

Aspergillus niger contig An16c0070, genomic contig

GenBank: AM270359.1

[FASTA Graphics](#)

[Go to:](#)

LOCUS AM270359 42878 bp DNA linear PLN 14-MAR-2015

DEFINITION Aspergillus niger contig An16c0070, genomic contig.

ACCESSION AM270359

VERSION AM270359.1

DBLINK BioSample: [SAMEA3283178](#)

KEYWORDS .

SOURCE Aspergillus niger

ORGANISM [Aspergillus niger](#)

Eukaryota; Fungi; Dikarya; Ascomycota; Pezizomycotina;
Eurotiomycetes; Eurotiomycetidae; Eurotiales; Aspergillaceae;
Aspergillus; Aspergillus subgen. Circumdati.

REFERENCE 1

AUTHORS Pel,H.J., de Winde,J.H., Archer,D.B., Dyer,P.S., Hofmann,G.,
Schaap,P.J., Turner,G., de Vries,R.P., Albang,R., Albermann,K.,
Andersen,M.R., Bendtsen,J.D., Benen,J.A., van den Berg,M.,
Breestraat,S., Caddick,M.X., Contreras,R., Cornell,M.,
Coutinho,P.M., Danchin,E.G., Debets,A.J., Dekker,P., van
Dijk,P.W., van Dijk,A., Dijkhuizen,L., Driessen,A.J.,
d'Enfert,C.,
Geysens,S., Goosen,C., Groot,G.S., de Groot,P.W., Guillemette,T.,
Henrissat,B., Herweijer,M., van den Hombergh,J.P., van den
Hondel,C.A., van der Heijden,R.T., van der Kaaij,R.M., Klis,F.M.,
Kools,H.J., Kubicek,C.P., van Kuyk,P.A., Lauber,J., Lu,X., van der
Maarel,M.J., Meulenbergh,R., Menke,H., Mortimer,M.A., Nielsen,J.,
Oliver,S.G., Olsthoorn,M., Pal,K., van Peij,N.N., Ram,A.F.,
Rinas,U., Roubos,J.A., Sagt,C.M., Schmoll,M., Sun,J., Ussery,D.,
Varga,J., Vervecken,W., van de Vondervoort,P.J., Wedler,H.,
Wosten,H.A., Zeng,A.P., van Ooyen,A.J., Visser,J. and Stam,H.

TITLE Genome sequencing and analysis of the versatile cell factory
Aspergillus niger CBS 513.88

JOURNAL Nat. Biotechnol. 25 (2), 221-231 (2007)

PUBMED [17259976](#)

REFERENCE 2 (bases 1 to 42878)

AUTHORS Pel,H.J.

TITLE Direct Submission

JOURNAL Submitted (01-MAY-2006) Pel H.J., DSM, 624-0295, P.O. Box 1, 2600
MA Delft, THE NETHERLANDS

FEATURES Location/Qualifiers

source 1..42878

/organism="Aspergillus niger"

```

/mol_type="genomic DNA"
/strain="CBS 513.88"
/db_xref="taxon:5061"
/clone="An16c0070"
gene      <1342..>3292
          /locus_tag="An16g01800"
mRNA      join(<1342..1683,1749..2292,2331..>3292)
          /locus_tag="An16g01800"
CDS       join(1342..1683,1749..2292,2331..3292)
          /locus_tag="An16g01800"
          /inference="profile:COGS:COG0724"
          /inference="profile:PFAM:PF00076"
          /inference="similar to AA sequence:PIR:S39836"
          /note="unnamed protein product; Title: strong similarity
to hypothetical protein YBL051c - Saccharomyces
cerevisiae"
          /codon_start=1
          /protein_id="CAK42797.1"
          /db_xref="GOA:A2R701"
          /db_xref="InterPro:IPR000504"
          /db_xref="InterPro:IPR001374"
          /db_xref="InterPro:IPR012677"
          /db_xref="InterPro:IPR024642"
          /db_xref="UniProtKB/TrEMBL:A2R701"
          /translation="MSYQQAQDMYHDNSAARSPGSQRHQQLHRQPSRQFDAYGMPMV
NLYDDMARYDTGRLERLNSSLHNNSYAYDLAGSQTNPNPGFANAQALGGIRSASTSLK
TTSRTGRAGLPTTWLDQQPGIPSAFNSNLGPGPLQSSAMRPEASGPSEAEDELIPTAIV
IKNIPFAVKKEQLVQLMTELNPLPYAFNYHFDNGVFRGLAFANFTSAEETATVIEVL
NHFELQGRKLRVEYKKMLPLQERERIEREKRRERGQLEEQHRPMATSQLQTQSSMSSL
TSHLPATSPSPVSQRGQKLVANPCRVSTEVDLNDSTTLSYYSQLLLKFEDTARDTLVF
PSNLSFPVQRRTVHTLAHNMGLGHASRGSGDQRQVHVFKVAPGTNVSPPLSSIPAAVQP
AETARRGLNRAATIDFSESRLNDGPTHFNTLRGQPSSGFLGVLDSPGNFGNTQNLRAAK
SFADLRSYTPSPVPSSASFPAALQSNGARLQHYDGATSGTSNTPTLTLSAPSGSSLGMQ
RDDSLLVNSLSSLSLGTGIGGPNASPRRLRGMFSWEQPEQPSSAGPIGSNRSIGVGF
DGQSQERMPIRQPRGPMPEKGPGRFRRQNGHQSRGSDELRTSSGVEIIVE"
exon      1342..1683
          /locus_tag="An16g01800"
          /number=1
intron    1684..1748
          /locus_tag="An16g01800"
          /number=1
exon      1749..2292
          /locus_tag="An16g01800"
          /number=2

```

```

intron      2293..2330
            /locus_tag="An16g01800"
            /number=2
exon        2331..3292
            /locus_tag="An16g01800"
            /number=3
gene        complement(<5428..>6282)
            /locus_tag="An16g01810"
mRNA        complement(join(<5428..5675,5759..5828,6184..>6282))
            /locus_tag="An16g01810"
CDS         complement(join(5428..5675,5759..5828,6184..6282))
            /locus_tag="An16g01810"
            /inference="profile:COGS:COG5262"
            /inference="profile:PFAM:PF00125"
            /inference="similar to AA sequence:PIR:S52560"
            /note="unnamed protein product; Function: the histone H2A
            variant encoded by the pht1 gene in S. pombe is important
            for chromosomal structure and function, possibly including
            a role in controlling the fidelity of chromosomal
            segregation during mitosis.;
            Title: strong similarity to histone H2A variant pht1p -
            Schizosaccharomyces pombe"
            /codon_start=1
            /protein_id="CAK42798.1"
            /db_xref="GOA:A2R702"
            /db_xref="InterPro:IPR002119"
            /db_xref="InterPro:IPR007125"
            /db_xref="InterPro:IPR009072"
            /db_xref="UniProtKB/Swiss-Prot:A2R702"
            /translation="MPGGKGKKSIGGKAGSKDSAGKSQKSHSAKAGLQFPCGRVKRFLK
            NNTQNKMVRVGA KA AVYVTAVLEYLTAEVLELAGNAAKDLKVKRITPRHLQLAIRGDEE
            LDTLIRATIAFGGVLPRINRALLLKVEQKKKNKTEA"
exon        complement(5428..5675)
            /locus_tag="An16g01810"
            /number=1
intron      complement(5676..5758)
            /locus_tag="An16g01810"
            /number=1
exon        complement(5759..5828)
            /locus_tag="An16g01810"
            /number=2
intron      complement(5829..6183)
            /locus_tag="An16g01810"
            /number=2

```

exon complement (6184..6282)
 /locus_tag="An16g01810"
 /number=3
 gene <7251..>8242
 /locus_tag="An16g01820"
 mRNA join(<7251..7267,7373..7533,7591..7972,8191..>8242)
 /locus_tag="An16g01820"
 CDS join(7251..7267,7373..7533,7591..7972,8191..8242)
 /locus_tag="An16g01820"
 /inference="profile:COGS:COG5065"
 /inference="similar to AA sequence:PIR:S48271"
 /note="unnamed protein product; Title: strong similarity
 to hypothetical inorganic phosphate transporter and
 regulator of Pho81p Pho88 -Saccharomyces cerevisiae;
 See PMID 8709965"
 /codon_start=1
 /protein_id="[CAK42799.1](#)"
 /db_xref="GOA:[A2R703](#)"
 /db_xref="InterPro:[IPR012098](#)"
 /db_xref="InterPro:[IPR019263](#)"
 /db_xref="UniProtKB/TrEMBL:[A2R703](#)"
 /translation="MVSPQVTNLAIIVMMQLAKKVPFEDPDVLMMLVRGMYILSNVII
 LGVYLYTQSKINKKNDLTTLKYVEPSPMGSGEEPKPVTTTNMEYDKGQLRQLMRSQLM
 GVGMMGVMHLYFKYTNPLLIQSIIPKGAFESNLVKIHVF GKPATGDIQRPFKANNSF
 LNQGQIKSDKASVENAEKNWRGGVKEETTSVVTSCPIITHAAA"
 exon 7251..7267
 /locus_tag="An16g01820"
 /number=1
 intron 7268..7372
 /locus_tag="An16g01820"
 /number=1
 exon 7373..7533
 /locus_tag="An16g01820"
 /number=2
 intron 7534..7590
 /locus_tag="An16g01820"
 /number=2
 exon 7591..7972
 /locus_tag="An16g01820"
 /number=3
 intron 7973..8190
 /locus_tag="An16g01820"
 /number=3
 exon 8191..8242

```

/locus_tag="An16g01820"
/number=4
gene      <10132..>11567
/locus_tag="An16g01830"
mRNA      join(<10132..10143,10219..10254,10314..10327,10429..10492,
10568..11252,11309..11479,11539..>11567)
/locus_tag="An16g01830"
CDS       join(10132..10143,10219..10254,10314..10327,10429..10492,
10568..11252,11309..11479,11539..11567)
/locus_tag="An16g01830"
/EC_number="1.2.1.12"
/inference="profile:COGS:COG0057"
/inference="profile:PFAM:PF00044"
/inference="profile:PFAM:PF02800"
/inference="similar to AA sequence:SWISSPROT:G3P.ASPNG"
/note="Catalytic activity: D-glyceraldehyde 3-phosphate +
phosphate + NAD(+) <=> 3-phospho-D-glyceroyl phosphate +
NADH;
Gene-ID: gpdA;gpd;
Mapping: gpdA from A. niger is mapped to chromosome V; see
list from DSM, EMBL X99652.;
Remark: a splice site was detected upstream of the START
codon.;
Remark: also acts very slowly on D-glyceraldehyde and some
other aldehydes. Thiols can replace phosphate.;
Similarity: identical to glyceraldehyde-3-phosphate
dehydrogenase of Aspergillus niger."
/codon_start=1
/product="glyceraldehyde-3-phosphate dehydrogenase
gpdA-Aspergillus niger"
/protein_id="CAK42800.1"
/db_xref="GOA:A2R704"
/db_xref="InterPro:IPR006424"
/db_xref="InterPro:IPR016040"
/db_xref="InterPro:IPR020828"
/db_xref="InterPro:IPR020829"
/db_xref="InterPro:IPR020830"
/db_xref="InterPro:IPR020831"
/db_xref="UniProtKB/TrEMBL:A2R704"
/translation="MAPKVGINGFGRIGRIVFRNAINHGEVDVVAVNDPFIETHYAAAY
MLKYDSTHGQFKGTIETIYEGLIVNGKKIRFFAERDPAAIPWGTTGADYIVESTGVFT

