

Hepatitis C virus genotype 7, complete genome

NCBI Reference Sequence: NC_030791.1

[FASTA Graphics](#)

[Go to:](#)

LOCUS NC_030791 9443 bp RNA linear VRL 28-MAY-2019
DEFINITION Hepatitis C virus genotype 7, complete genome.
ACCESSION NC_030791
VERSION NC_030791.1
DBLINK BioProject: [PRJNA485481](#)
KEYWORDS RefSeq.
SOURCE Hepatitis C virus genotype 7
ORGANISM [Hepatitis C virus genotype 7](#)
Viruses; Riboviria; Orthornavirae; Kitrinoviricota; Flasuviricetes;
Amarillovirales; Flaviviridae; Hepacivirus.
REFERENCE 1 (bases 1 to 9443)
AUTHORS Murphy,D.G., Sablon,E., Chamberland,J., Fournier,E., Dandavino,R.
and Tremblay,C.L.
TITLE Hepatitis C virus genotype 7, a new genotype originating from
central Africa
JOURNAL J. Clin. Microbiol. 53 (3), 967-972 (2015)
PUBMED [25520447](#)
REFERENCE 2 (bases 1 to 9443)
CONSRMT NCBI Genome Project
TITLE Direct Submission
JOURNAL Submitted (19-AUG-2016) National Center for Biotechnology
Information, NIH, Bethesda, MD 20894, USA
REFERENCE 3 (bases 1 to 9443)
AUTHORS Murphy,D.
TITLE Direct Submission
JOURNAL Submitted (03-MAR-2015) Biologie Moleculaire, Laboratoire de Sante
Publique du Quebec, Institut National de Sante Publique du Quebec,
20045 Chemin Sainte-Marie, Sainte-Anne-de-Bellevue, Quebec H9X 3R5,
Canada
REMARK Sequence update by submitter
REFERENCE 4 (bases 1 to 9443)
AUTHORS Murphy,D.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-2006) Biologie Moleculaire, Laboratoire de Sante
Publique du Quebec, Institut National de Sante Publique du Quebec,
20045 Chemin Sainte-Marie, Sainte-Anne-de-Bellevue, Quebec H9X 3R5,
Canada
COMMENT REVIEWED [REFSEQ](#): This record has been curated by NCBI staff. The
reference sequence is identical to [EF108306](#).
COMPLETENESS: full length.

FEATURES	Location/Qualifiers
source	1..9443 /organism="Hepatitis C virus genotype 7" /mol_type="genomic RNA" /isolate="QC69" /isolation_source="plasma" /host="Homo sapiens" /db_xref="taxon: 1544901 " /country="Canada" /subtype="7a" /note="emigrant from the Democratic Republic of Congo"
<u>5' UTR</u>	1..339
<u>gene</u>	340..9381 /gene="POLY" /locus_tag="BFX91_gp1" /db_xref="GeneID: 35984104 "
<u>CDS</u>	340..9381 /gene="POLY" /locus_tag="BFX91_gp1" /codon_start=1 /product="polyprotein" /protein_id=" YP_009272536.1 " /db_xref="GeneID: 35984104 " /translation="MSTNPKPQRLTKRNTVRRPQNVKFPGGGQIVGGVYLLPRRGPRLGVRGTRKSSERSQPRGRRQRIPKAASSQGKAWGKPGYPWPLYGNEGCGWAGWLLSPRGSRPTWGPTDPRHRSRNLGKVIDTMTCGFADLMGYIPVLGAPLGGVARALAHGVRVLEDGVNYATGNLPGCSFSIFLLALLSCLTVPASAYEVRNSSGVYHLTNDCPNASIVYETDNAILHEPGCVPCVREGNTSRCWEPVAPTLAVRYRGALTDDLRTIDLVASATLCSALYVGDICGAIFIASQAVLWKPGGGRIVQDCNCSIYPGHVTGHRMAWDMMQNAPALSMVAAYAVRVPGVIIITVAGGHWGVLFGLAYFGMAGNWAKVILIMLLMSGVDAETMAVGARAHTTGALVSLNPGPSQRLQLINTNGSWHINRTALNCNDSLQTFIAALFYTHRFNSSGCPERMASCKPLSDFDQGWGPLWYNSTERPSDQRPYCWYAPSPCGIVPAKDVCGPVYCFTPSPVVVGTTDRRGVPTYTWGENESDVFLNSTRPPQGSWFGCSWMNTTGFTKTCGGPPCKIRPQGAQSNTSLTCPTDCFRKHPRATYSACGSGPWLTPRCMVHYPYRLWHYPCTVNFTIHKVRLYIGGVEHRLDAACNWTRGERCDLEDNRDVRDMSPLLHSTELAILPCS FVPLPALSTGLIHLHQNIVDAQYLYGLSPAIIISWAIWEVWVLVFLLLADARICACLWMMMLMAQAEAALENLIHLNAASLAGTHGIWWLLLVCASWHLRGRVPLVPTYGICGMWPFFLMLLSLPPRAYALDREVSAAALGTGMLAIILLVTLGPHYKRLALILWWVTYFLTRCEAALQTWVPPLNPRGGRDGFILCVLLCYPGLVFDITKWLLVMMCPLYLLQLCLV RTPYFVRAQALIRVCSLFKTLAGGRYVQAALLTIGRWTGTIYNHLAPLETWAAGGLRDLA VAVEPVIFSPMEKKIIVWGAETTACGDIICGLPVSARLGREVLGPPADDYRSMGWQLL APISAYAQQTRGLISTLVVSLTGRDKNETAGEVQVLSTSTQTFLGTNVGGVMWGPYHG AGTRTVAGRGGPVLQMYTSVSDDLVGWPAPPGSKSLEPCSCGSADLYLVTRNADVLPL RRKGDGTASLLSPRPVSSSLKGSSGGPVLCPQSHCVGIFRAAVCTRGVAKAVQFVPIEK

