATGCGCGCCTCTACGAAGGCAGTGCTCAAGTCCATGACTGGCTCATGCGC

LeuProLeuProValMetProPheGlyLeuGlyLeuAspAlaArgLeuTyrGluGlySerAlaGlnValHisAspTrpLeuMetArg

CCGCTGCACGACACCTTAGTGCAGGCGTGTCATCGCCACGGGCTGCCACCGCGCAGCGGTGCCCATGAGTTTCTCTCGCCGCTGGCC

ProLeuHisAspThrLeuValGlnAlaCysHisArgHisGlyLeuProProArgSerGlyAlaHisGluPheLeuSerProLeuAla

CAGTTGAGCCAGACACCTCTGCATCTGGATTTCCCAAGAGCGGATCTACCCAAGCACTTTCACGCCGTCGGCCCTTTACGCAGCACG

GlnLeuSerGlnThrProLeuHisLeuAspPheProArgAlaAspLeuProLysHisPheHisAlaValGlyProLeuArgSerThr

CCCTCCCAGGCGAGCCCCAACTGGTCTTTCGAATCCGGCAAGCCTTTGGTATTCGCAAGCCTGGGAACGCTGCAGGGCCACCGGTTC

ProSerGlnAlaSerProAsnTrpSerPheGluSerGlyLysProLeuValPheAlaSerLeuGlyThrLeuGlnGlyHisArgPhe

GATCTGTTTGCACGGATTGCCAGGGCTTGCCGGGGCCTGGATGCCCAGTTGCTGGTAGCTCACTGCGGTGGGTTGGACGCGCATCAG

AspLeuPheAlaArgIleAlaArgAlaCysArgGlyLeuAspAlaGlnLeuLeuValAlaHisCysGlyGlyLeuAspAlaHisGln

CAGGCACGGCTACGCCAGATCGGTGCCACCGTCGTGACCGACTTCGCCCCCCAGCACTGGGCTGTGCGACGAGCCAATGTGGTGATC

GlnAlaArgLeuArgGlnIleGlyAlaThrValValThrAspPheAlaProGlnHisTrpAlaValArgArgAlaAsnValValIle

AGCCATGGCGGCTTGAACACGGTGCTCGATACCGTGTGCGCCGGCACAGCGCTGCTGGTCATGCCCATTGCATTCGATCAGCCCGGC

SerHisGlyGlyLeuAsnThrValLeuAspThrValCysAlaGlyThrAlaLeuLeuValMetProIleAlaPheAspGlnProGly

GTGGCCGCACGGGTGCGCCATCACCGCTTGGGCCGGGTGCTCTCCAAACGTGCCAGCGCATCGGCCATTGGCCAAGCGTTGCGTCAG

ValAlaAlaArgValArgHisHisArgLeuGlyArgValLeuSerLysArgAlaSerAlaSerAlaIleGlyGlnAlaLeuArgGln

CTGCTTGATGAGGCTGCACCACGTGATGCGGCGCTTGAAGCCGAGCTTGCCACCTGCGGTGGCGTACAGCGCGCTGCCAGCCTCATC

LeuLeuAspGluAlaAlaProArgAspAlaAlaLeuGluAlaGluLeuAlaThrCysGlyGlyValGlnArgAlaAlaSerLeuIle

GAAGCCGCTCTGCGCAGTGGCCAGCCCAGCATTTCGGAGAATGTACCGTGCCCTATGACCTGATCCTGGTAGGTGGGGGGTTGGCC

GluAlaAlaLeuArgSerGlyGlnProSerIleSerGluAsnValProCysProMetThr\*\*\*

-----------------------------------------------ValProTyrAspLeuIleLeuValGlyGlyGlyLeuAla

*crtY* →

AACGGCCTGATCGCCTGGCGACTGCAGCAGCTGCGCCCGCAATGGCGCGTGTTGTGCCTGGAGCAAGGGCCGGTCCTGGGCGGCAAC

AsnGlyLeuIleAlaTrpArgLeuGlnGlnLeuArgProGlnTrpArgValLeuCysLeuGluGlnGlyProValLeuGlyGlyAsn