```

TQEKAAHLKGGAKKVVISAPSADAPMFVMGVNNTSYTKDINVLSNASCTTNCLAPLA
KVINDKFGIVEGLMTTVHSYTATQKVVDAPSSKDWRGGRTAAQNIIPSSTGAAKAVGK
VIPTLNGKLTGMAMRVPTSNVSVVDLTCRLEKATSYDEIKKALKDASENELKGILGYT
EDDIVSSDLNGDDHSSIFDAKAGIALNSNFVKLVSWYDNEWGYSRRVVDLIAYISKVD
AQ"

exon	10132..10143 /gene="gpdA" /locus_tag="An16g01830" /number=1
intron	10144..10218 /gene="gpdA" /locus_tag="An16g01830" /number=1
exon	10219..10254 /gene="gpdA" /locus_tag="An16g01830" /number=2
intron	10255..10313 /gene="gpdA" /locus_tag="An16g01830" /number=2
exon	10314..10327 /gene="gpdA" /locus_tag="An16g01830" /number=3
intron	10328..10428 /gene="gpdA" /locus_tag="An16g01830" /number=3
exon	10429..10492 /gene="gpdA" /locus_tag="An16g01830" /number=4
intron	10493..10567 /gene="gpdA" /locus_tag="An16g01830" /number=4
exon	10568..11252 /gene="gpdA" /locus_tag="An16g01830" /number=5
intron	11253..11308 /gene="gpdA" /locus_tag="An16g01830"

```

        /number=5
    exon      11309..11479
              /gene="gpdA"
              /locus_tag="An16g01830"
        /number=6
    intron    11480..11538
              /gene="gpdA"
              /locus_tag="An16g01830"
        /number=6
    exon      11539..11567
              /gene="gpdA"
              /locus_tag="An16g01830"
        /number=7
    gene      <12671..>15750
              /locus_tag="An16g01840"
    mRNA      join(<12671..13310,13362..13650,13836..13906,14084..14244,
                  14294..14504,14548..14659,14703..14750,14801..14864,
                  14910..15002,15052..15137,15187..15266,15315..15421,
                  15478..>15750)
              /locus_tag="An16g01840"
    CDS       join(12671..13310,13362..13650,13836..13906,14084..14244,
                  14294..14504,14548..14659,14703..14750,14801..14864,
                  14910..15002,15052..15137,15187..15266,15315..15421,
                  15478..15750)
              /locus_tag="An16g01840"
              /inference="profile:COGS:COG0123"
              /inference="profile:PFAM:PF00850"
              /inference="similar to AA sequence:PIR:S62933"
              /note="unnamed protein product; Function: Hda1p of S.
                    cerevisiae is involved in telomeric silencing and
                    transcriptional regulation.;
                    Function: Hda1p of S. cerevisiae is subunit of the HDA
                    complex involved in histone deacetylation.;
                    Remark: the systematic gene name of S. cerevisiae HDA1 is
                    YNL021w.;
                    Title: strong similarity to histone deacetylase A Hda1 -
                    Saccharomyces cerevisiae;
                    nucleus;
                    See PMID 10856249;
                    See PMID 8663039;
                    See PMID 8962081"
              /codon_start=1
              /protein_id="CAK42801.1"
              /db_xref="GOA:A2R705"

```

/db_xref="InterPro:[IPR000286](#)"
/db_xref="InterPro:[IPR019154](#)"
/db_xref="InterPro:[IPR023801](#)"
/db_xref="UniProtKB/TrEMBL:[A2R705](#)"
/translation="MDEDEDTVMGEAGSHHPTSVSSHQWPSNGPSSQGAGQLNGSSDP
SIAPPLLGPTMKLPLSVHSGDPRPENTASAEDQGVSPPKSLPKSDSGDATTEDFDKQD
SLVEDNSDWTDEDTVARTGLPIASLPSGLCYDVQMRYHCEVRPTADVHPEDPRRIYYI
YKELCRAGLVDDPESCRPLVSRPLKRINARNATEAEISLIHTPDHFVAFVESTKDMSDD
ELIALEHTRDSIYFNKLTFASSILSVGGAIETCLAVATRKVKNAIAVIRPPGHAEHD
KTMGFCLFNNVSVAAARVCQNRLGDSCKILILDWMGNSTLEGKKVTGITVVKALVSEA
DVFISNQCEPVASGFDAAGDELGGCFVTPSCYAHMTHMLMTLANGKVAVCLEGGYNF
RSISKSALAVTKTLMGDPDPDLHTTSPSALATTTVRRVMMIQSHYWHCMYPKGPPQDG
LFTDRLHDVIRAYQSKQLYDNYKLTSLFIYRTAISRSFENQVLASSNYHQTVPLVVIF
HDPPEIMGLPHPVTNKLEAHNCWLADVMKDYIGWIVSKGYAVIDVNI PKHVTIDPSSG
KYEEEDENRPTATEELAGYLWDNYIEPSEATEIFFMGVGNAFYGVANLLINRETLYKR
VNSVVSFVAENPVRAIASHTQVWLSRWYKDNSLVFVSHTHGVWNNVETRKRPSKRYGQ
LVQSPKASLSEMLMHHKKEEVYQWISERADVQDSEETEDEKKPKSKRSRPTKENQTEPR
SVPG"

exon	12671..13310
	/locus_tag="An16g01840"
	/number=1
intron	13311..13361
	/locus_tag="An16g01840"
	/number=1
exon	13362..13650
	/locus_tag="An16g01840"
	/number=2
intron	13651..13835
	/locus_tag="An16g01840"
	/number=2
exon	13836..13906
	/locus_tag="An16g01840"
	/number=3
intron	13907..14083
	/locus_tag="An16g01840"
	/number=3
exon	14084..14244
	/locus_tag="An16g01840"
	/number=4
intron	14245..14293
	/locus_tag="An16g01840"
	/number=4
exon	14294..14504
	/locus_tag="An16g01840"

	/number=5
intron	14505..14547
	/locus_tag="An16g01840"
	/number=5
exon	14548..14659
	/locus_tag="An16g01840"
	/number=6
intron	14660..14702
	/locus_tag="An16g01840"
	/number=6
exon	14703..14750
	/locus_tag="An16g01840"
	/number=7
intron	14751..14800
	/locus_tag="An16g01840"
	/number=7
exon	14801..14864
	/locus_tag="An16g01840"
	/number=8
intron	14865..14909
	/locus_tag="An16g01840"
	/number=8
exon	14910..15002
	/locus_tag="An16g01840"
	/number=9
intron	15003..15051
	/locus_tag="An16g01840"
	/number=9
exon	15052..15137
	/locus_tag="An16g01840"
	/number=10
intron	15138..15186
	/locus_tag="An16g01840"
	/number=10
exon	15187..15266
	/locus_tag="An16g01840"
	/number=11
intron	15267..15314
	/locus_tag="An16g01840"
	/number=11
exon	15315..15421
	/locus_tag="An16g01840"
	/number=12
intron	15422..15477

```

        /locus_tag="An16g01840"
        /number=12
    exon        15478..15750
        /locus_tag="An16g01840"
        /number=13
    gene        <16979..>17458
        /locus_tag="An16g01850"
    mRNA        <16979..>17458
        /locus_tag="An16g01850"
    CDS        16979..17458
        /locus_tag="An16g01850"
        /inference="similar to AA sequence:UniProtKB:AF277088.1"
        /note="unnamed protein product; Function: bys1 in B.
        dermatitidis is expressed at very high levels in the yeast
        phase but whose levels diminish rapidly when yeast cells
        are transferred to 25 degrees C to promote conversion to
        the mycelial phase.;"
        Remark: synonymous name of Ajellomyces dermatitidis is
        Blastomyces dermatitidis.;
        Title: similarity to blastomyces yeast phase-specific
        protein 1 bys1 - Ajellomyces dermatitidis;
        See PMID 8188377"
        /codon_start=1
        /protein_id="CAK42802.1"
        /db_xref="GOA:A2R706"
        /db_xref="InterPro:IPR001938"
        /db_xref="InterPro:IPR006771"
        /db_xref="UniProtKB/TrEMBL:A2R706"
        /translation="MHLSTKSLLSLLPLPLASAVGSAIVQNNCTSPIYLWSVGGSVN
        SMQTIDPGSSYSETFYDDTSGGVALKITTTENGLYNGSPQTDYAYTLDTSSGNVFYD
        ISDVFSGDPFSGSVSLVSSDSSCESICWAGGVPPAGSVVRSCQDEADEVLTVCADSC"
    sig_peptide    16979..17047
        /locus_tag="An16g01850"
        /inference="protein motif:SignalP:2.0"
    mat_peptide    17048..17455
        /locus_tag="An16g01850"
        /product="unnamed"
    exon        16979..17458
        /locus_tag="An16g01850"
        /number=1
    gene        <18681..>22324
        /locus_tag="An16g01860"
    mRNA        join(<18681..19163,19241..19340,19404..>22324)
        /locus_tag="An16g01860"

```

CDS

```
join(18681..19163,19241..19340,19404..22324)
/locus_tag="An16g01860"
/EC_number="3.4.-.-"
/inference="profile:COGS:COG1025"
/inference="profile:PFAM:PF00675"
/inference="profile:PFAM:PF05193"
/inference="similar to AA sequence:PIR:S55945"
/note="unnamed protein product; EC:3.4.99.-;
Similarity: shows also strong similarity to human and rat
insulin degrading enzyme, and to several other zinc
proteases from different sources.;
Title: strong similarity to protease involved in a-factor
processing Ste23 - Saccharomyces cerevisiae"
/codon_start=1
/protein_id="CAK42803.1"
/db_xref="GOA:A2R707"
/db_xref="InterPro:IPR007863"
/db_xref="InterPro:IPR011237"
/db_xref="InterPro:IPR011249"
/db_xref="InterPro:IPR011765"
/db_xref="UniProtKB/TrEMBL:A2R707"
/translation="MPPRSPATRLWNSQVPLYLRSPAFAFRFSKLGASAAIAPHFIYP
SPRFPTPNERCSSFPALRSSHALVSLSLPSRARFFSNTATMGSIERIAEQLEKPELDD
RSYRVIRLPNKLEALLVHDPDPTDKASAAVNVNVGNFSDADDMPGMAHAVEHLLFMGTQ
KYPKENAYNQYLASHSGSSNAYTAATETNYFFEIDTPAKPNYPTSPLYGALDRFAQFF
VEPLFLESTLDRELRAVDSENKKNLQSDLWRLMQLNKTLSPAHPHYHHFSTGNLQTLK
EEPQKRGLQEVQRQEFIKFYQAHYSSNIMKLVVLGRDPLDEMEQWVGDLFKHVKNQDLQP
NRWDHAQPCLPEHLGKQIFAKPVMMDMRSLDIYFFPMDEESMFESQPSRYLSHLIGHEG
PGSILAYIKAKGWANGLSAGVMPVCPGSAFFTTISVRLTPEGLKQYREVTKVVFYIGM
IKEREPQQWIFDEMKNLAEEVEFRFKQKSPASRFTSRLSSVMQKPYPREWLLSGSLLRK
FEPELVKKALSYLRPDNFRMVIVAQDYPGDWNCREKWYGYTEYKVEDIPQDFMDSIRKA
VETSPESRSLSELHMPHKNEFVPTRLSVEKKEVSEPAKTPKLIRHDDHVRWLWYKKDDRF
WVPKATVHITLRNPLVWATPANLVKAKFYSELVRDALNEYSYDAELAGLDYSLSASLF
GLDISVGGYNDKMSVLEKVLTSMRDLVVKPDRFNI IKERMTRNYKNAEYQQPFYQVG
DYTRYLTAERTWLNEQYAAELEHIEAEDVSCFFPQILRQNHIEVLAHGNLYKEDALRM
TDLVESTLQSRTLPEQWYVRRNMILPPGANYIYERTLKDPANVNHCI EYYLFIGKID
DDALRAKLLLFAQMTDEPAFDQLRSKEQLGYVWVSGARYSATTMGYRVIIQSERNAPY
LESRIDSF LTGFGKSLQNMSEQDFENHKRSVVNKRLEKLNLSSETNRFWSHIGSEYF
DFVQNESDAANVRKLTQADLIQFFQQFVDPTSATRAKLA IHLKAQAGAHAAHATKPEEQ
KAQLVSLFGKQLEAAGFAADTERLSAAFENVVDVSSGDEAQILTTLKAFLESGLSLSEE
QTKPVIGQVESNLGLQLKQVGIEPKKSDGTAVQANGTNGDIQKSTYINNVPFEFKARLA
VSAGPSPVTDLSEFEFEDAKL"
```