MQVAQRSPSFSNSTPPAVPSTYQVGYLHAPTGS GKSTKVPAA YASQGYKVLVLNPSV
AATLGFGAYMSKAYGIDPSVRTGARTVTTGAPITYSTY GKFLADGGCSGGAYDIIICD
ECHAIDATTVVGIGTVLDQAETSGVRLVVLATATPPGSVTVPHPNIEEVALGNDGEIP
FYGKAIP LQH IKGGRHLIFCHSKKKCDELAKLTSLGLTAVAYYRGLDVSVIPTSGDV
VVVATDALMTGFTGDFDSVIDCNVAVTQTVDFSLDPTFTIETTTVPQDSVSRSQRRGR
TGRGRLGIYRYVSSGERPSGMFDTSVLCECYDLGCSWYELTPSETTTRLRAYLNCPL
PVCQDHLEFWEGVFTGLTHIDAHFLSQTKQEGQNYAYLTAYQATVCARAKAPPSWDV
QWKCLQRLKPLLVGPTPLLYRLGSVTNEVTFTHPITKYIATCMAADLEVTTSTWVIVG
GVLA AVAA YCMSTGSVVVGRVVLGSNVVTAPDREVL YQH FDEMEEC SKAPELLKHAQ
TIGGMFKDKALAVLDTLKPAQA AVPIVETNFQKVEKLWNQHMMWNFISGIQYLAGLST
LPGNPTVASLMAFTASVTSPLATSTTLVLN ILGGWFASQLAPPSAATTFVVSGLAGAA
VGSVGLGKVLVDVLAGYGAGIAGALVAFKIMS GEVPSTEDLANLLPAILSPGALVVG
VCAAI IKRHTGTSEGV TQWMNRLIAFASRGNHVSPTHYIQDD DASKRVMGILSSLTIT
SLIKRVLAWAQTDYSAPCAGSWLREVWDWCMVLSDFASWLKAKVLP SLPGIPFLSCQ
KGYKGEWRNDGIMNTKCPCGALIAGHVKN G SMRIVGPKTCRNTWWGTFFINSHTTGPS
SPVPSHCYQRALWRVSATEYVEILRHNDQHYVVGVTAE DLKCPCQVPSPEFFSFVDGV
RIHRFAPEPKPMIREEAA FVVGLHSYVVGSQLPCEPEPDVQTVS QLLTDP SHITAETA
ARRLRGSPPSNASSSASQLSAPSLKATHTTLPQH PDAELIEANLMWEHKVGAI RME
TDTKVIILDSFDSASSVEDDMEPSTAAECLRTRKVFPPAMPIWARPDYNPPVENWKD
PEYAPPQVSGCALPPAQTPFPVPPRRKRAVIQLTESAVSTALAE LERSFPKEEAPPS
DSAISLDSPAANDPPSDCDQGSEISFSMPPLEGEPGDPDLS DGSWSTVSTRSDVICC
SMSYSWTGALVTPSGPEEERLPINALSNTMLRHYNMVYSTTSRSASQRAKKVTFDRLQ
VLDDHYKRALADVKADASTVKAQLLSVEEAAALTPAHSARSKFGYGAKEVRS LAPKAM
SHIKEVWKDLLQDMTTP IPTTIMAKNEVFCVNPAGGKKPARLIVY PDLGVRVCEKRA
LYDIAQKLPKAIMGQAYGFQYSPSQRVEYL VKTWKSKRTPMGFSYDTRCFDSTVTEQD
IRTESEIYQCCNLDPEARTIINALTERLYVGGPMFNSKGQRVGYRRCRASGVFPTSMG
NTMTCYIKAKAAAAAAGLESTDFLVCGDDL VVICESK GVERDRADLQAF AAAMTRYSA
PPGDMPQPAYDLEHIDSCSSNVSVARDNSGKR VYYLTRDPTNPLSRAAWETARHSPVN
SWVGNIIMFAPTIWVRMVLMT HFFALLLNEERLNDPVSFEMYGATYTVCPDLPDI IQ
RLHGLRAFELHTYSPAELTRVAATLRKLGVPPLRTWRQ RARKVRAGLIGQGGRARICG
LYLFNWAVRTKIKLTPLAGAGRLDLSSWFSVCAGEADVDHSTPRAHPRPLLLCLLLLA
VGVGIFLLPAR"

