Copper Tolerance of *Listeria monocytogenes* strain DRDC8

Francesca Y Bell, B. Biotech. (Hons.) (Adelaide)

A thesis submitted for the Degree of Doctor of Philosophy
School of Molecular and Biomedical Science
Faculty of Sciences,
The University of Adelaide
Adelaide, South Australia, Australia

May, 2010

Appendix A

This sequence data was assembled as part of collaborative effort by F. Bell, M. Francis, M. Webster, A. Akya and C. Thomas. The sequence is available from the NCBI GenBank database (Accession Number U15554).

```
LOCUS       LMU15554               37279 bp    DNA     linear   BCT 25-SEP-2007
DEFINITION  Listeria monocytogenes strain DRDC8 hypothetical protein (pCT0030)
            gene, partial cds; hypothetical protein (pCT0031), hypothetical
            protein (pCT0032), hypothetical protein (pCT0001), hypothetical
            protein (pCT0002), unknown (pCT0003), hypothetical protein
            (pCT0004), hypothetical protein (pCT0005), hypothetical protein
            (pCT0006), hypothetical protein (pCT0007), and hypothetical protein
            (pCT0008) genes, complete cds; insertion sequence ISS1-like
            hypothetical protein (pCT0009) gene, complete cds; unknown
            (pCT0010) gene, complete cds; insertion sequence ISS1-like
            hypothetical protein (pCT0011) gene, complete cds; hypothetical
            protein (pCT0012), hypothetical protein (pCT0013), and hypothetical
            protein (pCT0014) genes, complete cds; insertion sequence ISS1-like
            hypothetical protein (pCT0015) gene, complete cds; hypothetical
            protein (pCT0016), hypothetical protein (pCT0017), hypothetical
            protein (pCT0018), hypothetical protein (pCT0019), and CtpA (ctpA)
            genes, complete cds; transposon pCT0021 gene, complete sequence;
            and hypothetical protein (pCT0022), hypothetical protein (pCT0023),
            hypothetical protein (pCT0024), hypothetical protein (pCT0025),
            hypothetical protein (pCT0026), hypothetical protein (pCT0027),
            hypothetical protein (pCT0028), and hypothetical protein (pCT0029)
            genes, complete cds.
ACCESSION   U15554
VERSION     U15554.3  GI:157429038
KEYWORDS    .
SOURCE      Listeria monocytogenes
  ORGANISM  Listeria monocytogenes
            Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
REFERENCE   1  (bases 1 to 37279)
  AUTHORS   Francis,M.S. and Thomas,C.J.
  TITLE     The Listeria monocytogenes gene ctpA encodes a putative P-type
            ATPase involved in copper transport
  JOURNAL   Mol. Gen. Genet. 253 (4), 484-491 (1997)
   PUBMED   9037109
REFERENCE   2  (bases 1 to 37279)
  AUTHORS   Francis,M.S. and Thomas,C.J.
  TITLE     Mutants in the CtpA copper transporting P-type ATPase reduce
            virulence of Listeria monocytogenes
  JOURNAL   Microb. Pathog. 22 (2), 67-78 (1997)
   PUBMED   9049996
REFERENCE   3  (bases 1 to 37279)
  AUTHORS   Thomas,C.J., Bell,F.Y., Akya,A., Francis,M. and Webster,M.
  TITLE     Partial sequence analysis of a large plasmid from Listeria
            monocytogenes strain DRDC8
  JOURNAL   Unpublished
REFERENCE   4  (bases 1 to 37279)
  AUTHORS   Thomas,C.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-OCT-1994) Microbiology and Immunology, University of
            Adelaide, Adelaide, South Australia 5005, Australia
REFERENCE   5  (bases 1 to 37279)
  AUTHORS   Thomas,C.J., Bell,F.Y., Akya,A., Francis,M. and Webster,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-NOV-2005) Molecular and Biomedical Science, The
            University of Adelaide, Adelaide, SA 5005, Australia
  REMARK    Sequence update by submitter
```

```
REFERENCE   6  (bases 1 to 37279)
  AUTHORS   Thomas,C.J., Bell,F.Y., Akya,A., Francis,M. and Webster,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-2007) Molecular and Biomedical Science, The
            University of Adelaide, Adelaide, SA 5005, Australia
  REMARK    Sequence update by submitter
COMMENT     On Sep 24, 2007 this sequence version replaced gi:83268869.

FEATURES             Location/Qualifiers
     source          1..37279
                     /organism="Listeria monocytogenes"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:1639"
                     /plasmid="pCT100"

     gene            complement(<1..450)
                     /locus_tag="pCT0030"
                     /evidence=not_experimental
     CDS             complement(<1..450)
                     /locus_tag="pCT0030"
                     /note="Similar to plasmid replication proteins.  pli0070
                     of Listeria innocua Clip11262 (NP_569223), Listeria
                     monocytogenes str. 4b H7858 (ZP_00231652)."
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="Hypothetical protein"
                     /translation="MFDNEAMYQYSSVIDAVVHEGITAFRKKGSKGVLAIRDHYAAVE
                     AASENRKGVQFVVRDKGHLQMEHGVKGFIVTSQEALLTEADKITHWTPNVFRWGTYTD
                     DKRQYIKGFDEQNLQQINTFVVDVDTQQVDVAKMLTASMKVLDQTPTF"


     gene            1118..2164
                     /locus_tag="pCT0031"
                     /note="Overlaps pCT0032."
                     /evidence=not_experimental
     CDS             1118..2164
                     /locus_tag="pCT0031"
                     /note="Similar to plasmid replication protein Rep63B of
                     Bacillus thuringiensis serovar kurstaki (CAB43193)"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="Hypothetical protein"
                     /translation="MSLSDKFIFLSFYLYYKYITGGERMFNSPAYIEEYKKVQKVLSD
                     NRCRVIINANQKGGVGKTTNTTMEAIILAHVFNKKVLLIDEDMQGNETSFMSKTYDVI
                     EFPMSLMKAIEEGDLTKAIVNLDENIDIIPGSYDMRKMVNFLIGKFKTEEAQTFYLSS
                     LIDKIRDDYDFIFIDVPPSTDLKVDNAVVAADYLVVVQETQQFALEGSNTFISTYLNT
                     LVDDFGSRFKTEIIGILPVLLQKKRKLHEKILETTKEKFGKENIFGTTINNHARLELY
                     GKFGVQFEDHWDRTMFALYTDIVQEMLERMQLIENGSDLDNYQYKPIFYNPKIGKVTA
                     LGKEIVVNGSINRS"
     misc_feature    1259..2035
                     /locus_tag="pCT0031"
                     /note="ATPases involved in chromosome partitioning [Cell
                     division and chromosome partitioning]; COG1192; Region:
                     Soj"
                     /evidence=not_experimental
     misc_feature    1262..>1423
                     /locus_tag="pCT0031"
                     /note="ParA and ParB of Caulobacter crescentus belong to a
                     conserved family of bacterial proteins implicated in
                     chromosome segregation; cd02042; Region: ParA"
                     /evidence=not_experimental
     misc_feature    <1532..1825
```

```
                /locus_tag="pCT0031"
                /note="ParA and ParB of Caulobacter crescentus belong to a
                conserved family of bacterial proteins implicated in
                chromosome segregation; cd02042; Region: ParA"
                /evidence=not_experimental

gene            2142..2438
                /locus_tag="pCT0032"
                /note="Overlaps pCT0031."
                /evidence=not_experimental
CDS             2142..2438
                /locus_tag="pCT0032"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MARLIDRKDKEQAQVKTQADTSRKVNTKSNSEKKVNAKDRKNIK
                VNPDDMLLIDTIANNFFETKLKHYELVNLLLENWVDTKLEPRQQKILKDLLKNR"

gene            complement(2489..3769)
                /locus_tag="pCT0001"
                /evidence=not_experimental
CDS             complement(2489..3769)
                /locus_tag="pCT0001"
                /note="Similar to UV Damage repair protein, Listeria
                innocua (pli0067, NP_569220.1  GI:18450349, lin2823,
                CAC98049.1 GI:16415359)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="hypothetical protein"
                /translation="MVQIEDYTHVPSRDILCVDVKSFFASVECVKRGLDPLQAFLVVM
                SNADRAGGLVLAASPQMKKVFRIKTGSRKFEVPEDDPRIIIAPPRMQLYLKVNELIQK
                IFLRYVSREDFHVYSIDEAFLDVTHSKKLFGSAEVIAKKIQADILKELKLYVTVGIGD
                NPLLAKLALDNEAKHQKSGIAEWRYKDVPETVWKIRPITDFWGIGKRTAFQLERMGIY
                SIHGLSQMLPSYLKRRLGVIGEQLYYHAHGIDYSRLSEKYKPLEKSYGKSQILEKDYH
                DPEQIAVIIQEMVEEVAMRLRNHHVDTSVIHLSAGYSRYSTRNGFSHQKKIMATDSSK
                ELVPYFLEMFWKYQENDAVRSVAVSCGGIKRKTSMQLSVFEDYTKTLQQEQLERTIDK
                IRDRYGFKALMHANSLVAGATGLKRSDLVGGHKG"
misc_feature    complement(3155..3727)
                /locus_tag="pCT0001"
                /note="PROSITE entry PS50173; Region: UmuC domain"
                /evidence=not_experimental

gene            complement(3757..4101)
                /locus_tag="pCT0002"
                /evidence=not_experimental
CDS             complement(3757..4101)
                /locus_tag="pCT0002"
                /note="Similar to hypothetical protein, Listeria innocua
                (pli0066, CAC42064.1 GI:16415849, lin2822, NP_472151.1
                GI:16801883)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Conserved hypothetical protein"
                /translation="MAYRDRGMQKWQGMILSEHTEQLEQAKRKLSWLPQQTEECLVQI
                LSYAVTKQHAITIQRLTKDENGYPLPFVMGRIAGVEDDRIYIRDMAGDITVLTFDDIN
                HISEGTIEKWYR"

gene            complement(4306..5040)
                /locus_tag="pCT0003"
                /evidence=not_experimental
CDS             complement(4306..5040)
```

```
                /locus_tag="pCT0003"
                /note="Similar to ParA-like partitioning proteins"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Unknown"
                /translation="MDEDLMSAAVASERWGFDSSYVRQLYRKYPNRFLDGTIRQFGKT
                YVITKEGMEYLTGQGELEAKISRNTLKGTKKIINILSFKGGVGKTTTAKIINEALYKS
                GKNSLYVDLDYEAFPAKIESDYVILDQSPYVKQDESLYLDKSDFIIFPYAENISNIKI
                LSLMKEVILRAKVHNESLQVIFVKTGLEKNNNFDKEAKSLKKQVPIIFADFSVDNSEE
                LTNINDQIKNERSEEIEKIIKNILGT"

gene            complement(5586..6377)
                /locus_tag="pCT0004"
                /evidence=not_experimental
CDS             complement(5586..6377)
                /locus_tag="pCT0004"
                /note="Similar to DNA transposition-associated proteins,
                Listeria innocua (pli0057, NP_569210.1  GI:18450339),
                Staphylococcus aureus (NP_932183.1  GI:37595765)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MKYKYIGICYGNPGVGKTISAKHYANWHRLETILTRKGNRLNPE
                LVIDKLKGIKTDALFVTAPAAKPSSLSSQLSSLAIDIARVNEIISSTENKVIPKADRI
                DCMATSIDRIVENIFYAVDLIIIDEIDRLKIQTLDLLRDIYDQNDVGMVFIGMQGIEK
                RLSKYPQLYSRIGFAHEFKKLSSSEIKHILEYRCQDIGLAIDYENFEDYDAINRIIKL
                TNGNFRLLQRLFTQIDRIMEINNLEKISSDVVQAARESLIIGFTE"

gene            complement(6436..7869)
                /locus_tag="pCT0005"
                /evidence=not_experimental
CDS             complement(6436..7869)
                /locus_tag="pCT0005"
                /note="Similar to transposase, Listeria innocua (pli0058,
                CAC42056.1 GI:16415841), Bacillus cereus (AAY60237.1
                GI:66970261)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MSNKKILANYSEEQRLKALKKFRVLQPYLKKEIPLSKVSEENDI
                STRTLRKWISNYKKSGLVGLITKERSDKNSMKISEDLCKIIKHIYLENKNVSITSIYR
                RTVSWCKENNIQEPSYYQVRKIIQLVPDKLRVLAHQGSKIYMNHYELVHIRETTAPNI
                IWQADHTLLDIEVLNAKGYPERPWLTIIMDDYSRAITGYFIAFESPNSLRTALTLRQA
                IWRKDQADWPICGIPRIFLQITVAILLLNILNRSLLIENRIKFSKVGVPRGRGKIERF
                FHTVNSMLLQDLPGYIKNRKAKSMLTLEEFSDIFRNWLLRIYHQKQHSTTKEKPIAMW
                NNYDFLPNMPNSLEDLDLLLIKVKKERVVHSDGIHLFGMKYVHPTLSAFVSEPVVIRY
                DPRDISDVRVFYKNVFLCTAVSTSFEQYAIGIREIEKERSKLKRELKRELIVSTNKVI
                EKLVGRQKENSSTVKNNVSSLRRYENE"
misc_feature    complement(6856..7410)
                /locus_tag="pCT0005"
                /note="PS50994; Region: Integrase catalytic domain"
                /evidence=not_experimental

gene            8028..8627
                /locus_tag="pCT0006"
                /evidence=not_experimental
CDS             8028..8627
                /locus_tag="pCT0006"
                /note="Similar to invertase/site-specific recombinase,
                resolvase family, Listeria innocua (pli0059, CAC42057.1
                GI:16415842), Staphylococcus aureus (GI:32470565
```

```
                NP_863235.1)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MKIAYARVSTLEQNLTRQIETLKSFGAEKIFTEQKSGASIKNRK
                AFQEALQFVRSGDTFMVEAIDRLGRNYDEIIQTVYFLKEKEVKLVITSLPIMAEAIGN
                PLLDKFIKDLIIQILAMIAEQERAESKRRQAQGIHSAKQKGIYKGRPILYSPTSKDPQ
                KRVIYQRVIDMLSKGESISKIAEESGITRQTVYRIKNSN"
misc_feature    8040..8066
                /locus_tag="pCT0006"
                /note="PS00397, YARVSTLEQ; Region: Site specific
                recombinase active site"
                /evidence=not_experimental
misc_feature    8193..8231
                /locus_tag="pCT0006"
                /note="PS00398, GDTFMVEAIDRLG; Region: Site specific
                recombinase signature"
                /evidence=not_experimental

gene            8901..9260
                /locus_tag="pCT0007"
                /evidence=not_experimental
CDS             8901..9260
                /locus_tag="pCT0007"
                /note="Similar to Cadmium resistance regulator/accessory
                protein CadC, Lactococcus lactis (AAB37344.1  GI:1699048),
                Listeria innocua (pli0060, NP_569213.1 GI:18450342),
                Streptococcus thermophilus (CAC87153.1 GI:22416342)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MNNEICEITCIHEDKVNRAKSKLANFDTPSVSGFFKILSDENRL
                KIVHALVHEDELCVCDIANIIDASVATTSHHLNSLKKLGVVDSHKDGKLVYYFIKNIK
                ILNLMELGVNFKEEVLA"
misc_feature    8967..9257
                /locus_tag="pCT0007"
                /note="PS50987, HTH_ARSR_2.
                LANFDTPSVSGFFKILSDENRLKIVHALVHEDELCVCDIANIIDASVATTSHHLNSLK
                KLGVVDSHKDGKLVYYFIKNIKILNLMELGVNFKEEVLA; Region: ArsR-type
                HTH domain"
                /evidence=not_experimental
misc_feature    9072..9143
                /locus_tag="pCT0007"
                /note="CvcDIaniidaSvattSHH; Region: bacterial regulatory
                protein arsR family signature"
                /evidence=not_experimental

gene            9257..11374
                /locus_tag="pCT0008"
                /evidence=not_experimental
CDS             9257..11374
                /locus_tag="pCT0008"
                /note="Similar to Cadmium efflux ATPase, Listeria innocua
                (pli0061, CAC42059.1  GI:16415844), Lactococcus lactis
                (AAF98300.1  GI:9789448), Streptococcus thermophilus
                (CAE52398.1 GI:46019872)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MSEKMTEKTYRIEGLSCTNCAGKFEKNVKQLPGVTSATVNFGAS
                RISVEGQTTIEELEEAGAFENLIIRDDQENDEQVRSKESFIKRNIALIISLGFILVAV
```

```
                ISQLSLGEDHLLTKALYILAIIIGGFDLFKEGFSDLIKLDFSMESLMTIAIIGAAFIG
                EWAEGSIVVILFAISEALERFSMDKARQSIRSLMDIAPKEALIRRNNVEQLVSVDKID
                IDDIMIIKPGQKIAMDGLVINGHSSVNQAAITGESVPVEKQLDDEVFAGTLNEEGVLE
                VKVTKKVTDTTIAKIIHLVEEAQGERAPAQAFVDKFAKYYTPFIIIMALLIVVVPPLF
                FGGDWNKWLYQGLSILVVGCPCSLVISTPVSIVSAIGNAAKNGVLVKGGVYLEEIGHL
                RAIAFDKTGTLTKGKPVVTDFIATSSETDINYLSIISSLESLSQHPLASAILNEADKT
                NVDYKSIQIEDFQSITGKGLTGIHQNIRYYIGSPKLFSASVIEETAVKVQYRQFQEQG
                KTAMYFGTDEQILGVIAVADEVRDSSAAVISELHKLSIEHTIMLTGDNTKTAESIGKQ
                LGVTEIKGDLMPQEKLDSIKALRTTYNKVAMVGDGINDAPALAASTVGIAMGGAGTDT
                ALETADVALMGDDLQKLPFIVRLSRQTLKVIKQNITFSLGIKLLALLLVIPGWLTLWI
                AIVADMGATLLVTLNGLRLMKVK"
misc_feature    9290..9379
                /locus_tag="pCT0008"
                /note="PS01047, HMA_1, IegLsCtNCagkFeknVkqlpgvtsat.VnF;
                Region: Heavy metal associated domain"
                /evidence=not_experimental
misc_feature    10448..10468
                /locus_tag="pCT0008"
                /note="PS00154, ATPase_E1_E2 DKTGTLT; Region: ATPase
                phosphorylation site"
                /evidence=not_experimental
repeat_unit     11912..11929
                /note="Inverted repeat (left) for ISS1 like element
                enclosing pCT0009"
                /rpt_type=Inverted
-10_signal      11951..11956
                /evidence=not_experimental
RBS             11972..11975
                /evidence=not_experimental



gene            11984..12664
                /locus_tag="pCT0009"
                /evidence=not_experimental
CDS             11984..12664
                /locus_tag="pCT0009"
                /note="Similar to transposase, Listeria innocua (pli0071,
                CAC42069.1  GI:16415854, pli0024, CAC42022.1  GI:16415808,
                pli0063, CAC42061.1  GI:16415846), Lactococcus lactis
                (Orf945, AAC64335.1  GI:2895550)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein. Similar to pCT0011,
                pCT0015."
                /translation="MSHFKGKQFKQDIIMVAVGYYLRYNLSYRDVSEILNERGIKVCH
                TTVYRWVQEYGSIIYCLWKKRNKSASDSWRMDETYIKVKGKWHYLYRAIDSTGLTLDI
                WLRKNRDKQAAYAFFKRLIKQFGEPRVFVTDKAPSLSSAFNRLKSEGLFSNTEHRTSK
                YLNNIIEQDHRPIKKRHKLYQSIRTASSTIKGIETIHALYKTSQRDFSLFGFSVIHEI
                NNLMGVPA"
repeat_unit     12701..12719
                /note="Inverted repeat (right) for ISS1 like element
                enclosing pCT0009"
                /rpt_type=Inverted

gene            13299..13964
                /locus_tag="pCT0010"
                /evidence=not_experimental
CDS             13299..13964
                /locus_tag="pCT0010"
                /note="Similar to exonuclease V gamma subunit"
                /codon_start=1
                /evidence=not_experimental
```

```
                /transl_table=11
                /product="Unknown"
                /translation="MTLEVYMQTDIEVNKQEELKEELKNFIVNVGWTHKIHIVRSDEY
                IYYSKLFKWIRIILAAFTSSGLIGILFDSQSYILKVVTAIFSFATVAMNGLDKAGDFE
                KLCLKEKNDANNFWKLRTDAESLLSKTVFSTESMGRIDQEFEELKRLRIQYNAGLLNV
                PAKTVKKASKLIHKRKDNGFCCKVQRSKMISQLFSNFKLEHPLDYLFHESLKSQREKS
                LFG"
repeat_unit     complement(13833..13851)
                /note="Inverted repeat (right) for ISS1 like element
                enclosing pCT0011"
                /rpt_type=Inverted

gene            complement(13888..14568)
                /locus_tag="pCT0011"
                /evidence=not_experimental
CDS             complement(13888..14568)
                /locus_tag="pCT0011"
                /note="Similar to transposase, Listeria innocua (pli0071,
                CAC42069.1  GI:16415854, pli0024, CAC42022.1  GI:16415808,
                pli0063, CAC42061.1  GI:16415846), Lactococcus lactis
                (Orf945, AAC64335.1  GI:2895550)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein.  Similar to pCT0009,
                pCT0015."
                /translation="MSHFKGKQFKQDIIMVAVGYYLRYNLSYRDVSEILNERGIKVCH
                TTVYRWVQEYGSIIYCLWKKRNKSASDSWRMDETYIKVKGKWHYLYRAIDFTGLTLDI
                WLRKNRDKQAAYAFFKRLIKQFGEPRVFVTDKAPSLSSAFNRLKSEGLFSNTEHRTSK
                YLNNIIEQDHRPIKKRHKLYQSIRTASSTIKGIETIHALYKTSQRDFSLFGFSVIHEI
                NNLMGVPA"
RBS             complement(14577..14580)
-10_signal      complement(14596..14601)
repeat_unit     complement(14623..14640)
                /note="Inverted repeat (left) for ISS1 like element
                enclosing pCT0011"
                /rpt_type=Inverted

gene            complement(15341..20014)
                /locus_tag="pCT0012"
                /evidence=not_experimental
CDS             complement(15341..20014)
                /locus_tag="pCT0012"
                /note="Similar to restriction modification system LlaG1
                from Lactococcus lactis W10 (AAK71920.1  GI:14669432).
                Comprises a single protein encoding a catalytic motif,
                seven helicase-like motifs characteristic of type I and II
                endonucleases and four methylase motifs characteristic of
                adenine-methylases"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MTKQTFDDLINQINTVSKDVQRERGTLFEKLTLAYLKNEPTYKA
                LYQNVWLLSEVPESYGIPKKDTGVDLVAEQKNGDLVAIQAKFYTNKVGKSEINSFVAE
                LGKSYYQRGLIVSTMDDWNSNARETIDQNEKGIEIIGLSDLRNSQIDWSQFNFERPEN
                VVVKKPKKLRDYQQTAKENALAHFKENERGQLIMAPGTGKTFTSLKISEALSKDKAGP
                FKVLYLVPSIQLLTQTLRGWNNDTELTITSMAVTSDRDASRGTDGTEDIKASDIGYPA
                TTSSKKILQNWHDFESLSKPTDMLVVFSTYQSIEVIGEAQKEGFPEFDFIISDEAHRT
                TGAHEAAKEASAFSKVHSNNNVKGLKRMYQTATPKIYGESAKKNAKDKSILLSSMDDE
                SKYGEVFFRMGFGQAVSRDTLTDYKVMVLAVDEAAIQKDMQRTLADPENGLNIDDVGR
                IVGIWNGMMRRNGYKNPIKNSPYDGAPLERAIAFTRTIEESKKVSSQFEEVVNEYISE
                AIEDESIHLSMRHADGQMNALQKGEVLDWLANPNKPADEARIVSNVRFLTEGIDIPTL
                DAVIFLSPKKSQVDIVQAVGRIMRKAEGKDYGYIILPIVIPTGEKPETILDNNKNYET
```

```
                VWQVINALRSVDERFEAMIDKLNMAKPKQLKVIGVGSAPDQVNDKDKSIENTPVQTEL
                EFEWDKFEGAIFGKIVQKVGDRKYLENWSKDVAKIAERQINWIKNKLSDKKDPISLEF
                KKFVSSLQHNINESIDENQAAEMLSQHLITKPIFEALFSEYSFVNQNPVSQAMESIVS
                ELEKAGFAKEQENLEPLYESVRMRAEGIEKAEDKQKIIVTLYDKFFKTAFKATTERLG
                IVFTPIEVVDFIVHSVDDVLKKHFGKSLASKDVHILDPFTGTGTFIVRTLTYLKEQMD
                AGEISLADITRKFMNELHANEIVLLSYYIAAINIEATFDEINGEEEGYVPFEGIVLTD
                TFESTETEDTLDDDYFGTNDERLKRQQDVQITAIIGNPPYSVGQSNANDDNKNVQYQK
                LNNRISETYALHSSATNKRNLYDDFIKAFRWTSDRLKGNGIIGFVSNASFINSQSADG
                LRYSLYKEFNHLYIFNLRGDQRTQGETSRKEGGKIFGSGSRTPIAISILVKDGSDNHE
                VHYHDIGDYLTREQKLLKLKDVKSVLNMEWKNIVPDKNNDWINQRDESYENYLALYRE
                IFAEKAVGVVSSRDAWVYGFSYAKVRDNSEKMINNYNQELDRLKNNYDISLLNHSGNF
                VKWSRGLKQKFSKGKKIDFLPDATIIGIYRPFTKKYLQYQVDIIESPGKYKDIFGDNN
                LSIYSTGAGASKGFSALITNAIPNFHFMDSGQGFYKNDNSSKDGFFVKDSSNLIENKL
                ELSQDDCFYYVYSVLHSSEYRKKYANDLQKTLPRIPLLKHREEYVEIGRKLADLHLNY
                EDQPIWEGVEVEISKPNYRVKKMKHPKKGILDTIIFNDSITIKNIPEKAYEYIVNGRP
                AIEWIIDQYQVKTDKKSGITDDPNDFSDDPKYILNLLLAVITVSMRTLELIDELPEFE