CATACGTGGTCCTTCCACGACGCCGACCTTACTGCCGAGCAACTGCGCTGGTTGGCCCCGCTGGTGGAGCAGCGCTGGCCTGCGTAC

HisThrTrpSerPheHisAspAlaAspLeuThrAlaGluGlnLeuArgTrpLeuAlaProLeuValGluGlnArgTrpProAlaTyr

CACGTGCACTTCCCGGCTTTTTCCAGGCGCATGGAAGGCGGCTATGCGAGCATTACCAGCGAGCGTTTCGCGCAGGTGATCACCCCA

HisValHisPheProAlaPheSerArgArgMetGluGlyGlyTyrAlaSerIleThrSerGluArgPheAlaGlnValIleThrPro

GCCCTCGGCCCGAGCTTGCGCACCGACTGCGCGGTCGCCGAAGTGGGTGCCAGCCATGTCACGCTCGCCAGTGGCGAAACCTTGCAC

AlaLeuGlyProSerLeuArgThrAspCysAlaValAlaGluValGlyAlaSerHisValThrLeuAlaSerGlyGluThrLeuHis

GCCGCGGTGGTGATCGATGGCCGGGGCATGCAGCACAGCCCTCACCGGGTCCTCGGCCAGCAGGCGTTCCTCGGGCAGGTCGTACGT

AlaAlaValValIleAspGlyArgGlyMetGlnHisSerProHisArgValLeuGlyGlnGlnAlaPheLeuGlyGlnValValArg

CTGACGGCCCCCCATGGCTTGGTCGCACCGATTCTGATGGACGCACGCGTACCTCAAGGCGAGGGCTATCGCTTCGTCTATGTCCTG

LeuThrAlaProHisGlyLeuValAlaProIleLeuMetAspAlaArgValProGlnGlyGluGlyTyrArgPheValTyrValLeu

CCGTTCACGGCTGACACCTTGCTGGTGGAGGACACCCATTACGTCGACAGCCAGGCGCCCGCCGCGCAGCAGCTGCGCCAGCACATC

ProPheThrAlaAspThrLeuLeuValGluAspThrHisTyrValAspSerGlnAlaProAlaAlaGlnGlnLeuArgGlnHisIle

GAGTGCTACCTGCAGCAGCGCGGCTGGCACGTCGCCGAGTGCCTGCGTGAAGAACAGGGCGTATTGCCGATTACCCTGGCCGACGAC

GluCysTyrLeuGlnGlnArgGlyTrpHisValAlaGluCysLeuArgGluGluGlnGlyValLeuProIleThrLeuAlaAspAsp

TTTCCTGCCTGGTGGAACGAACAGCACGGGCCGCTCGCCGTTGGCCTGCGGGCCGGATTGTTTCACTGCACCACGGGTTATTCCCTG

PheProAlaTrpTrpAsnGluGlnHisGlyProLeuAlaValGlyLeuArgAlaGlyLeuPheHisCysThrThrGlyTyrSerLeu

CCCGATGCCGTGCGCCTGGCCGACTGGTTGTCAACGCAAGCGCTGGAGGACACGGCGTCGCTCAAGCCCGCGTTGGGCGCCTACCTG

ProAspAlaValArgLeuAlaAspTrpLeuSerThrGlnAlaLeuGluAspThrAlaSerLeuLysProAlaLeuGlyAlaTyrLeu

CATCGACGCTGGCGTGAGCAAGGGTTCTATCGCCTGCTCAACCGAATGCTGTTTCTGGCCGGCCGGCCGAACCGTCGCTGGCAGGTG

HisArgArgTrpArgGluGlnGlyPheTyrArgLeuLeuAsnArgMetLeuPheLeuAlaGlyArgProAsnArgArgTrpGlnVal