mat_peptide 340..912
/gene="POLY"
/locus_tag="BFX91_gp1"
/product="core protein"
/protein_id="[YP_009272680.1](#)"

mat_peptide 913..1488
/gene="POLY"
/locus_tag="BFX91_gp1"
/product="E1 protein"
/protein_id="[YP_009272681.1](#)"

mat_peptide 1489..2589
/gene="POLY"

```

        /locus_tag="BFX91_gp1"
        /product="E2 protein"
        /protein_id="YP\_009272682.1"
mat_peptide    2590..2778
        /gene="POLY"
        /locus_tag="BFX91_gp1"
        /product="p7 protein"
        /protein_id="YP\_009272683.1"
mat_peptide    2779..3429
        /gene="POLY"
        /locus_tag="BFX91_gp1"
        /product="NS2 protein"
        /protein_id="YP\_009272684.1"
mat_peptide    3430..5322
        /gene="POLY"
        /locus_tag="BFX91_gp1"
        /product="NS3 protease/helicase"
        /note="NS3 proteinase / ATPase / helicase peptide"
        /protein_id="YP\_009272685.1"
mat_peptide    5323..5484
        /gene="POLY"
        /locus_tag="BFX91_gp1"
        /product="NS4A protein"
        /protein_id="YP\_009272686.1"
mat_peptide    5485..6267
        /gene="POLY"
        /locus_tag="BFX91_gp1"
        /product="NS4B protein"
        /protein_id="YP\_009272687.1"
mat_peptide    6268..7605
        /gene="POLY"
        /locus_tag="BFX91_gp1"
        /product="NS5A protein"
        /protein_id="YP\_009272688.1"
mat_peptide    7606..9378
        /gene="POLY"
        /locus_tag="BFX91_gp1"
        /product="NS5B RNA-dependent RNA polymerase"
        /note="RdRp"
        /protein_id="YP\_009272689.1"
CDS            join(340..368,370..706)
        /gene="POLY"
        /locus_tag="BFX91_gp1"
        /ribosomal_slippage

```

/note="Potential -2/+1 frameshift protein F; position
 predicted based on homology to protein F in HCV1 RefSeq
 NC_004102.1; The predicted ribosomal slippage site appears
 to be mutated in this virus."
 /codon_start=1
 /product="protein F"
 /protein_id="[YP_009305096.1](#)"
 /db_xref="GeneID:[35984104](#)"
 /translation="MSTNPKPQRLPKETPSVAHRTLSSRVAGRSLAESTCCRAGALDW
 VCAALGRVRSRSPGEDANVSPKLP LHRVKPGASPGTLGFCMVTRAVAGQGGSCPPAA
 LDLLGAPXTPGTARETSVR"