                LLEKE"
misc_feature    complement(19790..19849)
                /locus_tag="pCT0012"
                /note="PESYGIPKKDTGVDLVAEQK; Region: Catalytic
                endonuclease motif"
                /evidence=not_experimental
misc_feature    complement(19394..19450)
                /locus_tag="pCT0012"
                /note="RGQLIMAPGTGKTFTSLKI; Region: Helicase-like (DEAD
                motif) endonuclease motif I"
                /evidence=not_experimental
misc_feature    complement(19313..19360)
                /locus_tag="pCT0012"
                /note="FKVLYLVPSIQLLTQT; Region: Helicase-like (DEAD
                motif) endonuclease motif IA"
                /evidence=not_experimental
misc_feature    complement(19019..19042)
                /locus_tag="pCT0012"
                /note="FIISDEAH; Region: Helicase-like (DEAD motif)
                endonuclease motif II"
                /evidence=not_experimental
misc_feature    complement(18908..18934)
                /locus_tag="pCT0012"
                /note="RMYQTATPK; Region: Helicase-like (DEAD motif)
                endonuclease motif III"
                /evidence=not_experimental
misc_feature    complement(18476..18532)
                /locus_tag="pCT0012"
                /note="SSQFEEVVNEYISEAIEDE; Region: helicase-like (DEAD
                motif) endonuclease motif IV"
                /evidence=not_experimental
misc_feature    complement(18476..18532)
                /locus_tag="pCT0012"
                /note="SSQFEEVVNEYISEAIEDE; Region: helicase-like
                endonuclease motif IV"
                /evidence=not_experimental
misc_feature    complement(18305..18367)
                /locus_tag="pCT0012"
                /note="IVSNVRFLTEGIDIPTLDAVI; Region: Helicase-like (DEAD
                motif) endonuclease motif V"
                /evidence=not_experimental
misc_feature    complement(18245..18286)
                /locus_tag="pCT0012"
                /note="SQVDIVQAVGRIMR; Region: Helicase-like (DEAD motif)
                endonuclease motif V"
                /evidence=not_experimental
misc_feature    complement(17399..17449)
                /locus_tag="pCT0012"
```

```
                /note="GIVFTPIEVVDFIVHSV; Region: Methylase motif Is"
                /evidence=not_experimental
misc_feature    complement(17300..17344)
                /locus_tag="pCT0012"
                /note="ILDPFTGTGTFIVRT; Region: methylase motif CM I"
                /evidence=not_experimental
misc_feature    complement(16979..17008)
                /locus_tag="pCT0012"
                /note="ITAIIGNPPY; Region: Methylase motif Cm II"
                /evidence=not_experimental
misc_feature    complement(16784..16840)
                /locus_tag="pCT0012"
                /note="FRWTSDRLKGNGIIGFVSN; Region: Methylase motif CM
                III"
                /evidence=not_experimental

gene            complement(20011..20553)
                /locus_tag="pCT0013"
                /evidence=not_experimental
CDS             complement(20011..20553)
                /locus_tag="pCT0013"
                /note="Similar to gram positive plasmid replication
                protein B, Listeria innocua (pli0022, NP_569176.1
                GI:18450305) and cryptic protein often found adjacent to
                replication protein genes, Lactococcus lactis (AAK71921.1
                GI:14669433)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MMSEKLKTIKELADEIGVSKQAVWQKIKKESSIDLRQFTSKKGN
                TVYVDGQKVIKSAFFNKTSTKKRQQKVFVDDNVNNSVDDNPEGNEEILFLRNLVSELQ
                SEKKDLHKLLDQQQRLALQDKKLLEEYKEEIKELKSLMVPVRKDDKDPMVRKSKDIDN
                TKEELHTKNKKWWHFGRNVK"

gene            complement(20550..21701)
                /locus_tag="pCT0014"
                /evidence=not_experimental
CDS             complement(20550..21701)
                /locus_tag="pCT0014"
                /note="Similar to plasmid replication initiation protein,
                Listeria innocua (pli0023, CAC42021.1  GI:16415807),
                Lactococcus lactis (AAD43535.1  GI:5453328) and repB
                replication protein, Lactococcus lactis (AAB52513.1
                GI:1944625)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MTSIPKNEQNQKQVQILNELSKRKVVEHNSLITSIAKMDKTPLK
                MFELAVSCIDTEEPPKDHTVYLSKEELFAFFKVSDNDKHSRFKQAVENMQKQAFFKIK
                EKKEHGFEFENIVPIPYVKWTDYHDEVTIRFSPEILPYLINLKKNFTQHALSDISELN
                SKHSIILYRWLSMNYNQYEHYSYKGGRREEQVEAYRNLSISIRELREMTDIVNEYKLF
                ADLEKWILKKPLEEINDHTSFTVTYDKVKKGRSIDSIVFHITKKRRADDNSYKLEDKV
                YQEDKARKAETEDMLTVQALKSPYTKLLMEHFLLSYLDLTDTKILSGLQAHVYPLYDE
                LKDLRGLDGVNDHLGYVSSKREEYSKHNIAKYLKKAIEQYLSTVKRQDL"
RBS             complement(21712..21715)
                /evidence=not_experimental
-10_signal      complement(21776..21781)
                /evidence=not_experimental
repeat_unit     21787..21799
                /rpt_type=Inverted
-35_signal      complement(21797..21802)
                /evidence=not_experimental
```

```
repeat_unit     complement(21809..21839)
                /note="30 base direct repeat typical for promoter regions
                associated with repB genes"
                /rpt_type=Direct
repeat_unit     complement(21832..21861)
                /note="30 base direct repeat typical for promoter regions
                associated with repB genes"
                /rpt_type=Direct
repeat_unit     complement(21832..21843)
                /rpt_type=Inverted
repeat_unit     complement(21854..21868)
                /note="15 base imperfect repeat typical for promoter
                regions associated with repB genes"
                /rpt_type=Direct
repeat_unit     complement(21927..21944)
                /note="Inverted repeat (right) for ISS1 like element
                enclosing pCT0015"
                /rpt_type=Inverted

gene            complement(21981..22661)
                /locus_tag="pCT0015"
CDS             complement(21981..22661)
                /locus_tag="pCT0015"
                /note="Similar to transposase, Listeria innocua (pli0024,
                pli0071, CAC42022.1 GI:16415808, pli0063 CAC42061.1
                GI:16415846), Lactococcus lactis (Orf946, AAC64335.1
                GI:2895550)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein.  Similar to pCT0009,
                pCT0011."
                /translation="MSHFKGKQFKQDIIMVAVGYYLRYNLSYRDVSEILNERGIKVCH
                TTVYRWVQEYGSIIYCLWKKRNKSASDSWRMDETYIKVKGKWHYLYRAIDSTGLTLDI
                WLRKNRDKQAAYAFFKRLIKQFGEPRVFVTDKAPSLSSAFNRLKSEGLFSNTEHRTSK
                YLNNIIEQDHRPIKKRHKLYQSIRTASSTIKGIETIHALYKTSQRDFSLFGFSVIHEI
                NNLMGVPA"
RBS             complement(22670..22673)
                /evidence=not_experimental
-10_signal      22689..22694
                /evidence=not_experimental
repeat_unit     22716..22733
                /note="Inverted repeat (left) for ISS1 like element
                enclosing pCt0015"
                /rpt_type=Inverted

gene            23209..24534
                /locus_tag="pCT0016"
                /evidence=not_experimental
CDS             23209..24534
                /locus_tag="pCT0016"
                /note="Similar to Na+-driven multidrug efflux pump, MATE
                efflux family protein, Streptococcus suis (ZP_00332543.1
                GI:50591223), Streptococcus thermophilus (norN,
                ZP_00388590.1 GI:62527305)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MTRMLKNLIKLALPATVENILETSVGFVDSLMISQIGLFAVAGI
                GIANAILNVYIAVFIALGIGTSSLISRNIGAKNNEKAKLLLTIFTISNRLGLFWVGES
                SNRSTTLKLMGATKQTLNYSLQFFSIVGGGSVAIATMIILGSMLRAIGDTKTPMKIGL
                ITNILNISLDYILIFGLGPIPALGVVGTAIGTLVARLLGTILLYRKVHLSVLSFPLLS
                IFQRSNYQELLKLSFPAALERLVMRLGQVLYFGVIVAIGAKTYSAHSIAGSIESFVYM
```

```
                PAYGLATAAATLAGNSIGKKDYAETKRVALYAVKYGVIVLSILGIGLFFGTPYIAPWF
                TIDAEAINQIVIALKIDAFNQPGLAISLILAGVLQGMGDTKTPLYSTAFGMWITRILG
                CILLGSIMGLGIAGVWLAIGLDLYVRSLFLTYHFKRNLQILNSNNELSS"

gene            25295..25789
                /locus_tag="pCT0017"
                /evidence=not_experimental
CDS             25295..25789
                /locus_tag="pCT0017"
                /note="Similar to negative transcriptional regulators
                associated with copper transport operons copR, Lactococcus
                lactis  (AAK04930.1  GI:12723756), CopY, Streptococcus
                mutans  (AAN58178.1  GI:24376801, AAG10085.1  GI:9965434)
                Streptococcus pneumoniae R6 copY (NP_358233.1
                GI:15902683)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MENKHQSNSNEVLMSQLPSDPKILKELNISDSELVIMRVVWSLG
                STTADEIGRELSETYQWSPSTIKTFLARLIKKGLLKNSRDGRKYVYIATCSEDEAICQ
                MTLSFLNKICAHKHANVILEMIDASSITAENKEAISEKLSSKNVVDEVTCDCINRLNC
                CDNN"
misc_feature    25379..25417
                /locus_tag="pCT0017"
                /note="Region: ISDSELVIMRVVW; consensus N-terminal DNA
                binding domain"
                /evidence=not_experimental
misc_feature    25475..25525
                /locus_tag="pCT0017"
                /note="Region: WSPSTIKTFLARLIKKG; N-terminal consensus DNA
                binding domain"
                /evidence=not_experimental
misc_feature    25748..25777
                /locus_tag="pCT0017"
                /note="Region: CDCINRLNCC; putative copper-binding motif"
                /evidence=not_experimental

gene            25804..26175
                /locus_tag="pCT0018"
                /evidence=not_experimental
CDS             25804..26175
                /locus_tag="pCT0018"
                /note="Similar to conserved hypothetical proteins,
                Lactobacillus johnsonii (LJ1836, AAS09781.1  GI:41584175),
                Streptococcus pneumoniae R6 (spr0640, NP_358234.1
                GI:15902684)"
                /codon_start=1
                /transl_table=11
                /translation="MSQIIVLIIGLTLIAFIGWWFFGKHEVHQETAVISSDGQTATIV
                VNGGYNPAVLNLKKDVPVNLIFNRKDASSCLEKVIFPDFGVDADLPLNQDVSILIDTS
                KSGEFIYSCGMNMFHGKIIIK"

gene            26218..26457
                /locus_tag="pCT0019"
CDS             26218..26457
                /locus_tag="pCT0019"
                /note="Putative copper binding protein, similar to
                N-terminal region of P-type ATPase - probable copper
                transporter ctpA, Streptococcus pneumoniae R6  (AAK99445.1
                GI:15458226)"
                /codon_start=1
                /transl_table=11
                /product="Hypothetical protein"
```

```
                /translation="MFGSKINNKVEVVVDGGYSPSKFKLKAGEPAEVSFTRVSDKGCA
                QQIIFNGELRNLPLNESVTFNFTPVEKGRHNWSAV"

gene            26528..28489
                /gene="ctpA"
                /locus_tag="pCT0020"
CDS             26528..28489
                /gene="ctpA"
                /locus_tag="pCT0020"
                /function="involved in cation transport"
                /note="copper-translocating ATPase: similar to copA,
                Lactobacillus plantarum (NP_786368.1  GI:28379476),
                Streptococcus pneumoniae R6 ctpA (AAK99445.1
                GI:15458226). Identified by similarity to Enterococcus
                hirae copper-translocating ATPase: Swiss-Prot Accession
                Number P32113, and Synechococcus sp. cation-transporting
                ATPase: Swiss-Prot Accession Number P37279"
                /citation=[1]
                /citation=[2]
                /codon_start=1
                /evidence=experimental
                /transl_table=11
                /product="CtpA"
                /translation="MSIKNRFIIGVIGSVPLLINMFMSLGGSMLGGDKYGVWILFAFG
                SLVYWFSGLPFLRTAVASFKNHHANMDTLVGLGTTIAYVYSLYAMFARPNETYFEAVA
                VVITLILLGSYFEERMKASASSAVDKLMGLQAKDAEVLRDGEFIKLPIEEIIVGDLIR
                VKPGEKVAVDGQIVEGTSTLDESMVTGESMPVEKGPGDNVIGATLNNTGSFTFEVTKV
                GADTMLSNIAEMVRHAQNSRAPIQKTVDRISNIFVPIVLMISILTFIVWYVFLGSTLV
                TAMIFSVSVMIIACPCALGIATPTALMVGTGRSAKLGILIKNAEVLEATHDIKTVVMD
                KTGTITVGKPQVTDIISIGRISENEILRIAAGLEDSSEHPLALAVINEAKDKKITPAV
                AKNFTAISGKGVQALIDGKQAFIGNDRLSDDFNMTDDLKVKMTSLQAQAKTVVLVGYD
                GQIIALIGIQDAPKSSSKAAIRAMQKSGFHTVMLTGDNRLVAQAIADDIGIDEVIADV
                MPGDKAQHIRKLQEKGAVAFVGDGINDAPALSTATVGIAMGSGSDIAIESGGIVLVKN
                DLMDVVTSLVLARKTYSRILINLFWAFIYNVIGIPVAAGIFSALGFTLSPELAGLAMA
                LSSITVVLSSLLLNYVRLPKSSETLIGNS"
misc_feature    27017..27124
                /gene="ctpA"
                /locus_tag="pCT0020"
                /note="Region: GEKVAVDGQIVEGTSTLDESMVTGESMPVEKGPGDN;
                Putative transduction domain"
                /evidence=not_experimental
misc_feature    27395..27424
                /gene="ctpA"
                /locus_tag="pCT0020"
                /note="Region: CPCALGIATP; Putative ion channel"
                /evidence=not_experimental
misc_feature    27494..27652
                /gene="ctpA"
                /locus_tag="pCT0020"
                /note="Region:
                EATHDIKTVVMDKTGTITVGKPQVTDIISIGRISENEILRIAAGLEDSSEHPL;
                Putative aspartyl kinase domain"
                /evidence=not_experimental
misc_feature    28055..28171
                /gene="ctpA"
                /locus_tag="pCT0020"
                /note="Region: PGDKAQHIRKLQEKGAVAFVGDGINDAPALSTATVGIAM;
                Putative hinge and ATP binding domain"
                /evidence=not_experimental
repeat_unit     28620..28650
                /note="Left Inverted repeat enclosing pCT0021"
                /evidence=not_experimental
                /rpt_type=Inverted
```

```
gene            28770..30071
                /locus_tag="pCT0021"
                /note="CDS inactivated by mutations introducing
                termination codons located at positions 25810, 25840,
                25906, 26173, 26221"
                /evidence=not_experimental
CDS             28770..30071
                /locus_tag="pCT0021"
                /note="Similar to transposases used to facilitate
                insertion sequence transposition, Enterococcus faecium
                (EAN11060.1  GI:68196635), Staphylococcus aureus
                (CAD33702.1  GI:21436688)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MYHCTRKLLGLTDENLFFEEEWLETVEEDGFRTNLIHAKLSYIL
                SHCRKCGIKNEGQIIKNGSHKTKVQLLPYRATKTELRLVRTRFYCKECQSTFNAQTNL
                VDENCYLSKELKVQIALELAKNTIKKELPIAILYQT*SFCVSCTPV*KRIILVLIRYL
                RSFVLMSTNR*SPAVEKMSFVFMNGQTQQLIGVLENRRLTFLKPYFLNFTRKARANVK
                YVVMDMNAPYFELVKAVFPNAKIVTNRFHIVKQITRTLNQL*IKTMNRFQKTEPTKY*
                RLKRFWKLLLTHAYDLDSSDYQYDRFFRRPMTQKAMVDELLSYDEQLTRAYETCQLLL
                YHFKHKDNQSFFDTINSLDQCLPQWFCKKLTFLNKYKLGIQYALKPRYSNGALERTNN
                KIKVIKRVAYGYRNFHNFRARIYLIQGLIFQVKQKPVKHSA"
repeat_unit     complement(30104..30134)
                /note="Right Inverted repeat enclosing pCT0021"

gene            complement(30269..30865)
                /locus_tag="pCT0022"
                /evidence=not_experimental
CDS             complement(30269..30865)
                /locus_tag="pCT0022"
                /note="Hypothetical protein similar to YdhK, Bacillus
                subtilis (G69784  GI:7474963)"
                /codon_start=1
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MNTVKKMIKLVFASLGIIVFLGACSNQSESNNSKSTNEESTSIA
                SSEMNSMEGMNHEGMVPSSMKDAANPKFPVGSNVILLGDHMEGMKGANAQVVGAFDTT
                MYEVSYEPKTGGPMVNNHRWVVQEELKDTETVANEGDTVILNADHMDGMMGAEAKVDK
                SITGTVYVVNYTPTDGQEEVKNHMWVTEDEMEYDENNE"

gene            complement(31107..32219)
                /locus_tag="pCT0023"
                /evidence=not_experimental
CDS             complement(31107..32219)
                /locus_tag="pCT0023"
                /note="Similar to two-component sensor histidine kinase,
                Clostridium perfringens (NP_563247.1 GI:18311313) and
                ATP-binding region, ATPase-like:Histidine kinase, HAMP
                region:Histidine kinase A, N-terminal, Clostridium
                thermocellum (EAM44605.1 GI:67848970)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MVWELTGKDLKSAQEKLKKNALKVSEKNSVKLDETIEVKKKLIN
                DGNEFGKVTFYYFGPFAYTEHDALFISSMKQSLMYVAIAALLVSFILASWISARLGLP
                LKHVSDFTHKLTRGEYADKIPQETSIIEINSLIDSLNDLSNQLEKQHGLRKRLTTDIS
                HELRTPLATLKGNVEGMIDGVWKITPERLQSCYDEIDRLTRLIGNIEIINKIEAKYDH
                LNKTEFNIYKLIESVIENFASKIESKNLHVEIQGDNINISADKDKMNQVVINLLNNAI
                KFTQKEGTIKFSISKNKDHVLLIVEDNGIGIEKDQQLHIFDRFYMADPSRSRALGGQG
                IGLAIVKSVVEAHKGSITVKSKLGLGTKFVIKLPFQ"
misc_feature    complement(31710..31757)
```

```
                /locus_tag="pCT0023"
                /note="LTTDISHELRTPLATL; Region: Histidine kinase H-box
                motif"
                /evidence=not_experimental
misc_feature    complement(31383..31442)
                /locus_tag="pCT0023"
                /note="DKDKMNQVVINLLNNAIKFT; Region: Histidine kinase
                N-box motif"
                /evidence=not_experimental
misc_feature    complement(31263..31334)
                /locus_tag="pCT0023"
                /note="VLLIVEDNGIGIEKDQQLHIFDRF; Region: Histidine kinase
                D-box motif"
                /evidence=not_experimental
misc_feature    complement(31119..31223)
                /locus_tag="pCT0023"
                /note="QGIGLAIVKSVVEAHKGSITVKSKLGLGTKFVIKL; Region:
                Histidine kinase G-box motif"
                /evidence=not_experimental

gene            complement(32502..33182)
                /locus_tag="pCT0024"
                /evidence=not_experimental
CDS             complement(32502..33182)
                /locus_tag="pCT0024"
                /note="Complete; Similar to prokaryotic cytoplasmic
                two-component response regulators, Thermoanaerobacter
                tengcongensis (AAM25604.1 GI:20517480), Clostridium
                perfringens (BAB82038.1 GI:18145997)"
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MKILIVDDEEKILEIIDAYLVANHYSVYKAISGAMALEKFEKNN
                PDLVVLDLMLPDTDGLTVCRKIREISNVPIIMLTAKSDEDDILTGLKLGADDYMLKPF
                SPKELVARIQTVLRRTETLSNPNKLSSNNKELIVYPDSRQVYLHQVELNLTTTEFDIL
                HALMSTPNKVFSRSDLIEKVKGIEFDGLDRSIDSHIKNLRHKIENDAKSPMYILTVHG
                TGYRFGGH"
misc_feature    complement(32838..33179)
                /locus_tag="pCT0024"
                /note="Conserved N-terminal regulatory domain; PS50110;
                Region: Response regulatory domain"
                /evidence=not_experimental
misc_feature    complement(33015..33017)
                /locus_tag="pCT0024"
                /note="Phosphorylatable aspartate residue; phosphorylation
                site"
                /evidence=not_experimental
misc_feature    complement(32514..32786)
                /locus_tag="pCT0024"
                /note="C-terminal effector domain contains DNA and RNA
                polymerase binding sites; cd00383; Region: Response
                Regulator trans_reg_C effector domain"
                /evidence=not_experimental

gene            complement(33210..34796)
                /locus_tag="pCT0025"
CDS             complement(33210..34796)
                /locus_tag="pCT0025"
                /note="Similar to multi-copper oxidases. Bacillus
                coagulans 36D1 (ZP_01696949), Staphylococcus haemolyticus
                JCSC1435 (YP_252021)."
                /codon_start=1
                /evidence=not_experimental
```

```
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MWGDRHEISTYFFYGKYIIDMGDDGLKNNKYIWLAATTISIFTL
                GGIFILIQQQNKEEHAEVIYQEEQRQISSRMSNMMGDNQEADLKVVHQTDRKLLIPPV
                LEPSSKSKENVTYDIVTQNGEVQIMDGEKTETLGYNGDFLGPVIRLKKGQKVTINTTN
                NLDASTSFHWHGLKVASDADGGPHQIIEAGQKKSVTFEVDQEASTLWFHPHPEGETAS
                QVYKGLAGLMYIDDGNSKLLDLPSKYGVDDIPLIVQDKSFSSTNQINYENDFNSDGTK
                GETLLTNGTINPYVEIKSRWMRYRIVNGSNARNFTFNLDNDESFYQIATDGGFLNTSV
                KLSKLLLAPGERAEILVDTQNYKKGKVIHLLANNLVALTMRIENTIDNKEFNPTDSLN
                TISTLDEKKLEDLTRQSINLSGMSHMVNINNKQFDMERIDLYKKLGTQEIWEVNNISS
                MMGGMIHPFHIHGVQFQILSRDGNQPALNEQGWKDTVLVNPDETVELLVKFDREGIFM
                YHCHILEHEEYGMMGQMEIK"
misc_feature    complement(34092..34445)
                /locus_tag="pCT0025"
                /note="Copper oxidase-like domain not recognised by the
                pfam00394 model; pfam07732; Region: Cu-oxidase_3,
                Multicopper oxidase."
                /evidence=not_experimental
misc_feature    complement(33219..33533)
                /locus_tag="pCT0025"
                /note="Divergent copper oxidase-like domain not recognised
                by the pfam00394 model; pfam07731.  Contains the consensus
                signature pattern: H - C - H - x(3) - H - x(3) - [AG] -
                [LM]. The first 2 H's are copper type 3 binding residues]
                [The C, the third H, and L or M are copper type 1 ligands;
                Region: Cu-oxidase_2, Multicopper oxidase"
                /evidence=not_experimental

gene            complement(34807..35238)
                /locus_tag="pCT0026"
                /evidence=not_experimental
CDS             complement(34807..35238)
                /locus_tag="pCT0026"
                /note="Truncated.  Probably inactivated open reading
                frame.  Similar to heavy metal translocating P-type
                ATPase.  Desulfitobacterium hafniense Y51 (YP_518797),
                Desulfitobacterium hafniense DCB-2 (ZP_01369576),
                Pyrococcus furiosus DSM ~3638 (NP_578469)."
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MISKYNLYGMTCAVCATTIEKKIHELDGVYFAKVNLTTEVLKLE
                YDEGVLSNHTVITAIQDIGYDAEIRKKTEIKVFGISGMNCSGCATKVRNALEAEPTVS
                VKIVDLARGTVTFDADSKLTLTFLNQLLKNTKYTVTRDISI"
misc_feature    complement(35038..35229)
                /locus_tag="pCT0026"