ATGCAGCGCTTTTATCGATTGCCTGAGGGGCTGATCGCCCGTTTCTACGCAGGCCAGAGTACGCTGCGTGATCGCCTGCGGATCCTG

MetGlnArgPheTyrArgLeuProGluGlyLeuIleAlaArgPheTyrAlaGlyGlnSerThrLeuArgAspArgLeuArgIleLeu

TGCGGCAAACCGCCTGTGCCCGTTGGCGAGGCCATGAAGGCTGCCTTGCGACACTCCCCTGCCCATTTCAAGGATTATCCATGAGC

CysGlyLysProProValProValGlyGluAlaMetLysAlaAlaLeuArgHisSerProAlaHisPheLysAspTyrPro\*\*\*

--------------------------------------------------------------------------------MetSer

*crtI* →

AACGAACGATCCGCCCTGGTCATCGGCGCAGGCCTTGGCGGCCTGGCCCTGGCCATCCGCCTGCAGGCCGCGGGTATCCCCACGACC

AsnGluArgSerAlaLeuValIleGlyAlaGlyLeuGlyGlyLeuAlaLeuAlaIleArgLeuGlnAlaAlaGlyIleProThrThr

TTGCTCGAGGCGCGCGACAAACCAGGTGGCCGCGCTTACGTCTATCACGACCAAGGCTTCACCTTCGACGCAGGCCCCACGGTGATC

LeuLeuGluAlaArgAspLysProGlyGlyArgAlaTyrValTyrHisAspGlnGlyPheThrPheAspAlaGlyProThrValIle

ACCGACCCGAGCGCCATCGAAGCCCTGTTCACCCTCGCCGGCCGGCCCATGCATGAGTACGTGGAATTGCTGCCGGTAACGCCTTTC

ThrAspProSerAlaIleGluAlaLeuPheThrLeuAlaGlyArgProMetHisGluTyrValGluLeuLeuProValThrProPhe

TACCGCCTGTGCTGGGAGGATGGCCGGGCATTCGACTATGGCAATGACCAAGCCGACCTGGAGCGCCAGATCGAGGCCTTCGAGCCC

TyrArgLeuCysTrpGluAspGlyArgAlaPheAspTyrGlyAsnAspGlnAlaAspLeuGluArgGlnIleGluAlaPheGluPro

AAGGACGTGGAAGGCTATCGACAGTTCCTGGCCTATTCCAAGGCCGTGTTGCAGGAAGGCTATATCAAGCTTGGCGCAGTGCCTTTC

LysAspValGluGlyTyrArgGlnPheLeuAlaTyrSerLysAlaValLeuGlnGluGlyTyrIleLysLeuGlyAlaValProPhe

CTGTCTTTTCGCGACATGCTGCATGCCGGCCCGCAGTTGGCTCGCCTGCAAGCCTGGCGCAGCGTCTACAGCATGGTGGCGCGGTTC

LeuSerPheArgAspMetLeuHisAlaGlyProGlnLeuAlaArgLeuGlnAlaTrpArgSerValTyrSerMetValAlaArgPhe

ATCAAAGATGACCAACTGCGCCAGGCGTTCTCGTTCCACGCCTTGCTGGTCGGGGGTAATCCCTTCGCCACGTCCTCGATCTACACG

IleLysAspAspGlnLeuArgGlnAlaPheSerPheHisAlaLeuLeuValGlyGlyAsnProPheAlaThrSerSerIleTyrThr

CTGATCCATGCGCTGGAGCGCACCTGGGGTGTCTGGTTCCCCAAAGGCGGCACGGGCGCGCTGGTGCAGGGCATGGCCAAGTTGTTC

LeuIleHisAlaLeuGluArgThrTrpGlyValTrpPheProLysGlyGlyThrGlyAlaLeuValGlnGlyMetAlaLysLeuPhe