3'UTR

9382..9443

ORIGIN

```

1 gccagcccct tacggggcga cactccgcc tgaatcactc ccctgcgagg aaccactgtc
61 ctcacgcaga aagcgtctag ccatgacggt agtatgagtg tcgtacagcc tccaggaccc
121 cccctcccgg gagagccata gtggtctgcg gaaccggtga gtacaccgga attgccggga
181 agactgggtc ctttcttgga tcaaccactc ctatgcccg agatttggtc gtgccccgcg
241 gagactgcta gccgagtagt gttgggtcgc gaaaggcctt gtggtactgc ctgatagggt
301 gcttgcgagt gccccgggag gtctcgtaga ccgtgcacca tgagcacgaa tcctaaacct
361 caaagattaa caaaagaaa caccgtccgt cgccacaga acgttaagtt cccgggtggc
421 gggcagatcg ttggcggagt ctacttggtg ccgcgcaggg gccctagatt ggggtgtcgc
481 ggcactagga agagttcgga gcgatcgag ccaggggaa gacgccaacg tatcccaaaa
541 gctgcctctt cacagggtaa agcctggggc aagcccggt acccttggtc cctgtatggt
601 aacgagggct gtggctgggc aggggtggtc ctgtccccc gcggctctcg acctacttgg
661 ggccccacyg acccccggca ccgctcgcga aacctcggt aggtgatcga caccatgacc
721 tgcggggttg ccgacctcat ggggtacatc cctgtcctag gcgccccct agggggcggt
781 gccagggctc tggcacatgg tgttagagtt ctggaggacg ggggtcaacta tgcaacaggg
841 aacttgcttg gttgctcctt ttctatcttc ytactagccc tcctgtcatg tctaacagtc
901 ccggcatcgg cttatgaagt ccgcaactcc agtggrgtct atcatctcac caatgactgc
961 cccaacgcta gtatagtcta tgaaacagac aacgccatcc tacacgagcc tgggtgctg
1021 ccttgcggtc gcgagggtaa tactagcagg tgttggaac cagtggcccc cactttggcg
1081 gtccgctatc gcggagcgt tactgacgat ttgcggacgc atattgacct agtgggtggc
1141 tcagctaccc tgtgtccgc cttgtacgtg ggggacattt gtggagccat cttcattgcc
1201 agccaagctg ttctctggaa gcccgggggg ggtcggatag tgcaagattg caattgttcg
1261 atctaccgga gccacgtcac cggccacagg atggcgtggg acatgatgca gaactgggcg
1321 ccggccttgt caatggttgc cgcttacgct gtgagagtgc ccggtgtcat cattaccact
1381 gtagcggggc gccactgggg tgtgttattt ggctcgtctt actttggtat ggcgggaaac
1441 tgggcaaagg taatactcat catgtactc atgtccggcg tcgacgcgga aaccatggcc
1501 gtcggggcta gggcgctca caccactggc gcccttgtca gcctgtcaa tccaggggcc
1561 agtcagcgcc tgcagctgat caacaccaat gggatcatgc acatcaaccg gaccgcttgc
1621 aactgcaatg actctttgca gacagggttc atagcggccc tcttctacac acatagggtc
1681 aattctagtg gctgtcccga gaggatggct tcttgtaaac ctctcagtga ctttgaccag
1741 ggggtggggc cgctgtggta caattcaaca gaaagacctt cggaccagcg accctattgc
1801 tggcactacg cgccatcgcc gtgtggtatt gtgccggcta aggatgtttg cgggtccggtc