                /note="Heavy-metal-associated domain (HMA) is a conserved
                domain of approximately 30 amino acid residues found in a
                number of proteins that transport or detoxify heavy
                metals, for example, the CPx-type heavy metal ATPases and
                copper chaperones. HMA domain contains two cysteine
                residues that are important in binding and transfer of
                metal ions, such as copper, cadmium, cobalt and zinc;
                Region: HMA Heavy metal associated domain"
                /evidence=not_experimental

gene            complement(35222..35587)
                /locus_tag="pCT0027"
                /evidence=not_experimental
CDS             complement(35222..35587)
                /locus_tag="pCT0027"
                /note="Similar to CopY transcriptional regulator protein.
                Lactobacillus acidophilus NCFM (YP_194787) and Clostridium
```

```
                beijerinckii NCIMB 8052 (YP_001310197)."
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MVLVEHKIQVSNSELDVLKFIWRYEPVTCGLITHGMQERNSWHP
                STTKTLIRRLLDKNVITFNTSKNQRYYTSLIKKQEFLETEIQRLLSGMDEGCISEVSY
                YLNGLVKRSDTEEFNDFKI"
misc_feature    complement(35288..35560)
                /locus_tag="pCT0027"
                /note="Penicillinase repressor. The penicillinase
                repressor negatively regulates expression of the
                penicillinase gene. The N-terminal region of this protein
                is involved in operator recognition, while the C-terminal
                is responsible for dimerisation of the protein; pfam03965;
                Region: Pencillinase_R"
                /evidence=not_experimental
misc_feature    complement(35417..35464)
                /locus_tag="pCT0027"
                /note="Region: WHPSTTKTLIRRLLDK; N-terminal DNA binding
                domain"
                /evidence=not_experimental

gene            complement(35603..36217)
                /locus_tag="pCT0028"
                /evidence=not_experimental
CDS             complement(35603..36217)
                /locus_tag="pCT0028"
                /note="Similar to integral membrane proteins from Gram
                positive bacteria.  Enterococcus faecium DO (ZP_00603327),
                Lactobacillus salivarius subsp. salivarius UCC118
                (YP_535330), Clostridium phytofermentans ISDg
                (ZP_01354135)."
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="Hypothetical protein"
                /translation="MKKSMLFLLFFVFALSLAITLTINFPGLLKINLFFQSDINFQNW
                SRSQVNQDFRNLMGYLNNPFQKELVFDNLYVSDRGINHFKDVKFLFQLNYSLLLSTSS
                ILLYLNRKKLVTRDQVREITSLIKWMIISVCVMALLFFDKAFVLFHQVFFDNDDWMFD
                YRTDPIISFLPETFFFLCFLLIVTISVSTLTTIHHLFNKEERTL"
sig_peptide     complement(36164..36217)
                /locus_tag="pCT0028"
                /note="Putative signal peptide cleaveage site between
                position 18 and 19"
misc_feature    complement(35693..36163)
                /locus_tag="pCT0028"
                /note="Protein of unknown function (DUF1461). This family
                contains a number of hypothetical bacterial proteins of
                unknown function approximately 200 residues long. These
                are possibly integral membrane proteins; pfam07314;
                Region: DUF1461"
                /evidence=not_experimental

gene            36787..37176
                /locus_tag="pCT0029"
                /evidence=not_experimental
CDS             36787..37176
                /locus_tag="pCT0029"
                /note="probably truncated or inactivated.  Similar to
                prolipoprotein diacylglyceryl transferases. Bacillus
                cereus ATCC 10987 (NP_982198), Bacillus amyloliquefaciens
                FZB42 (YP_001422779), Bacillus sp. NRRL B-14911
                (ZP_01171649)."
```

```
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="Hypothetical protein"
                     /translation="MNQEAYGEIVDIEFLQSLKLPSFIIDQMYIDGAYHHPTFLYESL
                     WNIGVLIILLLVSRNRMFFGQIFLIYLSLYSVGRFWIEGLRTDSLMLTANLRMAQVLS
                     IVLLIGSILTYIYLKKSKEEDLHGSIT"

BASE COUNT    11763 a   6600 c   6101 g  12812 t      3 others
ORIGIN
        1 aaaagtcggc gtttgatcca gcactttcat ngatgccgtc agcatcttcg caacatcgac
       61 ttgttgtgta tcgacatcga ccacgaatgt attgatttgt tgcaagtttt gttcgtcaaa
      121 accttttatg tattgccgtt tatcgtccgt atatgttccc catctaaaaa cgtttggtgt
      181 ccaatgcgtg attttatctg cttctgtcag cagcgcttct tggctcgtaa caataaaccc
      241 tttgacgcca tgctccatct gcagatgacc tttatcgcga acgacaaatt gcacgccttt
      301 acgattttcg cttgctgctt cgacagctgc atagtgatct cgaatagcca aaacgccttt
      361 acttcccttt ttccgaaaag cagtgattcc ttcatgcacg accgcatcga tcacgctgct
      421 atattgatac atcgcttcgt tgtcaaacat ctgatcctct ctcctctctt cgttgattct
      481 cacttttagc aaacaaaaaa ggacaagatc cctccctgca gagaaatctt gtccttttaa
      541 acaaaaatct cgcgcaggaa aaggaaactt tttcacatgc acgaaacgtt caaaaacctt
      601 ataaaaatgc ttgaaaactt gcttttaagc attgtcaaat aagcctactt atgctataat
      661 tttttatca attttaaatt aatagcttgt agggtttcgt ctttttaatg tgtttttggt
      721 cgaacacatt aaagtgcatt ttggtcaatg catgggcgat tcctttttta tttacttgtt
      781 gtttacacaa tacaacaagt tttttttttt gtaaactatc atttttctc acaataagct
      841 tactcctata taatatttga taaaaagata taaatatctt tttatctttt tatctatata
      901 tacttttatc tttttatctc tttatctttt tatctatata tctcattatc tttttatttt
      961 tttatctata tatcttttta tctcctgtaa agtattgctg ttatgcaatt tacaattaat
     1021 tttataagtt actatgttac tttatactac aaataattta tataacctcc gaatatttat
     1081 ctttttttcgc tttacataaa attcaaggaa tgttatcatg agtttatcag ataaatttat
     1141 ctttttatcc ttttatcttt attacaaata tattacagga ggcgagagaa tgtttaatag
     1201 tccggcttat atagaagaat ataaaaaagt gcagaaagta ctaagtgata atcgatgcag
     1261 agttattatt aatgcaaatc aaaaaggtgg agtagggaaa actaccaaca ccactatgga
     1321 ggcaattatt cttgcgcatg tctttaataa aaaagtttta ttgattgatg aagatatgca
     1381 aggtaatgaa actagtttca tgagtaaaac ttatgatgtt atagagttcc ccatgtcact
     1441 aatgaaagca attgaagaag gagacttaac aaaagcgatt gttaatttag acgaaaatat
     1501 tgatattatc cctggtagtt acgatatgag aaaaatggtt aattttttaa ttggtaaatt
     1561 caaaactgaa gaagcacaaa ctttttatct ctcttcactt atagacaaaa ttcgtgatga
     1621 ctatgatttt atttttattg atgttccacc ttctacagat ttaaaagtag ataatgctgt
     1681 agttgctgct gattacttag ttgtggttca agaaactcaa caatttgcat tagaaggctc
     1741 aaacaccttt atttcaactt atcttaatac cttagtagat gattttggca gtcgcttcaa
     1801 aacagaaata attggtattt taccagtttt actacagaaa aaaagaaaat tgcatgaaaa
     1861 aatactagaa actacaaaag aaaaatttgg taaagaaaat atatttggaa ctacaataaa
     1921 caatcacgca agactagagt tatatggtaa atttggcgtt caatttgaag atcattggga
     1981 tcgaacaatg tttgcactct atactgatat agtacaagag atgcttgaac gtatgcagct
     2041 tattgaaaac ggcagtgatt tagataacta tcaatataaa ccaatttttt acaatccaaa
     2101 aattggtaaa gtaactgctt tgggaaagga gattgttgtt aatggctcga ttaatagatc
     2161 gtaaagacaa ggaacaagca caagtaaaaa ctcaagcaga tacatcgaga aaagttaata
     2221 caaagtccaa ctccgaaaaa aaagttaatg caaaagacag gaaaaatatt aaggttaacc
     2281 ctgacgatat gttattaatt gatacaatag ctaataattt ctttgaaaca aaattaaaac
     2341 attatgaatt ggtaaatctc ttattagaaa actgggtgga tactaagcta gaaccaagac
     2401 aacaaaagat acttaaggat ttattgaaaa acagatagtt ttatctaaag ataattagat
     2461 aaaaagatag ttttatctaa ttatcttttt atcctttatg cccaccgacc aaatcagatc
     2521 gtttcaatcc tgtagccccg gctactaagc tattggcatg catcaatgct ttaaatccat
     2581 aacgatcacg aatcttatca atcgtacgct ctagctgttc ttgctggagc gtttttgtat
     2641 agtcttcaaa tacgctcagc tgcatcgatg tttttctttt gatgccacca caactaaccg
     2701 caactgatcg caccgcatca ttttcctgat acttccaaaa catttctaag aagtagggga
     2761 caagctcctt gctactgtcg gttgccatta tctttttttg gtggctaaaa ccgttccgag
     2821 tggaatatcg gctgtagccg gcgcttagat gaatgacgct ggtatcaaca tgatgattgc
     2881 gtaggcgcat cgctacttct tccaccattt cttgaataat cacggcgatc tgttccggat
     2941 catggtaatc cttttctagg atttggctct ttccatacga cttctcaagt ggcttatact
     3001 tttctgacag acggctataa tcgatcccat gcgcatgata atacaattgt tccccaatca
     3061 cacctaaacg tctttttaag taggagggga gcatctggct caatccatgg atgctataga
     3121 tccccattcg ctccagctga aaagcagtcc tttttccaat tccccaaaaa tcagtgatcg
     3181 gtcggatctt ccacactgtt tccggtacat ctttatatcg ccattccgca ataccagatt
```

```
3241 tttgatgttt ggcttcatta tccaatgcta attttgcaag aaggggattg tctccaatac
3301 caaccgtgac atacagcttt aactccttta aaatatctgc ttgtatcttt tttgcgatga
3361 cttcagcgga tccaaaaagc tttttactgt gcgtgacatc caaaaaggct tcgtcgatag
3421 aatagacatg aaaatcttcc ctggatacat atcgaaggaa gattttctgg atgagctcat
3481 tcacttttaa gtatagttgc attctaggtg gagcaataat gatgcgagga tcgtcctctg
3541 gcacttcaaa ttttctagat cctgtcttaa ttcggaaaac ttttttcatt tgaggacttg
3601 ctgcaagcac caatcctcca gcgcgatctg cattactcat gacaacaaga aaagcctgga
3661 gtggatctaa cccacgctta acgcattcca cactggcgaa aaacgacttc acatcgacac
3721 aaagaatatc gcgagagggt acgtgggtgt aatcttctat ctgtaccatt tttctattgt
3781 cccttctgaa atatgattaa tatcgtcaaa tgtcaacacg gtgatatctc cggccatatc
3841 tctaatataa attcgatcat cttccacacc agctattctt cccatgacaa aggggagggg
3901 atagccgttt tcatcctttg ttagtcgttg aatcgttata gcatgctgct ttgtaaccgc
3961 atagcttaga atttgtacca ggcattcttc agtttgctgc ggtagccagg ataattttct
4021 ttttgcttgt tctagttgtt ctgtatgttc agatagaatc atgccttgcc acttttgcat
4081 gcctcgatct cgataagcca ttgaaatgac ctcctccaac aagaacaaat gttctattta
4141 tgtattaaat tatacaaaca aacgttctat ttggcaagtg aattatatag aaaagcgtgt
4201 ttttaaaata attacttcaa catgtctgct tattagtgat aggcaaacaa aaaacactcc
4261 attaggtgga gtgttttcag tcagagtata atagtaactt atagattatg ttcccaatat
4321 attcttaatg atttttttcta tttcttcaga cctctcgttt ttaatttgat catttatatt
4381 agtcaattct tcactgttat ccacgctaaa atccgcaaaa ataatgggaa cttgcttctt
4441 caaacttttt gcttctttat caaagttgtt atttttctcc aatcctgttt tcacaaaaat
4501 aacttggagt gattcattgt gcacttttgc tctaagaata acttctttca ttagacttag
4561 aatctttata ttacttatat tttcagcata tggaaagata atgaaatccg acttatcaag
4621 gtataagctt tcatcttgtt ttacatatgg actttggtct aatataacat aatcactctc
4681 tatcttagca gggaaggcct catagtcgag atcgacatat agactatttt tcccagactt
4741 atatagtgct tcattaataa tctttgctgt cgtcgtttta cctacgcccc ctttgaaaga
4801 tagaatattg ataatcttct ttgtaccttt taaggtgttt cttgaaattt ttgcttctaa
4861 ttccccttga ccagtcaagt actccattcc ttccttcgta atgacatatg tcttcccaaa
4921 ttgtctaatt gttccgtcta aaaatcgatt cggatatttt ctatataact gccttacata
4981 actgctatca aaccccccaac gttcactcgc aactgccgca ctcattaaat cctcgtccat
5041 taaatcatat ttccttctca attcattcac ctcatcataa ctatacccta tatagggtat
5101 tgaatcagcc tattttgaat ctaatagaga acgattcatt tttttcgagc aataaaaaag
5161 aaagaaatcc aattatacga cttcttcctt tttcttttaa gcgatctttt caagaacttg
5221 tacctaaacc agctgtaact ttgaattaaa tatcacattt gcaccattga atttttccta
5281 ttagccaaaa aaaataatat ggcaaagtct actttcttta tattataagt aattaggcaa
5341 acaattaagc cctgtcaatt taaataggag gtggtaataa ctaatattta gaaatacaga
5401 aaggagggaa tacttcattg cttttattta tcgcatggat gcttatctta atatatttttg
5461 atgggaaatt ttaggtgtga ctttaaccga taattttcac tattttatat agctataatt
5521 catgattgtt tatagccgaa gccaacttct tgtaagtata aaggtagttg gcttctagcc
5581 aaatattatt cagtaaatcc aataattagg gactctctag cagcttgaac tacatctgaa
5641 gatatcttct caagattatt gatttccatt attctatcaa tttgagtaaa taagcgctgg
5701 agtaatcgga aattgccatt agttagctta ataatgcgat taatagcatc ataatcttca
5761 aaattttcat aatcaattgc tagacctata tcttggcatc tatattctaa tatatgcttt
5821 atttctgaag aacttaactt tttaaactca tgtgcaaatc caattctaga gtataattga
5881 ggatatttag atagtctttt ttcaattcct tgcattccta taaagaccat accgacatca
5941 ttttgatcat aaatgtctct tagaagatcc aacgtttgga tcttcaaacg atctatttca
6001 tcaataatta ttaaatcaac cgcataaaat atattttcca cgattcgatc tatactagtt
6061 gccatgcaat ctattctatc agctttaggg attactttat tttctgtact tgaaattatc
6121 tcatttacgc gtgctatatc tattgcaaga ctgcttaatt gagaactcaa tgaactgggt
6181 ttagcggcag gtgctgtcac aaatagagca tcggttttta tcccttttaa cttatctatg
6241 actaattccg gattgagtct attacctttt ctagttaaaa tagtttctaa tctatgccag
6301 tttgcataat gctttgctga aatggtcttc ccgacacctg gattgccata acatatgcca
6361 atatatttat atttcatgca tgctctacaa aattcggcaa atttctttta ttctttagtt
6421 tcaataaagt tgagttcatt cattttcata cctcctcaaa cttgatacat tatttttaac
6481 agtactacta ttctcttttt gtctcccaac caatttttct atcactttat tagtggagac
6541 gattagttcc ctttttaact ctcttttcag tttgctcctt tctttttcta tttctctaat
6601 tccgatagcg tattgttcaa aagatgtaga cactgcggta caaaggaata cattttttata
6661 aaagactcta acatcactaa tatctctagg atcatatcta atcaccacag gttcactaac
6721 aaaagctgaa agagttgggt ggacatactt catcccaaat aagtgaatac cgtcagagtg
6781 taccacccgt tctttcttaa cttttattaa aagaagatct aaatcctcca aagagtttgg
6841 catattcggt agaaaatcat agttattcca catagctata ggtttttctt tagttgtact
6901 atgctgtttt tgatggtata ttcttaacaa ccaatttctg aaaatatcgg aaaactcctc
6961 caatgttagc attgactttg ctttccgatt cttaatataa cctggaagat cttgtagtaa
7021 catagaattt acagtatgaa aaaaccgttc gattttccct ctacctctag gcaccccaac
```

```
 7081 cttagaaaac tttattctat tttcaatcag cagagacctg ttcaagatgt tgagaagtaa
 7141 aatcgctacc gtgatctgta aaaaaattct agggatacca caaataggcc agtcagcctg
 7201 atcttttctc catattgctt gtcttaaagt tagagctgtt cgtagtgagt ttggtgattc
 7261 aaacgcaata aagtacccag taattgctcg actgtagtca tccattatta ttgttagcca
 7321 aggcctttca ggataacctt tcgcattcaa tacctcaatg tctaataatg tatgatcagc
 7381 ttgccatatg atattgggag ctgtagtttc cctaatatga actaattcgt aatgattcat
 7441 atatatttta gatccctgat gcgctagtac tctcaattta tcgggaacta attgaataat
 7501 tttccttacc tgatagtaag aaggctcctg aatattattt tctttacacc aagacacagt
 7561 tctcctataa atggaagtta tagaaacatt tttattttct aagtaaatgt gtttaatgat
 7621 tttacaaagg tcttcagaaa ttttcattga atttttatcc gacctttctt ttgtaattaa
 7681 tcccaccaat cccgattttt tataattcga aatccacttt ctaagtgtac gtgttgaaat
 7741 atcattttct tcacttactt tggaaagtgg aatttctttt tttaagtatg gctgaagtac
 7801 ccgaaatttt tttaatgctt ttaatctctg ctcctctgag taattagcta aaattttttt
 7861 attactcatt tcctcacctc ctgatatgaa caaattattt catatatcag attatatgtc
 7921 taattagggt ctaccctaat tagacatata aaacagattg aaaacaaacg atttttagtg
 7981 tatgtttagg gtgtgtgata aacagtcaga aaataggaag tgggaaaatg aaaattgctt
 8041 atgctcgtgt ttcaacttta gaacaaaatt taacacgaca aatcgaaaca ctcaaaagtt
 8101 ttggtgctga aaaaattttt acggaacaaa agtcaggtgc ttctattaaa aatagaaaag
 8161 catttcaaga agctttgcaa tttgttcgaa gtggtgacac ttttatggta gaagcgattg
 8221 atcgattggg tcgtaattat gatgaaatca tccaaaccgt ttattttcta aaagaaaaag
 8281 aggtcaaact ggttattacg agtctgccta ttatggctga agcaatagga aatccattat
 8341 tggataagtt tattaaagac ttgattatcc agatattagc catgattgcc gagcaagaac
 8401 gagcagagag taagagacgg caggcacaag gcattcacag tgccaagcaa aaaggaattt
 8461 acaaaggaag acctattcta tattcgccta cttcaaaaga tccacagaaa cgagttatct
 8521 atcagagagt cattgacatg ttaagtaaag gagagtctat ttcaaaaata gctgaagaaa
 8581 gtggtattac aagacaaacc gtatatcgaa ttaaaaattc gaactaataa actaaaagga
 8641 tctcttcagc caataaaaag tagccctttt gatctaacag caaaagggct acttatatta
 8701 ttatacgtta ttatatttag atgaaaatta tcaataaaaa gcggtgaaat ttattaatgg
 8761 aagtcacacg aaatttaata ggaaaagaag agtgttcaga aaatgaccat tttatgaaca
 8821 ttcaaacatt cacttgaata tagatggatg ttctgatata atacattcaa acaaacattt
 8881 gaatgtaaag gaatgatcgt atgaataatg aaatttgtga aataacttgt attcatgaag
 8941 acaaagttaa tcgtgctaaa tctaaactag caaactttga tactccgtca gtaagtggtt
 9001 tctttaaaat attgtcagat gaaaatcgac ttaaaattgt acatgcttta gttcatgaag
 9061 atgagctatg tgtttgtgat attgctaata tcattgatgc ttcggtcgca acaacttcac
 9121 accatctgaa ctccttaaaa aaattgggag tagttgatag tcacaaggat ggtaagttag
 9181 tctactattt tattaaaaat atcaaaatct taaatcttat ggagcttggc gtaaatttta
 9241 aagaagaggt gctagcatga gtgagaaaat gaccgaaaaa acctatcgaa tagaaggatt
 9301 aagctgtaca aattgtgcag gtaaatttga aaaaaatgta aagcagttgc caggagttac
 9361 gagtgcaaca gtaaactttg gtgcaagtag gatttcagta gaaggccaaa caacaattga
 9421 agaactcgaa gaagcgggag ccttcgaaaa tcttatcatt cgtgatgatc aagaaaatga
 9481 tgaacaagtt agatcaaaag agtcttttat taaacggaat atcgctttaa tcatctctct
 9541 tggttttata cttgttgctg tgatttccca actatcattg ggcgaagatc atttattaac
 9601 gaaagctctt tatatattgg caataataat tggtggattt gatttattta aagaagggtt
 9661 ctctgatttg atcaaattgg acttctccat ggaatcttta atgactattg caattattgg
 9721 ggccgcattt attggagaat gggccgaagg ttctattgta gtaatactat ttgctattag
 9781 cgaagctttg gaacgatttt caatggataa agcaagacag tcgattcgtt ctcttatgga
 9841 tatagctccc aaagaagctc ttatcaggag aaacaatgta gagcaattag tcagtgtcga
 9901 taaaatcgat atcgacgata ttatgatcat taaacccgga caaaaaatag ccatggatgg
 9961 gcttgttatc aacggtcatt cttcagtaaa tcaagcagca ataacaggag aatcggttcc
10021 tgttgaaaaa cagttagatg atgaagtgtt tgcaggaaca ctaaatgaag aaggtgttct
10081 tgaagtcaaa gtaacaaaaa aagtaacgga tacaacgatt gccaaaatca ttcacttagt
10141 tgaagaagcg caaggagaac gagctcctgc acaagccttt gtagataaat ttgcaaaata
10201 ctatacgcca ttcataatta ttatggccct tttgatagtg gttgttcctc cgctattctt
10261 tggtggtgac tggaataaat ggctttacca agggttatcc attttagttg tcggatgtcc
10321 ttgttcatta gtgatctcaa cacctgtttc aattgtttcg gctattggta atgcagcaaa
10381 aaatggtgtt ttagtaaaag gtggagtcta tctggaagaa attggtcatt tgagagcaat
10441 tgctttcgat aagactggaa ccttaacaaa aggaaagcct gtcgtaacgg atttcattgc
10501 aaccagttct gaaactgata taaattatct atcgattatt tcatcgttgg aatctctttc
10561 tcaacaccca ttagcttctg ctattttaaa cgaagcggat aaaactaatg tggattacaa
10621 gtcaatacaa atcgaggatt tccaatctat cacaggtaaa ggacttacag gtatacatca
10681 aaacatacgt tactatattg gaagccctaa actgttttct gcatcagtaa tcgaagaaac
10741 agcggtcaaa gtacaataca gacaatttca agaacaaggg aaaaccgcca tgtatttcgg
10801 tactgacgaa caaattttag gagtaattgc tgtcgcggat gaagtaagag attcaagtgc
10861 tgccgttatt tctgaactcc acaagttgtc aattgaacac actattatgc ttactggtga
```

```
10921 taatactaaa accgctgaat ctattggtaa gcaattaggt gtgactgaaa taaaaggaga
10981 tctaatgcct caagaaaaat tagattcgat taaagcgttg cggacaactt ataataaagt
11041 agcaatggtg ggcgacggca ttaatgatgc accagcatta gctgcatcga cggttggtat
11101 tgcaatgggc ggtgcaggaa cagatacggc tctcgaaact gctgatgttg ctctaatggg
11161 cgacgattta caaaaattac cattcattgt tagacttagc cgtcaaactc taaaagtaat
11221 taagcaaaat atcaccttct cattaggtat aaaactatta gctctactac tagtcatacc
11281 aggttggcta acactatgga ttgctattgt agcagatatg ggagcaacgc ttttagtaac
11341 attaaacgga ttacgattga tgaaagttaa ataacacaac aaataaaaga aggcatgttt
11401 ataagcgtgc cttcttttta tgatactttt tagttttttt gaacgctagt tatgaagcaa
11461 actatgcgcc atttaaatat taaaatgaca gtctaacctt attaagtagg tgttcccatt
11521 aatactttga tttccttaga taatagaaag ctccagagaa ataaaataag atcttcaaat
11581 tataaatttc atctctttaa aaattattat tttgaatttt taaagagagc ttttatgagc
11641 accaccatta taaaagctaa aaacaaagct actacaccaa cgacaacaaa agaaagtata
11701 atcagcatag ctatcatcac taaacatttg gaagagtaaa aacgataaat gctgcaataa
11761 agctaattaa aaattgaagc catatctggt ggatatattt ttggaggaga aaaatatatc
11821 catacatgcc tattaaaagt aaagtaaaat ctatattttg aagtatagga tttccttgga
11881 tgatatttat aattagagtg gtaaatgtga cggttctgtt gcaaagtttt atttaactga
11941 ttatttagtg tagaatggaa gaaactacac gaggagaata acaatgtccc attttaaagg
12001 taaacaattc aaacaggata tcatcatggt agccgtgggc tactacttaa gatacaattt
12061 gagctatcgt gatgtttcag aaatccttaa tgagcgtgga atcaaggttt gtcacacaac
12121 tgtttatcga tgggtacaag aatatggttc gattatttat tgcctatgga agaaaagaaa
12181 taagtctgct agcgattctt ggagaatgga tgaaacatat attaaagtaa aaggaaaatg
12241 gcattatttg tatcgtgcca tcgattctac aggtttgaca ttagatattt ggttgcgaaa
12301 gaatcgggat aaacaggcgg cctatgcttt ctttaaacga ttaatcaagc aatttggaga
12361 acctagagtt tttgtaacgg acaaagctcc ttcgctaagc agtgcgttca atcgacttaa
12421 atctgaaggt ttattttcta acactgaaca tcgaacgagt aaatatttga ataatatcat
12481 tgaacaagat cataggccta tcaagaaaag gcacaaatta tatcagtcca taaggactgc
12541 ctcatccaca atcaaaggaa ttgaaactat acatgcttta tacaaaacta gccaaagaga
12601 cttttctctc tttggctttt cagtgattca tgaaataaat aatctaatgg gtgttccagc
12661 ttaaaattag aaaatagttg agatatcatt ttacttcttt gaactttgca acagaaccca
12721 tcattttact tctttgaact ttgtaacaga acccaaataa aatataataa aacgcatcat
12781 agcgtgatta ttttcctagt gttttctttc cgaacagagc ggtggtggct ggaacgccgt
12841 gaggaaagta tcctatggtc ttgcctttaa ataacgaaag aagaaaatat tgtttttaaa
12901 atggctattg tggtataata acaatgaaaa gagagctatg cttcaacata gctctccagc
12961 agtagcaccg ttaaaaagac ggtaaccata gatatattaa gaagtaaccc tctaacacac
13021 ccggtcaaaa gcgagttaga tggttatttt ttttgctcaa agcctagtat cagcacgact
13081 aatgtcgcaa aagcgatcat cagcgtaagt gcttcaaaca cagacaagtc tagctccttt
13141 ccaggagtaa aactatacat tcaagcatag gcatcacccc ctaatgagga tttagccacc
13201 atctttctca ctttgctact gtatcaatta tagcatatca aatatttaat gataagtaat
13261 aaaatagtag aatggatcag agataataaa aaatataaat gactttggag gtatatatgc
13321 aaacggatat agaagttaat aagcaagaag aacttaaaga agaattgaaa aacttcattg
13381 ttaatgtggg gtggacgcat aagatacata tagtaagatc agatgagtat atatattatt
13441 ctaaattatt taaatggata cgtattatat tagctgcttt tacttcttct ggtttgattg
13501 gtatattatt tgacagtcaa tcatatattc ttaaggtggt aacagctata ttttcattcg
13561 ctacagtagc tatgaatggt ttagataaag caggagattt tgagaaactt tgtctaaagg
13621 aaaagaatga tgctaataat ttttggaagt tacgaactga tgcagaatcg ttattaagta
13681 agactgtatt tagcacggaa agtatgggac gcattgacca ggagtttgaa gagctaaaga
13741 gactaagaat acaatataat gcaggattat taaatgtgcc tgccaaaaca gtaaaaaaag
13801 ctagtaagct aattcataaa cgtaaagata acgggttctg ttgcaaagtt caaagaagta
13861 aaatgatatc tcaactattt tctaatttta agctggaaca cccattagat tatttatttc
13921 atgaatcact gaaaagccaa agagagaaaa gtctctttgg ctagttttgt ataaagcatg
13981 tatagtttca attcctttga ttgtggatga ggcagtcctt atggactgat ataatttgtg
14041 ccttttcttg ataggcctat gatcttgttc aatgatatta ttcaaatatt tactcgttcg
14101 atgttcagtg ttagaaaata aaccttcaga tttaagtcga ttgaacgcac tgcttagcga
14161 aggagctttg tccgttacaa aaactctagg ttctccaaat tgcttgatta atcgtttaaa
14221 gaaagcatag gccgcctgtt tatcccgatt ctttcgcaac caaatatcta atgtcaaacc
14281 tgtaaaatcg atggcacgat acaaataatg ccattttcct tttactttaa tatatgtttc
14341 atccattctc caagaatcgc tagcagactt atttcttttc ttccataggc aataaataat
14401 cgaaccatat tcttgtaccc atcgataaac agttgtgtga caaaccttga ttccacgctc
14461 attaaggatt tctgaaacat cacgatagct caaattgtat cttaagtagt agcccacggc
14521 taccatgatg atatcctgtt tgaattgttt acctttaaaa tgggacattg ttattctcct
14581 cgtgtagttt attccattct acactaaata atcagttaaa taaaactttg caacagaacc
14641 ggaattattt agtaactcaa taatatgatc aattgcgcta ttaaaagtaa tcgtatctgc
14701 aatattaaaa tgtgtcccat tatcagtact catattttat tttctccata aatgtaattt
```