exon

```
18681..19163
/locus_tag="An16g01860"
```

```

        /number=1
intron      19164..19240
            /locus_tag="An16g01860"
            /number=1
exon        19241..19340
            /locus_tag="An16g01860"
            /number=2
intron      19341..19403
            /locus_tag="An16g01860"
            /number=2
exon        19404..22324
            /locus_tag="An16g01860"
            /number=3
gene        <22761..>23796
            /locus_tag="An16g01870"
mRNA        join(<22761..22801,22875..23126,23182..23429,
23504..>23796)
            /locus_tag="An16g01870"
CDS         join(22761..22801,22875..23126,23182..23429,23504..23796)
            /locus_tag="An16g01870"
            /note="unnamed protein product; Complex: in human, three
major polypeptides of 25,59, and 68 kD and a less abundant
polypeptide of 72 kD have been shown to copurify with CF
Im (CFIM) activity.;
Function: cleavage factors (CFs) Im and IIm, both of which
are required in order for cleavage and polyadenylation to
occur, are involved in the first step of 3'- processing.;
Function: modifications include the generation of a new
3'-end by endonucleolytic cleavage followed by
polyadenylation of the upstream cleavage product.;
Function: most eukaryotic mRNA primary transcripts are
processed both internally and at their termini in the
nucleus, before they are exported to the cytoplasm.;
Title: strong similarity to mRNA cleavage factor I 25 kDa
subunit CFIM25 - Homo sapiens;
See PMID 9659921"
            /codon_start=1
            /protein_id="CAK42804.1"
            /db_xref="GOA:A2R708"
            /db_xref="InterPro:IPR016706"
            /db_xref="UniProtKB/TrEMBL:A2R708"
            /translation="MSTTTATAIQSSRPPIIPKDFSAQQPKTIRLYPLSNYTFGKET
QPEEDPSVLARLKRLEEHYDLHGMRRRTCEGVLVCHEHNHPHVLMLQIANAFFKLPGDY
LHFEDDEVEGFKKRLNERLAPVGSQFSGEGVNEDWEIGDTLAQWWRPNFETFMYPFLP

```

```

GHVTRPKECKKLYFIQLPKKKVLSVPKNMKLLAVPLFELYDNTARYGPQLSAIPHLLS
RYNFEFVDENDNVVAATPGTPLPDHQIPRTKVLAGDSGDGQDEGMTDYPSPGEGENGAS
A"
exon      22761..22801
          /locus_tag="An16g01870"
          /number=1
intron    22802..22874
          /locus_tag="An16g01870"
          /number=1
exon      22875..23126
          /locus_tag="An16g01870"
          /number=2
intron    23127..23181
          /locus_tag="An16g01870"
          /number=2
exon      23182..23429
          /locus_tag="An16g01870"
          /number=3
intron    23430..23503
          /locus_tag="An16g01870"
          /number=3
exon      23504..23796
          /locus_tag="An16g01870"
          /number=4
gene      <25910..>26953
          /locus_tag="An16g01880"
mRNA      join(<25910..25991,26046..26211,26257..26593,
26645..>26953)
          /locus_tag="An16g01880"
CDS       join(25910..25991,26046..26211,26257..26593,26645..26953)
          /locus_tag="An16g01880"
          /EC_number="3.1.1.3"
          /inference="profile:COGS:COG3675"
          /inference="profile:PFAM:PF01764"
          /note="unnamed protein product; Similarity: sequence is
almost identical to sequence 1 from Patent WO9831790,
where the protein is described having a phospholipase
activity.;"
          Title: strong similarity to lysophospholipase -Aspergillus
foetidus"
          /codon_start=1
          /protein_id="CAK42805.1"
          /db_xref="GOA:A2R709"
          /db_xref="InterPro:IPR002921"

```

```

/db_xref="InterPro:IPR005592"
/db_xref="InterPro:IPR029058"
/db_xref="UniProtKB/TrEMBL:A2R709"
/translation="MFSGRFGVLLTALAALSAAAPTPLDVRSVSTSTLDELQLFSQWS
AAAYCSNNIDSDSNVTCTADACPSVEEASTKMLLEFDLTNNFGGTAGFLAADNTNKR
LVVAFRGSSTIKNWIADLDFILQDNDDLCTGCKVHTGFWKAWEEAADNLTSIKSAMS
TYSGYTLYFTGHSLGGALATLGATVLRNDGYSVELYTYGCPRVGNALAEHITSQGSG
ANFRVTHLNDIVPRLPPMDFGFSQPSPEYWITSGTGASVTASDIELIEGINSTAGNAG
EATVDVLAHLWYFFAISECLL"

sig_peptide    25910..25966
/locus_tag="An16g01880"
/inference="protein motif:SignalP:2.0"

mat_peptide    join(25967..25991,26046..26211,26257..26593,26645..26950)
/locus_tag="An16g01880"
/product="unnamed"

exon           25910..25991
/locus_tag="An16g01880"
/number=1

intron         25992..26045
/locus_tag="An16g01880"
/number=1

exon           26046..26211
/locus_tag="An16g01880"
/number=2

intron         26212..26256
/locus_tag="An16g01880"
/number=2

exon           26257..26593
/locus_tag="An16g01880"
/number=3

intron         26594..26644
/locus_tag="An16g01880"
/number=3

exon           26645..26953
/locus_tag="An16g01880"
/number=4

gene           <27489..>28640
/locus_tag="An16g01890"

mRNA           join(<27489..27880,27936..28378,28432..>28640)
/locus_tag="An16g01890"

CDS            join(27489..27880,27936..28378,28432..28640)
/locus_tag="An16g01890"
/note="unnamed protein product; Title: strong similarity
to hypothetical protein encoded by An14g02280 -

```

```

Aspergillus niger"
/codon_start=1
/protein_id="CAK42806.1"
/db_xref="GOA:A2R710"
/db_xref="InterPro:IPR021838"
/db_xref="UniProtKB/TrEMBL:A2R710"
/translation="MLFFRGGVNSPRRRNQEKAI FCTILIIYALYFLFFAGSSDPKRS
ETSAEDAVKGGPHRDSRSGKASPSASASRQRQALDKDMVVASMGDDVSWLYTYFPEW
HKSIYVADDKKAPLTVALNKGRESMVLYTYIIDNYDNLPSVLFHISQRYQWHNDPPY
YDGVFVLRNFQVPYLQKQGYVNLRCAWVLGCPAEIHPLSDTHRNDVHAGEYFKTGME
LFPGAEVPEEVGVSCCAQFAVTRWKIRERP KSDYVRYRKWLTETTL PDDL SGRIMEYS
WHMIFGKGPVYCPTAE ECYCKVFGLCNLDCPTEGDCAGRYVLPPYSSLPKGWPYVGWK
GQKQDPSKGLPEA"
exon      27489..27880
/locus_tag="An16g01890"
/number=1
intron    27881..27935
/locus_tag="An16g01890"
/number=1
exon      27936..28378
/locus_tag="An16g01890"
/number=2
intron    28379..28431
/locus_tag="An16g01890"
/number=2
exon      28432..28640
/locus_tag="An16g01890"
/number=3
gene      complement(<29882..>30545)
/locus_tag="An16g01900"
mRNA      complement(join(<29882..30374,30430..>30545))
/locus_tag="An16g01900"
CDS       complement(join(29882..30374,30430..30545))
/locus_tag="An16g01900"
/note="unnamed protein product; Title: similarity to
hypothetical protein encoded by An04g06980 - Aspergillus
niger"
/codon_start=1
/protein_id="CAK42807.1"
/db_xref="UniProtKB/TrEMBL:A2R711"
/translation="MDSIELMLFLFSVLLVKSLQMPTRKLVVLSSFAVRLPLLVFT
SLHLYHLQKYAYSMNPTLDIIDGYVWTQMGMNYALIVCTSFCLAPFMRAISV TYGNAG
EVTLTGTSSARSKSGKYARDVSKQSQSYVLQSM DNQEEAGHGMNNPRHSQPTFVPERYG
GMLSTANDGSSRTDGD SVASSESTKMI IKKDVEYTVTHEPR"

```

```

sig_peptide    complement (30483..30545)
                /locus_tag="An16g01900"
                /inference="protein motif:SignalP:2.0"
mat_peptide    complement (join (29885..30374,30430..30482))
                /locus_tag="An16g01900"
                /product="unnamed"
exon           complement (29882..30374)
                /locus_tag="An16g01900"
                /number=1
intron         complement (30375..30429)
                /locus_tag="An16g01900"
                /number=1
exon           complement (30430..30545)
                /locus_tag="An16g01900"
                /number=2
gene           <31261..>31684
                /locus_tag="An16g01910"
mRNA           join (<31261..31296,31409..>31684)
                /locus_tag="An16g01910"
CDS            join (31261..31296,31409..31684)
                /locus_tag="An16g01910"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="CAK42808.1"
                /db_xref="GOA:A2R712"
                /db_xref="UniProtKB/TrEMBL:A2R712"
                /translation="MTRATQRRQQATVLGAFRSERNKTISSGSRPWWQGQLGYQGSGP
SLRGLVKLRISIIGGSNSTKGSARAYGRIFGPDSESGQNVLRGTKGALFITADVRNLKP
L"
exon           31261..31296
                /locus_tag="An16g01910"
                /number=1
intron         31297..31408
                /locus_tag="An16g01910"
                /number=1
exon           31409..31684
                /locus_tag="An16g01910"
                /number=2
gene           complement (<32581..>34386)
                /locus_tag="An16g01920"
mRNA           complement (join (<32581..33015,33079..33404,33456..33872,
33915..33997,34050..34163,34214..34284,34336..>34386))
                /locus_tag="An16g01920"
CDS            complement (join (32581..33015,33079..33404,33456..33872,

```

```

33915..33997,34050..34163,34214..34284,34336..34386))
/locus_tag="An16g01920"
/inference="profile:COGS:COG0833"
/inference="similar to AA sequence:UniProtKB:ENI131668.1"
/note="unnamed protein product; Function: gabA of A.
nidulans is subject to carbon catabolite and nitrogen
metabolite repression, omega-amino acid induction and
regulation in response to ambient pH (being
acid-expressed).;
Similarity: shows also similarity to several putative
fungal amino acid permeases.;
Title: strong similarity to GABA permease gabA
-Aspergillus nidulans;
See PMID 10320578"
/codon_start=1
/protein_id="CAK42809.1"
/db_xref="GOA:A2R713"
/db_xref="InterPro:IPR002293"
/db_xref="UniProtKB/TrEMBL:A2R713"
/translation="MADQRDMQALGKEQVLRNFRFISIVAFGCTLIASWEVVLTMMS
SPLVDGGTAGFIWGFIVSVGVLLLLFASLAEMASMAPTAGGQYHWVSEFAPRKYQKFL
SYITARSHLGWLCATGWECAIVSIAFLAGTIIQGLVILNYPDYVSEWHGTLITIAIT
LFSVLENTFLAKKLPLVEAFILILHVAGFFIVIIPLWVLAPIKDAKTVFTQFTNGGDW
NSNGTATMVGLITAITAMIGYDCTVHMSEEVVDASTTLPKAIMAACSLNGVLAFFVMV
TLCFTLGNVDEVLD SATGYPFIQVFFNATKSYAATNAMTAIVTMSLIASVITEVATAS
RQLWSFARDGGVPLATFFGHVNPWNIPNLA VLVSLVLTMLLSLINIGSTVALDALIT
LTVGALLNSYIVTVSCVSWKRIRGEPLPPHRWSLGRWGLWVNLGALAFIPEFVFVMF
PLYSTVDPENMNWSSLMYGGMLIFSIIYYVLYGRKTYVPPVALVKREI"
exon      complement (32581..33015)
/locus_tag="An16g01920"
/number=1
intron    complement (33016..33078)
/locus_tag="An16g01920"
/number=1
exon      complement (33079..33404)
/locus_tag="An16g01920"
/number=2
intron    complement (33405..33455)
/locus_tag="An16g01920"
/number=2
exon      complement (33456..33872)
/locus_tag="An16g01920"
/number=3
intron    complement (33873..33914)