```

1861 tactgcttta caccaagccc ggttgtggtg ggcaccacg atcgccgggg ggtgcccacg
1921 tatacttggg gtgaaaatga gtctgatgtc ttcytgctca acagcacaa ggcgccgcaa
1981 ggcagttggt ttgggtgctc atggatgaac acaacggggt tcacgaagac ctgcgaggt
2041 cctccgtgca agatacgtcc ccagggtgcc cagagtaaca cctctctcac ttgtcccaca
2101 gactgcttca ggaaacatcc gcgtgccaca tactccgctt gcggatctgg tccgtggttg
2161 acacctagat gcatgggtcca ttacccttat agactgtggc actaccctg tacagtcaac
2221 ttcaccatac acaaagtcag gttatacata gggggtgtag aacataggct cgatgcagcg
2281 tgcaattgga cgcggggtga gcgatgcgac ctggaggacc gagacagggt ggacatgtcc
2341 cccctgctcc attccactac ggagctcgca atacttccgt gtccctttgt gccgttccg
2401 gccttatcta cgggactgat ccacctgcac caaaacatcg ttgacgcca gtacctttat
2461 ggtctttctc ccgtataat aagctgggcc atcagatggg agtgggtagt cctcgtttct
2521 ctactcctgg cggacgcgcg gatctgcgcc tgcccttgga tgatgatgct tatggcccag
2581 gctgaagccg ctctggagaa cttgatccac ctcaacgcg ccagccttgc gggaacccat
2641 ggtatctggt ggctcctttt agtcttttgt gcctcttggc atctacgagg cagggttgtc
2701 cctctggtga cgtatgggat atgcgggatg tggcccttct tctcatgtt gctgagcctc
2761 cccccacgag cgtatgctct ggacaggga gtgagcgag cgttgggaac gggcatgctc
2821 gccatcatcc tattagttac cttgggaccg cactacaaga gacttctagc ccttattctc
2881 tgggtgggtca catatttcct tacaagggtg gaagcagcac tccaaacgtg ggtccctcct
2941 ctcaaccctc gggggggcag ggacggttct atcctgtgtg tgctgctgtg ctatccaggc
3001 cttgtctttg acatcacaaa atggttgcgt gtcgatgatg gccctctcta cctcctccag
3061 ttgtgttttg tgaggactcc atactttgtg agggcccagg ccctcatcag agtgtgttct
3121 ctcttcaaaa cgctagctgg gggacggtac gtgcaggccg cgctgctcac tattggccgc
3181 tggaccggca cttatattta taaccatctc gccccctg aaacatgggc cgccggcggc
3241 ctacgggatt tggccgttgc agtcgagccc gtgatattct ccccatgga gaagaagatc
3301 atagtttggg gggcggagac cactgcttgt ggcgacattc tttgtggcct gcctgtctca
3361 gctcggctcg gcagggaagt cctgctaggg ccgcgggatg actacaggtc catgggatgg
3421 caactcctgg ctccaatctc agcatacgca cagcagaccc gcggtctcat ctctaccctt
3481 gtcgtgagcc tcactggccg tgataagaac gagactgccg gcgaggtgca ggtgttgtcc
3541 acctcaacgc agaccttctt ggggaccaat gtgggcgggt tcatgtgggg accttatcac
3601 ggagcgggca cgcgactgt ggccggccg ggtgggcctg tctccagat gtacacgtca
3661 gtcagtgacg atctgggtgg ctggcctgcc ccgcctggct ccaagtccct cgagccctgc
3721 tcgtgcgggt cagcagacct ttacttgggt acgcgcaatg ctgacgtcct tctcttagg
3781 aggaagggtg acggtacagc gtccctactg tccccgcgcc ctgtgtcctc ccttaaaggc
3841 tctcgggag ggctgtctt gtgccccaa agtcattgcg ttggcatctt cagagctgct
3901 gtctgcaccc gcggtgtagc caaagcagtt cagtttgttc ccatcgagaa gatgcagggtg
3961 gccagaggt cgccatcatt cagtgacaat agcaccctc cagcgggtgcc cagcacctat
4021 caagtaggct atttacatgc cccactggc agcggaaga gactaagggt cctgcggcg
4081 tatgccagcc agggctacaa agtgctcgtg cttaatccat cagtcgcggc cacccttggg
4141 tttggtgcct acatgtccaa ggcttatgga attgatccaa gtgtgcgcac cggcgcgcgc
4201 acggtgacca cgggggctcc cataacttac tcaacctatg ggaagtccct tgccgacggg
4261 ggttgttcag gtggtgccta tgatataatc atatgtgatg aatgccatgc catcgatgct
4321 accaccgtgg tcggcatagg caccgtgttg gaccaggcag agactagtgg tgtccgactt
4381 gtggtgctag ctacggccac ccccccgga agcgttacg tgccccaccc caacatcgag
4441 gaggtcgccc ttggcaacga tggggaaatt cccttctatg ggaaagccat cccctgcaa