14761 ggatctatta tatcatttaa attaaaaaca gcccccatta tctattgatt agataatggg
14821 ggctgtctta atagtgtttt agaggggcga ttatctaatt atagaccctg tttttttgtg
14881 ttgagtttat tttttaagaa atcagaagtt atttgtgtta gtggttcccc gattgccttc
14941 ttataattat atccaagaag ctgaagtgtc tttaatgtat ctgaattcaa taaattcagt
15001 aatacatcga tagatttttc atctgtttta ctttcagatt gagacaatgg cataacaagt
15061 tttataactg tactgacagt gttatgcaaa ctttcgatat cagaatcgct aattgtcacc
15121 ttgtttctag tcgttctttg taatttcgtg ctgtaagttc taatttattt ttgttgtcta
15181 acaaatcatt tatcagttgt tcataatctg ttacttgctt ttttaaatca tgatttgatt
15241 taatttgtgt aattttgttg gtaacttttt ctacagttac ttcagtaatt gatttagcaa
15301 aagccgtatt taatgctaat agtgtctcat ccattgaaca tcattccttt tctaataatt
15361 caaattctgg taattcatca attaattcaa gtgttcgcat acttactgta ataacagcta
15421 gtaagagatt taaaatgtat ttaggatcat cgctgaaatc attaggatca tcagttattc
15481 ctgatttttt atctgttttt acttgatatt gatcaataat ccattcaata gcaggtcgac
15541 cattaacaat atactcatag gctttttctg gaatattttt aatcgtgatg ctatcgttaa
15601 aaataatagt atctaatata cccttttttag gatgtttcat cttcttaaca cgataatttg
15661 gtttagaaat ttcaacttca acaccttccc agataggctg atcttcataa ttcaaatgta
15721 aatcagcaag tttacgtccg atctcaacat actcttctct atgttttaaa agaggaatac
15781 gaggtaatgt tttctgcaaa tcattagcgt actttttacg atattcagaa gaatgaagta
15841 cagaataaac gtagtaaaag caatcatctt gagatagttc aagtttattt tctattaaat
15901 tactagaatc tttaacaaaa aagccatctt ttgagctatt atcgtttttg taaaacccct
15961 gtcccgagtc cataaaatga aagtttggaa ttgcattagt aatcaaagct gaaaaacctt
16021 tgcttgcacc ggcaccagta gaatatatag ataaattatt atctccaaaa atatctttgt
16081 attttcccgg actttctata atatctactt ggtattgtaa atatttttta gtaaatggac
16141 gatatatccc aataattgta gcatctggta aaaaatctat ttttttccct tttgagaatt
16201 tttgctttaa acccctactc cattttacaa agttacctga atggttcaat aaactaatat
16261 catagttgtt ttttaatcta tccaattcct gattatagtt attaatcatc ttttcagaat
16321 tatctcgaac ttttgcataa ctaaatccat aaacccatgc atcacgactg cttaccacac
16381 ctacagcctt ttcagcaaag atttctctat acagtgccaa ataattttca taggattcat
16441 cacgctgatt tatccaatca ttatttttat cgggaacaat atttttccat tccatattca
16501 aaactgattt cacatctttt agttttaaca gtttttgttc acgagttaga taatccccaa
16561 tatcatgata atgaacctca tgattatcac taccatcttt tactaaaata gaaatagcga
16621 taggtgttct acttcctgag ccaaaaattt ttccgccctc cttacgtgaa gtttctcctt
16681 gagtacgttg gtctccacgt aagttaaaga tataaaggtg attaaattcc ttatatagag
16741 aataacgaag accatctgct gactggctat taatgaatga agcgtttgat acaaaaccaa
16801 taataccatt acctttttaat ctatcgctag tccaacgaaa agcctttata aaatcatcat
16861 ataagtttct cttatttgtt gctgatgaat gtagagcata ggtttcagaa atcctattat
16921 ttaattttttg atattgaaca tttttattat catcatttgc atttgactgc ccaaccgaat
16981 aaggtggatt cccaataatt gcagtaatct gcacgtcttg ttgacgtttc aatcgttcgt
17041 cgttagtgcc aaagtaatca tcatctaatg tatcttcagt ctctgtactc tcaaaagtat
17101 ctgttaaaac aatcccctca aaaggaacat atccttcttc ctcaccgttt atttcatcaa
17161 aagttgcttc aatgttaatg gccgcaatgt aataacttaa caaaacaatc tcattcgcat
17221 gtaattcgtt catgaactta cgtgtaatat cagctaatga tatttctcca gcatccattt
17281 gttcttttag gtatgtcaac gttcggacaa taaacgtacc cgtacctgta aatggatcta
17341 aaatatggac atccttgctc gctaaagatt ttccaaagtg ctttttaagc acatcatcaa
17401 cggaatgaac aatgaaatcc acaacttcaa ttggcgtaaa gacaatccct aatctttcag
17461 tcgttgcttt aaaagcagtt ttgaagaatt tatcatacaa agtaacaata atttttttgtt
17521 tgtcttctgc tttttcaata ccttcagctc gcattctgac agattcatat aaaggctcta
17581 ggtttttcttg ttctttggca aaaccagctt tttctagctc tgatacaatc gattccatcg
17641 cttgggagac gggattttga tttacaaaac tatattccga aaatagtgcc tcaaatatag
17701 gttttgtaat taaatgctgt gagagcattt ccgcagcttg attttcatca atactttcgt
17761 tgatattatg ttgtaatgat gaaacgaact ttttaaattc taaactaatc ggatcttttt
17821 tatctgataa cttattttta atccagttaa tttgacgctc ggcaatttta gcaacatctt
17881 tcgaccaatt ctcaagatat ttacgatctc caactttttg aacaattttt ccaaaaatag
17941 ccccttcaaa tttatcccac tcaaattcaa gttccgtttg gacaggagta ttttcaatag
18001 atttatcttt gtcattcact tgatctggag cacttcctac accaataact tttaactgtt
18061 taggttttgc catgtttaac ttatcaatca ttgcttcaaa tcgttcatcg acagaacgta
18121 aagcattaat cacttgccaa accgtttcgt aattcttgtt attgtctaat attgtttctg
18181 gtttttctcc agttggaata acgattggta agataatata tccatagtct ttaccttcag
18241 ctttacgcat gatacgacca acagcttgaa caatatcaac ttgagatttc ttaggagata
18301 aaaatataac tgcatcaagc gtcggaatat cgataccctc tgttaagaag cggacattag
18361 aaacaattcg agcttcatcc gctggtttat taggatttgc taaccaatct aagacctctc
18421 ctttttggag tgcattcatt tggccatccg catgacgcat tgataaatgg atagattcgt
18481 cttcaatagc ttcgctaatg tactcattga caacttcttc aaattgactc gatacttttt
18541 tagattcttc aattgttcgt gtaaaggcaa tcgctcgttc taaaggggct ccatcgtaag

```
18601 gactattttt gataggattc ttatagccat tacggcgcat cattccgttc caaatcccta
18661 caatacgacc aacatcatca atatttaatc cattttctgg atcagcgagc gttcgttgca
18721 tatctttttg aatggctgct tcatcaactg ctaaaaccat cactttataa tcggttaggg
18781 tatcacgtga aactgcttga ccaaatccca tacggaagaa tacttccccg tacttacttt
18841 catcatccat tgatgaaagt aagatactct tgtctttggc atttttttc gcactttccc
18901 catatatctt tggtgtagcc gtttggtaca ttcttttaag accttttaca ttgttattgc
18961 tgtgaacttt agagaatgca ctggcttctt tagctgcttc gtgagcacct gtcgttctgt
19021 gagcctcatc agaaataata aaatcaaact ctggaaaccc ttctttttgc gcttcaccaa
19081 taacttctat actctgatac gtgctaaaaa caacgagcat atcggttggt tttgacaatg
19141 attcaaaatc atgccagttt tgtaatattt ttttactaga ggtggttgct ggatagccaa
19201 tatcagatgc ttttatatcc tcggttccat ctgtaccacg tgacgcatca cggtctgatg
19261 taacggccat acttgttatc gttagttcag tatcattgtt ccaaccacgc aatgtttgcg
19321 tcaataattg aatactagga acaagataca acactttaaa aggaccagct ttatctttag
19381 atagagcttc agaaatttta agactagtaa atgtttttcc tgtacctggt gccataatta
19441 gttgtccacg ctcattttct ttaaaatgag caagtgcatt ttctttagct gtttgttggt
19501 aatcccttaa ttttttaggc tttttaacga ctacattttc tggacgttca aaattaaatt
19561 gagaccagtc aatttgagag tttctaaggt ctgataatcc aatgatttca attccttttt
19621 cattctgatc aatcgtttca cgtgcattgg aattccaatc atccatcgta gatacaatca
19681 gtcctcgttg atagtaactt ttacctaact cagcaacaaa ggaattaatt tcagacttac
19741 caaccttatt tgtataaaat ttggcttgta tcgcaaccaa atcaccattt ttttgttcag
19801 caactaagtc aacacctgta tctttttttgg gaataccata agattcagga acttctgaca
19861 ataaccatac attttgatac aaggctttgt aagttggttc atttttttaaa taagccaacg
19921 ttagtttctc aaaaagcgtc ccacgttctc tttgtacatc tttagaaaca gtattaatct
19981 gattaattaa atcatcaaaa gtttgtttag tcattttaca ttccttccaa aatgccacca
20041 tttttttattt tttgtatgaa gttcttcttt tgtattatct atgtctttag attttcgtac
20101 cataggatct ttatcatctt ttcgtactgg taccatcaaa gattttaatt ccttgatttc
20161 ttctttatat tcttcaagta attttttatc ttgtaaggct agacgttgtt gttggtctaa
20221 tagcttatga aggtcttttt tttctgattg aagttcggac actagattcc ttaaaaataa
20281 aatctcttca ttcccctctg gattgtcatc aactgaatta ttcacgttgt cgtcaacaaa
20341 aactttttgt tgacgttttt tagttgacgt tttgttaaaa aaagcactct taataacttt
20401 ttgaccgtca acgtaaacgg tatttccttt ttttgacgta aattgacgta aatctattga
20461 cgattctttt tttatttttt gccatacagc ttgctttgaa acacctattt catccgcaag
20521 ctctttgatt gtctttaact tttcactcat cataagtcct gccttttaac ggttgatagg
20581 tattgctcga ttgctttttt gaggtactta gcgatattgt gcttagaata ttcttcacgc
20641 ttgctagaga cataacctag gtggtcgttg acaccatcca accctcgtaa gtctttaagt
20701 tcgtcataga gagggtacac gtgtgcttgt aagcctgaaa gaattttagt atccgtcaag
20761 tctagataag acaataagaa atgttccata agcaacttgg tatacggact ttttagggct
20821 tgtaccgtca gcatgtcttc ggtttctgct ttacgtgctt tatcttcttg gtagacttta
20881 tcttccaact tgtagctgtt atcgtctgct cgacgttttt tcgtaatatg aaagacgata
20941 gaatcaatcg agcgcccctt tttgactttg tcatacgtca cggtaaaaga agtatgatca
21001 ttgatttctt ccaaaggttt ttttaatatc cattttttcta aatctgcaaa aagtttatat
21061 tcattgacta tatccgtcat ttctcgtaat tctcggattg aaattgaaag attgcggtag
21121 gcttccactt gttcttctct ccgtccaccc ttataactat aatgctcgta ttggttataa
21181 ttcatggata accaacgata caagatgata gaatgtttac tattgagttc tgaaatatcg
21241 gacaaggcat gttgtgtgaa attcttttta agattgatta gataaggcaa gatttcagga
21301 ctaaatcgaa tagtaacttc atcatgataa tcggtccact ttacataggg aatagggacg
21361 atattttcaa actcaaaccc atgttctttt ttttctttaa ttttgaaaaa agcttgtttc
21421 tgcatattct caactgcctg cttgaaacga ctatgtttgt cattatcaga aactttaaaa
21481 aacgcaaaaa gttcttcttt tgagagatag actgtatgat ctttaggcgg ttcttcggta
21541 tcaatgcacg atacggctaa ctcaaacatt ttgagtgggg ttttatccat tttggcaatg
21601 ctggtaatta aagaattatg ctccactact tttcgttttg ataattcgtt caagatttgc
21661 acctgcttct gattttgttc atttttttggt atactagtca tagaaaactg ctcctttgag
21721 tgtttttaga tacgcatgaa gattctctcg acgaaaagtg tttcttcatg cttttattat
21781 accataatac atggagaagg taatagacat tctccatgta taaacagttt tttctccatg
21841 tataaacagt tttttctcca tatattattc tttaacccct tgaaaacact cacttttttt
21901 tgacgcaaaa gaataaaaga ttaaaaggtt ctgttgcaaa gttcaaagaa gtaaaatgat
21961 atctcaacta ttttctaatt ttaagctgga acacccatta gattatttat ttcatgaatc
22021 actgaaaagc caaagagaga aaagtctctt tggctagttt tgtataaagc atgtatagtt
22081 tcaattcctt tgattgtgga tgaggcagtc cttatggact gatataattt gtgccttttc
22141 ttgataggcc tatgatcttg ttcaatgata ttattcaaat atttactcgt tcgatgttca
22201 gtgttagaaa ataaaccttc agatttaagt cgattgaacg cactgcttag cgaaggagct
22261 ttgtccgtta caaaaactct aggttctcca aattgcttga ttaatcgttt aaagaaagca
22321 taggccgcct gtttatcccg attctttcgc aaccaaatat ctaatgtcaa acctgtagaa
22381 tcgatggcac gatacaaata atgccatttt ccttttactt taatatatgt ttcatccatt
```

```
22441 ctccaagaat cgctagcaga cttatttctt ttcttccata ggcaataaat aatcgaacca
22501 tattcttgta cccatcgata aacagttgtg tgacaaacct tgattccacg ctcattaagg
22561 atttctgaaa catcacgata gctcaaattg tatcttaagt agtagcccac ggctaccatg
22621 atgatatcct gtttgaattg tttaccttta aaatgggaca ttgttattct cctcgtgtag
22681 tttcttccat tctacactaa ataatcagtt aaataaaact ttgcaacaga acccataatt
22741 aaccttatat gaataccaaa aaaataatat atctttctct attcaactta tcatttacat
22801 ctgtaaattt tttgttgaca atttaaatct acaagagtag acttttctca tactattaaa
22861 gaagagatag atttatagga ataggaggca tctcattaat atatgatttg gggttatatg
22921 aagagttgat taagaaacca aataaaattg aagggcaagc agccaatatc aaaagaataa
22981 taaatgcgaa taaatattat acccatattt taattcagct aaattctgcg caatcggcta
23041 ttcaaaagat aagtcaagtt aatttagaag ctcaagcaga tcattcattg atgcacgtga
23101 gtgatggttc tgatttaatg agagagatag aaatgcgaaa acgtgtaatc aatataataa
23161 attgaattga gtggttgtat atgaagacaa tgaaaatcga atctaacgat gaccagaatg
23221 ttaaaaaacc tgattaagtt agcacttcct gctacggtcg aaaatatatt agaaacatct
23281 gtaggatttg ttgattcgtt gatgatttct cagattggat tatttgctgt agcagggatt
23341 gggatagcaa atgcaatact aaatgtttat attgctgtat ttattgcttt aggaatcgga
23401 acatcttctc tgatttctag aaatattgga gcaaaaaata atgaaaaggc gaagttattg
23461 ctaacaatct ttactattag caatcgtctc gggttattct gggttggtga gtcttctaat
23521 aggtccacaa ctcttaaact tatgggtgca acaaaacaaa cacttaacta ttccttacaa
23581 tttttagta tagtaggtgg cggatcggtg gctattgcca caatgattat attaggaagt
23641 atgttaaggg cgataggtga tacgaaaact cctatgaaaa ttggacttat cacgaatatt
23701 ctaaatattt ccctagatta tattctgatt tttggtttag gacctatacc agctttgggt
23761 gtagtaggaa cagcaatagg aacgttagtt gcgagattac taggtactat tttgttatat
23821 cgaaaagtac atttatctgt tctttctttc cctttactct caatctttca acgaagtaat
23881 tatcaagaat tattaaaact atcgtttcct gctgcattag aacgtttagt gatgcggtta
23941 gggcaagttt tgtattttgg agtgatcgta gcaattggcg caaaaactta ttcagctcac
24001 tccatcgcag gcagcattga aagttttgtc tatatgccag cctatgggtt agctactgct
24061 gcagcaactt tagcaggaaa tagtattgga aaaaaagact atgctgaaac aaaacgtgtt
24121 gcactttatg cggttaaata cggtgttatt gttttaagta ttctaggtat aggattgttt
24181 tttggtacgc catatattgc accttggttt accatagatg ctgaggcaat caatcaaatt
24241 gtcattgcac tcaaaattga cgcttttaat caaccaggat tagctatcag tctaatatta
24301 gcaggagtac tccaaggaat gggagacact aaaaccccat tatatagcac cgcatttggg
24361 atgtggatca cacgaatact agggtgtata cttttaggga gtataatggg tttagggatt
24421 gctggagtgt ggttggcaat cggcttggat ctatatgtgc gttctctttt tttaacctat
24481 catttcaaac gaaatcttca aattttaaac agtaataacg aactttcttc ttaagcaagt
24541 gatttttat cacttgtttt tttcgtttaa cgataaattt ctttgccatt agtaggtaga
24601 aaaatgacac atatggatct tatattaaat gaatctcacc atgtctcctc cctctcgctc
24661 ttcacagact tccctttgag cccttagatt tagagagtat atctagaagc gtggtttgaa
24721 gcatcagaat cacttccgct aagaatcttc gattattttt tctctattta gtaactagtt
24781 actccatggg caatcttcta atatcctgca acagcaattt tagaatcata accagtcgtt
24841 ttttatattc accgtgaata tttgattcta agatattaac ataccttagt cagcttggtt
24901 tttcctgagt cagtcctttc tggtaaaaac aacatatttt aaaaaaattt tatttctgtc
24961 caaaacaact ctagcattaa tatctatcat tattttaaca atatcatttt tcagaaagaa
25021 tgatattgaa aatttaactg aacacgttgt tactgaatag attgaaatat aaataaataa
25081 atttaacctg ctactaaaaa tttactttgt gtagtagcat agatgattat ctaatatcag
25141 caatatgtta attcgtaatc aaccttatat gactaaacga aaaaaggttt ttttctctat
25201 taaatttatc atctacaaat gtaaattttt agttgacaaa ttaaatctac aagagtagac
25261 tttggtcata ccataaaaat tgagaggaga aatcatggaa aataaacatc aaagtaattc
25321 aaatgaagta ctaatgagcc aattaccttc tgatcctaaa atattaaaag aattaaatat
25381 ttcggacagt gaattagtta ttatgcgtgt cgtttggtct ctagggagta ctacagcaga
25441 tgaaattgga cgtgagctta gtgaaactta tcagtggagc ccttccacaa taaaaacatt
25501 tttagcgcgt ttaattaaaa aaggattatt aaaaaacagt cgtgatggtc gtaaatatgt
25561 ctatattgca acttgttctg aagatgaagc gatttgccaa atgacactat ctttttttgaa
25621 taaaatttgt gcacataaac atgccaatgt tattttagaa atgattgacg cgagtagtat
25681 tactgctgaa aataaagaag ctatcagcga aaaactaagt agcaaaaatg ttgttgacga
25741 agtaacttgt gattgtataa acagattaaa ttgttgtgat aataattaga aaaggagaat
25801 aacatgtctc aaattatcgt acttattatt ggattgactc taattgcttt tattggatgg
25861 tggttctttg gcaaacatga ggtccatcag gaaactgccg ttatcagttc tgatggacaa
25921 acagctacta ttgtggtaaa tggaggatat aatccagcag tcttaaattt aaaaaaagat
25981 gtcccagtta acttgatctt taaccggaaa gatgcctcat cttgtcttga aaaagttatt
26041 tttccagatt ttggagtgga tgcggattta ccactcaatc aggatgtgtc tattttgatt
26101 gatacatcta agtcaggcga gtttatatat tcatgtggga tgaatatgtt ccatggaaaa
26161 atcatcatca aatagaatca caaattaaat actaatatac cctaaggagg atataaaatg
26221 ttcggatcaa aaattaataa taaagtagaa gtagttgttg acggaggtta ctctcctagt
```

```
26281 aaatttaagt taaaagctgg agagccagct gaagtctctt ttactcgtgt ctctgataaa
26341 ggttgtgctc agcaaattat cttcaatgga gaacttcgaa atttaccctt aaatgaatct
26401 gtcactttta atttcactcc agttgaaaaa ggacgtcata attggtctgc ggtatgaaaa
26461 tgatcctggg gagctattca gttaaataaa aaactatttt taatgtagtt acaaaaggag
26521 atactttatg tcaataaaaa atcgttttat cataggtgtc atcggatcag tcccattgct
26581 tatcaatatg tttatgagct tgggcggttc catgcttgga ggcgataaat atggtgtttg
26641 gattctgttt gcctttggct cattagttta ctggttctca ggattgccat tcttgcgtac
26701 tgcagtcgct tcgttcaaaa atcatcatgc caatatggac acgcttgttg gattaggcac
26761 aactattgca tatgtttaca gtctttacgc tatgtttgct cgtccaaatg aaacatattt
26821 cgaagctgtt gctgttgtta ttacattaat cttgttagga tcatattttg aggaacgtat
26881 gaaagcttca gcttcatctg ctgttgacaa attgatgggt ttgcaagcaa aagatgcaga
26941 agttctgcgt gatggtgagt ttattaaact tcctattgaa gaaattattg taggtgattt
27001 aatccgcgtc aaaccaggtg aaaaagttgc tgttgatggc caaattgttg aaggaacttc
27061 tacattagat gaatctatgg taactggaga atccatgcct gtagaaaagg gccctggcga
27121 taatgtaatt ggagcaaccc ttaataacac tgggtcattt acattcgaag ttaccaaagt
27181 tggggctgat accatgttgt caaatattgc tgaaatggtt cgtcatgcac aaaactcccg
27241 agcacctata caaaaaactg ttgatcggat ttcaaatatt tttgtaccta tagttttaat
27301 gatttcaatt ttaactttta ttgtatggta tgtgtttcta ggatcaactc ttgtcactgc
27361 gatgatattt tcggtatcag ttatgattat tgcttgtcca tgtgcattag ggatcgcaac
27421 tccaacagca ttgatggttg gaaccggacg ttctgctaaa ctgggaattt tgataaaaaa
27481 tgctgaggtt cttgaagcga cccacgatat aaaaaccgtc gtcatggata agactggaac
27541 aattactgtt ggcaaaccac aagtgaccga tattatctct atcggaagaa ttagtgagaa
27601 cgagattctg cgaatcgctg ctggacttga ggattcatca gaacacccat tagctttagc
27661 agtaattaat gaagcaaagg acaagaaaat tactcctgcc gtagctaaaa atttcactgc
27721 tatttctggt aaaggggtac aagctttgat tgatggtaag caggctttta ttggtaatga
27781 tcgtttatcc gatgacttta acatgacaga tgatcttaag gttaaaatga catctttaca
27841 ggcgcaggcg aaaactgtgg tattagttgg ttacgatggt caaataattg ctttaattgg
27901 gattcaagat gcacctaagt ccagctctaa agctgctatc agggcaatgc aaaaatcagg
27961 atttcacact gtaatgttaa ctggggacaa ccgtttggtc gcacaagcca tagcagatga
28021 tattgggatt gacgaggtca tagcagatgt tatgcctggg gacaaagcac aacatattag
28081 aaagttgcaa gaaaaaggag cagtcgcctt tgtaggtgat ggaatcaatg atgcccctgc
28141 attatccacg gcaacagtag gtattgctat gggatcgggg agtgatattg caattgaatc
28201 tggaggtatt gtactagtca aaaatgattt gatggatgtt gtaacttctt tagtattagc
28261 acgaaaaaca tatagtcgga ttttgattaa cttattttgg gctttcatat ataacgtgat
28321 aggtattcca gtcgcagcgg gtatattctc agcacttgga tttacactat ctccagagtt
28381 agctggtctt gcaatggcac ttagctcaat tactgttgtt ttgagctcac tattattgaa
28441 ctatgtgcgc ttgccaaaaa gtagtgagac acttataggt aatagttagt atttgtaata
28501 tacaaaagtg acgcttttaa ttataatagt aattttcctt tttcaaaatt tgaatcagag
28561 tcttaaatac taggcataca ataagagaac caaaaatcag aactacatcc agtagttctg
28621 atttttgatt ctacaatatt tagtgttggt ggaaaatcta cagattttct actgacactt
28681 tttttgtaaa aaaagataga aaaaagagat agcctaactg gtggaattta gtcgtctaaa
28741 acaaaaccat accaagaaag gatatctcca tgtatcattg tactagaaaa ttgttaggtt
28801 taacagatga aaatttattc ttcgaagaag aatggttaga gacagtagaa gaagatggtt
28861 ttcgcaccaa tcttatccat gctaaactta gctatatact aagccattgc cgtaaatgtg
28921 gcatcaaaaa tgaaggacaa atcattaaga atggctcaca taaaacaaaa gtccaacttt
28981 tgccttaccg agcaaccaaa acagaattac gcttggtacg tacgcgcttt tactgtaaag
29041 aatgtcagtc gacattcaac gctcagacca atttggtaga tgaaaactgc tatctctcaa
29101 aggaattaaa ggtacaaatt gccttggaac tagcgaagaa tacaataaaa aaagaattgc
29161 caatcgctat tttgtatcag acgtaatcgt tttgcgtgtc ctgcacacct gtttaaaaac
29221 gtatcatcct cgttttgata cgttaccttc ggtcctttgt tttgatgagt acaaatcgat
29281 gaagtcctgc agtggaaaaa atgagttttg tttttatgaa tgggcagacc cagcaattaa
29341 tcggtgtatt ggagaatcgt cgccttactt tcttaaaacc atatttcctt aatttcacac
29401 gaaaagcacg agcgaacgta aaatacgttg tgatggatat gaatgcccct tattttgaac
29461 ttgtaaaagc ggtcttccca aacgccaaaa tcgtgaccaa tcgctttcac attgtcaaac
29521 aaatcactcg tacgttaaat cagttgtgaa tcaaaacaat gaacaggttt caaaaaacgg
29581 aaccgacaaa gtattgacga ttgaagcgat tctggaaatt actccttacg catgcctatg
29641 atctggatag ttccgattat caatatgatc gattctttcg cagaccaatg acgcaaaaag
29701 cgatggtcga tgagcttttg agttacgacg agcaattaac cagggcttac gagacttgcc
29761 agctcctcct ctatcacttt aaacacaagg acaaccaaag ttttttcgat acgatcaaca
29821 gcttagatca atgccttcca caatggtttt gtaagaagct gacgttttta aataaataca
29881 aattagggat tcaatatgct ttaaaaccca ggtatagtaa tggtgcattg gaaagaacaa
29941 ataataagat aaaagtgata aaacgagtag cctacggcta tcgaaatttc cacaactttc
30001 gagcaagaat ttatctcatt caaggcttga tcttccaagt aaaacaaaaa ccagttaagc
30061 actctgccta actggttatg actaaatttt cagttgtctc tctaccagca ctaaatgtcg
```

```
30121 aagagccaaa aatctgagcc agaatgaggt tcgttctgac ccagatttga ttttgcacaa
30181 ctcatcagtc cttattaaca aagagctcat attcctcatg ttgagaatac gaaggccatg
30241 taccagctta ttggtttagt atataaaatt actcattatt ttcgtcatat tccatctcat
30301 cttcagttac ccacatatga ttcttaactt cttcttgacc atcggttgga gtgtaattaa
30361 caacgtaaac agttcctgta atagatttat ctacttttgc ttcagctccc atcatcccat
30421 ccatgtgatc ggcattcaag ataacagtat cgccctcatt tgctacagtc tcagtatctt
30481 ttaactcttc ctgtactacc catcgatgat tgttaaccat tggtcccccca gtttttggct
30541 catagcttac ttcatacatt gtagtatcaa aggcccctac aacttgagca tttgcacctt
30601 tcatcccttc catatgatcg cctagcagta taacattact tcccaccggg aatttagggt
30661 tcgcagcatc tttcattgaa gaaggaacca taccttcgtg attcattccc tccatactat
30721 tcatttcact agacgctata gaagtcgatt cctcgtttgt acttttagaa ttattcgatt
30781 cactttgatt gctacacgct cctaaaaata caataattcc aagacttgca aatactaatt
30841 ttatcatctt tttcactgtg ttcatcctcc tttaattaat cacatttagg tcatctcata
30901 aatttgtaat ccaaaaactt ctttttaat tgtagctatt gataagtcct tcttagaaaa
30961 gattactaca gaatttgata agatatctaa atgaactttt tcgacactat ctatcgaaat
31021 cagttgaaaa aatatttttt ttgcccaaca ttcacatact agtttagcaa tttttatctg
31081 tttttgaaac aaataactct cctcctttat tgaaatggta actttataac aaatttagtt
31141 cctagaccta atttactttt tacagttata ctacctttat gcgcttcaac aacgctcttt
31201 acaatagcta atccgatacc ttgcccccct aaagctctac ttctggaggg gtcagccata
31261 taaaatctat caaaaatatg aagttgctga tctttttcta taccgatgcc gttgtcttca
31321 acaattaaaa gaacatggtc tttattttta cttatagaaa acttaatggt tccctccttt
31381 tgcgtaaatt taatagcatt attcaaaaga ttgataacaa cttggttcat tttatcttta
31441 tctgcagata tattgatgtt atccccttgg atttcaacat gaaggttttt tgactcaatt
31501 ttgctggcga aattttcaat aacggattca atcaacttat atatgttaaa ctcagtttta
31561 tttaaatgat catattttgc ttcaattttg ttgattattt ctatatttcc aattaaacga
31621 gttagacgat ctatttcgtc ataacaagat tgtaggcgct cgggagttat tttccagact
31681 ccgtcaatca ttccttccac attgccttta agcgtagcta atggagttct taattcatgt
31741 gaaatatcgg tagttaatct ttttcgaagc ccatgctgtt tctccaactg attacttaaa
31801 tcatttaggg aatcaattaa ggaatttatt tcaataattg aagtttcttg tgggatttta
31861 tcagcatatt ccccacgagt taatttatgt gtaaaatcgc ttacatgctt tagtggcaat
31921 cctaatctag cagagatcca cgaagcaagt atgaatgata caagaagagc tgcaattgca
31981 acatacatta aactttgttt catactggag ataaacaagg catcgtgctc tgtataagca
32041 aaaggtccga aataataaaa cgttactttt ccaaactcgt ttccatcatt aattaacttc
32101 tttttcactt ctattgtttc gtctagtttc acactgtttt tctcagaaac tttcaatgca
32161 tttttttca gtttctcctg tgctgacttc aaatcttttc cagtaagttc ccaaaccatg
32221 ttgccttgtt cgtcttctat agaaaaataa atattattct ctaatacttt ctcagatagt
32281 gctgcaagtt cttcactatt ccatgtttcc ccactattta accacaataa atcaattgta
32341 tatacatact gattcaacaa atcttcttgc cgctcttcca cgtatttact aaaatgattg
32401 ttcgtcaaac ttaatgcgac taaactaatc gatcctatta ttaatgttgc gattgctaaa
32461 aaggaaacag cctaattgag tttttatcgt tctttttctc attagtggcc accaaatctg
32521 tatcctgtac catggactgt caaaatatac ataggtgatt tagcatcatt ctcaatttta
32581 tgtcttaagt tttttatatg ggaatcgatg ctcctatcta gtccatcaaa ttctattccc
32641 ttaactttct ctatcaaatc actacgagaa aagactttat taggtgtgga cattagagcg
32701 tgtagaatat caaattcagt tgttgtaagg tttaactcca cttgatgtag atacacttgt
32761 ctactatcag ggtatactat caattcttta ttatttgaag aaagcttatt aggatttgat
32821 aaggtctcag ttctgcgcaa tacagtttgt attctcgcaa ctaattcttt gggactaaat
32881 ggtttaagca tatagtcatc tgcacctaac ttgagaccag taagaatatc atcctcatca
32941 gacttagctg ttaacattat aattggtaca tttgatattt ctctgatttt tctacataca
33001 gtcaatccat cagtatccgg aagcattaga tctagcacaa ctaaatcagg attatttttc
33061 tcgaattttt ccaatgccat tgcgccactt attgctttat aaactgaata atgattagca
33121 actagatagg catctataat ttctaatatt ttttcttcat cgtccacaat aagaattttc
33181 atagacaatc atccttctta aattatttcc tatttaattt ccatttgtcc catcattcca
33241 tactcttcat gttctaatat atgacaatga tacatgaaaa taccttcacg gtcgaatttc
33301 acaagaagtt ctactgtttc atccggattt accagcacag tatctttcca tccttgctcg
33361 tttagtgcgg gctgatttcc atctctagat aaaatttgaa actgaacacc atggatatgg
33421 aacggatgaa tcatgcctcc catcatacta ctaatattgt taacttccca aatctcttga
33481 gtacctaatt ttttatataa atctattctt tccatatcaa attgcttatt gttgatgttt
33541 accatgtgag acattccact taaattaata ctctgtcttg ttaaatcttc caatttcttt
33601 tcatctaaag tactaatagt atttagacta tcagtgggat taaactcttt gttgtctatc
33661 gtattctcaa ttctcattgt cagtgcaacc aaattatttg ctagtaaatg aatcacctta
33721 cctttttat aattctgagt atccactaat atttctgccc tttccccggg cgctaaaagt
33781 agtttactaa gtttaacaga agtattcaga aaaccaccat ctgtagcaat ttggtagaat
33841 gattcgtcat tatccaaatt aaaagtgaag tttcttgcat tggatccatt aacaattcta
33901 tacctcatcc atcggctttt tatctcaacg tatggattga ttgtgccatt agtaagcaac
```

```
    33961 gtctcccctt ttgttccatc agagttaaaa tcattttcat agttaatttg attagtagaa
    34021 ctgaaagatt tatcttgcac aatcaaggga atgtcgtcga caccatattt ggaaggtaaa
    34081 tctaataatt ttgagttgcc gtcatcaata tacatcaaac ctgctaaccc cttataaact
    34141 tgagatgctg tttctccttc cggatgggga tgaaaccaga gtgtagaagc ttcttgatca
    34201 acctcgaatg ttacactttt tttctgtcct gcttctataa tctgatgtgg acctccatct
    34261 gcatcagaag ctactttcaa accgtgccaa tgaaaagatg tgctagcatc aagattattg
    34321 gttgtattga tagtaacttt ttgacctttt ttcagtctaa taactggccc taaaaaatcc
    34381 ccgttatagc ctaatgtttc agtttttctt ccatccatga tttgaacttc cccgttttga
    34441 gtaactatat catatgtcac attttctttg ctttttgaag atggttcaag aacaggggga
    34501 atcaatagtt ttctatctgt ttggtgtaca accttcaaat cagcttcttg attatccccc
    34561 atcatattgg acattctact tgagatttgt cgttgttctt cttgatagat aacttcagca
    34621 tgctcctcct tgttctgctg ttgaatcaat atgaaaatac cgcctagggt aaatattgat
    34681 attgttgttg ctgctaacca aatatattta ttattcttca acccatcatc tcccatatct
    34741 attatgtatt taccataaaa aaaatatgtg gagatttcgt gtcgatcccc ccacatatat
    34801 ttatatttat atagatatat ccctagttac agtatacttt gtatttttta gtagttgatt
    34861 taagaatgtt aatgttaatt ttgaatcagc atcgaaagta actgtacctc tagccaagtc
    34921 aactattttt acactgactg ttggctctgc ttccaaagca tttcttactt ttgttgcaca
    34981 ccctgagcaa ttcattccac ttattccaaa tacctttatc tcagtttttt ttcgaatctc
    35041 agcgtcataa ccaatgtctt gaatagctgt tataactgta tgattagaca aaacaccttc
    35101 atcatattcc aattttagaa cctcagttgt taagttaact ttagcgaaat atactccatc
    35161 caactcatga atttttttct ctattgttgt tgcgcaaact gcacaagtca ttccataaag
    35221 attatatttt gaaatcattg aattcctccg tgtcagaacg tttaaccaat ccgtttaaat
    35281 agtagcttac ttctgagata caaccttcat ccattccaga aagtaacctt tgtatttcgg
    35341 tttctaaaaa ttcttgcttc ttaatcagtg aagtgtagta tctttggttt tttgaagtgt
    35401 taaacgtgat cacattttta tctaacaacc ttcttattaa cgtttttgtt gtggagggat
    35461 gccaactatt acgctcttgc attccatgtg ttattaaacc acatgttacc ggttcatatc
    35521 tccagataaa ctttaaaaca tctaactctg aattcgatac ctgtattttg tgttcgacca
    35581 agaccatccc tccttaggct attcataaag tacgttcttc tttattaaac aaatgatgaa
    35641 ttgttgtcaa agtagaaaca cttatagtta caattaataa aaagcataaa aagaaaaatg
    35701 tttccggtaa aaaactgata attggatctg tacgatagtc aaacatccaa tcatcattgt
    35761 caaaaaatac ttgatgaaaa agcacaaatg ctttatcaaa aaacagaagt gccattacac
    35821 agactgatat tatcatccac ttgattaaac tagtgatttc tctgacttga tcacgcgtta
    35881 ctaatttttt cctgtttagg tataaaagaa tgctactcgt agaaagaagc aatgaatagt
    35941 taagctggaa caaaaattta acatctttaa aatggtttat accgcgatct gagacgtata
    36001 gattatcaaa aaccaattct ttttgaaaag gattatttaa atagcccatt aaatttctaa
    36061 aatcttgatt aacttgagat ctactccagt tttgaaaatt tatatcgctt tgaaagaata
    36121 aattgatttt gagcaaaccc gggaagttga ttgtcaatgt aatcgcgagg gataaagcaa
    36181 agacaaaaaa taatagaaat agcatgcttt ttttcattaa attaatattt atattagatc
    36241 actcctttta atataatgat aaaaaaactt gtgtagaaaa tgtgtaccaa tgtatctgtc
    36301 aaaaagtcta cataaaattt ccataatagc ctgctatagt caaaacaagt tcattaggaa
    36361 ggagggaaga atttgagaga accctacgac agagtattta tttcgtttgg tccatttact
    36421 atttattggt atgccatttt catagtgttt gggatcataa tcggatattt tgttgctaat
    36481 agaagagcaa agcgtgcggc attgcctgag gatacaattg ttgatctgtt gttatatgga
    36541 ttgccagttt ctattatatc ggccagatta tactatgtac tttttgaatt acctttgtac
    36601 attgatgatc ctttaagtat cttaaaaatt tgggaaggtg gtctagcaat acatggtgga
    36661 ttaattggtg ctgtattgac aggggttatt tattgcaaga agaaaaatct ttctttttgg
    36721 aatgttgctg acgtagttgc tcccagtata gtaattggac aaatcattgg acgatgggga
    36781 aattttatga atcaagaggc ctatggagaa attgtagaca tagagttcct gcaatctctg
    36841 aaactcccat ctttcataat tgatcaaatg tatattgatg gggcctacca tcatccgact
    36901 tttttatatg agtctttatg gaatatagga gtactcatta ttttgcttct ggtgtctaga
    36961 aatagaatgt tctttgggca aatattttta atttatcttt cactatattc cgtcgggaga
    37021 ttctggattg aaggattacg gaccgacagt ttaatgctaa cagcaaactt gagaatggct
    37081 caagtcctgt ccatagtcct attaattggc tctatattaa cttacatcta tctcaaaaaa
    37141 tctaaggagg aagatttaca tggaagcatt acttaatctc ttaccnctac tattactact
    37201 aatttgtcct atctcaatga tgnttatgca taaagggcac ggacatgctg ggcatcatca
    37261 tatgcatgga gaagatcat
//
```