ACCGACCTGGGCGGCACCCTGGAGCTCAATGCCGAGGTGGCCGCGTTCGAAACCCAAGGCAGCCGCGTGACCGGCGTGCGCACCCAG

ThrAspLeuGlyGlyThrLeuGluLeuAsnAlaGluValAlaAlaPheGluThrGlnGlySerArgValThrGlyValArgThrGln

GACGGGCGGCGCTGGCGGGCAGACGCAGTCGCCTCGAATGCCGATGTGGTGCACACCTACTCCGACCTGCTGGCCGAACACCCTCGC

AspGlyArgArgTrpArgAlaAspAlaValAlaSerAsnAlaAspValValHisThrTyrSerAspLeuLeuAlaGluHisProArg

GGTGAACAGGAAGGCAAGCGGCTCAAGGGCAAGCGCTTTAGCAACTCGCTGTTCGTGGTGCACTTCGGGCTCAAGCGCACGCAGCCG

GlyGluGlnGluGlyLysArgLeuLysGlyLysArgPheSerAsnSerLeuPheValValHisPheGlyLeuLysArgThrGlnPro

CAACTGCGTCATCACACGGTGTGCTTTGGCCCGCGCTACCGGGAGCTGATTGACGAAATCTTCAACGGTCACGAGCTTGCCGAAGAC

GlnLeuArgHisHisThrValCysPheGlyProArgTyrArgGluLeuIleAspGluIlePheAsnGlyHisGluLeuAlaGluAsp

TTTTCACTCTACCTGCATACCCCCTGCGCCACCGACCCTAGCCTGGCCCCGGAAGGATGCTCGAGCCACTACGTATTGGCGCCCGTA

PheSerLeuTyrLeuHisThrProCysAlaThrAspProSerLeuAlaProGluGlyCysSerSerHisTyrValLeuAlaProVal

CCGCACATGGGCAACGCGGCCATCGACTGGGAGGTGGAAGGGCCACGGTTCCGGGACAAGATTCTCGACTACCTGGAACGTCACTAC

ProHisMetGlyAsnAlaAlaIleAspTrpGluValGluGlyProArgPheArgAspLysIleLeuAspTyrLeuGluArgHisTyr

ATTCCAGGCCTGCGCGAGGACTTGGTCACCTGCCGCACCTTCACGCCGCTGGACTTCCGCGAAGAACTTAATGCCCACCTGGGGTCG

IleProGlyLeuArgGluAspLeuValThrCysArgThrPheThrProLeuAspPheArgGluGluLeuAsnAlaHisLeuGlySer

GCGTTCTCTCTGGAACCCATCCTGACCCAGAGCGCCTGGTTCCGCCCGCACAACCGTGACAGCGAAGTGAGCAATCTGTACCTGGTG

AlaPheSerLeuGluProIleLeuThrGlnSerAlaTrpPheArgProHisAsnArgAspSerGluValSerAsnLeuTyrLeuVal

GGGGCAGGCACGCACCCAGGGGCCGGTGTGCCAGGCGTGATCGCCTCGGCCAAGGCCACCGCAGGCCTGATGCTCGAGGACCTGGGC

GlyAlaGlyThrHisProGlyAlaGlyValProGlyValIleAlaSerAlaLysAlaThrAlaGlyLeuMetLeuGluAspLeuGly

GCATGACCGCACCGAACCACGACCAGTTGCTTGCCCGTGCCGAACACAGCATTGCCGTGGGTTCGAAAAGCTTTGCCGCAGCAGCC

Ala\*\*\*

--MetThrAlaProAsnHisAspGlnLeuLeuAlaArgAlaGluHisSerIleAlaValGlySerLysSerPheAlaAlaAlaAla

*crtB* →

CGGCTGTTCGATGCCCATACCCGGCGCAGCGCTGTGCTGTTGTATGCCTGGTGCCGCCATTGCGATGACGTGATCGATGGTCAGCAA

ArgLeuPheAspAlaHisThrArgArgSerAlaValLeuLeuTyrAlaTrpCysArgHisCysAspAspValIleAspGlyGlnGln