4501 cacatcaagg ggggaaggca tctcattttc tgtcattcca aaaagaaatg tgatgagctt
4561 gccgggaagc tcacctcatt ggggctgact gccgttgctt actatagggg gttggatggt
4621 tcggtgatcc caacgtctgg cgatgtagtc gtggtggcta ctgatgcctt gatgacgggc
4681 tttaccggag acttcgactc cgtcatagac tgcaacgtgg cggttaccca gactgtagac
4741 ttctccttgg accccacctt cactattgag accaccacag tcccgcaaga ttctgtctcg
4801 cgctctcaga gacgtgggag gacaggtagg ggcaggcttg gcacttacag atacgtgtcc
4861 agtggcgagc ggccatccgg tatgttcgac accagcgtgc tctgtgagtg ctatgacctg
4921 gggtgctcgt ggtatgaact aacgcccagt gaaaccacca cgcgcctcag agcctacctc
4981 aattgccctg gtcttccagt gtgccaaagc catctggagt tctgggaggg cgtctttacg
5041 ggacttactc atattgatgc acacttcctg tcccagacca aacaagaggg ccaaaattat
5101 gcgtacctga ctgcatacca agctacagtg tgcgcccggg cgaaagcccc cccgcctcc
5161 tgggacgtcc agtggaagtg tctccaaaga ctaaagccac tgtagtagg gccacgccc
5221 ttgctataca ggctaggcag tgttaccaat gaagtcacat tcactcatcc catcaccaaa
5281 tacattgcca cgtgcatggc tgccgacttg gaggtcacta cgagcacctg ggtgatcgta
5341 ggtggcgtcc tcgcccgtt ggccgcatac tgcatgtcta ctggcagtg tggtgtggtt
5401 ggccgtgtgg tccttgatc caacgtggtg acggcccccg acagagaggt ttataaccag
5461 cactttgatg agatggaaga gtgctccaaa gctcctgaat tgctgaaaca cgcccaaacc
5521 atcggcggaa tgttcaagga caaggccctc gccgtcttgg atacgtcaa gccggccgcc
5581 caggctgctg ttcccattgt ggagaccaac ttccaaaagg tggagaagct gtggaatcag
5641 catatgtgga acttcataag cggcattcaa tacctagctg ggctatccac actgccaggt
5701 aatccaactg tggcctcact gatggccttc accgcctcag ttactagccc ccttgctaca
5761 tcgaccacac tactagtcaa catcctgggt ggggtggttcg caagccagtt agcgccgccg
5821 tctgctgcca cgactttcgt ggtctcgggc ctccgagggg cggcagtggt tagtggtggc
5881 ttgggcaagg tcttagttga tgtcctggcc ggctatggcg ccggcatagc tggcgccctg
5941 gtggccttca aaatcatgtc tggagaggtc cctagcacag aggatctggc aaacctctta
6001 cctgccatac tctcacctgg agcttttagt gtgggagttg tttgtgccgc gatcatcaaa
6061 cggcataccg gcactagtga gggagtcacg cagtggatga accgtctcat tgcggttcgcc
6121 tcccgtggaa atcacgtgtc gccgaccac tacatacagg atgatgatgc gtcaaaacgc
6181 gtcatgggga tactcagctc ccttaccatc accagcctca taaaacgagt actggcctgg
6241 gcgcagactg aytattctgc gccctgtgcc gggagctggc tccgggaggt gtgggactgg
6301 gtttgcatgg tcctatccga ctttgcatct tggttgaagg ccaaggtagt tccagcctt
6361 ccagggatcc cattcctatc ctgtcagaaa ggggtataaag gtgagtggcg gaacgatggc
6421 atcatgaaca ccaagtgcc gtgtggggcc ctgatagcag ggcacgtcaa gaacggatcg
6481 atgcgcatcg tgggccccaa gacctgccga aacacctggt ggggaacctt ccctattaat
6541 tcccatacca cgggaccaag ttctcctgta cctcccayt gttatcaaag ggccttgtgg
6601 agggatatct cracggagta cgtcgagatc ctctgcacaa atgaccagca ttaygtggtg
6661 ggagtgcagg cggaggacct caagtgcccg tgtcaggttc cttccccga attcttcagc
6721 tttgtggacg gggcgcgat acacagattc gccctgaac ccaagcccat gatccgggaa
6781 gaagccgctt tcgtgggtgg tctgcactca tatgtggtgg ggtctcagct cccttgtgaa
6841 ccggagccgg acgtccaaac tgtctcgag ctcttaaccg accctcaca catcaccgca
6901 gaaaccgcgg ctagaaggct caggagaggt tccccccat ccaacgccag ttcttctgct
6961 agtcaactct cagctccatc tctaaaagcc acccacacaa ctctgccaca acatcctgac
7021 gcagaactga tagaggcaa cctcatgtgg gaacataagg ttgggtgcat cagaaggatg
7081 gagacagaca ccaaagtcac catacttgac tcctttgatt cagcttcttc cgtygaggat