```

```

        /locus_tag="An16g01920"
        /number=3
    exon        complement(33915..33997)
        /locus_tag="An16g01920"
        /number=4
    intron       complement(33998..34049)
        /locus_tag="An16g01920"
        /number=4
    exon        complement(34050..34163)
        /locus_tag="An16g01920"
        /number=5
    intron       complement(34164..34213)
        /locus_tag="An16g01920"
        /number=5
    exon        complement(34214..34284)
        /locus_tag="An16g01920"
        /number=6
    intron       complement(34285..34335)
        /locus_tag="An16g01920"
        /number=6
    exon        complement(34336..34386)
        /locus_tag="An16g01920"
        /number=7
    gene        <35568..>36098
        /locus_tag="An16g01930"
    mRNA        join(<35568..35662,35783..35957,36003..>36098)
        /locus_tag="An16g01930"
    CDS         join(35568..35662,35783..35957,36003..36098)
        /locus_tag="An16g01930"
        /codon_start=1
        /product="hypothetical protein"
        /protein_id="CAK42810.1"
        /db_xref="GOA:A2R714"
        /db_xref="UniProtKB/TrEMBL:A2R714"
        /translation="MAETYPIGTRKECERLVKDWGFKHVFTWADGNSNAHYPPHSHGGV
TTHLIRQGTLAITYPDDNAKLHNGEVKKETFGVGARLDVPAGKLHEVGADVGYGLSLT
DYDHLSDVLYYEFALLSNL"
    exon        35568..35662
        /locus_tag="An16g01930"
        /number=1
    intron       35663..35782
        /locus_tag="An16g01930"
        /number=1
    exon        35783..35957

```

```

/locus_tag="An16g01930"
/number=2
intron    35958..36002
/locus_tag="An16g01930"
/number=2
exon      36003..36098
/locus_tag="An16g01930"
/number=3
gene      complement(<36203..>38137)
/locus_tag="An16g01940"
mRNA      complement(join(<36203..36227,36284..36367,36423..36541,
36595..37438,37490..37818,37871..>38137))
/locus_tag="An16g01940"
CDS       complement(join(36203..36227,36284..36367,36423..36541,
36595..37438,37490..37818,37871..38137))
/locus_tag="An16g01940"
/inference="profile:COGS:COG0477"
/note="unnamed protein product; Similarity: shows
similarity to several probable transporter proteins of
different species.;
Title: similarity to allantoin permease Dal5
-Saccharomyces cerevisiae;
See PMID 10869563"
/codon_start=1
/protein_id="CAK42811.1"
/db_xref="GOA:A2R715"
/db_xref="InterPro:IPR011701"
/db_xref="InterPro:IPR020846"
/db_xref="UniProtKB/TrEMBL:A2R715"
/translation="MTTDGVKEVPADTVSANEDTKKNPDVSTSSASSSSSLTDSEHRVA
KVTDCAFDTTEDPRFYKPIPEYESAHRWDYPFEWTEEEEKQLIRKIDIRVCTFACVTF
FALQLDRGNIVQANSDSMLGLGMTTNDYNTGQTIFYLTFLFSELPSSLKKGVPDR
WIPIQMVAWSFVAAFQAFLKNRAGYFACRALLGLLEGGFADTILFLSFFYKSKELPQ
RLSYFWVSYEATSIVGAFLAFGLHIHGSNGYTGGWRYLFAFEGLITGIIGIIAAFWM
PASPTQTKGGFRGKDGWFNEREEKILVNRVLRDDPSKGSNMNRQAVTPKMLWHALCDY
DMWPIYLLGLTWMIPNTPATNYITLKLKSLGFGTFETNLLTIPAYVIFINLLFWTWI
SERFNQRLLLGVVSEFWCLVLLIALEVLDPGASGWARWIINALLIGAPYVHAIIVAMT
SRNAGSVRTRTVASAVYNNMMVQASSIIASQIYQEKDKPYRVGNKVLIALSVFSMALF
IGAKFYIWRNKYNAARWNAMSSEEREYLAANKDLGNKRLDFRFTH"
exon      complement(36203..36227)
/locus_tag="An16g01940"
/number=1
intron    complement(36228..36283)
/locus_tag="An16g01940"

```

```

        /number=1
    exon      complement (36284..36367)
              /locus_tag="An16g01940"
        /number=2
    intron    complement (36368..36422)
              /locus_tag="An16g01940"
        /number=2
    exon      complement (36423..36541)
              /locus_tag="An16g01940"
        /number=3
    intron    complement (36542..36594)
              /locus_tag="An16g01940"
        /number=3
    exon      complement (36595..37438)
              /locus_tag="An16g01940"
        /number=4
    intron    complement (37439..37489)
              /locus_tag="An16g01940"
        /number=4
    exon      complement (37490..37818)
              /locus_tag="An16g01940"
        /number=5
    intron    complement (37819..37870)
              /locus_tag="An16g01940"
        /number=5
    exon      complement (37871..38137)
              /locus_tag="An16g01940"
        /number=6
    gene      <39723..>41143
              /locus_tag="An16g01950"
    mRNA      join (<39723..39928,39979..40184,40240..40694,40826..40975,
41030..>41143)
              /locus_tag="An16g01950"
    CDS       join (39723..39928,39979..40184,40240..40694,40826..40975,
41030..41143)
              /locus_tag="An16g01950"
              /EC_number="1.-.-.-"
              /inference="profile:COGS:COG2175"
              /inference="profile:PFAM:PF02668"
              /inference="similar to AA sequence:PIR:S50963"
              /note="unnamed protein product; Function: disruption of
the YLL057c of S. cerevisiae encoding this enzyme
negatively affects the use of isethionate and taurine as
sulfur sources by S. cerevisiae,providing strong evidence

```

that YLL057c plays a role in sulfonate catabolism.;
Function: purified recombinant protein of YLL057c of *S. cerevisiae* is shown to be an Fe(II)-dependent sulfonate/alpha-ketoglutarate dioxygenase; although taurine is a poor substrate, a variety of other sulfonates are utilized, with the best natural substrates being isethionate and taurocholate.;
Title: strong similarity to Fe(II)-dependent sulfonate/alpha-ketoglutarate dioxygenase YLL057c -*Saccharomyces cerevisiae*;
See PMID 10482536"

/codon_start=1
/protein_id="[CAK42812.1](#)"
/db_xref="GOA:[A2R716](#)"
/db_xref="InterPro:[IPR003819](#)"
/db_xref="UniProtKB/TrEMBL:[A2R716](#)"
/translation="MAPSLVEDTPLSLRAAPKSTLKT DAGFNKENVIGYGEAYKHENE
LKGTEKQPPASFNPYLPVWDNETERYPPLQPF EHYDHGKDADPTFPDLLPKDKVEVDD
ITPTIGTEVRGVQLSQLSKEGKDQLALYVAQRKVVAFRDQDFAHLPIEKALEFGGYFG
RHHIHQTSGAPKGFPEIHLVHRGADDRSGAEFLSQRNTITWHS DVTFEKQPPGTTFL
YILDGPSSGGDTL FADMAQAYRRLSPEFRKRLHGLKAVHSGIEQVNNSLNKGGIARRD
PITSEHPIV RTHPTLTETVTRYIVGYKKEESDYLLKFLYDHIALS QDLQTRVRWRPGT
VVVWDNRVVTHSALFDWEDGQRRHHARITPQAEAPYETPYEG"

exon	39723..39928 /locus_tag="An16g01950" /number=1
intron	39929..39978 /locus_tag="An16g01950" /number=1
exon	39979..40184 /locus_tag="An16g01950" /number=2
intron	40185..40239 /locus_tag="An16g01950" /number=2
exon	40240..40694 /locus_tag="An16g01950" /number=3
intron	40695..40825 /locus_tag="An16g01950" /number=3
exon	40826..40975 /locus_tag="An16g01950" /number=4

```

intron      40976..41029
            /locus_tag="An16g01950"
            /number=4
exon        41030..41143
            /locus_tag="An16g01950"
            /number=5
gene        complement(<41403..>42817)
            /locus_tag="An16g01960"
mRNA        complement(join(<41403..41807,42311..42472,42566..>42817))
            /locus_tag="An16g01960"
CDS         complement(join(41403..41807,42311..42472,42566..>42817))
            /locus_tag="An16g01960"
            /codon_start=1
            /product="hypothetical protein [truncated ORF]"
            /protein_id="CAK42813.1"
            /db_xref="GOA:A2R717"
            /db_xref="InterPro:IPR019193"
            /db_xref="UniProtKB/TrEMBL:A2R717"
            /translation="PTILTCTTCTTQIGLPDIANGYRLFKTSLSANTTSSSEDNNPP
EETWESHPPETILTAQLLELIERESYRRFVIHCGAKDGLLIWPFNPSLRYSTTSASHS
TAAQLAMKVFYQHVDNVEEILHPEVGKGNQSLAIEELLIHPPPSSTMPPTTSQSQSHH
PPRSFSTSSDENNTVTWEHLEHVPPPSHSTNPTTTTPATTHNKNLSTADNHPERILD
PETHPESEAFHMDDVMGAGEGTGTGTGTCMAPAIRIHPPGDAVSASTKSAAFG"
exon        complement(41403..41807)
            /locus_tag="An16g01960"
            /number=1
intron      complement(41808..42310)
            /locus_tag="An16g01960"
            /number=1
exon        complement(42311..42472)
            /locus_tag="An16g01960"
            /number=2
intron      complement(42473..42565)
            /locus_tag="An16g01960"
            /number=2
exon        complement(42566..>42817)
            /locus_tag="An16g01960"
            /number=3

```

ORIGIN

```

1 catccatcct tttttccccc ccattcctccc catcttcccc tactacccat tcagcatccg
61 attcgggtag ggtggccctt atttaacatt cctggctatt gtccccctac cacctccaac
121 atccctctga cagtttcatc tattgttgga gaattactga gggagtcaga tcgggtgttc
181 ctgacttttg gggcaagcct cactctctcg ttgctttctc tccatcacat ttccgcacg
241 atctatccag tgcagtgcag cagagtgcgt tcagcccatc cccatcgatc ttttttattt