Appendix B

ClustalX multiple sequence alignment of PCR products generated using the CuTrR/CuTrF oligonucleotide pair from a putative EGD-Rif transconjugant (Transconjugant) and strains DSE201 and EGD-Rif. ‘*’ indicates identical base pairs.

```
CLUSTAL X (1.81) multiple sequence alignment

Transconjugant  ATGCGGGTTATGACGCGGCGGAAAAAATGTCTAAAGAAGAAAAAGATGCTGTGTTAGAAA
DSE201          ATGCGGGTTATGACGCGGCGGAAAAAATGTCTAAAGAAGAAAAAGATGCTGTGTTAGAAA
EGD-Rif         ATGCGGGTTATGACGCTGCGGAAAAAATGTCCAAAGAAGAAAAAGATGCTGTATTAGAAA
                **************** ************** ******************** *******

Transconjugant  AGAATTTCAAAAAAGAAGTGAGACGTTTTATTCTTTCGGCGGTTCTTTCCTTACCATTAC
DSE201          AGAATTTCAAAAAAGAAGTGAGACGTTTTATTCTTTCGGCGGTTCTTTCCTTACCATTAC
EGD-Rif         AGAATTTCAAAAAAGAAGTGAGACGTTTTATTCTTTCGGCGCTTCTTTCCTTGCCATTAC
                ***************************************** ********** *******

Transconjugant  TTCTTACCATGGTGACACATATTCCGTACATTCATGAAATGGTGTTTGCGGAAACGATTG
DSE201          TTCTTACCATGGTGACACATATTCCGTACATTCATGAAATGGTGTTTGCGGAAACGATTG
EGD-Rif         TTCTTACCATGGTGACACATATTCCGTATATCCATGAAATGGCGTTTGCGGAAACGATAG
                **************************** ** ********** *************** *

Transconjugant  GTAACTGGATTAATCCAACCATCCAATTAGTACTCGCAACGATTGTTCAGTTTTATATTG
DSE201          GTAACTGGATTAATCCAACCATCCAATTAGTACTCGCAACGATTGTTCAGTTTTATATTG
EGD-Rif         GTAACTGGATTAATCCAACGATTCAATTAGTTCTTGCAACTATTGTTCAGTTTTATATTG
                ******************* ** ******** ** ***** *******************

Transconjugant  GTTGGCGGTTTTATGATGGGGCTTATAAAGCGCTTCGAGGTAAAAGTGCAAATATGGATG
DSE201          GTTGGCGGTTTTATGATGGGGCTTATAAAGCGCTTCGAGGTAAAAGTGCAAATATGGATG
EGD-Rif         GTTGGCGGTTTTACGATGGGGCTTATAAAGCGCTGCGAGGTAAAAGTGCGAATATGGATG
                ************* ******************** ************** **********

Transconjugant  TATTAGTCGCACTTGGGACTTCTGCTGCTTATTTCTATAGTGTAGTGGAATACATTCGTC
DSE201          TATTAGTCGCACTTGGGACTTCTGCTGCTTATTTCTATAGTGTAGTGGAATACATTCGTC
EGD-Rif         TATTAGTCGCACTTGGGACATCTGCTGCATACTTCTATAGTGTGGTAGAGTATGTCCGCC
                ******************* ******** ** *********** ** ** **  * ** *

Transconjugant  ACATGATTGATCCGAGCGTGATGCCGCATTACTACTTTGAAACAAGTGCTGTGTTAATCA
DSE201          ACATGATTGATCCGAGCGTGATGCCGCATTACTACTTTGAAACAAGTGCTGTGTTAATCA
EGD-Rif         ATATAATTGACCCAAGCGTGATGCCACATTACTACTTTGAAACAAGTGCCGTGCTAATCA
                * ** ***** ** *********** *********************** *** ******

Transconjugant  CCTTAATCTTGTTAGGTAAATTACTTGAATCATACGCAACATCGAGAACAACCGAATCTA
DSE201          CCTTAATCTTGTTAGGTAAATTACTTGAATCATACGCAACATCGAGAACAACCGAATCTA
EGD-Rif         CATTAATCTTATTAGGTAAACTGCTTGAATCATACGCAACTTCCAGAACAACGGAATCTA
                * ******** ********* * ***************** ** ******** *******
```

Appendix C

Nucleotide sequence (5’ – 3’) (incomplete) for a 1883 bp region of ORF *cutR* amplified and sequenced using the CuTrM2F/CuTrM2R oligonucleotide pair.

```
   1  GCGAATGTCA ACTTAGTAAC GGAAAATGCA GCCGTATATT ATGACCCGGA
  51  AGTTACATCG ACAGAAGATT TGATTAAAGT TGTTAAACAT GCGGGTTATG
 101  ACGCGGCGGA AAAAATGTCT AAAGAAGAAA AAGATGCTGT GTTAGAAAAG
 151  AATTTCAAAA AAGAAGTGAG ACGTTTTATT CTTTCGGCGG TTCTTTCCTT
 201  ACCATTACTT CTTACCATGG TGACACATAT TCCGTACATT CATGAAATGG
 251  TGTTTGCGGA AACGATTGGT AACTGGATTA ATCCAACCAT CCAATTAGTA
 301  CTCGCAACGA TTGTTCAGTT TTATATTGGT TGGCGGTTTT ATGATGGGGC
 351  TTATAAAGCG CTTCGAGGTA AAAGTGCAAA TATGGATGTA TTAGTCGCAC
 401  TTGGGACTTC TGCTGCATAT TTCTATAGTG TAGTGGAATA CATTCGTCAC
 451  ATGATTGATC CGAGCGTGAT GCCGCATTAC TACTTTGAAA CAAGTGCTGT
 501  GTTAATCACC TTAATCTTGT TAGGTAAATT ACTTGAATCA TACGCAACAT
 551  CGAGAACAAC CGAATCTATT GCTGGTTTAC TTGAATTACA AGCAAAAGAA
 601  GCGACCGTTA TTCGAGAAGG AAAAGAATGG TTAGTGCCAG TAGATTCATT
 651  GAAAATTGGC GATATTATCC TTGTTCGTCC GGGTGAAAAA GTTCCGATGG
 701  ATGCCGAAAT TATTTCTGGT GAAACGAGTA TTGATGAAGC GATGATTACT
 751  GGGGAACCTG TACCAGTAGA GAAGAAACCA GGCGATTCTG TTATTGGCGC
 801  AACGATTAAC TTTGACGGGG CTTTCCAAGC AAAAATTACG AAACGAATGG
 851  AAGAAACCGT TTTAGAATCC ATTATTCGTT TGGTGGAAGA AGCGCAAGGT
 901  ATTAAAGCGC CAATTCAACG TTTGGCAGAT AGAATTTCCG GCATATTTGT
 951  ACCAATTGTA CTTGGGATTG CTGCTGTAAC CTTTATTATT TGGTATCTTG
1001  TTACTGGAAC GGTGGATGGT TCACTTGAAG CTGCGATTGC GGTATTAGTT
1051  ATCGCCTGTC CCTGTGCGCT TGGTCTTGCA ACGCCAACCG CTATCATGGC
1101  TGGAACTGGT AAAGGCGCCG AAAGTGGGAT ATTATTTAAA GGTGGCGAAC
1151  ATTTAGAACG TACTTCCAAA GTGGATACTA TCGTTTTTGA CAAAACTGGT
1201  ACTTTAACAG AAGGTAAACT AGAAGTGAGT GATAAAAAAG CAGCCAATGA
1251  TCACTTTTTC CCTTATTTAT TCTTAATGGA ACAACAATCA GAGCATCCGA
1301  TTGCGAAAGC GATTATTAAG ATGTTAGAGC CAGAGAATAT AGATGCTTCC
1351  GCAGTGAAAC AAGGGAAAAT CCGCGCGAAA GCTGGGCACG GCATGACCGG
1401  TAATCTGGAT GATAGTAAAG TGGAACTAGG TGCTTACCGC TATGTTTCTT
1451  CCCTTACAAC GATTCCAAAA GAAGATGATG AATTAATCGA AAGTTGGATG
1501  CATGCAGGAA AAACAGTCGT AGCAATGGCA ATTGATGGTG TATACGCAGG
1551  TGCCCTCGCT TTATCTGACA CACCACGACC AGAAGCAAAA GAAGCCATCC
1601  AAAAACTAAA AGCACAAGGT ATTAAAACAG CAATTTGTTC TGGGGACCAA
1651  TCTGTTGTCG TAGAAAATAT GGCTAAAGAT TTAGGTATTG ATATGTTCTT
1701  TGCCGAACAA CTACCAAATG ATAAGAGCGC CTTAGTCGAG AAATTACAGC
1751  AAGACGGTCA TATCGTTGCA TTCGTTGGTG ATGGCATTAA TGATGCTCCA
1801  GCTCTTGCGG CAAGTGATAT TGGGATTAGT ATCGGTACTG GAACTGACAT
1851  TGCGATTGAA ACAGGAGATG TAACACTTGT AAG
```

Appendix D

ClustalX multiple sequence alignment of the deduced polypeptide sequence of ORFs pCT0009, pCT0011 and pCT0015. ‘*’ indicates identical amino acids.

```
CLUSTAL X (1.81) multiple sequence alignment

pCT0009        MSHFKGKQFKQDIIMVAVGYYLRYNLSYRDVSEILNERGIKVCHTTVYRWVQEYGSIIYC
pCT0011        MSHFKGKQFKQDIIMVAVGYYLRYNLSYRDVSEILNERGIKVCHTTVYRWVQEYGSIIYC
pCT0015        MSHFKGKQFKQDIIMVAVGYYLRYNLSYRDVSEILNERGIKVCHTTVYRWVQEYGSIIYC
               ************************************************************

pCT0009        LWKKRNKSASDSWRMDETYIKVKGKWHYLYRAIDSTGLTLDIWLRKNRDKQAAYAFFKRL
pCT0011        LWKKRNKSASDSWRMDETYIKVKGKWHYLYRAIDFTGLTLDIWLRKNRDKQAAYAFFKRL
pCT0015        LWKKRNKSASDSWRMDETYIKVKGKWHYLYRAIDSTGLTLDIWLRKNRDKQAAYAFFKRL
               ********************************** *************************

pCT0009        IKQFGEPRVFVTDKAPSLSSAFNRLKSEGLFSNTEHRTSKYLNNIIEQDHRPIKKRHKLY
pCT0011        IKQFGEPRVFVTDKAPSLSSAFNRLKSEGLFSNTEHRTSKYLNNIIEQDHRPIKKRHKLY
pCT0015        IKQFGEPRVFVTDKAPSLSSAFNRLKSEGLFSNTEHRTSKYLNNIIEQDHRPIKKRHKLY
               ************************************************************

pCT0009        QSIRTASSTIKGIETIHALYKTSQRDFSLFGFSVIHEINNLMGVPA
pCT0011        QSIRTASSTIKGIETIHALYKTSQRDFSLFGFSVIHEINNLMGVPA
pCT0015        QSIRTASSTIKGIETIHALYKTSQRDFSLFGFSVIHEINNLMGVPA
               **********************************************
```

Appendix E

ClustalX multiple sequence alignment of the deduced polypeptide sequence of ORFs pCT0017 and pCT0027. ‘*’ indicates positions which have a single, fully conserved residue shared between all sequences, ‘**:**’ indicates amino acids that are in a 'strong' similarity group are fully conserved and ‘**.**’ indicates that the amino acids of a 'weaker' similarity group are fully conserved.

```
CLUSTAL X (1.81) multiple sequence alignment

pCT0017         MENKHQSNSNEVLMSQLPSDPKILKELNISDSELVIMRVVWSLGSTTADEIGRELSETYQ
pCT0027         -------------------MVLVEHKIQVSNSELDVLKFIWRYEPVTCGLITHGMQERNS
                                      : :::::*:*** :::.:*   ..*.. * : :.*  .

pCT0017         WSPSTIKTFLARLIKKGLLKNSRDGRKYVYIATCSEDEAICQMTLSFLNKICAHKHANVI
pCT0027         WHPSTTKTLIRRLLDKNVITFNTSKNQRYYTSLIKKQE-------------FLETEIQRL
                * *** **:: **:.*.::. . . .:  * :  .::*               ... : :

pCT0017         LEMIDASSITAENKEAISEKLSSKNVVDEVTCDCINRLNCCDNN
pCT0027         LSGMDEGCISEVS-------YYLNGLVKRSDTEEFNDFKI----
                *. :* ..*:  .          :.:*..   : :* ::
```