7141 gacatggaac cgtccaccgc cgcggagtgt ctgcgtacgc gtaaggtctt cccaccggcc
7201 atgcccataat gggcgaggcc agactacaat cctcctgttg tggagaattg gaaggacccc
7261 gagtacgcac ctccccaggt gtctgggtgc gctctaccac ctgcgcagac cctcctgtt
7321 ccaccacca ggcggaagcg ggcggttatc caactcacag agtccgcggt ttccacggca
7381 ctggcggaac tcgctgagag gtccctccca aaggaggagg cccccccag cgactccgcc
7441 attagtctgg attcaccggc ggcaaatac cctccatctg attgtgacca aggcagttag
7501 atctctttct catccatgcc acccctagag ggggagccg gtgatccaga cttgtctgat
7561 ggctcttggc caacagtcag caccagatct gacgtgattt gctgttccat gtcctactca
7621 tggacgggag ccttagtgac gccctccggt ccggaggagg aaaggctccc aataaatgcc
7681 ctyagtaaca ccatgcttcg gcattataac atgggtttaca gcacaacatc acgctcagct
7741 tcacaaagag caaaaaaagt gacttttgac agattgcaag tcctagacga cactacaaa
7801 agggcgctcg ccgacgtgaa agcggatgct tccacggta aggcgcagtt gctctcagta
7861 gaggaagctg ctgccctcac cctgcccac tcagccagat ctaagtttg gtacggggca
7921 aaggaagtca gatctcttgc gcccaaggca atgtctcaca tcaaagaggt ctggaaggac
7981 ttgctacagg acatgaccac cccgataccg accaccataa tggccaagaa tgaagtcttc
8041 tgcgtaaatc ctgccaaagg cgggaaaaag ccagctagac tgattgttta ccctgaccta
8101 ggtgtgaggg tatgcgagaa gcgggctttg tacgacatag ctcaaaagct tccgaaggcc
8161 ataatggggc aggcgtacgg gttccaatat tcaccatcgc agcgggtcga rtacytagtt
8221 aagacgtgga agtccaaacg gactcccacg ggcttctcat atgacacccg ctgctttgat
8281 tccactgtca cagaacaaga catccgcact gagagcgaga tctatcagtg ctgcaacctc
8341 gaccccgagg ctgcacccat cattaacgcc cttaccgaga ggttgtagct ggggggaccc
8401 atgttcaact ctaaaggatc gagggtcggt tatcgaaggt gcagagccag cggagttttc
8461 cccaccagca tgggaaacac tatgacgtgc tacatcaaag ccaaagcggc tgcagcagct
8521 gcagggctag agagcactga tttcctggtc tgcggcgatg acttggtggt catttgcgag
8581 agcaagggag tcgaacggga tagggcgagc ctgcaagcct ttgcagcggc aatgaccagg
8641 tactccgctc caccggggga catgccccag cccgcttatg accttgagca cattgactct
8701 tgcctcatcca atgtctcagt cgcgggggac aactcgggga aaagagtcta ctacctact
8761 agagacccta ccaaccact atccagggca gcttgggaga ccgccgtca ctcacctgta
8821 aactcctggg tgggcaacat aatcatgttt gccccacca tatgggtgag gatggttctc
8881 atgactcact ttttcgctct cctcctcaat gaggagcgat tgaacgatcc agtgagtttt
8941 gagatgtatg gtgcaacctc cacggtctgc ccaacagact taccagacat tatacaaaga
9001 ctccatgggc tccgcgcatt tgaactccat acttactctc cagcagaact gacgcgggtg
9061 gcagctaccc tcaggaagct tggggtgcc cccctcagga cctggagaca acgggctcgc
9121 aaggctccgc cggggctgat aggccagggg ggcagggctc gcatatgcg tctctacctc
9181 ttcaactggg ctgtcaggac caagatcaaa ctactccac ttgcaggcgc tggccggctc
9241 gatctttcga gctggttcag cgtttgcgcc ggcgaggcgc acgtggatca cagcacgccc
9301 agagcccatc ctgcgccatt actcctgtgc ctactcctac ttgcgtagg ggtaggcatt
9361 ttctcctac ccgctcggta agcgggtaac actccaggct tttaagccat aaattctggt
9421 tttttttttt tttttttttt ttt