```

301 tttttccctc tctcttctcc cttcctttgg ggtgtggatt tgcggaagag cagggaaaaa
361 aagaccagac cagacccgac ccgactaagc ggtcgcggtg ttgccggctt ttctctctgc
421 gtgccgtttt ggcgggcggtg cctttccctt tccctttctt ccttttcccc ccagcagca
481 acaagcagca ataagcagca gccgcagcag cagcaccacc acccaaagag tgagattcct
541 ttggtgttcc tctcctgga ccacctctg ctacttgctc gttcctcggt ttccctcctt
601 tccccctcac tctcccatc ctcttctctc cctccctctc tctctgttcg gccgtcaaaa
661 agccttgggc actgcttata catcccaggt agtggcagct cgtttccttt tccctttttg
721 cccttttccc cttatcttcc tctaccctg ttaaagctcc gtttattttc ctttgctgcc
781 ttctcttggt cttcaatttt tcccgacccc cttttctttt ttcccagatt ctccctcccc
841 tgccctcttc tgatgcccc aggggggttg agaagcgaaa gacaaacata acaaaaaaaaa
901 gaaaaaaaac aaaaaaaaaa ttcagttcct aacattttcg tgctaaagga tacgacggcc
961 attgatgacc caatagtcac gagagaatta tttcagcagc tgtgagcgat tttgcttggt
1021 tttcttcggt ttttgctgct tccaccgcgc ttttcctctc cgcaattacc aacgatcgg
1081 cactccctcc caccctcgg aaagttgacc tgttttgcta cgatccgctc gaccttgcaa
1141 ataccctcga acaagcgtgg tttgagttcc tcttctccgc tagtatcaag cgctgccccg
1201 gtgctccta tgtttgagta gttccgctg acacaacagt tgcatctca ttgaacgaac
1261 gggttttggt gtgtggtgga tctcagtgcc cacattcctc ggtttggtg cttctagga
1321 ctctgcaatc ttttcgccgc catgagctat caacaagccc aggacatgta ccacgataac
1381 tccgcggggc gttctcctgg ctcccaacgc catcagcagc cgctgcaccg tcaaccttct
1441 cgctcagttt acgcgtacgg ccctatgcct gtgaacctgt atgacgacat ggctcggtac
1501 gacaccggtc ggctggagcg cctgaattcg tcgctgcaca acaactctta cgctacgat
1561 cttgcgggct ccagacctg gaatcccaac ggctttgcca atgcgcaagc tcttgggggc
1621 attcggtcgg cttcgaccag cctcaagact acttcgcgca ctggcagggc gggcttgccg
1681 acggtacgta ttcaattgag tgggtggaat aaaattgatg gaacctatcg ttgctgactg
1741 acgagtagac atggcttgat cagcagcctg gcattcccag cgccttctcc aacctcgcc
1801 ccggtccctc ccaaagcagc gcgatgcgcc cggaagcctc tggctcgtcg gaagcggagg
1861 atgagttgat tccaccgcc attgtcatca agaacttcc ttttgcggtg aagaaggagc
1921 agctggtgca gcttatgact gagctcaact tgccgttgcc ttacgcctc aactatcact
1981 tcgacaacgg cgtcttccgg ggattggctt tcgccaaact cactctgctg gaggagactg
2041 caaccgtcat cgaggtcctc aatcatttcg agctccaggg tcgcaaactg cgggttgaat
2101 acaagaagat gctgccgctg caggagcggg agcggattga acgtgagaag cgcgagcgtc
2161 gtggtcagct ggaagagcag catcgcccca tggcgacctc gcaacttcag actcagagtt
2221 ccatgtcgtc tctcacctcg caccttcccg caacctcccc gtctccgctc tccagcgtg
2281 gccaaaagtt gggtaggtgt cttgacttcc ttgtgctctt ttatggcaag tcgctaacc
2341 atgccgcgta tctacagagg tcatctcaa cgatagcaca acgcttctgt actactctca
2401 gcttctctc ttcaaggagg atacggcgcg tgacacctc gtgttccct ccaacctttc
2461 tctgtccag cgcgcaccg ttcatacact tgctcacaac atgggcctag gccatgcgtc
2521 tcggggatcg ggtgatcagc ggcaggtgca tgttttcaag gtcgtcccg gcaccaatgt
2581 tagccctccc ctgtcgtcca tccctgccgc tgtccaacca gcggagacag ctgctcgtgg
2641 actgaaccgc gcgcgaacca ttgatttcag tgagtctcgt aacgacggac ccaactttt
2701 taacactttg cgcggccagc cttcttcggg cttcctggga gtcttggtg ccccgggcaa
2761 ctttggaac acgcagaacc ttcagccgc caagtcgttt gcagacctc gtcctacac
2821 tccatcccc gtccctctt cggccagctt ccctgcagcc ctgcagtcta acggagcgcg
2881 cctgcaacac tacgatggtg ccacgagcgc cacttccaac acccgcgcg ttacgtccgc

2941 gccatccggt tcctccctcg gtatgcagcg cgacgacagt ctgttggtga acagcctcag
3001 tagcctctct ctgggactg gaatcgggtg gccgaatgcg agcccagga gattgcgggg
3061 tatgttctct tgggagcagc cagaaagtca gccctctagc gctgggtccca tcggtagtaa
3121 ccgatccatc ggagttggat ttgacggtca gtcacaggag cgcattgccca ttaggcagcc
3181 gagaggcccg atgcccgaga aagggtccagg atttcgacgc cagaacggac accaatcgcg
3241 cggaagcgac gagttacgca ccagctctgg tgtagagatt attgtggagt aaacactttg
3301 atgggaagac gaagacgacg aaacttggtta gattgctttg ccgcaagatg catacgactc
3361 ttatatcacg aggcctgatc ctgttctgcg attattggct tacgattcga tgagggatga
3421 cgagatgaca ctcaaccgag agaaccaaaa gaatatttct ctgatctagc caagttaga
3481 cgttattttg gggcagccca gcaggcacgc tgtttgaacg caaccttttt atgcttttct
3541 ccttgtttcg cggaaaataa acacctcccc aaaaagaaag ctgtggactt tttgagacgc
3601 gggcgccttg catgcaggaa catttttttt ttttttacct atattttttc ttttactttt
3661 actttacctt ctacaaattg tgtcggcgga aagcaatata aaaccaatca tcaacacctg
3721 gatggctacg aaggacgaca acaaacggg atccggatgt ctttgtgtgc ccatggcaag
3781 agaaatgact caccagcatt tgcatttaac ctttacttat ctcttctat ctacttgttt
3841 gtcactttgt atatggggca atttgcatca tttaatgttt ttatttcttt gtcagacgtg
3901 tccaatgtgc tttgttctct catctctct aatcgtgttt ttttatggaa cccctgtgat
3961 agtggggagt ggaggaaatg aacggaacgt tgtcggcata ccacacacac gtatgcaatg
4021 tggctgttgc cgaggacgag ccattcattc tcttgcgatt tttatttccc tgacctgcag
4081 gtgtaattca aaaacttgat gacttgcagc gcgtgcgtta ggggagacac ggggtggcgaa
4141 tcgaaatggg cggatatttg cattctttgc agcttcttga aggggagagt ttggcttgtg
4201 ggtatttctt agggatgaga atgtccaggc aactatctag gtggacaaat gacaacttct
4261 tgcaaaattg tccgactgga tgacgtctcg tgatttatcc cttttccgac ggcggtctgg
4321 ccggattata tttcccttga tttggcattg gaccattagg tgtgctcttc atcctcttcc
4381 cttatctgct gctctctctc tctctctccc ttgccatcta gatcgccata ctacttacta
4441 ctactagctg ttgttttctg tatttatatg ttggctctcc atccttctct ctacacact
4501 ctcttctcaga acccccacag aacttgattg atccgtttac taccttcagt ttgtttcatc
4561 atggctcgctt cgggacttag actgtagtat atagtactag gtatagagac gtcattcagc
4621 tgctcaattg catcattccg gtttaagtag tactactagt aagtaattat tagtgtttcg
4681 atatatgtat gtagatcgca ttgaaattac tactaaagca acgtgataga taatatgata
4741 tcaggcttcg aacatatacg tacgataata ctaatgttta tttccataca agcgggttca
4801 agtaatgtta ctgaaggtag taagatctac tcacaccata ctcatgtagt acgactccta
4861 taaaccccat ccactcacc attcacttaa tatatcacca gggcatggat gaacaggatca
4921 caatagaaag aagataagag gaaaggaggt gaagaaaaag gtagactcgt acagaccctg
4981 taatatataag agaagacaca agagtacaga aacaaaaggg gtaaaatgaa agaaaaaaaa
5041 atcatggaaa taataataac tttcccacta ccgaaacata tccccttttt cccattacag
5101 cgcccttgga aaatttgaca tgaaacacgg aatacacaaa agacatcaag ctgtcagggg
5161 agataaataa taatgataaa ataaaagatg gggggaagca aggaagaact tggaaaaaat
5221 catctcgagc ttgcgacgt caccagaata gaagggtggc tttttttgtt cgtcagaagg
5281 acatgacctt gcctccctgt tcgggtgtgg ttttgagag ggggtggagg agcatcatga
5341 tcacagtaaa gaacccttcc ctagcttttc ccttttaacg atcattcaat caccaaacgc
5401 cggagtctcg cgggtatcca ccatcattta agcctcagtc ttgttcttct tcttttgctc
5461 caccttcagc aagagcgcg gggtgatgcg cggcagaaca cctccgaagg caatagtggc
5521 gcggatcagg gtatccagct cctcatctcc acggatggcg agctgcagat gacgcggggg

5581 gatacgcttg actttgagat ccttggcagc gtttccggcc aattcgagaa cttctgcggt
5641 cagatattcg agaacagcgg tgacatacac ggcggtcgcg atcgggattt gtcagcatgg
5701 ttgaaagagg agacaaagga aggtgcgcct tctgtaggct taatggaaag ggctgtacct
5761 ttggcgccaa cgcgcattct gttctgggtg ttgttcttca aaaagcgctt aacacgaccg
5821 caagggaact gtgatcgaag gataatgtta gcgaatgaaa ctgatgataa tagataggca
5881 tgaatgtgtg aagtgaagct ttgaaggtaa atacgggagg gatcatgacc gccacccctc
5941 gcaaccctgt cgcctgaact gtgactagcc ggtgcggtat agctcttaaa ggggtattat
6001 tcaatataaa agagctatac gcgtgctttt ggcaacatag aggcgatcgc gaggggcaaa
6061 ggaggcaaaa agcaaaatcc caggccgtaa cggcttcagc ggcattgtcg cgccagaacc
6121 catgagatgg gtcaagcgtc atatgcctgc caggaggaga aaaaaagaac gactaatacg
6181 gacctgcaga ccagcctttt cgctgtggga cttctggctc ttccccgcag agtccttgga
6241 gccggccttt ccaccgatag actttccctt tccaccaggc attttgtaag atcgctcgaga
6301 gttggggccg caggggcaaa aatgtcgcgt cagccgagga ttagaagtgg tagagcgtga
6361 tgggtggatgt tggcgcgacg cgatgcagca caagcacgcg cgttgttgat ggatgaagga
6421 gtcagaaatg tcagagccat cctggtcgct aaagacgggc gttgaccgtg aagcggcgct
6481 agtccgcttt gggaaacctg ctcagctccc aaccacggag atgtcataac gggcgaggtc
6541 gttgttggtc gccgttgacg aggcagcgat ggacaagatt cgaatccttg tctcttgcca
6601 tgatccagaa aagcatttca gccaaagcac tggatgatccg gcttcacgag gcgcattggg
6661 ctggggctgg gcgtccacgc cattgctggg tagagtact cagtagtca atccccaat
6721 gcccgacaca ccacactggt aagcagattt aattgcttca ttatacttgc accagtaggc
6781 ctaaacatg gaaaggggac agagaggcct ttaaacgta acgtgggggg cttgcacacg
6841 ttcgtagcga tatcaaagag gttggcagga gagtgtgcag accatgatag ttgctcagcg
6901 gattatctaa gcaaccagggt gatctgtgcc tttggccttg tccgggggac tgtctgccta
6961 aggtgcact agcccttgac ctgcccattc aggttaataaa tctgcgtatc ccgtcttctc
7021 gctgctatct tccatcacgc cttacactga tacggctagg ccgaggtagg taaagaataa
7081 gcacgtgacc ctcactatct ggacagctgt ccgacttccc cggggctcca gttccgcagt
7141 tcgaggaagt ttctgtcctt gcgcacgcat tccccatcct cccagtttc acctatcgt
7201 gtgctgctcc tcactttcca aggtccctg gtctcagtag taccggcaag atgggtgtctc
7261 ctcaagtgtg agtcaatcct gcgttattcc tgccgctgac atgcacaacc gtgggtctgct
7321 gtctcgtgat gctactcgag gaataaaaac tgatttgtcg tctgtcttgc agtaccaacc
7381 tggccatcat cgtggtgatg atgcagctag cgaagaaggt tcccttcgag gaccccgatg
7441 tccttatgct tgtccgtggc atgtacatcc tgtccaacgt gatcattctg ggagtgtact
7501 tgtacactca gtggaagatc aacaaaaaga atggatgct tcgtgtcctt gctttcttta
7561 agctacggct ctctaataat ctgtgatcag atctgactac cctgaaatac gtggagcctt
7621 ctctatggg cagcggcgag gagcccaagc ccgtcaccac taccaacatg gaggtagaca
7681 agggacagct gcgccagctg atgaggagcc agctgatggg tgtgggtatg atgggtgtga
7741 tgcacctgta ctttaagtac accaaccctt tcctgatcca gtccatcatt cctctcaagg
7801 gtgccttcga gtggaacctg gtcaagatcc acgtcttttg caagcctgcg accggtgacc
7861 ttcagcgccc cttcaaggcc aacaacagct tcctgaacca gggccagatc aagagcgaca
7921 aggcctccgt tgagaacgct gagaagaact ggagggggcg tgtcaaggag gagtaagttg
7981 acagaacctg gggaaggagc gcgagggtta ttgctaggcc ttgcctccaa tgtatttata
8041 gatgcggttt caatggagtg agcgcgaatc cgctatggat tcgcgcggtt ctggcttcac
8101 actagcttct tggctgctgt gatgggagac tttatgacta tgcatgtcgt cgatatcgtt
8161 cgcaatgacc tcttactatc aatataatag aaccacgtca gtgggtgacct catgtcctat