Appendix F

Nucleotide sequence and deduced amino acid sequence of ORF pCT0017 carrying an introduced *Bam*HI site. The nucleotide sequence is numbered relative to pCT100 sequence shown in Appendix A, page 310. Nucleotide changes (CGATTT → GGATCC) introduced using overlap extension PCR are shown (in bold). The *Bam*HI site (GGATCC) introduced as a result of these changes is indicated (underlined). These bases changes introduced two missense mutations resulting in the following amino acid changes: A → G and R → C (underlined and bold).

```
25295 - ATGGAAAATAAACATCAAAGTAATTCAAATGAAGTACTAATGAGCCAATTACCTTCTGAT
        M  E  N  K  H  Q  S  N  S  N  E  V  L  M  S  Q  L  P  S  D

25355 - CCTAAAATATTAAAAGAATTAAATATTTCGGACAGTGAATTAGTTATTATGCGTGTCGTT
        P  K  I  L  K  E  L  N  I  S  D  S  E  L  V  I  M  R  V  V

25415 - TGGTCTCTAGGGAGTACTACAGCAGATGAAATTGGACGTGAGCTTAGTGAAACTTATCAG
        W  S  L  G  S  T  T  A  D  E  I  G  R  E  L  S  E  T  Y  Q

25475 - TGGAGCCCTTCCACAATAAAAACATTTTTAGCGCGTTTAATTAAAAAAGGATTATTAAAA
        W  S  P  S  T  I  K  T  F  L  A  R  L  I  K  K  G  L  L  K

25535 - AACAGTCGTGATGGTCGTAAATATGTCTATATTGCAACTTGTTCTGAAGATGAAGGGATC
        N  S  R  D  G  R  K  Y  V  Y  I  A  T  C  S  E  D  E  G  I

25595 - CGCCAAATGACACTATCTTTTTTGAATAAAATTTGTGCACATAAACATGCCAATGTTATT
        C  Q  M  T  L  S  F  L  N  K  I  C  A  H  K  H  A  N  V  I

25655 - TTAGAAATGATTGACGCGAGTAGTATTACTGCTGAAAATAAAGAAGCTATCAGCGAAAAA
        L  E  M  I  D  A  S  S  I  T  A  E  N  K  E  A  I  S  E  K

25715 - CTAAGTAGCAAAAATGTTGTTGACGAAGTAACTTGTGATTGTATAAACAGATTAAATTGT
        L  S  S  K  N  V  V  D  E  V  T  C  D  C  I  N  R  L  N  C

25775 - TGTGATAATAATTAGAAAAGGAGAATAACATGTCTCAAATTATCG
        C  D  N  N  *
```

Appendix G

Nucleotide sequence and deduced amino acid sequence of ORF pCT0018 carrying an introduced *Bam*HI site. The nucleotide sequence is numbered relative to pCT100 sequence shown in Appendix A, page 310. The nucleotide change (GGATGC → GGATCC) introduced using overlap extension PCR is shown (in bold). The *Bam*HI site (GGATCC) introduced as a result of these changes is indicated (underlined). This base change introduced a missense mutation resulting in the following amino acid change: A → P (underlined and bold).

```
25804 - ATGTCTCAAATTATCGTACTTATTATTGGATTGACTCTAATTGCTTTTATTGGATGGTGG
        M  S  Q  I  I  V  L  I  I  G  L  T  L  I  A  F  I  G  W  W

25864 - TTCTTTGGCAAACATGAGGTCCATCAGGAAACTGCCGTTATCAGTTCTGATGGACAAACA
        F  F  G  K  H  E  V  H  Q  E  T  A  V  I  S  S  D  G  Q  T

25924 - GCTACTATTGTGGTAAATGGAGGATATAATCCAGCAGTCTTAAATTTAAAAAAAGATGTC
        A  T  I  V  V  N  G  G  Y  N  P  A  V  L  N  L  K  K  D  V

25984 - CCAGTTAACTTGATCTTTAACCGGAAAGATGCCTCATCTTGTCTTGAAAAAGTTATTTTT
        P  V  N  L  I  F  N  R  K  D  A  S  S  C  L  E  K  V  I  F

26044 - CCAGATTTTGGAGTGGATCCGGATTTACCACTCAATCAGGATGTGTCTATTTTGATTGAT
        P  D  F  G  V  D  P  D  L  P  L  N  Q  D  V  S  I  L  I  D

26104 - ACATCTAAGTCAGGCGAGTTTATATATTCATGTGGGATGAATATGTTCCATGGAAAAATC
        T  S  K  S  G  E  F  I  Y  S  C  G  M  N  M  F  H  G  K  I

26164 - ATCATCAAATAG
        I  I  K  *
```

Appendix H

Nucleotide sequence and deduced amino acid sequence of ORF pCT0019 carrying an introduced *Bam*HI site. The nucleotide sequence is numbered relative to pCT100 sequence shown in Appendix A, page 310. The nucleotide change (GGATC<u>A</u> → GGATC<u>C</u>) introduced using overlap extension PCR is shown (in bold). The *Bam*HI site (GGATCC) introduced as a result of these changes is indicated (underlined). This base change did not result in an amino acid change.

```
26218 - ATGTTCGGATCCAAAATTAATAATAAAGTAGAAGTAGTTGTTGACGGAGGTTACTCTCCT
        M  F  G  S  K  I  N  N  K  V  E  V  V  V  D  G  G  Y  S  P

26278 - AGTAAATTTAAGTTAAAAGCTGGAGAGCCAGCTGAAGTCTCTTTTACTCGTGTCTCTGAT
        S  K  F  K  L  K  A  G  E  P  A  E  V  S  F  T  R  V  S  D

26338 - AAAGGTTGTGCTCAGCAAATTATCTTCAATGGAGAACTTCGAAATTTACCCTTAAATGAA
        K  G  C  A  Q  Q  I  I  F  N  G  E  L  R  N  L  P  L  N  E

26398 - TCTGTCACTTTTAATTTCACTCCAGTTGAAAAAGGACGTCATAATTGGTCTGCGGTATGA
        S  V  T  F  N  F  T  P  V  E  K  G  R  H  N  W  S  A  V  *
```

Appendix I

Nucleotide and deduced amino acid sequence for ORF *cutR* carrying an introduced *Bam*HI site. The nucleotide sequence is numbered relative to the known *cutR* sequence shown in Appendix C, page 327. A nucleotide change (GGATGC → GGATCC) introduced using overlap extension PCR is shown (in bold). The *Bam*HI site (GGATCC) introduced as a result of this change is indicated (underlined). This base change also introduced a missense mutation resulting in the following amino acid change: A → P (underlined and bold).

```
  1 - GCGAATGTCAACTTAGTAACGGAAAATGCAGCCGTATATTATGACCCGGAAGTTACATCG
      A  N  V  N  L  V  T  E  N  A  A  V  Y  Y  D  P  E  V  T  S

 61 - ACAGAAGATTTGATTAAAGTTGTTAAACATGCGGGTTATGACGCGGCGGAAAAAATGTCT
      T  E  D  L  I  K  V  V  K  H  A  G  Y  D  A  A  E  K  M  S

121 - AAAGAAGAAAAAGATGCTGTGTTAGAAAAGAATTTCAAAAAAGAAGTGAGACGTTTTATT
      K  E  E  K  D  A  V  L  E  K  N  F  K  K  E  V  R  R  F  I

181 - CTTTCGGCGGTTCTTTCCTTACCATTACTTCTTACCATGGTGACACATATTCCGTACATT
      L  S  A  V  L  S  L  P  L  L  L  T  M  V  T  H  I  P  Y  I

241 - CATGAAATGGTGTTTGCGGAAACGATTGGTAACTGGATTAATCCAACCATCCAATTAGTA
      H  E  M  V  F  A  E  T  I  G  N  W  I  N  P  T  I  Q  L  V

301 - CTCGCAACGATTGTTCAGTTTTATATTGGTTGGCGGTTTTATGATGGGGCTTATAAAGCG
      L  A  T  I  V  Q  F  Y  I  G  W  R  F  Y  D  G  A  Y  K  A

361 - CTTCGAGGTAAAAGTGCAAATATGGATGTATTAGTCGCACTTGGGACTTCTGCTGCATAT
      L  R  G  K  S  A  N  M  D  V  L  V  A  L  G  T  S  A  A  Y

421 - TTCTATAGTGTAGTGGAATACATTCGTCACATGATTGATCCGAGCGTGATGCCGCATTAC
      F  Y  S  V  V  E  Y  I  R  H  M  I  D  P  S  V  M  P  H  Y

481 - TACTTTGAAACAAGTGCTGTGTTAATCACCTTAATCTTGTTAGGTAAATTACTTGAATCA
      Y  F  E  T  S  A  V  L  I  T  L  I  L  L  G  K  L  L  E  S

541 - TACGCAACATCGAGAACAACCGAATCTATTGCTGGTTTACTTGAATTACAAGCAAAAGAA
      Y  A  T  S  R  T  T  E  S  I  A  G  L  L  E  L  Q  A  K  E

601 - GCGACCGTTATTCGAGAAGGAAAAGAATGGTTAGTGCCAGTAGATTCATTGAAAATTGGC
      A  T  V  I  R  E  G  K  E  W  L  V  P  V  D  S  L  K  I  G

661 - GATATTATCCTTGTTCGTCCGGGTGAAAAAGTTCCGATGGATCCCGAAATTATTTCTGGT
      D  I  I  L  V  R  P  G  E  K  V  P  M  D  P  E  I  I  S  G

721 - GAAACGAGTATTGATGAAGCGATGATTACTGGGGAACCTGTACCAGTAGAGAAGAAACCA
      E  T  S  I  D  E  A  M  I  T  G  E  P  V  P  V  E  K  K  P

781 - GGCGATTCTGTTATTGGCGCAACGATTAACTTTGACGGGGCTTTCCAAGCAAAAATTACG
      G  D  S  V  I  G  A  T  I  N  F  D  G  A  F  Q  A  K  I  T
```

Continued next page ……

```
 841 - AAACGAATGGAAGAAACCGTTTTAGAATCCATTATTCGTTTGGTGGAAGAAGCGCAAGGT
       K  R  M  E  E  T  V  L  E  S  I  I  R  L  V  E  E  A  Q  G

 901 - ATTAAAGCGCCAATTCAACGTTTGGCAGATAGAATTTCCGGCATATTTGTACCAATTGTA
       I  K  A  P  I  Q  R  L  A  D  R  I  S  G  I  F  V  P  I  V

 961 - CTTGGGATTGCTGCTGTAACCTTTATTATTTGGTATCTTGTTACTGGAACGGTGGATGGT
       L  G  I  A  A  V  T  F  I  I  W  Y  L  V  T  G  T  V  D  G

1021 - TCACTTGAAGCTGCGATTGCGGTATTAGTTATCGCCTGTCCCTGTGCGCTTGGTCTTGCA
       S  L  E  A  A  I  A  V  L  V  I  A  C  P  C  A  L  G  L

1081 - ACGCCAACCGCTATCATGGCTGGAACTGGTAAAGGCGCCGAAAGTGGGATATTATTTAAA
       T  P  T  A  I  M  A  G  T  G  K  G  A  E  S  G  I  L  F  K

1141 - GGTGGCGAACATTTAGAACGTACTTCCAAAGTGGATACTATCGTTTTTGACAAAACTGGT
       G  G  E  H  L  E  R  T  S  K  V  D  T  I  V  F  D  K  T  G

1201 - ACTTTAACAGAAGGTAAACTAGAAGTGAGTGATAAAAAAGCAGCCAATGATCACTTTTTC
       T  L  T  E  G  K  L  E  V  S  D  K  K  A  A  N  D  H  F  F

1261 - CCTTATTTATTCTTAATGGAACAACAATCAGAGCATCCGATTGCGAAAGCGATTATTAAG
       P  Y  L  F  L  M  E  Q  Q  S  E  H  P  I  A  K  A  I  I  K

1321 - ATGTTAGAGCCAGAGAATATAGATGCTTCCGCAGTGAAACAAGGGAAAATCCGCGCGAAA
       M  L  E  P  E  N  I  D  A  S  A  V  K  Q  G  K  I  R  A  K

1381 - GCTGGGCACGGCATGACCGGTAATCTGGATGATAGTAAAGTGGAACTAGGTGCTTACCGC
       A  G  H  G  M  T  G  N  L  D  D  S  K  V  E  L  G  A  Y  R

1441 - TATGTTTCTTCCCTTACAACGATTCCAAAAGAAGATGATGAATTAATCGAAAGTTGGATG
       Y  V  S  S  L  T  T  I  P  K  E  D  D  E  L  I  E  S  W  M

1501 - CATGCAGGAAAAACAGTCGTAGCAATGGCAATTGATGGTGTATACGCAGGTGCCCTCGCT
       H  A  G  K  T  V  V  A  M  A  I  D  G  V  Y  A  G  A  L  A

1561 - TTATCTGACACACCACGACCAGAAGCAAAAGAAGCCATCCAAAAACTAAAAGCACAAGGT
       L  S  D  T  P  R  P  E  A  K  E  A  I  Q  K  L  K  A  Q  G

1621 - ATTAAAACAGCAATTTGTTCTGGGGACCAATCTGTTGTCGTAGAAAATATGGCTAAAGAT
       I  K  T  A  I  C  S  G  D  Q  S  V  V  V  E  N  M  A  K  D

1681 - TTAGGTATTGATATGTTCTTTGCCGAACAACTACCAAATGATAAGAGCGCCTTAGTCGAG
       L  G  I  D  M  F  F  A  E  Q  L  P  N  D  K  S  A  L  V  E

1741 - AAATTACAGCAAGACGGTCATATCGTTGCATTCGTTGGTGATGGCATTAATGATGCTCCA
       K  L  Q  Q  D  G  H  I  V  A  F  V  G  D  G  I  N  D  A  P

1801 - GCTCTTGCGGCAAGTGATATTGGGATTAGTATCGGTACTGGAACTGACATTGCGATTGAA
       A  L  A  A  S  D  I  G  I  S  I  G  T  G  T  D  I  A  I  E

1861 - ACAGGAGATGTAACACTTGTAAG
       T  G  D  V  T  L  V
```

Appendix J

Nucleotide sequence (5’ – 3’) for the *Bam*HI fragment of plasmid pCT800. The amino acid translation of the *erm*R gene is shown. Start and stop codons are indicated.

```
   1 - GATCCGGACTAGTCCGGGCCCTCGAGCTCAAGCTAGCTTGGTACCAGATCTGAGATCACG

  61 - CGTTCTAGAGGTACGTACCGCATGCGATGCGGCCGCATCGGCCTAGGCGATTTAGAAGCA

 121 - AACTTAAGAGTGTGTTGATAGTGCAGTATCTTAAAATTTTGTATAATAGGAATTGAAGTT

 181 - AAATTAGATGCTAAAAATTTGTAATTAAGAAGGAGTGATTACATGAACAAAAATATAAAA
   1 -                                           M  N  K  N  I  K
 241 - TATTCTCAAAACTTTTTAACGAGTGAAAAAGTACTCAACCAAATAATAAAACAATTGAAT
   7 - Y  S  Q  N  F  L  T  S  E  K  V  L  N  Q  I  I  K  Q  L  N

 301 - TTAAAAGAAACCGATACCGTTTACGAAATTGGAACAGGTAAAGGGCATTTAACGACGAAA
  27 - L  K  E  T  D  T  V  Y  E  I  G  T  G  K  G  H  L  T  T  K

 361 - CTGGCTAAAATAAGTAAACAGGTAACGTCTATTGAATTAGACAGTCATCTATTCAACTTA
  47 - L  A  K  I  S  K  Q  V  T  S  I  E  L  D  S  H  L  F  N  L

 421 - TCGTCAGAAAAATTAAAACTGAATACTCGTGTCACTTTAATTCACCAAGATATTCTACAG
  67 - S  S  E  K  L  K  L  N  T  R  V  T  L  I  H  Q  D  I  L  Q

 481 - TTTCAATTCCCTAACAAACAGAGGTATAAAATTGTTGGGAGTATTCCTTACCATTTAAGC
  87 - F  Q  F  P  N  K  Q  R  Y  K  I  V  G  S  I  P  Y  H  L  S

 541 - ACACAAATTATTAAAAAAGTGGTTTTTGAAAGCCATGCGTCTGACATCTATCTGATTGTT
 107 - T  Q  I  I  K  K  V  V  F  E  S  H  A  S  D  I  Y  L  I  V

 601 - GAAGAAGGATTCTACAAGCGTACCTTGGATATTCACCGAACACTAGGGTTGCTCTTGCAC
 127 - E  E  G  F  Y  K  R  T  L  D  I  H  R  T  L  G  L  L  L  H

 661 - ACTCAAGTCTCGATTCAGCAATTGCTTAAGCTGCCAGCGGAATGCTTTCATCCTAAACCA
 147 - T  Q  V  S  I  Q  Q  L  L  K  L  P  A  E  C  F  H  P  K  P

 721 - AAAGTAAACAGTGTCTTAATAAAACTTACCCGCCATACCACAGATGTTCCAGATAAATAT
 167 - K  V  N  S  V  L  I  K  L  T  R  H  T  T  D  V  P  D  K  Y

 781 - TGGAAGCTATATACGTACTTTGTTTCAAAATGGGTCAATCGAGAATATCGTCAACTGTTT
 187 - W  K  L  Y  T  Y  F  V  S  K  W  V  N  R  E  Y  R  Q  L  F

 841 - ACTAAAAATCAGTTTCATCAAGCAATGAAACACGCCAAAGTAAACAATTTAAGTACCGTT
 207 - T  K  N  Q  F  H  Q  A  M  K  H  A  K  V  N  N  L  S  T  V

 901 - ACTTATGAGCAAGTATTGTCTATTTTTAATAGTTATCTATTATTTAACGGGAGGAAATAA
 227 - T  Y  E  Q  V  L  S  I  F  N  S  Y  L  L  F  N  G  R  K  *

 961 - TTCTATGAGTCGCTTTTGTAAATTTGGAAAGTTACACGTTACTAAAGGGAATGTAGATAA

1021 - ATTATTAGGTATACTACTGACAGCTTCCAAGGAGCTAAAGAGGTCCCTAGCGCCTACGGG

1081 - GAATTTGTATCGATGATAAGCTGTCAAGCATGAGAATTATCGCATGCGCCTCCCCG
```

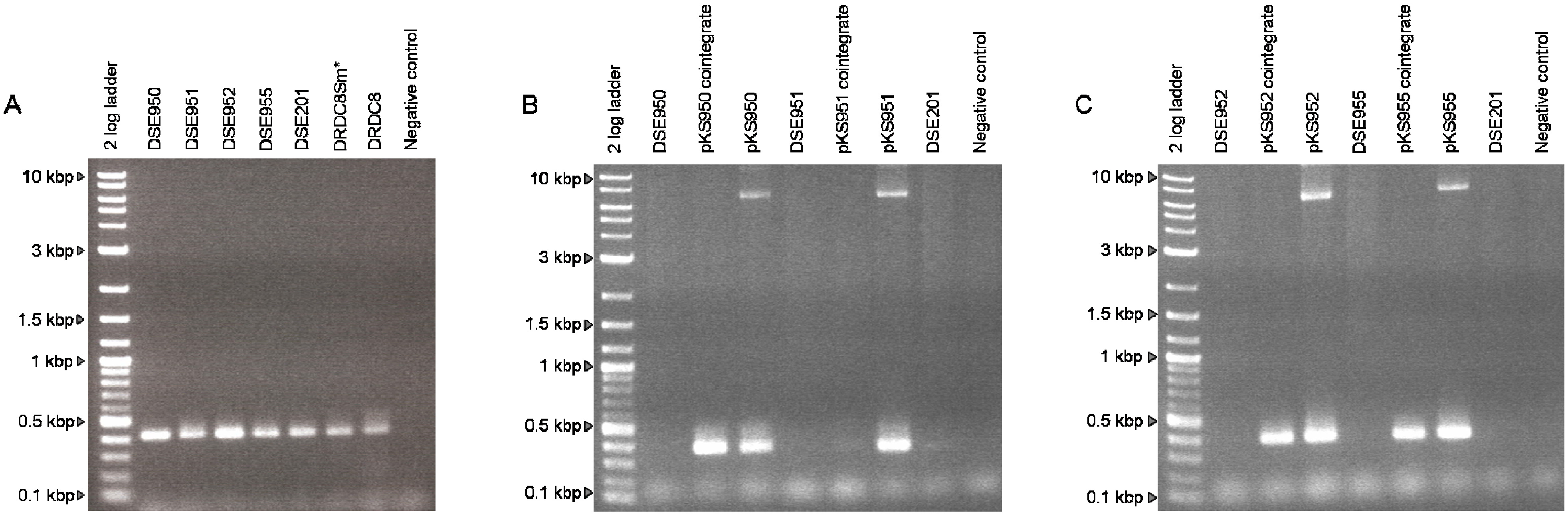
Appendix K

Panel A.

PCR analysis using the p234 and p319 oligonucleotide pair showed amplification of a 417 bp fragment of the *hly* gene from DNA extracted from the putative mutant strains DSE950, DSE951, DSE952 and DSE955, as well as the positive control strains DRDC8Sm* and DRDC8. No amplicon was obtained for the no DNA negative control.

Panel B and C.

PCR analysis using the catPF and catPR oligonucleotide pair showed amplification of a 391 bp fragment of the catP gene from DNA extracted from *L. monocytogenes* strains pKS950 cointegrate, pKS951 cointegrate, pKS952 cointegrate and pKS955 cointegrate and from plasmids pKS950, pKS951, pKS952 and pKS955. Amplicons were not obtained from DNA extracted from the putative mutant *L. monocytogenes* strains DSE950, DSE951, DSE952 and DSE955. PCR products were also absent from the negative control strain DSE201 and the no DNA negative control.

A
2 log ladder
DSE950
DSE951
DSE952
DSE955
DSE201
DRDC8Sm*
DRDC8
Negative control
10 kbp
3 kbp
1.5 kbp
1 kbp
0.5 kbp
0.1 kbp
B
2 log ladder
DSE950
pKS950 cointegrate
pKS950
DSE951
pKS951 cointegrate
pKS951
DSE201
Negative control
10 kbp
3 kbp
1.5 kbp
1 kbp
0.5 kbp
0.1 kbp
C
2 log ladder
DSE952
pKS952 cointegrate
pKS952
DSE955
pKS955 cointegrate
pKS955
DSE201
Negative control
10 kbp
3 kbp
1.5 kbp
1 kbp
0.5 kbp
0.1 kbp

Appendix L

Nucleotide sequence (5' – 3') for the 1496 bp PCR product generated using the p2037/ermR2 oligonucleotide pair from strain DSE950 DNA. The *Bam*HI site flanking the 3' end of the *erm* insertion (shaded in grey) in ORF pCT0017 is underlined. ORF pCT0017 stop codon is shown in bold.

1 GTACCGTTAC TTATGAGCAA GTATTGTCTA TTTTTAATAG TTATCTATTA
51 TTTAACGGGA GGAAATAATT CTATGAGTCG CTTTTGTAAA TTTGGAAAGT
101 TACACGTTAC TAAAGGGAAT GTAGATAAAT TATTAGGTAT ACTACTGACA
151 GCTTCCAAGG AGCTAAAGAG GTCCCTAGCG CCTACGGGGA ATTTGTATCG
201 ATGATAAGCT GTCAAGCATG AGAATTATCG CATGCGCCTC CCCGGATCCG
*Bam*HI
251 GATCCGCCAA ATGACACTAT CTTTTTTGAA TAAAATTTGT GCACATAAAC
301 ATGCCAATGT TATTTTAGAA ATGATTGACG CGAGTAGTAT TACTGCTGAA
351 AATAAAGAAG CTATCAGCGA AAAACTAAGT AGCAAAAATG TTGTTGACGA
401 AGTAACTTGT GATTGTATAA ACAGATTAAA TTGTTGTGAT AATAAT**TAG**A
pCT0017 Stop
451 AAAGGAGAAT AACATGTCTC AAATTATCGT ACTTATTATT GGATTGACTC
501 TAATTGCTTT TATTGGATGG TGGTTCTTTG GCAAACATGA GGTCCATCAG
551 GAAACTGCCG TTATCAGTTC TGATGGACAA ACAGCTACTA TTGTGGTAAA
601 TGGAGGATAT AATCCAGCAG TCTTAAATTT AAAAAAAGAT GTCCCAGTTA
651 ACTTGATCTT TAACCGGAAA GATGCCTCAT CTTGTCTTGA AAAAGTTATT
701 TTTCCAGATT TTGGAGTGGA TGCGGATTTA CCACTCAATC AGGATGTGTC
751 TATTTTGATT GATACATCTA AGTCAGGCGA GTTTATATAT TCATGTGGGA
801 TGAATATGTT CCATGGAAAA ATCATCATCA AATAGAATCA CAAATTAAAT
851 ACTAATATAC CCTAAGGAGG ATATAAAATG TTCGGATCAA AAATTAATAA
901 TAAAGTAGAA GTAGTTGTTG ACGGAGGTTA CTCTCCTAGT AAATTTAAGT
951 TAAAAGCTGG AGAGCCAGCT GAAGTCTCTT TTACTCGTGT CTCTGATAAA
1001 GGTTGTGCTC AGCAAATTAT CTTCAATGGA GAACTTCGAA ATTTACCCTT
1051 AAATGAATCT GTCACTTTTA ATTTCACTCC AGTTGAAAAA GGACGTCATA
1101 ATTGGTCTGC GGTATGAAAA TGATCCTGGG GAGCTATTCA GTTAAATAAA
1151 AAACTATTTT TAATGTAGTT ACAAAAGGAG ATACTTTATG TCAATAAAAA
1201 ATCGTTTTAT CATAGGTGTC ATCGGATCAG TCCCATTGCT TATCAATATG
1251 TTTATGAGCT TGGGCGGTTC CATGCTTGGA GGCGATAAAT ATGGTGTTTG
1301 GATTCTGTTT GCCTTTGGCT CATTAGTTTA CTGGTTCTCA GGATTGCCAT
1351 TCTTGCGTAC TGCAGTCGCT T

Appendix M

Panel A.