GCCGGGCATGAGGCGGTGGTCATCAGCCAGGAAGAAGCGCAGGTGCGCCTGTTGGATCTGCAGCGTCAGACGGAGGCCGCCTGCGCG

AlaGlyHisGluAlaValValIleSerGlnGluGluAlaGlnValArgLeuLeuAspLeuGlnArgGlnThrGluAlaAlaCysAla

GACTTGCCCGTGAACCAGCCGGCCTTCGCCGCACTGCGCGAAGTGGCCAGGCGCCACGGCATTCCCCGCGCCGAAGCCTTCGACTTG

AspLeuProValAsnGlnProAlaPheAlaAlaLeuArgGluValAlaArgArgHisGlyIleProArgAlaGluAlaPheAspLeu

CTCGAAGGCTTCGCCCAGGACGTGTCCGCCTGCCGCTACCAGACCCTGGATGAGGTCATGGTCTACTGCTATCACGTAGCAGGTGTG

LeuGluGlyPheAlaGlnAspValSerAlaCysArgTyrGlnThrLeuAspGluValMetValTyrCysTyrHisValAlaGlyVal

GTGGGCCTGATGATGGCCAGGGTCATGGGTGTTCAAGATGCCCGCGTGCTGGACCGGGCCTGCGACCTGGGCCTGGCCATGCAGCTG

ValGlyLeuMetMetAlaArgValMetGlyValGlnAspAlaArgValLeuAspArgAlaCysAspLeuGlyLeuAlaMetGlnLeu

ACTAACATCGCCCGGGACATCGTCGATGATGCCAAGGTCGGGCGCTGCTACTTGCCTCAGGCTTGGCTGTCGGAATTCGACCTCAGT

ThrAsnIleAlaArgAspIleValAspAspAlaLysValGlyArgCysTyrLeuProGlnAlaTrpLeuSerGluPheAspLeuSer

GCAGACGACCTGGCCGATGTGCGTCATCGCGCAGCCGTCGCTACCTTGGCCAGGCGCCTGGTCGCCCACGCCGAACCCTACTACGCC

AlaAspAspLeuAlaAspValArgHisArgAlaAlaValAlaThrLeuAlaArgArgLeuValAlaHisAlaGluProTyrTyrAla

AGCGCCCAGGCCGGGCTGCAGGCCTTGCCGCTTCGCTCGGCGTGGGCAGTGGCCACGGCAGGCCAGGTTTACCGGCGCATCGGCCTG

SerAlaGlnAlaGlyLeuGlnAlaLeuProLeuArgSerAlaTrpAlaValAlaThrAlaGlyGlnValTyrArgArgIleGlyLeu

AAGGTCGACAGCGCCGGGGTCGACGCCTGGGACGGCCGGCTGTCTACGTCAAAACTCGAAAAGGTCGGTGCCTTGGTCACGGGGTTT

LysValAspSerAlaGlyValAspAlaTrpAspGlyArgLeuSerThrSerLysLeuGluLysValGlyAlaLeuValThrGlyPhe

TGGCAGGCGCTGAGCTGTCGGCTGCATCGCCACCCTCAGCGCCCAGCTCGCCTGTGGCAGCGCCCTTCGCCCGCAGCGGCCCCCCAT

TrpGlnAlaLeuSerCysArgLeuHisArgHisProGlnArgProAlaArgLeuTrpGlnArgProSerProAlaAlaAlaProHis

GCAGCGCCCGAAGTTGCTGGCGCAGCCGCTCGATAG

AlaAlaProGluValAlaGlyAlaAlaAlaArg\*\*\*

(C)