8221 aatcacccat gcagcagcat gactcaggag gcgaatagat aattttgaaa tccctactga
8281 tacggcttcc caacgaggtg ggagcggaaa ggatgatgag tggccaagta cctgccgatg
8341 ctttgttgct tcacgacttg agtctcctga tataccaaca tcggtggccg gtgaagacaa
8401 tgaagacata tttctaagca atatgggctg tggccactcc gtgccacttg ctggaagtaa
8461 cctgttgcat ttaccccatg taaggctgga ggctggatgg gggcactttg cagcagtatg
8521 tagaaagtac tagaaccatc ttccggtctc cgagacactg gtcaatattg acacggcagc
8581 atgatcatga aacccgagtc agaccagagc cgttcggatt gtgctatacc ccaagtcacg
8641 ttgtccataa ttgaataaat atgagcagtc ctttgtggcg tggaaacata cttagcatgt
8701 agagacaaac cttggtgcgc ggcttcaggc gggcatagtt agtatgctac ggtaccaccg
8761 atcttattgt accagaaaaa gtcccagcca gtccaatccc cattctaagc cacatgcatc
8821 cgttcgcgat catctgacat atcagattcg tccatctggt gcagtatcta acagaggcca
8881 gagcatcacc aacatgggta cctcagcaa taatatgcat gcattgtgcc cccctatgg
8941 agccgtagct ttcaagcaat tagacacgcg cccggccgaa tgagatgaac cgttggagcc
9001 atcatccac tcattccgct ccagaaagga gagaaagaaa aaaaaaaat atgaccgagc
9061 gcgtgatgac cggtaggac tccggtgaat tgatttgggt gacgggagag acccaagagg
9121 ggccagaata ataagaatgg ggaaggcgaa ggtaccgcct ttggggcca gccacgcgac
9181 tccaacatgg aggggactg gactaacatt attccagcac cgggatcacg ggccgaaagc
9241 ggcaaggccg cgcactgccc ctcttttttg gtgaaagagc tggcagtaac ttaactgtac
9301 tttctggagt gaataatact actactatga aagaccgca tgggccgata gtagtagtta
9361 cttccattac atcatctcat ccgcccgtt cctcgcctcc gcggcagtct acgggtagga
9421 tcgtagcaaa aaccggggg atagaccgt cgccccgagc tggagtccg tataacctag
9481 gtagaaggta tcaattgaac ccgaacaact ggcaaacat tctcgagatc gtaggagtga
9541 gtaccggcg tgatggagg ggagcacgct catttgtccg tacggcagct gccgagggg
9601 agcaggagat ccaaatatcg tgagtctcct gctttgccg gtgtatgaaa ccggaaagga
9661 ctgctgggga actggggagc ggcgaagcc gggaatcca gctgacaatt gacctcct
9721 catgccgtg cagagcttga ggtagctttt gcccgtctg tctccccgt gtgcgcttc
9781 gactgggctg ggcattctgt cctcctccag gagcggagga ccagtagta agtaggcctg
9841 acctggtcgt tgcgtcagtc cagaggttcc ctcccctacc ctttttctac tcccctccc
9901 ccgcccgtca acttttcttt ccttttact ttctctctct cttcctcttc atccatcctc
9961 tcttcattac ttccctcttc cttcatcca attcatcttc caagtgagtc ttccctccca
10021 tctgtccctc catctttccc atcatcatct cccctcccag ctccctccct cctctcgtct
10081 cctcacgaag cttgactaac cattaccccg ccacatagac acatctaaac aatggctccc
10141 aaggtagcc aactaccagg cgcggttcag ctatctacca gcctgatgat ctcgataact
10201 cacatcaatg ctatttaggt cggatcaac ggcttcggtc gtatcgagc tatcgtaagt
10261 tgctcttccc gcattactgc tttcagttgg tagagactaa ctgggtgta caggctcttc
10321 gtaacgcgtg agtgctgtcc ttcaccgta acttgacggg cccaccggg gcccatggag
10381 tcaaacgcac attgcgagaa tgttatactg acatgtttgg ctacacagca tcaaccacgg
10441 cgagggtgac gtcgttgctg tcaacgacc cttcatcgag accactatg ctgtacgttc
10501 tgccctcatg cattgtctgc tacagatcaa ccagaagcaa gtggctaacg gagtgaatct
10561 gcaataggcc tacatgctca agtacgacag caccacggc cagttcaagg gcaccatcga
10621 gacctacgag gaggtctga ttgtcaacgg caagaagatc cgcttcttcg ctgagcgtga
10681 ccccgctgcc atcccctggg gcaccaccgg cgctgactac atcgctgagt cactgggtgt
10741 cttcaccacc caggagaagg ccgcccgtca cttgaagggt ggtgccaaga aggtcgtcat
10801 ctctgctcct tccgctgatg ccccatggt cgctatgggt gtcaacaaca cctctacac

10861 caaggacatc aacgtcctct ccaacgcttc ttgcaccacc aactgccttg ctcccctcgc
10921 caaggtcatc aacgacaagt tcggcatcgt tgaggggtctc atgaccaccg tccactccta
10981 caccgctacc cagaaggctcgc tcgatgcccc ctccagcaag gactggcgtg gtggccgtac
11041 tgcggcccgag aacatcatcc cctcttccac cgggtgctgcc aagggtgtcg gcaaggctcat
11101 ccctaccctt aacggcaagc tcaccggtat ggccatgcgt gtccccacct ccaacgtctc
11161 cgttgtcgac ttgacctgcc gcctcgagaa ggccaccagc tacgacgaga tcaagaaggc
11221 cctcaaggac gcttccgaga acgagctcaa ggggtgagttt cccaagatga aatgacaccg
11281 tctggcttga ccactaatct cttcccaggc atcctcggct acactgagga cgacatcgtc
11341 tcctccgacc tgaacgggtga cgaccactcc tccatcttcg atgccaaggc cggtatcgcc
11401 cttaactcca acttcgtcaa gctcgtctcc tggtagcaca acgagtggg tttactcccg
11461 cgtgtcgtcg acctcattgg taagaacctc ctattctcct ataagaaatt ggtgacatct
11521 cactaacctc tcttttagcc tacatctcca aggttgatgc ccagtaggaa tcaggacggc
11581 aaactgaatt cagaagtgtg ctgtgagtga gactgattgc cgagcgcaga cgactctcgt
11641 ggaacccggc ttgtggagaa gcttgagaag gtcttaactc ctagcgtaaa agctcatgat
11701 gacgtacaat ttaatgaaat gatacaatgt tcatatttcc cgttcaaatt tccggccttg
11761 gtcagtgcgt aagatgtcca cgattgaata ctaactcagt atgggttttg tagcattggc
11821 aatgtagtta taagcatgca ccggttgaag acgtcggccc cagatgcaat gctgcggtgg
11881 tgactaagct ctgcagtga tggaaatgcgt ttctttgatc gacttcggcg tgccgcggga
11941 ttttctcggc gcttctactg gtgcagaaag gacgatacca ctggctttcg gtccatgcca
12001 catcccagtc tcccgggaaa ttcattgcat actttaagaa acaaactgat ctccataatt
12061 tccgtcttta gagttcactt ggtacttttg ggtggatcga ggggtgtccg cggccatcca
12121 agtcacgtgg agggcagcta gaccacggat tttagagcta cattgatcca agactcctgg
12181 accagatctc ctctgctggt gtgcgcgttg cggtgactac cgtaccaaca gccgcccggtg
12241 gccttttaga cttggcctga agttgtattt tgcggcactt tcaattgctt ctgatacggc
12301 ctggtttttc ttgcattggt cgttgtcgcg gctgagcctc tgagcttcgg acctagtact
12361 atctactagt tcccttgttt tgtgctctcg caactctcca tacagcgacc tcggttctgc
12421 ttgtgtgctg gctcgcaagg ctccaacccc taccatctta gtctaactct ttaacttatcc
12481 ctcaattgtc tgccacctgt tccttgggtc atctgcagtc gtcgtattcc ggggtgtttct
12541 ttgagggaa cccacgctt gttgaacatt atctgaccct aatctgcga tttgaccatg
12601 acatttgccg gctgatagat tatcagtcag gaagaaatta ggccgacttt cacacttaat
12661 ctctatggca atggatgagg acgaagacac cgtcatggga gaggcaggct cccaccaccc
12721 cacctctggt tcttcacatc aatggccaag caatggaccc agtagccaag gtgctggcca
12781 gctcaatggt agttcggatc catctattgc accgccactg ctgggtccga ccatgaagct
12841 gccgttatct gtacactcgg gagacccaag gcctgagaac acagcatcgg cagaagatca
12901 gggagtgtcg ccaccgaaaa gcctccctaa gtcagactcc ggtgacgcaa ctactgaaga
12961 ctttgacaag caagatagcc ttgtcgaaga caactcggac tggactgatg aggatactgt
13021 agcccgccact gggcttccca tcgcctcact cccttcgggg ttatgctatg acgttcagat
13081 gcgctatcac tgtgaagtgc ggccaacagc ggacgttcat ccagaggatc ccagacgcat
13141 ctactacatt tacaaggaac tttgtcgagc aggactagtc gatgatccc aatcgtgccg
13201 tccgcttgtc tcaagaccac tgaagcgcat caacgctcgc aacgctactg aagcggagat
13261 atcgcttatt catacgctg accattttgc atttgtagaa agcaccaaag gtagaactat
13321 gttcattgtc tttcaagttt cgatattcta acggatgaca gacatgtctg atgatgagct
13381 gattgcattg gaacacacac gtgactcgat atactttaat aaattgacgt ttgcatcttc
13441 gattctgtca gttggaggcg cgattgagac atgcttggct gtcgcaaccc ggaagggtgaa