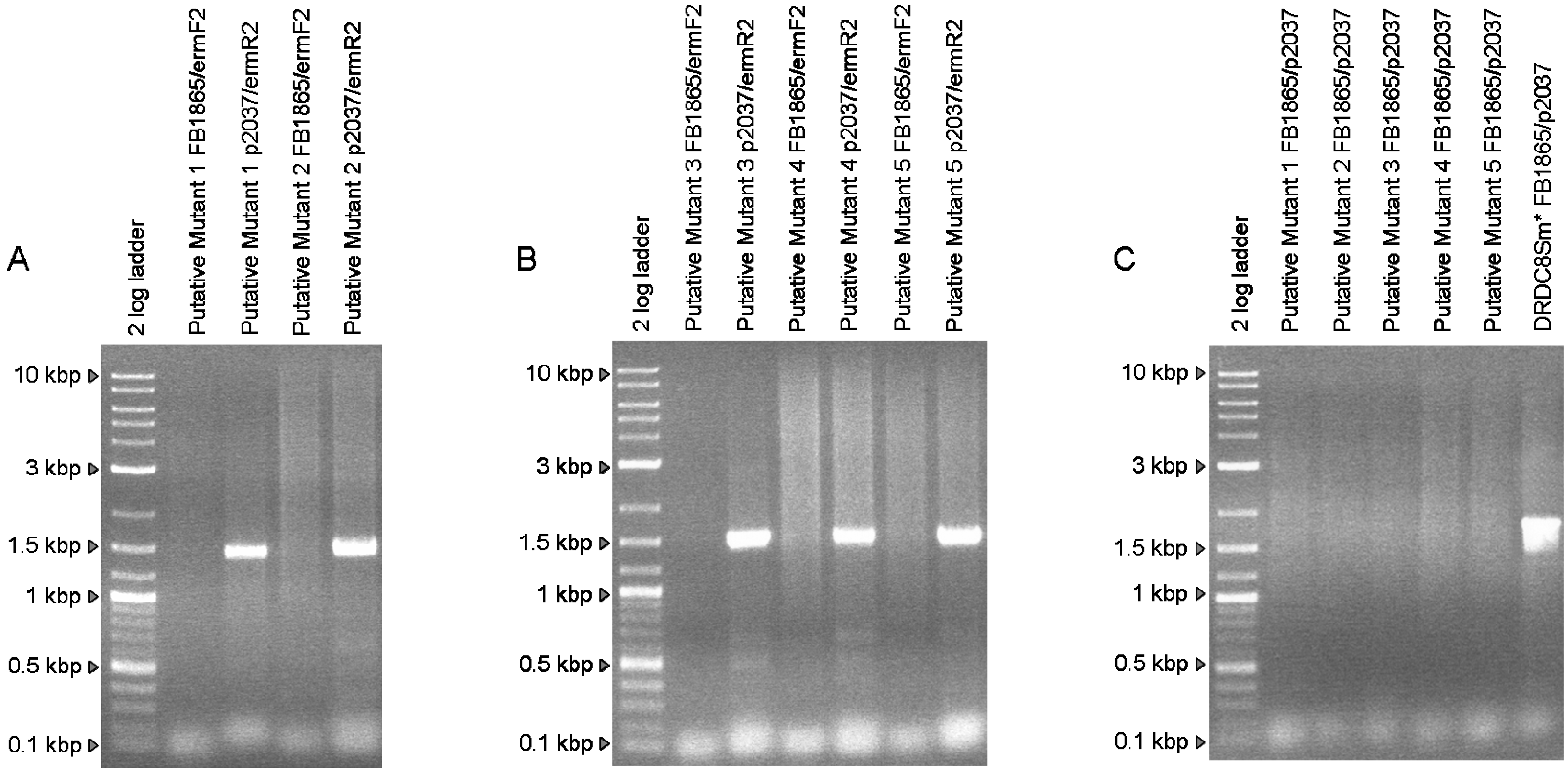
PCR analysis of DNA extracted from Putative Mutant 1 and 2 showed no amplification of DNA using the FB1865/ermF2 oligonucleotide pair. A 1496 bp product was amplified using the p2037/ermR2 oligonucleotide pair for Putative Mutant 1 and 2.

Panel B.

PCR analysis of DNA extracted from Putative Mutant 3, 4 and 5 showed no amplification of DNA using the FB1865/ermF2 oligonucleotide pair. A 1496 bp product was amplified using the p2037/ermR2 oligonucleotide pair for Putative Mutants 3, 4 and 5.

Panel C.

No PCR products were amplified from DNA extracted from Putative Mutant 1, 2, 3, 4, or 5 using the FB1865/p2037 oligonucleotide pair. The expected 1723 bp bp PCR product was amplified from DNA extracted from the positive control strain DRDC8Sm*.


A
2 log ladder
Putative Mutant 1 FB1865/ermF2
Putative Mutant 1 p2037/ermR2
Putative Mutant 2 FB1865/ermF2
Putative Mutant 2 p2037/ermR2
10 kbp
3 kbp
1.5 kbp
1 kbp
0.5 kbp
0.1 kbp
B
2 log ladder
Putative Mutant 3 FB1865/ermF2
Putative Mutant 3 p2037/ermR2
Putative Mutant 4 FB1865/ermF2
Putative Mutant 4 p2037/ermR2
Putative Mutant 5 FB1865/ermF2
Putative Mutant 5 p2037/ermR2
10 kbp
3 kbp
1.5 kbp
1 kbp
0.5 kbp
0.1 kbp
C
2 log ladder
Putative Mutant 1 FB1865/p2037
Putative Mutant 2 FB1865/p2037
Putative Mutant 3 FB1865/p2037
Putative Mutant 4 FB1865/p2037
Putative Mutant 5 FB1865/p2037
DRDC8Sm* FB1865/p2037
10 kbp
3 kbp
1.5 kbp
1 kbp
0.5 kbp
0.1 kbp

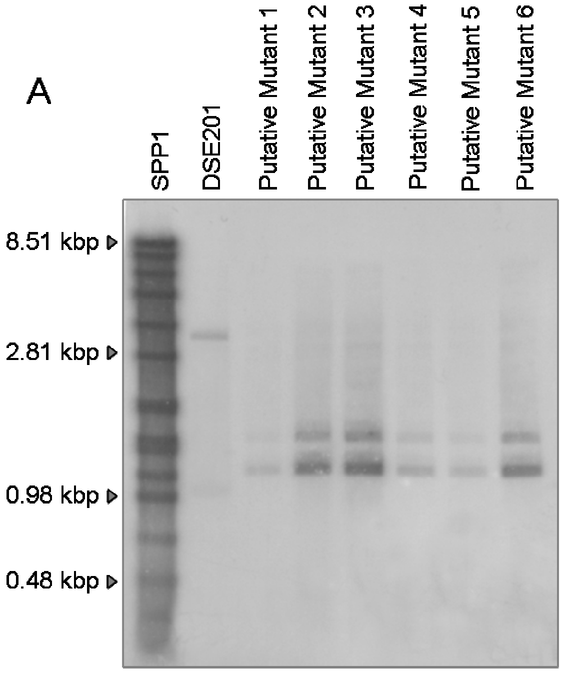
Appendix N

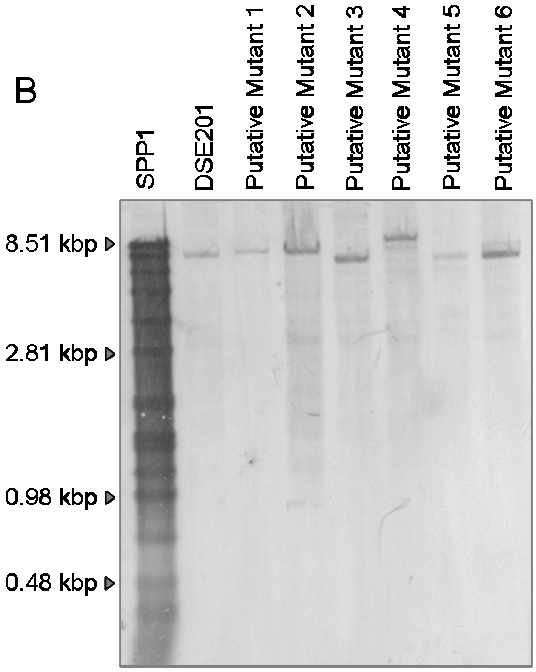
Panel A.

Southern hybridisation analysis of *Dra*I digested DNA extracted from Putative Mutants 1 – 6. A digoxigenin-labelled 645 bp *erm*-specific DNA probe hybridised to two *Dra*I DSE201 DNA fragments 1014 bp and 3347 bp in size. Probe DNA also hybridised with two *Dra*I DNA fragments *ca*. 1215 bp and 1718 bp in size for Putative Mutants 1 – 6.

Panel B.

Southern hybridisation analysis of *Eco*RV digested DNA extracted from Putative Mutants 1 – 6. Digoxigenin-labelled *erm*-specific probe DNA hybridised to a 7273 bp *Eco*RV DSE201 DNA fragment. Probe DNA also hybridised with a single *Eco*RV DNA fragment ranging from *ca*. 6900 to 8600 bp in size for Putative Mutants 1 – 6.





Appendix O

Nucleotide sequence (5' – 3') for the PCR product amplified using the FB1865/p2037 oligonucleotide pair from strain DSE951 DNA. The *Bam*HI sites flanking the 1136 bp *erm* insertion (shaded in grey) in ORF pCT0018 are underlined. ORF pCT0018 start and stop codons are shown in bold.

1 TACTTTGTGT AGTAGCATAG ATGATTATCT AATATCAGCA ATATGTTAAT
51 TCGTAATCAA CCTTATATGA CTAAACGAAA AAAGGTTTTT TTCTCTATTA
101 AATTTATCAT CTACAAATGT AAATTTTTAG TTGACAAATT AAATCTACAA
151 GAGTAGACTT TGGTCATACC ATAAAAATTG AGAGGAGAAA TCATGGAAAA
201 TAAACATCAA AGTAATTCAA ATGAAGTACT AATGAGCCAA TTACCTTCTG
251 ATCCTAAAAT ATTAAAAGAA TTAAATATTT CGGACAGTGA ATTAGTTATT
301 ATGCGTGTCG TTTGGTCTCT AGGGAGTACT ACAGCAGATG AAATTGGACG
351 TGAGCTTAGT GAAACTTATC AGTGGAGCCC TTCCACAATA AAAACATTTT
401 TAGCGCGTTT AATTAAAAAA GGATTATTAA AAAACAGTCG TGATGGTCGT
451 AAATATGTCT ATATTGCAAC TTGTTCTGAA GATGAAGCGA TTTGCCAAAT
501 GACACTATCT TTTTTGAATA AAATTTGTGC ACATAAACAT GCCAATGTTA
551 TTTTAGAAAT GATTGACGCG AGTAGTATTA CTGCTGAAAA TAAAGAAGCT
601 ATCAGCGAAA AACTAAGTAG CAAAAATGTT GTTGACGAAG TAACTTGTGA
651 TTGTATAAAC AGATTAAATT GTTGTGATAA TAATTAGAAA AGGAGAATAA
701 C**ATG**TCTCAA ATTATCGTAC TTATTATTGG ATTGACTCTA ATTGCTTTTA
pCT0018 Start
751 TTGGATGGTG GTTCTTTGGC AAACATGAGG TCCATCAGGA AACTGCCGTT
801 ATCAGTTCTG ATGGACAAAC AGCTACTATT GTGGTAAATG GAGGATATAA
851 TCCAGCAGTC TTAAATTTAA AAAAAGATGT CCCAGTTAAC TTGATCTTTA
901 ACCGGAAAGA TGCCTCATCT TGTCTTGAAA AAGTTATTTT TCCAGATTTT
951 GGAGTGGATC CGGGGAGGCG CATGCGATAA TTCTCATGCT TGACAGCTTA
*Bam*HI
1001 TCATCGATAC AAATTCCCCG TAGGCGCTAG GGACCTCTTT AGCTCCTTGG
1051 AAGCTGTCAG TAGTATACCT AATAATTTAT CTACATTCCC TTTAGTAACG
1101 TGTAACTTTC CAAATTTACA AAAGCGACTC ATAGAATTAT TTCCTCCCGT
1151 TAAATAATAG ATAACTATTA AAAATAGACA ATACTTGCTC ATAAGTAACG
1201 GTACTTAAAT TGTTTACTTT GGCGTGTTTC ATTGCTTGAT GAAACTGATT
1251 TTTAGTAAAC AGTTGACGAT ATTCTCGATT GACCCATTTT GAAACAAAGT
1301 ACGTATATAG CTTCCAATAT TTATCTGGAA CATCTGTGGT ATGGCGGGTA
1351 AGTTTTATTA AGACACTGTT TACTTTTGGT TTAGGATGAA AGCATTCCGC
1401 TGGCAGCTTA AGCAATTGCT GAATCGAGAC TTGAGTGTGC AAGAGCAACC
1451 CTAGTGTTCG GTGAATATCC AAGGTACGCT TGTAGAATCC TTCTTCAACA
1501 ATCAGATAGA TGTCAGACGC ATGGCTTTCA AAAACCACTT TTTTAATAAT
1551 TTGTGTGCTT AAATGGTAAG GAATACTCCC AACAATTTTA TACCTCTGTT
1601 TGTTAGGGAA TTGAAACTGT AGAATATCTT GGTGAATTAA AGTGACACGA
1651 GTATTCAGTT TTAATTTTTC TGACGATAAG TTGAATAGAT GACTGTCTAA
1701 TTCAATAGAC GTTACCTGTT TACTTATTTT AGCCAGTTTC GTCGTTAAAT
1751 GCCCTTTACC TGTTCCAATT TCGTAAACGG TATCGGTTTC TTTTAAATTC
1801 AATTGTTTTA TTATTTGGTT GAGTACTTTT TCACTCGTTA AAAAGTTTTG
1851 AGAATATTTT ATATTTTTGT TCATGTAATC ACTCCTTCTT AATTACAAAT
1901 TTTTAGCATC TAATTTAACT TCAATTCCTA TTATACAAAA TTTTAAGATA
1951 CTGCACTATC AACACACTCT TAAGTTTGCT TCTAAATCGC CTAGGCCGAT
2001 GCGGCCGCAT CGCATGCGGT ACGTACCTCT AGAACGCGTG ATCTCAGATC
2051 TGGTACCAAG CTAGCTTGAG CTCGAGGGCC CGGACTAGTC CGGATCCGGA
*Bam*HI

Continued next page ……

2101 TTTACCACTC AATCAGGATG TGTCTATTTT GATTGATACA TCTAAGTCAG
2151 GCGAGTTTAT ATATTCATGT GGGATGAATA TGTTCCATGG AAAAATCATC
2201 ATCAAA**TAG**A ATCACAAATT AAATACTAAT ATACCCTAAG GAGGATATAA
pCT0018 Stop
2251 AATGTTCGGA TCAAAAATTA ATAATAAAGT AGAAGTAGTT GTTGACGGAG
2301 GTTACTCTCC TAGTAAATTT AAGTTAAAAG CTGGAGAGCC AGCTGAAGTC
2351 TCTTTTACTC GTGTCTCTGA TAAAGGTTGT GCTCAGCAAA TTATCTTCAA
2401 TGGAGAACTT CGAAATTTAC CCTTAAATGA ATCTGTCACT TTTAATTTCA
2451 CTCCAGTTGA AAAAGGACGT CATAATTGGT CTGCGGTATG AAAATGATCC
2501 TGGGGAGCTA TTCAGTTAAA TAAAAAACTA TTTTTAATGT AGTTACAAAA
2551 GGAGATACTT TATGTCAATA AAAAATCGTT TTATCATAGG TGTCATCGGA
2601 TCAGTCCCAT TGCTTATCAA TATGTTTATG AGCTTGGGCG GTTCCATGCT
2651 TGGAGGCGAT AAATATGGTG TTTGGATTCT GTTTGCCTTT GGCTCATTAG
2701 TTTACTGGTT CTCAGGATTG CCATTCTTGC GTACT

Appendix P

Nucleotide sequence (5’ – 3’) for the PCR product amplified using the FB1865/p2037 oligonucleotide pair from strain DSE952 DNA. The *Bam*HI sites flanking the 1136 bp *erm* insertion (shaded in grey) in ORF pCT0019 are underlined. ORF pCT0019 start and stop codons are shown in bold.

1 ACCTGCTACT AAAAATTTAC TTTGTGTAGT AGCATAGATG ATTATCTAAT
51 ATCAGCAATA TGTTAATTCG TAATCAACCT TATATGACTA AACGAAAAAA
101 GGTTTTTTTC TCTATTAAAT TTATCATCTA CAAATGTAAA TTTTTAGTTG
151 ACAAATTAAA TCTACAAGAG TAGACTTTGG TCATACCATA AAAATTGAGA
201 GGAGAAATCA TGGAAAATAA ACATCAAAGT AATTCAAATG AAGTACTAAT
251 GAGCCAATTA CCTTCTGATC CTAAAATATT AAAAGAATTA AATATTTCGG
301 ACAGTGAATT AGTTATTATG CGTGTCGTTT GGTCTCTAGG GAGTACTACA
351 GCAGATGAAA TTGGACGTGA GCTTAGTGAA ACTTATCAGT GGAGCCCTTC
401 CACAATAAAA ACATTTTTAG CGCGTTTAAT TAAAAAAGGA TTATTAAAAA
451 ACAGTCGTGA TGGTCGTAAA TATGTCTATA TTGCAACTTG TTCTGAAGAT
501 GAAGCGATTT GCCAAATGAC ACTATCTTTT TTGAATAAAA TTTGTGCACA
551 TAAACATGCC AATGTTATTT TAGAAATGAT TGACGCGAGT AGTATTACTG
601 CTGAAAATAA AGAAGCTATC AGCGAAAAAC TAAGTAGCAA AAATGTTGTT
651 GACGAAGTAA CTTGTGATTG TATAAACAGA TTAAATTGTT GTGATAATAA
701 TTAGAAAAGG AGAATAACAT GTCTCAAATT ATCGTACTTA TTATTGGATT
751 GACTCTAATT GCTTTTATTG GATGGTGGTT CTTTGGCAAA CATGAGGTCC
801 ATCAGGAAAC TGCCGTTATC AGTTCTGATG GACAAACAGC TACTATTGTG
851 GTAAATGGAG GATATAATCC AGCAGTCTTA AATTTAAAAA AAGATGTCCC
901 AGTTAACTTG ATCTTTAACC GGAAAGATGC CTCATCTTGT CTTGAAAAAG
951 TTATTTTTCC AGATTTTGGA GTGGATGCGG ATTTACCACT CAATCAGGAT
1001 GTGTCTATTT TGATTGATAC ATCTAAGTCA GGCGAGTTTA TATATTCATG
1051 TGGGATGAAT ATGTTCCATG GAAAAATCAT CATCAAATAG AATCACAAAT
1101 TAAATACTAA TATACCCTAA GGAGGATATA AA**ATG**TTCGG ATCCGGGGAG
pCT0019 start *Bam*HI
1151 GCGCATGCGA TAATTCTCAT GCTTGACAGC TTATCATCGA TACAAATTCC
1201 CCGTAGGCGC TAGGGACCTC TTTAGCTCCT TGGAAGCTGT CAGTAGTATA
1251 CCTAATAATT TATCTACATT CCCTTTAGTA ACGTGTAACT TTCCAAATTT
1301 ACAAAAGCGA CTCATAGAAT TATTTCCTCC CGTTAAATAA TAGATAACTA
1351 TTAAAAATAG ACAATACTTG CTCATAAGTA ACGGTACTTA AATTGTTTAC
1401 TTTGGCGTGT TTCATTGCTT GATGAAACTG ATTTTTAGTA AACAGTTGAC
1451 GATATTCTCG ATTGACCCAT TTTGAAACAA AGTACGTATA TAGCTTCCAA
1501 TATTTATCTG GAACATCTGT GGTATGGCGG GTAAGTTTTA TTAAGACACT
1551 GTTTACTTTT GGTTTAGGAT GAAAGCATTC CGCTGGCAGC TTAAGCAATT
1601 GCTGAATCGA GACTTGAGTG TGCAAGAGCA ACCCTAGTGT TCGGTGAATA
1651 TCCAAGGTAC GCTTGTAGAA TCCTTCTTCA ACAATCAGAT AGATGTCAGA
1701 CGCATGGCTT TCAAAAACCA CTTTTTTAAT AATTTGTGTG CTTAAATGGT
1751 AAGGAATACT CCCAACAATT TTATACCTCT GTTTGTTAGG GAATTGAAAC
1801 TGTAGAATAT CTTGGTGAAT TAAAGTGACA CGAGTATTCA GTTTTAATTT
1851 TTCTGACGAT AAGTTGAATA GATGACTGTC TAATTCAATA GACGTTACCT
1901 GTTTACTTAT TTTAGCCAGT TTCGTCGTTA AATGCCCTTT ACCTGTTCCA
1951 ATTTCGTAAA CGGTATCGGT TTCTTTTAAA TTCAATTGTT TTATTATTTG
2001 GTTGAGTACT TTTTCACTCG TTAAAAAGTT TTGAGAATAT TTTATATTTT
2051 TGTTCATGTA ATCACTCCTT CTTAATTACA AATTTTTAGC ATCTAATTTA
2101 ACTTCAATTC CTATTATACA AAATTTTAAG ATACTGCACT ATCAACACAC

Continued next page ……

2151 TCTTAAGTTT GCTTCTAAAT CGCCTAGGCC GATGCGGCCG CATCGCATGC
2201 GGTACGTACC TCTAGAACGC GTGATCTCAG ATCTGGTACC AAGCTAGCTT
2251 GAGCTCGAGG GCCCGGACTA GTCCGGATCC AAAATTAATA ATAAAGTAGA
*Bam*HI
2301 AGTAGTTGTT GACGGAGGTT ACTCTCCTAG TAAATTTAAG TTAAAAGCTG
2351 GAGAGCCAGC TGAAGTCTCT TTTACTCGTG TCTCTGATAA AGGTTGTGCT
2401 CAGCAAATTA TCTTCAATGG AGAACTTCGA AATTTACCCT TAAATGAATC
2451 TGTCACTTTT AATTTCACTC CAGTTGAAAA AGGACGTCAT AATTGGTCTG
2501 CGGTA**TGA**AA ATGATCCTGG GGAGCTATTC AGTTAAATAA AAAACTATTT
pCT0019 stop
2551 TTAATGTAGT TACAAAAGGA GATACTTTAT GTCAATAAAA AATCGTTTTA
2601 TCATAGGTGT CATCGGATCA GTCCCATTGC TTATCAATAT GTTTATGAGC
2651 TTGGGCGGTT CCATGCTTGG AGGCGATAAA TATGGTGTTT GGATTCTGTT
2701 TGCCTTTGGC TCATTAGTTT ACTGGTTCTC AGGATTGCCA TTCTTGCGTA
2751 CTGCAGTCGC TTCGTTCAAA AATCATCATG

Appendix Q

Nucleotide sequence (5’ – 3’) for the 3019 bp PCR product generated using the CuTrMF2/CuTrMR2 oligonucleotide pair from strain DSE955 DNA. The *Bam*HI sites flanking the 1136 bp *erm* insertion (shaded in grey) in ORF *cutR* are underlined.

```
   1  TGACCCGGAA GTTACATCGA CAGAAGATTT GATTAAAGTT GTTAAACATG
  51  CGGGTTATGA CGCGGCGGAA AAAATGTCTA AAGAAGAAAA AGATGCTGTG
 101  TTAGAAAAGA ATTTCAAAAA AGAAGTGAGA CGTTTTATTC TTTCGGCGGT
 151  TCTTTCCTTA CCATTACTTC TTACCATGGT GACACATATT CCGTACATTC
 201  ATGAAATGGT GTTTGCGGAA ACGATTGGTA ACTGGATTAA TCCAACCATC
 251  CAATTAGTAC TCGCAACGAT TGTTCAGTTT TATATTGGTT GGCGGTTTTA
 301  TGATGGGGCT TATAAAGCGC TTCGAGGTAA AAGTGCAAAT ATGGATGTAT
 351  TAGTCGCACT TGGGACTTCT GCTGCATATT TCTATAGTGT AGTGGAATAC
 401  ATTCGTCACA TGATTGATCC GAGCGTGATG CCGCATTACT ACTTTGAAAC
 451  AAGTGCTGTG TTAATCACCT TAATCTTGTT AGGTAAATTA CTTGAATCAT
 501  ACGCAACATC GAGAACAACC GAATCTATTG CTGGTTTACT TGAATTACAA
 551  GCAAAAGAAG CGACCGTTAT TCGAGAAGGA AAAGAATGGT TAGTGCCAGT
 601  AGATTCATTG AAAATTGGCG ATATTATCCT TGTTCGTCCG GGTGAAAAAG
 651  TTCCGATGGA TCCGGACTAG TCCGGGCCCT CGAGCTCAAG CTAGCTTGGT
                BamHI
 701  ACCAGATCTG AGATCACGCG TTCTAGAGGT ACGTACCGCA TGCGATGCGG
 751  CCGCATCGGC CTAGGCGATT TAGAAGCAAA CTTAAGAGTG TGTTGATAGT
 801  GCAGTATCTT AAAATTTTGT ATAATAGGAA TTGAAGTTAA ATTAGATGCT
 851  AAAAATTTGT AATTAAGAAG GAGTGATTAC ATGAACAAAA ATATAAAATA
 901  TTCTCAAAAC TTTTTAACGA GTGAAAAAGT ACTCAACCAA ATAATAAAAC
 951  AATTGAATTT AAAAGAAACC GATACCGTTT ACGAAATTGG AACAGGTAAA
1001  GGGCATTTAA CGACGAAACT GGCTAAAATA AGTAAACAGG TAACGTCTAT
1051  TGAATTAGAC AGTCATCTAT TCAACTTATC GTCAGAAAAA TTAAAACTGA
1101  ATACTCGTGT CACTTTAATT CACCAAGATA TTCTACAGTT TCAATTCCCT
1151  AACAAACAGA GGTATAAAAT TGTTGGGAGT ATTCCTTACC ATTTAAGCAC
1201  ACAAATTATT AAAAAAGTGG TTTTTGAAAG CCATGCGTCT GACATCTATC
1251  TGATTGTTGA AGAAGGATTC TACAAGCGTA CCTTGGATAT TCACCGAACA
1301  CTAGGGTTGC TCTTGCACAC TCAAGTCTCG ATTCAGCAAT TGCTTAAGCT
1351  GCCAGCGGAA TGCTTTCATC CTAAACCAAA AGTAAACAGT GTCTTAATAA
1401  AACTTACCCG CCATACCACA GATGTTCCAG ATAAATATTG GAAGCTATAT
1451  ACGTACTTTG TTTCAAAATG GGTCAATCGA GAATATCGTC AACTGTTTAC
1501  TAAAAATCAG TTTCATCAAG CAATGAAACA CGCCAAAGTA AACAATTTAA
1551  GTACCGTTAC TTATGAGCAA GTATTGTCTA TTTTTAATAG TTATCTATTA
1601  TTTAACGGGA GGAAATAATT CTATGAGTCG CTTTTGTAAA TTTGGAAAGT
1651  TACACGTTAC TAAAGGGAAT GTAGATAAAT TATTAGGTAT ACTACTGACA
1701  GCTTCCAAGG AGCTAAAGAG GTCCCTAGCG CCTACGGGGA ATTTGTATCG
1751  ATGATAAGCT GTCAAGCATG AGAATTATCG CATGCGCCTC CCCGGATCCC
                                                  BamHI
1801  GAAATTATTT CTGGTGAAAC GAGTATTGAT GAAGCGATGA TTACTGGGGA
1851  ACCTGTACCA GTAGAGAAGA AACCAGGCGA TTCTGTTATT GGCGCAACGA
1901  TTAACTTTGA CGGGGCTTTC CAAGCAAAAA TTACGAAACG AATGGAAGAA
1951  ACCGTTTTAG AATCCATTAT TCGTTTGGTG GAAGAAGCGC AAGGTATTAA
2001  AGCGCCAATT CAACGTTTGG CAGATAGAAT TTCCGGCATA TTTGTACCAA
2051  TTGTACTTGG GATTGCTGCT GTAACCTTTA TTATTTGGTA TCTTGTTACT
2101  GGAACGGTGG ATGGTTCACT TGAAGCTGCG ATTGCGGTAT TAGTTATCGC
2151  CTGTCCCTGT GCGCTTGGTC TTGCAACGCC AACCGCTATC ATGGCTGGAA
```