ACCCGGGAGGCATCGGTAGGTGCCATGCCCGGGACCGTCTGGGTCATCAGCACGAACAGGCAGGCACCCGTGCTGACAGGCCATTGG

TGCGCAGGCCAGTCAAGCGCTGACGCAGTTGACGCTCGGCGCCGATCAAGGCCCCTAGCAGCAGGGCGGTGCTGACACGCAGCAAGA

ATACGTTCCAATCCATAACAAACCTCTTCCGCGGGCACAGGCCCGCTTCAACGCCATCACTCAGGAGGGCGCGGTGGTAACAACCAG

AGATGCGAAAGGACCACGGGAAAGGAGGGGGCGATGCTGTGCATCGGCTACCTTGGCAACTCGGCCAGGCAGCATGGGAGACTGACT

GGCTAGCGCGCTGCGTGGGCCTGTCGCAAGCGACTGGGGCTAGTGTCCAATGCAGGGCTCCTGTTGATGAATGCAAAAGTGCGGCGC

GGACGATAAGCCCGACCCTTGATCAGGTCAAGGGTAAAAACATGACAATGGCTTCATCCGGAAAAGCCTGACAAGCCGGCGGGTACG

ATGGAAAAATCCACGTTCAATGGTTAACCTTTTGCTTTTTACGGACCGTCGCGGGCTGATTCATGCTGTTCAATATCCTCGTCCTG

--------------------------------------------------------------MetLeuPheAsnIleLeuValLeu

CTGGGCACACTGGTGGCCATGGAGGGTGTCGGCACCTTGGCGCACAAGTACATCATGCATGGCTGGGGCTGGTGGCTACATCGCTCG

LeuGlyThrLeuValAlaMetGluGlyValGlyThrLeuAlaHisLysTyrIleMetHisGlyTrpGlyTrpTrpLeuHisArgSer

CACCACGAACCCCACCTGGGCGTGCTTGAAACCAACGACGTTTACCTGCTGGCACTGGCGGTCATTTCGCTGGCACTGGTGGCGGTT

HisHisGluProHisLeuGlyValLeuGluThrAsnAspValTyrLeuLeuAlaLeuAlaValIleSerLeuAlaLeuValAlaVal

GGCAAGGCCGGGCACGCGCCATTGCAATGGGTCGGAGGCGGGGTGGCGGCTTACGGCGCGCTGTACGTGATTGCGCATGATGGCTTT

GlyLysAlaGlyHisAlaProLeuGlnTrpValGlyGlyGlyValAlaAlaTyrGlyAlaLeuTyrValIleAlaHisAspGlyPhe

TTTCACCGTCATTGGCCGCGCAAGATGCGCCCGGTGGGGCGCTACCTGACGCGCTTACACCGCGCCCATCGGTTGCACCATGCCGTC

PheHisArgHisTrpProArgLysMetArgProValGlyArgTyrLeuThrArgLeuHisArgAlaHisArgLeuHisHisAlaVal

AAAGGGCGCGAGGGTGGTGTGTCGTTCGGTTTTTTCTACGCACCTCCGCTAAAGCGCCTGGTGCGCGAGCTGCGGCGCAACAGGGAC

LysGlyArgGluGlyGlyValSerPheGlyPhePheTyrAlaProProLeuLysArgLeuValArgGluLeuArgArgAsnArgAsp

GCCGGTCACTGAGATCTCACGCGCACAGGCGCTTTAGCTGAGCGAGTTCCCCCGCGATCCTACCGGTTCGGGCATACATTGCACTGT

AlaGlyHis\*\*\*

CTGGGTCTGCGCAGTAGCATGCGCGCTTTTCAGCCCAGGACTCGCTCTTCATGTCGTATCAGCAAGGTCTGCTTGCCACGCCCGTGC

CGGTGCACGCGCGTCACCTGTTCTTTTCCATC

Supplementary Figure 1. Genomic sequence and its deduced amino acid sequence of the carotenoid biosynthesis genes.

1. The region around *crtE* gene; (B) the region including *idi*, *crtX*, *crtY*, *crtI* and *crtB* genes; (C) the region around *crtZ* gene. Red and blue letters indicate the start and stop codons, respectively. Purple letters indicate the overlapping sequences of start and stop codons. The sequences of the shaded regions in (A) and (B) are complement to each other.