13501 gaatgccatt gcagtcatcc ggccctccggg gcatcatgca gagcacgaca aaaccatggg
13561 attttgtctg ttttaacaatg tctctgtagc ggcgcggtgt tgccagaacc gactaggtga
13621 cagctgcaga aagattttta ttctcgactg gtatgctagg ccctctagac ttogatttgc
13681 agagatactg accgactagg gatgtccatc acggtaatta ggggccctga caagaagaga
13741 acgaaagcgt ctgactcggg tttctaggga atgggtattca gaaagccttc tatgatgatc
13801 ccaatatcct atacatctca ctgcacgtgt atcaggatgg gaaattctac cctggagggg
13861 aagaaggtga ctgggatcac tgtggtgaag gccctgggtg cggaaggtat gatttgcagt
13921 tctctgtttg gaagcttcta acaagtttag aaatgtgaat ataccttggc cgagccaagg
13981 gatgggcgat ggtgactata tgtatgcctt tcaacaagtt gtcatgcaa ttgcccacga
14041 atttaacccg gatcttgtca tcggtagcga acacttcaat tagcggacgt gttcatctct
14101 aaccaatgtg aaccagttgc ctctgggttt gatgcggctg caggcgatga gcttgaggga
14161 tgtttcgtta caccttcgtg ctacgctcac atgacacaca tgctcatgac gcttgccaat
14221 ggcaagggtg ctgtttgtct cgaagtatgc acatgcggta acaattgagg attgtaggct
14281 aatgatataa tagggtggtt ataacttcag gtctatctcc aagtcagcac tcgcagtcac
14341 aaagacactg atgggggacc cgccggaccg gcttcatacc acctcacctt cagccctggc
14401 caccacgact gtccggcgtg tgatgatgat tcagtctcac tattggcatt gcatgtatcc
14461 taaggggcct cccaagatg ggctcttcac cgataggctt catggtaggt tgtcgctctc
14521 tacatccggc aatgctaacc tctgcagatg tcatacgcg ctaccagtcg aagcaactat
14581 acgacaacta caagctcaca tctctcttca tatacggac cgccatctcc cggtcattcg
14641 aaaatcaggt tctagcaagg taagtagaat cgcaatgatt ttggctgtgc taaccctgtg
14701 agttccaatt atcaccaaac cgtgccactt gtagtgatat tccacgatcc gtaagttcga
14761 tatgctgtcg actagtcaag cttcttgtta acataagcag gcctgagatc atgggcttgc
14821 cccatccagt caccaacaag ctggaagccc ataactgctg gctggtatgt gtctgcgtca
14881 taatatagca catgtactga cagtggtagg ctgatgtcat gaaggactac attgggtgga
14941 tcgttagcaa ggggtatgca gtcacgatg tcaacatccc taagcatgtg acaattgatc
15001 ctgtgagctc agtggcctct ccaattgttg aatgaacact gacctatgca gtcacccgtg
15061 aaatatgagg aagaagacga gaatcgtccc acggctactg aggaactggc gggctacctg
15121 tgggacaact atatcgagtt agtgagccga aactaccaa aggagccgg ggactgacaa
15181 ttgcagacca agcgaggcca ccgaaatttt cttcatgggc gtaggcaatg ctttctacgg
15241 agtagccaac ctactgatca atcgaggtga gtaattatag tcgagaatag aggagcagc
15301 tgacgtgtga acagaaacat tgtataaacg agtgaacagc gttgtttcct ttgttgccga
15361 gaaccgggtt cgcgccatag cctcacacac tcaggtttgg ctgtcgcggg ggtacaaaga
15421 tgtaagctgc cttacccctt tttgtcctga tattcaatca tactcacatg ccaatagaac
15481 tcccttgtat tcgtctcgca cagcacggc gtctggaaca acgtcgaaac ccggcgcaag
15541 ccatccaagc gctacggcca actcgtccag tctoccaaag ccagtctgag tgagatgctc
15601 atgcaccaca aggaagaggt ataccaatgg atatccgagc gtgcggatgt gcaggatagc
15661 gaggagacgg aagatgagaa gaagcccaag tcgaagagcc ggtccccgac gaaggaaaac
15721 cagacggagc cccgttcggg gccgggctga gcgggttatt aagtggatgt gcgagctgta
15781 attttgtttc tgttggatct atgcataggt gggaagtctg gttttgtctt ggtgatcgtc
15841 tgcaattgta cctagtagag acgcagggta ctogttgagc atacttcata agacatattg
15901 ctgtcattat tcacatatcc aacactgata atagggtagc tctataatag attgttagtc
15961 taccgggcag ggaggacagt gtctcatctc acccaatgta gtgtgaatta ccgcgatggt
16021 accacccata aacaggcatc atcaccgatt tcgcatccat tcacaaactc cctcatacaa
16081 catcttcaac ccacgtaagg ccccccatac atatattaac ccatttcaac tcaatacaat

16141 gctatctcat cccctagcca ctgcccgtg taatacgagc aacctacata cagtcttggt
16201 cttctgcaa ggaaatccca tccgatcatc cgaacctccg caaccgaact gagacataca
16261 tgagtagttc aaccatgact cctccacagc cgaatctcat gactggatga ttgtaactta
16321 taccgatata catactgctt atgtgcttat gcacctgaaa aatgagagta tagagtgggt
16381 ggagtaacca taagcgtctg tctctctcac aaatgtaact caaatgtctc acgcacgaac
16441 cccgcggggt aagccgcaca gtactatcat cataccttct tattactatg catgtattcg
16501 ggtgtatgta cggaccagac agaacagaaa tatacatacc caattatagc tgggcagatt
16561 gcttgattgc cggtttatct atgtgggacg ccttagatgt ggtgagatga gttccattgg
16621 ttgggtttta gtcgaggtca gttgattcat tgaggaatgt ttgcttcggc ttgatttgat
16681 tcggttcgat cgagatcgaa ggtgggtatt attttgggat ggcactatgc gctgctaattg
16741 tacttggcga tggagatgaa gaggagatgt gagacattat gacctatgcc tatgccattg
16801 atcgaaattg ggtgggatga tgtagtttct atattctctt ttcgtaagga tatgttcaaa
16861 gactagaaag gataaatggg gtatataaag caccctgact cccttcctcc aagtctctatc
16921 caaccagcca tcctacaatc tacacatcca caccaatcta ctacaatcaa caacccaaat
16981 gcatctctcc accaagagcc tcctctctct cctccctctt cccccctag catctgctgt
17041 cggtccgcc atcgtgcaaa acaactgcac ctccccatt tacctctggt cggttggcgg
17101 ctccgtcaat agcatgcaaa ccatcgacct cggatccagc tacagcgaga ccttctacta
17161 cgatacaacc tccggcggcg ttgccctgaa gatcaccacc acggagaatg gtctgtacaa
17221 cggtctctcc cagacggatt acgectatac gctggacaca tcctcgggta atgtcttcta
17281 tgatatttgc gatgtctttg gggatccgtt ctccggcagt gtctcagcc ttgtctcgag
17341 tgattccagc tgcgagagta tttgttgggc tgggggagtg ccgccgctg gcagtgttgt
17401 gagaagttgt caggatgagg cggatgaggt gttgacgggt tgtgcggata gttgctaggt
17461 tattgggtcg ggaaggattg atttgtggga attgaaggga caactgggtt gattggggat
17521 ttaatgttgt tgttatttgt gggttctctt atggattaat tgagggatgt ctggtgaagt
17581 gtgtctgtgt ttagatgttg gtgactggac tgggaagata tgtgatgtgt ttatgaggat
17641 ataggaagta taagccta atgtacattgc ttgttgatt ggacagtgat cccctctacg
17701 tttgggtata tatgttccaa atctggagca gtgactatac ggtctagtcg tcagtgaagc
17761 attggtctat tcctgtggca ttgcttgcat gagagccaat cattggatcg atgacctgt
17821 ggtagggact ggtggctgat tcaactacat tatcatcgcg tatttgtag tgtgttcgca
17881 gttgtcgcgt gcatatggtt gtagtagtat acagggtgct gcgtaatcg gctagctgta
17941 ctccacagtg agcggagaca atgcaatctg ggtgaagaaa actagatgag ccaccacaat
18001 taaggctact tagaagagct tttcccaata catcagcata caatattact aggaccaggt
18061 aaaataaaaa tttagactca ttgatcccg actggaccga tcattggtga agctgcactg
18121 aacttggtta tgtaaaactg cggagttgcc gaggtatata tagtagcaga aatcgtaacta
18181 cagtagtgcc cgggtgaagg gcaagcagtt ggcgttgaaa ttcaacaact tgaaagaccg
18241 aggttgaatc ttgaccataa atagataacg agctatacaa attaaagaaa cagcaattat
18301 tcaactaagta gcaacagata gtatacaagg tgaactactc aagcatcccc actggtgacc
18361 gccttatcgc ggggcctcac cgctatgac tccaagcatc cattgtgcat cgctcccg
18421 cgatggccac cgacaagtca ctctccctcc ttcgtgccac aggatatact tcgaccattg
18481 acccggtaat tgcctcgcat tcccttcac ctcttcttac gtggcgtttt gctgacccca
18541 ttcactctgc tcagttttta aactgcta atttcaaatag ggatctccct acccctttca
18601 cttttcacgc cggcaacttt ccaacaccgc ctgogccccg agcggtgctt cccggcaacc
18661 gatccctcct ttctctctgg atgcctccaa ggtccctgc aacacgcgc ttgtggaact
18721 ctcaagtgcc tttatatctt cgctccccag cttttcgcgc gttctccaag ttaggcgcct

18781 ctgcggcgat tgccctcat tttatttacc cctcgccacg cttccccacg ccgaacgagc
18841 gctgctcctc atttcctgcc ctgcgctcgt ctcacgctct tgtctcgctt tctctcccct
18901 ctcggtgctcg ctttttctct aataccgcaa ccatgggacg tatcgagcgc atcgctgagc
18961 agttggagaa gccggagctc gatgatcgct cctatcggtg cattcgctcg ccgaacaaac
19021 tagaggctct tctcgtacat gatccagaca cggacaaggc tagtgcggcc gtcaatgtca
19081 atgtcggcaa cttttcagac gcagacgaca tgccgggcat ggcgcatgcg gtagaacacc
19141 tgctttttat gggaacccag aaggtaggtg gtgcgtcccct ctgaaacatt agacccttg
19201 cccgcttggtg tttatgcggc aagatgttaa ccctgtgtag tatcccaagg agaatgccta
19261 caatcagtac ctagcatccc actccggttc ctccaacgcc tatactgcag ccaccgagac
19321 gaactacttc tttgagatcg gtgccacacc tgccctctacc gaggacacac ccaacggcgc
19381 caatggaacc tcgaacggca cagacacacc cgcaaagccg aactatccca cgtctccact
19441 gtacggcgct ctggatcgct ttgcgcaatt cttcgtagag ccccttttcc tcgagtcgac
19501 cttggaccgc gagctgcggg cagttgatcc ggaaaataag aagaacctgc agagcgacct
19561 gtggcggttg atgcaattga acaagacttt gtccaaccct gccaccccg accaccactt
19621 ttcgaccggg aacctgcaga ctctgaaaga ggagccgcaa aagcggggac tggaggtccg
19681 acaggagttt atcaaattct accaggcgca ctattcgctc aacataatga aactggtggt
19741 actgggacgg gaccctctgg atgaaatgga gcagtgggtg ggcgacttgt tcaaacatgt
19801 gaagaaccaa gatcttcccc agaaccgctg ggaccacgct caaccatgcc tgccggagca
19861 tttgggcaaa cagatcttcg ccaagccagt catggacatg cgctcgctcg atatctactt
19921 cccgttcatg gatgaagaat cgatgtttga atcccaaccg agccggtatc ttagccacct
19981 gatcggccat gaagggtccg gcagcatcct agcctacatc aaagcgaagg gttgggcaaa
20041 cggtttgtcc gcaggcgtga tgctgtctg cccgggctcc gccttcttca ccatctccgt
20101 tcggctgaca cccgagggtc tgaaacagta ccgcgaggtc accaaggtgg tgtttgaata
20161 cattggcatg atcaaggagc gggagcccca gcagtggatc ttcgatgaaa tgaagaacct
20221 ggctgaagtc gagttccgat tcaagcaaaa gtccccggcc agtcgcttca cgagtcgcct
20281 tagcagcgtg atgcagaagc cttacccag ggaatggctt ctcagcggca gcctcttgcg
20341 gaagttcgaa cccgagctgg tcaagaaggc actgtcatac ctgcgacccg acaacttcag
20401 aatggtcatt gtcgcccagg attaccctgg cgattggaac tgccgtgaga aatggtacgg
20461 aactgagtat aagggtcgaag atattccgca agacttcatg gacagcatcc gcaaggctgt
20521 ggaaacctcc cccgagagca gactgtctga gctccacatg ccacacaaga acgagtttgt
20581 tcctaccagg ctctcggttg agaagaagga ggtgtctgag ccggccaaga ccccaagct
20641 gatccgtcac gatgaccacg tccgattgtg gtacaagaag gacgatcgct tttgggtccc
20701 caaggcgact gttcatatca ctttgaggaa tcccttggtg tgggctacac cggcaaacct
20761 ggtcaaggcc aagttctaca gcgagctggt ccgcgacgct ctgaatgagt actcctacga
20821 cgctgagctt gctggtctcg actattcggt gtctgccagc ctcttcggcc ttgatattct
20881 agttggcgga tacaatgata agatgtcagt tctccttgaa aagggtgctta cgagtatgcg
20941 ggacctggtc gtcaaaccgg atcgtttcaa catcatcaaa gagcgtatga ctcggaatta
21001 caagaatgcg gagtaccagc aaccattcta ccaggttggc gactacactc gctacttgac
21061 tgccgaaagg acgtggctca acgagcagta tgccgctgag ttggaacata tcgaggcaga
21121 ggacgtttca tgcttcttcc ctcaaattct ccgtcagaac cacattgaag tgctggctca
21181 cggtaaacttg tacaaggagg atgcgctgcg catgaccgac ttggtggagt cgactcttca
21241 gagccgtact ctgctgaat ccagtggtgta cgttagacgc aacatgatcc tcccccggg
21301 cgccaattat atctacgagc gcacctgaa ggatcccgcc aacgttaacc actgcattga
21361 atactatctg ttcacggta agattgacga cgatgctctg cgggccaagc tgctgctctt