Continued next page ……

2201 CTGGTAAAGG CGCCGAAAGT GGGATATTAT TTAAAGGTGG CGAACATTTA
2251 GAACGTACTT CCAAAGTGGA TACTATCGTT TTTGACAAAA CTGGTACTTT
2301 AACAGAAGGT AAACTAGAAG TGAGTGATAA AAAAGCAGCC AATGATCACT
2351 TTTTCCCTTA TTTATTCTTA ATGGAACAAC AATCAGAGCA TCCGATTGCG
2401 AAAGCGATTA TTAAGATGTT AGAGCCAGAG AATATAGATG CTTCCGCAGT
2451 GAAACAAGGG AAAATCCGCG CGAAAGCTGG GCACGGCATG ACCGGTAATC
2501 TGGATGATAG TAAAGTGGAA CTAGGTGCTT ACCGCTATGT TTCTTCCCTT
2551 ACAACGATTC CAAAAGAAGA TGATGAATTA ATCGAAAGTT GGATGCATGC
2601 AGGAAAAACA GTCGTAGCAA TGGCAATTGA TGGTGTATAC GCAGGTGCCC
2651 TCGCTTTATC TGACACACCA CGACCAGAAG CAAAAGAAGC CATCCAAAAA
2701 CTAAAAGCAC AAGGTATTAA AACAGCAATT TGTTCTGGGG ACCAATCTGT
2751 TGTCGTAGAA AATATGGCTA AAGATTTAGG TATTGATATG TTCTTTGCCG
2801 AACAACTACC AAATGATAAG AGCGCCTTAG TCGAGAAATT ACAGCAAGAC
2851 GGTCATATCG TTGCATTCGT TGGTGATGGC ATTAATGATG CTCCAGCTCT
2901 TGCGGCAAGT GATATTGGG

Appendix R

Amino acid translation of ORF pCT0017 from plasmid pETCF confirmed by protein mass spectrometry sequencing (see Appendix S, page 349). The introduced *Nde*I (CATATG) site at the 5’ end and the *Xho*I (CTCGAG) at the 3’end are indicated (underlined and bold). Introduction of the *Xho*I site at the 3’ end of the ORF resulted in an animo acid change from N → L (underlined) and the addition of an E (underlined). Translation of the His-Tag (HHHHHH) is indicated in bold.

```
  1 - CCCGCGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTA

 61 - GAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGAAAATAAACATCAAAGTA
  1 -                                       M  E  N  K  H  Q  S  N

121 - ATTCAAATGAAGTACTAATGAGCCAATTACCTTCTGATCCTAAAATATTAAAAGAATTAA
  9 -   S  N  E  V  L  M  S  Q  L  P  S  D  P  K  I  L  K  E  L  N

181 - ATATTTCGGACAGTGAATTAGTTATTATGCGTGTCGTTTGGTCTCTAGGGAGTACTACAG
 29 -   I  S  D  S  E  L  V  I  M  R  V  V  W  S  L  G  S  T  T  A

241 - CAGATGAAATTGGACGTGAGCTTAGTGAAACTTATCAGTGGAGCCCTTCCACAATAAAAA
 49 -   D  E  I  G  R  E  L  S  E  T  Y  Q  W  S  P  S  T  I  K  T

301 - CATTTTTAGCGCGTTTAATTAAAAAAGGATTATTAAAAAACAGTCGTGATGGTCGTAAAT
 69 -   F  L  A  R  L  I  K  K  G  L  L  K  N  S  R  D  G  R  K  Y

361 - ATGTCTATATTGCAACTTGTTCTGAAGATGAAGCGATTTGCCAAATGACACTATCTTTTT
 89 -   V  Y  I  A  T  C  S  E  D  E  A  I  C  Q  M  T  L  S  F  L

421 - TGAATAAAATTTGTGCACATAAACATGCCAATGTTATTTTAGAAATGATTGACGCGAGTA
109 -   N  K  I  C  A  H  K  H  A  N  V  I  L  E  M  I  D  A  S  S

481 - GTATTACTGCTGAAAATAAAGAAGCTATCAGCGAAAAACTAAGTAGCAAAAATGTTGTTG
129 -   I  T  A  E  N  K  E  A  I  S  E  K  L  S  S  K  N  V  V  D

541 - ACGAAGTAACTTGTGATTGTATAAACAGATTAAATTGTTGTGATAATCTCGAGCACCACC
149 -   E  V  T  C  D  C  I  N  R  L  N  C  C  D  N  L  E  H  H  H

601 - ACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTG
169 -   H  H  H  *

661 - CCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTT
```

Appendix S

Protein mass spectrometry sequencing results for His-tagged pCT0017 protein.

The MS spectra obtained for the band of interest. The peak masses were matched to the provided sequence as follows:

Peak number	Meas. M/z	Intensity	Ion score	Range	Missed cleavage	Sequence
peak 46	2026.878	142.654	-	5 – 22	0	HQSNSNEVLMSQLPSDPK 10: Oxidation (M)
peak 36	1872.976	118.6	-	23 – 38	1	ILKELNISDSELVIMR
peak 39	1888.981	1459.776	-	23 – 38	1	ILKELNISDSELVIMR 15: Oxidation (M)
peak 9	1518.724	1889.144	-	26 – 38	0	ELNISDSELVIMR
MSMS 11	1534.727	14943.89	532	26 – 38	0	ELNISDSELVIMR 12: Oxidation (M)
MSMS 17	1590.763	24263.278	274	39 – 53	0	VVWSLGSTTADEIGR
MSMS 27	1668.748	1547.37	1083	54 – 67	0	ELSETYQWSPSTIK
peak 52	2257.087	545.001	-	54 – 72	1	ELSETYQWSPSTIKTFLAR
peak 72	2900.259	223.246	-	87 – 110	1	KYVVYIATCSEDEAICQMTLSFLNK 17: Oxidation (M)
peak 68	2772.171	797.402	-	88 – 110	0	YVVYIATCSEDEAICQMTLSFLNK 16: Oxidation (M)
peak 47	2071.965	824.969	-	116 – 134	0	HANVILEMIDASSITAENK 8: Oxidation (M)
peak 63	2729.277	480.044	-	116 – 140	1	HANVILEMIDASSITAENKEAISEK 8: Oxidation (M)
peak 18	1593.649	6970.638	-	145 – 157	0	NVVDEVTCDCINR 8: Carbamidomethyl (C) 10: Carbamidomethyl (C)
MSMS 35	1859.693	5038.121	540	158 – 171	0	LNCCDNLEHHHHHH 3: Carbamidomethyl (C) 4: Carbamidomethyl (C)

MS/MS analysis of the 1534.727, 1590.763 and 1668.748 parent ions produced clear spectrum that confirmed the peptide assignments (Figure 2). The majority of the peaks in the spectra could be assigned to the protein sequence, thus this appears to be the only protein present in the band. The level of coverage of the modified protein sequence was 84.2% by MS analysis and 32.7% by MS/MS analysis.

Figure 1.
The MS spectra obtained for the band of interest. The peaks corresponding to the peptides that were matched to the provided sequence are shown. Blue peak masses indicate that the peptide contains an oxidized methionine residue.

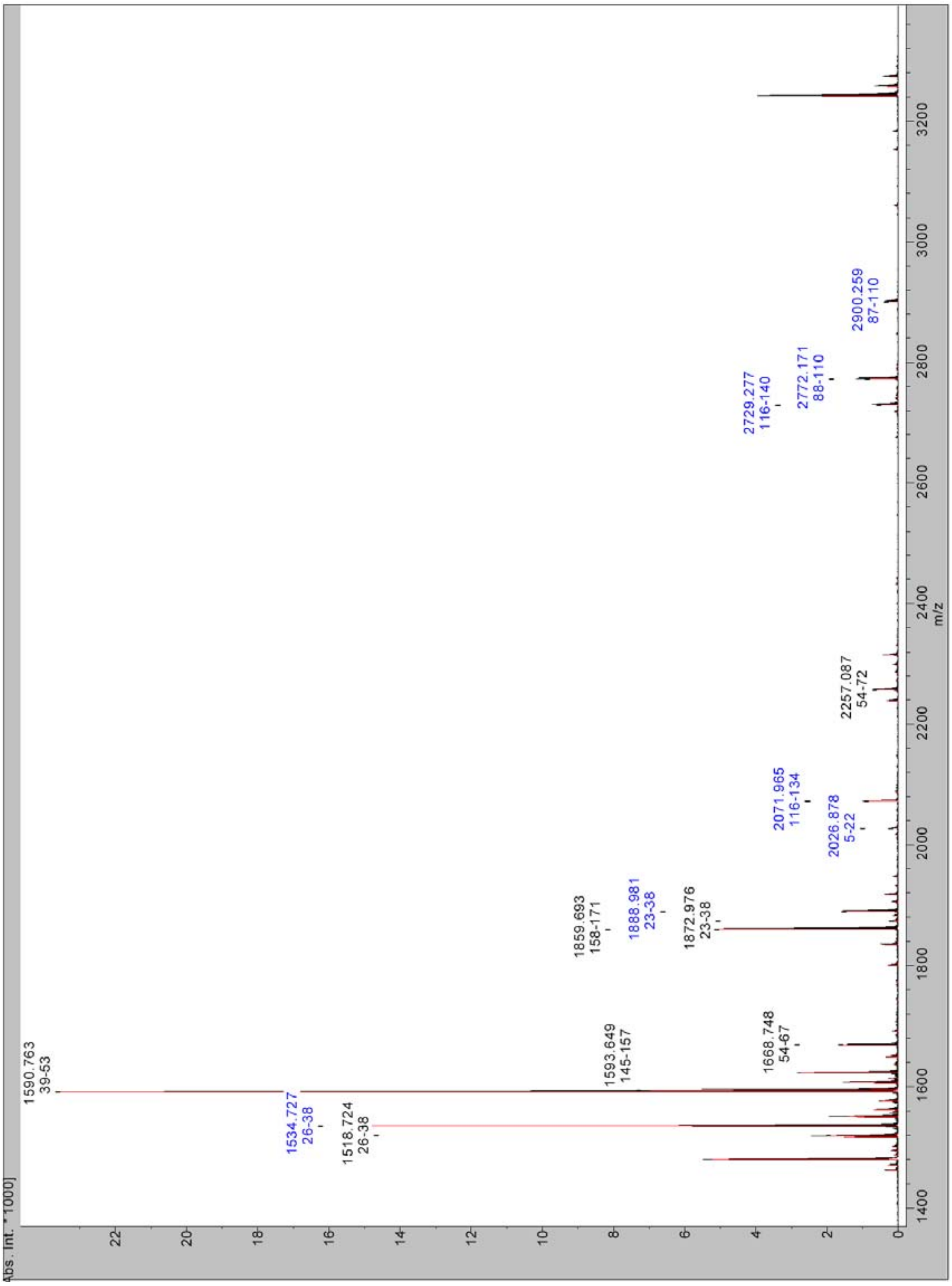
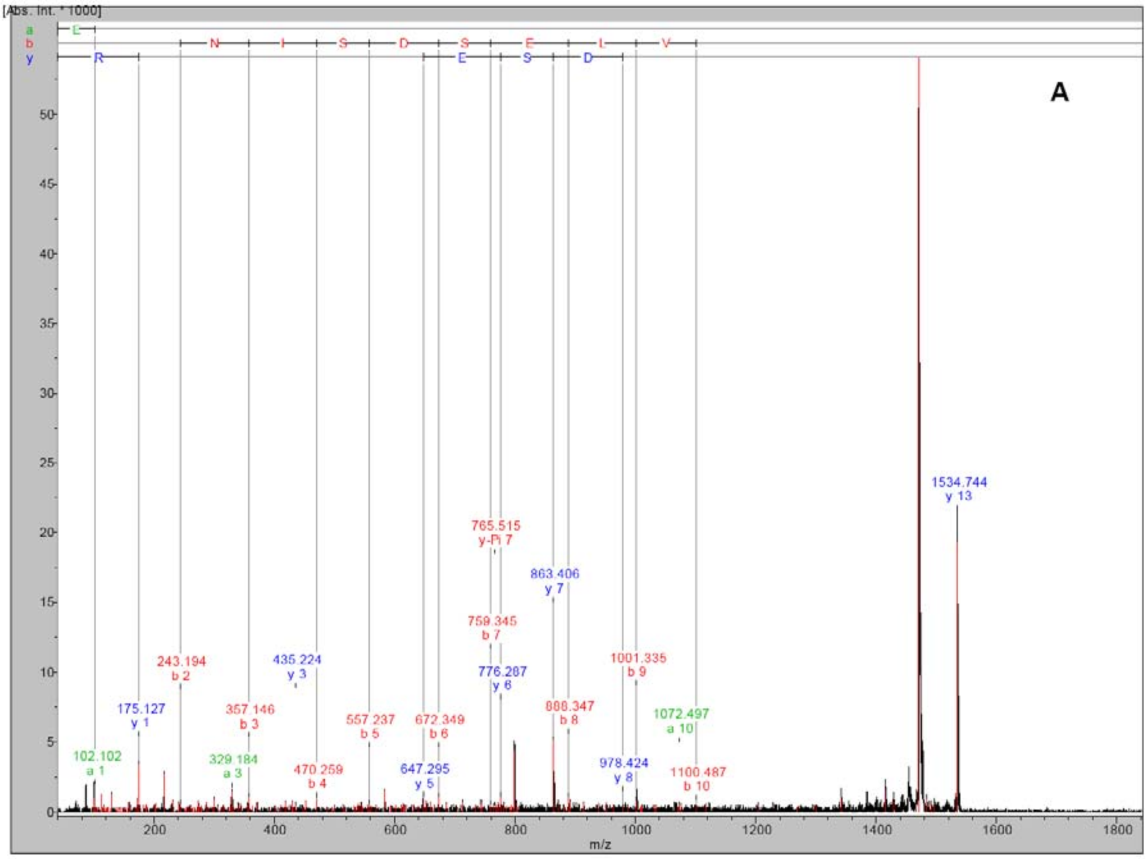
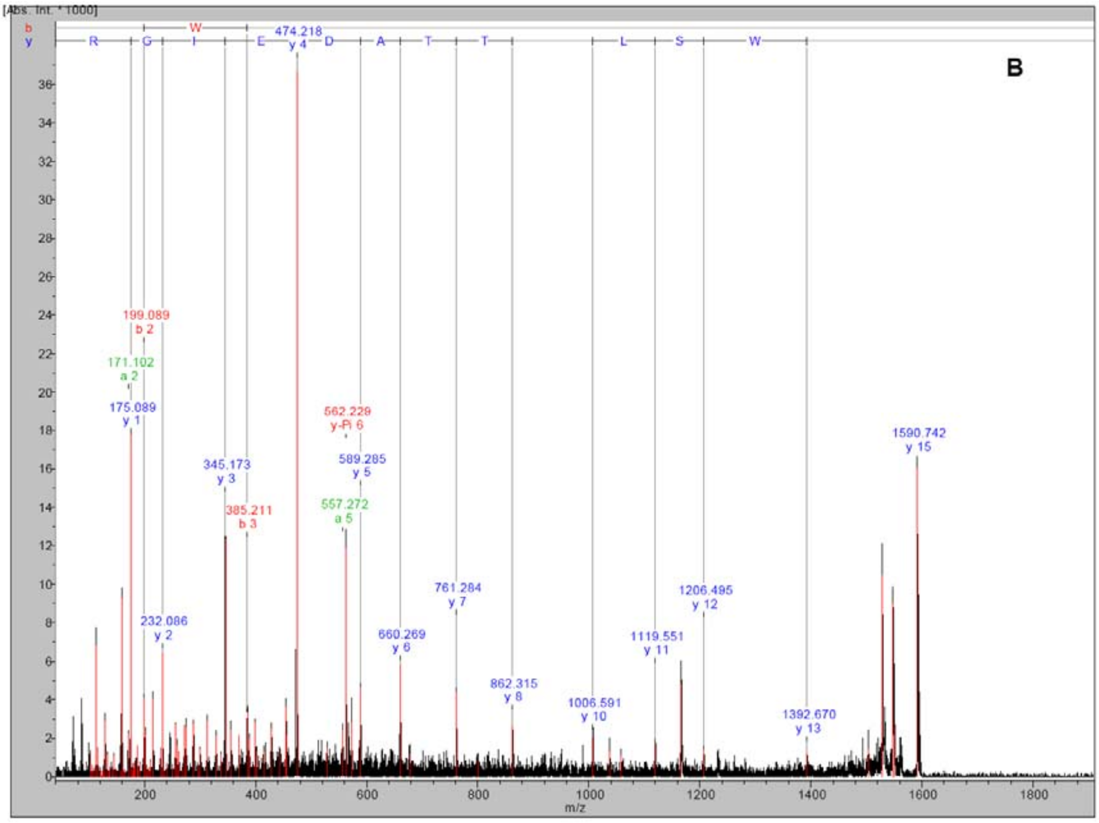


Figure 2.
The MS/MS spectra of the 1534.727 (A), 1590.763 (B) and 1668.748 (C) parent ions.




[Abs. Int. * 1000]
b
y
W
R G I E D A T T L S W
B
474.218
y 4
199.089
b 2
171.102
a 2
175.089
y 1
562.229
y-Pi 6
589.285
y 5
345.173
y 3
385.211
b 3
557.272
a 5
1590.742
y 15
761.284
y 7
1206.495
y 12
232.086
y 2
660.269
y 6
1119.551
y 11
862.315
y 8
1006.591
y 10
1392.670
y 13
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
30
32
34
36
200
400
600
800
1000
1200
1400
1600
1800
m/z

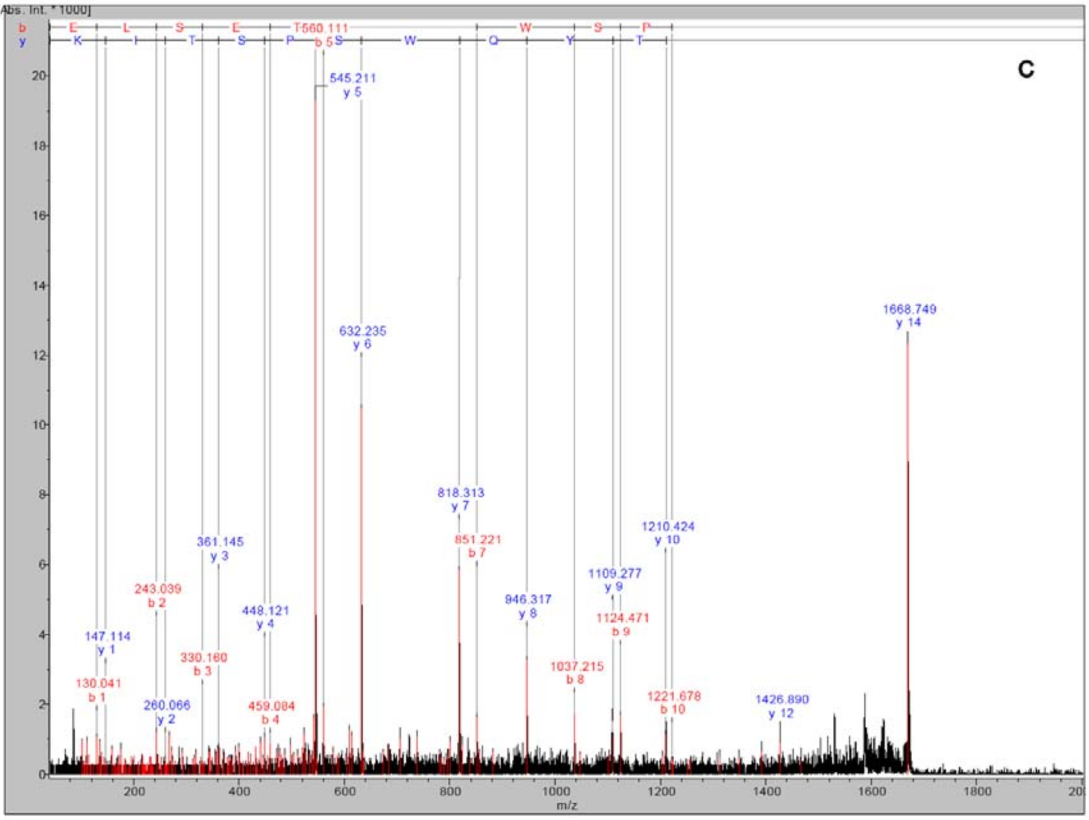

[Abs. Int. * 1000]
b
y
E L S E T W S P
K I T S P S W Q Y T
560.111
b 5
C
545.211
y 5
632.235
y 6
1668.749
y 14
818.313
y 7
851.221
b 7
1210.424
y 10
361.145
y 3
1109.277
y 9
243.039
b 2
946.317
y 8
448.121
y 4
1124.471
b 9
147.114
y 1
330.160
b 3
1037.215
b 8
130.041
b 1
1221.678
b 10
260.066
y 2
459.084
b 4
1426.890
y 12
0
2
4
6
8
10
12
14
16
18
20
200
400
600
800
1000
1200
1400
1600
1800
m/z

Figure 3.
The MS/MS spectra of the 1859.693 parent ion showing the sequence of the C-terminal peptide.

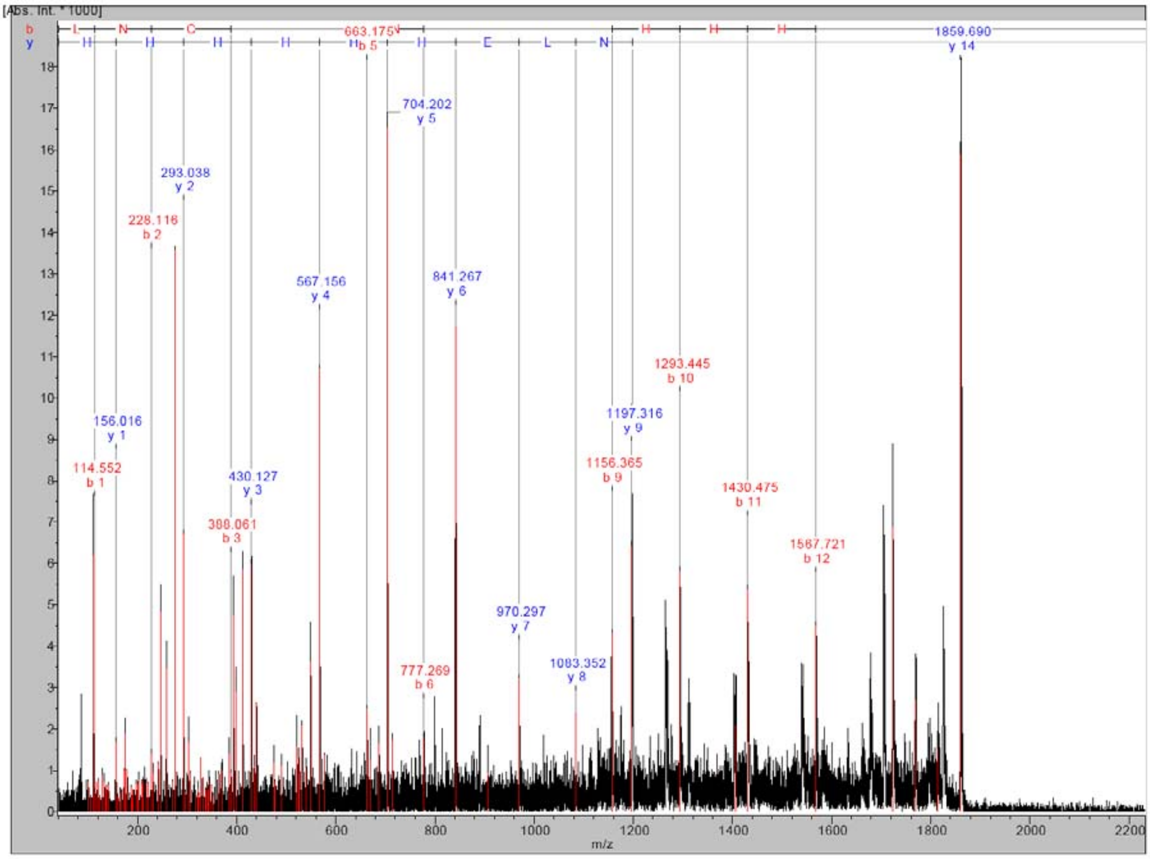


Figure 4.

The regions covered by the MS data are shown with the grey bars, whereby the depth of the grey colour corresponds to the peak's intensity in the spectra. The red boxes show the amino acids that were sequenced by MS/MS analysis, whereby the upper row of boxes represents the b-ion and the lower the y-ion series. The regions of sequence that were not identified contained peptides that were too small to be detected by the mass spectrometer (i.e. <800).