21421 cgctcagatg acggatgaac ctgcctttga ccaactgcga agcaaggagc agcttggtta
 21481 tgttgtttgg agtgggtgcc gctattctgc gacgaccatg ggctaccgtg tgatcatcca
 21541 gagtgagcgc aatgctcctt atctcgagtc tcgcatcgat tctttcctga ccggtttcgg
 21601 aaagtctctg cagaatatgt cagagcaaga ttttgaaaac cacaagcgca gtgttgtcaa
 21661 caagcgactg gagaagctta agaacttgag ttctgagacc aatcgcttct ggtcgcacat
 21721 cggctcagaa tacttcgact ttgtgcagaa cgaaagcgat gctgcgaacg tccggaagtt
 21781 gacccaggcc gatctcatcc agtttttcca gcagttcgtc gatcctacgt ccgcaactcg
 21841 ggctaagctg gcgattcacc tcaaagccca ggccgggtgcc catgcgcatg ccaccaagcc
 21901 cgaggagcag aaagctcagc tggtttctact cttcggcaag cagctggagg cggtggcctt
 21961 cgccgcggac acggaacggt tgagcgccgc ttttgagaat gtcgacgtgt cgagcggcga
 22021 cgaggctcag atcctgacca ctctcaaggc gttcctcgaa tctggcctga gtctttccga
 22081 ggagcagacg aagcccgta tcgggcaggt cgaaagcaac ctgggcctgc agttgaagca
 22141 agtgggcatc gagcctaaga agagtgatgg aactgccgtc caggcgaatg gcacgaatgg
 22201 cgacattcag aagtctacct acatcaacaa tgttcccag ttcaaggctc gtctggccgt
 22261 tagtgctggt ccagtcctgg tgaccgacct tagcgaattc gaggattttg acgccaagct
 22321 ctaaacttca tgattgtgtt ggatactggg gcggaagtgc gtctcatcga tagacatgat
 22381 gttgtattgt tgggtttgca tagttggcgt ggacttggtg catgggaggg atggggctgt
 22441 gcttgggggt aggttgtgat ttgagttaga tttaccctga taaatgagaa tagaatcagc
 22501 aaagaataaa attcatatga tgattccaag atatattctt ccaatggttg ctagctacat
 22561 tgtacggagt agagtgatta tggcgcggtg cgggcggtc gaatgttttc gcgggaggac
 22621 cgctccgacg ccgagcccag ctccgccatc atcggcgatc caccaccgag agctccagcc
 22681 gctcttgact tcccgactat ccagcattcg agtccccag cctgtttact cgctgcgcat
 22741 tggttacat cgcgttcaac atgtcgacca ccaccgacg ggcgatccaa agctcgcgcc
 22801 cgtacgtcct gacatcgacc tacctcacta ccgcttctcc acgcaatcac cactgacat
 22861 gtcataaata atagcccat catcccaaa gacttctccg cgcaacaacc caaaacaatc
 22921 cgctgtatc cctctccaa ctacacattc ggcaccaagg aaacgcaacc ggaggaggac
 22981 ccgtccgttt tggcgcgctc aaacgtctc gaggagcact acgacctaca tggcatgaga
 23041 cgcacctgcg agggcgctct cgtctgccac gaacacaacc atcctcatgt gctgatgctg
 23101 cagattgcca acgcgttttt caagctgtac gtttgatac cctctgcaat ctggagcttc
 23161 tacttaccct ggtggataca gcccggtga ctacctccac ttogaagacg acgaagtcga
 23221 aggttcaag aaacgcctca atgaacgtct cgccccgtc ggctcccagt tctcgggcga
 23281 aggcgttaac gaagactggg agattggtga cacgttggcg caatggtgga ggccgaactt
 23341 tgagacgttc atgtaccctt tcctccggg acatgtgact cggccgaagg agtgcaagaa
 23401 gttgtatttc atccaattgc cgaagaagag taagccttcc ttcatatatc atgcacagtc
 23461 aggccatcag ggaggaatgg tatgattaat ttgtgttgtg cagaggctct ctcggtcccg
 23521 aagaacatga aactcctcgc cgtccactc ttcgaactat acgacaacac ggcccggat
 23581 ggccctcagc tctcggctat ccgcacttg ctgtcgagat acaattttga gtttgtcgat
 23641 gagaacgaca acgtggttgc cgcgacgccg ggacgccc tgcgggatca tcagatcccc
 23701 cggacgaagg tgctggctgg tgatagcggg gatggacagg atgaggggat gacggattac
 23761 ccttctggag agggcgaaaa cggggctagt gcgtagtttg gggttggtcg tgtgcgtcct
 23821 gtgctgtcct gtcttgaacc gaaccagact ggcagggact attattccga gtattccccg
 23881 gcgaaggatg gcagtgatga gacattcggc tggattggat cggaccctg ctgatcagct
 23941 cggggaattc tcctattccg tgggattgat tcttttctt tcttttctt ttctatgcta
 24001 tgcaaaagac gtgatgtgtg tctggttggc gttccaggag cgaagattgc ttttcttact

24061 atattattcc ttcttctatg ttcttgaatt agctatgcag gtagcggcca atatatcctt
24121 taagcagata cctttactct tgctatggat atcataatcc tagaaggtag atccaccata
24181 gactacctag tccccggcca catcatgaaa tatcactggg acattattct ccgagaacca
24241 cagaaccact ccatagcaca ggcacataat ggcacctgcc cacaccaca ctcacggctc
24301 ctagtactaa cccctcgtag tccccagtag tccacaggta cctaatacac ttgtcccaca
24361 ccctcaccaa gaccaactaa ccaccaaagc tactcgcaga tagcaagata gtcacctgcc
24421 acaccgcaac tcactaaca acgcagatca caagcctaaa tattcgaacc aaacctgaag
24481 aatcctttcc aaccaacaca cacacggtag gggttggaag gaagggaaag aaagtggggt
24541 taagcaaact aactaacaaa ctctgcaagg tagcaaacgc tacactaacc aaccaaacta
24601 gttaaactaa tattaataat aataataaga agaagaagat gatgatgaag gagtcatcgt
24661 tctcccgctc ccggcttata gtacgggatg ctaaggctgc accccgcagc acgggaagcc
24721 cggacgttgg tggtttgctg gtgggtgact ggattagact aggggactcg taaggcagcg
24781 gactagttaa attaatgac tagtagtctg tgctgggctg ggcttatcta gtactgtgaa
24841 gtgccgggat gtgccgggtg agtgaagagt tcattaatga atgattgatt gatttgcggg
24901 ggtggttaggt gtagcgtagt gtaatggaac cacttcttat aattgtatta ctctactaac
24961 tgtaaagtgt tgtgttttat ggggacagga gaagttaagg tagtattggg ggcgggaaaa
25021 ggggatgggt gtctaattca gtttgtttca atagtttggt gaggtatttt cttgtcaaag
25081 ttacttgttc agattatggg cacgtgtacg tgcttagttt tcatagattg tgtgtatgtg
25141 tgtgtgtgct cgtgcgtgct gacctatcta tagcatgtat cgcggctctga atgtgtataa
25201 gcgctgcaag tgatcggcaa agtgacaaga agtctttcta cgaccacta gttatttgca
25261 tgtcacgaac cacaacaccg ccagtcacag cgacagcatg tgaagaagtc tcagatttgc
25321 taacactctc ccgtctactg ctctcctact tgatagactg acttacctcg gcctgtgcac
25381 atgccatgtc cctatcccaa tctactggct catgtgcaga ccacgcctcc acgtttctcc
25441 gccagtagat ccttcaacac cgctcatcct cagacaacat ctactttatt tgacttttgc
25501 gagaagacca tgcttgtcta cggccagcta tagcatgcta tagcagtcgg ttaatctcca
25561 ccgggctccc gctccgaacg gagatggggc caaactgcc ctcagttgc gcaacggaca
25621 ggcaccgaac cggacaacag gatgcggaag aggagacatg gtgcctgatt gcatgtgctg
25681 gcttcatctg ctatcgtgag ggtgctgtgc tggagaaatt tgttgtctga cttacccgc
25741 ttcttgcttt ttttccccct gatgccctg atggggaatt ggggtgggta atatgataca
25801 ggtataaaag ggggctcgga ggtgcagttg gatagaagca ttgtgtgtgc attgcagcag
25861 tccgttggtc tcacgtctct ggttgctcg attgtatata tactgcagga tgttctctgg
25921 acggtttgga gtgcttttga cggcgctcgc tgcgctgagt gctgcggcac cgacaccact
25981 tgatgtgcgg agtaggtgtg cctgatttga agtggctgga tagcactgat gaaggttttg
26041 aataggtgtc tcgacttcca cgttggtatga gctgcaattg ttctcgcaat ggtctgccgc
26101 agcttattgc tcgaacaata tcgactcgga cgactctaac gtgacatgca cggccgacgc
26161 ctgtccatca gtcgaggagg cgagcaccaa gatgctgctg gagtttgacc tgtatgttgc
26221 tccagtgaag tggatagaac acagctgatt gaataggaca aataactttg gaggcacagc
26281 cggtttctct gcccgggaca acaccaacaa gcggctcgtg gtcgccttcc gaggcagtag
26341 caccatcaag aactggattg ctgatctcga cttcatcctg caagataacg atgacctctg
26401 tactggctgc aaggttcaca ctggattctg gaaggcatgg gaagccgctg cagacaatct
26461 gacgagcaag atcaagtccg cgatgagcac gtattcgggc tataacctct acttcaccgg
26521 gcacagcttg ggcggcgcat tggctacact gggagcaacg gtcttgcgaa atgacggtta
26581 tagcgttgaa ctggtgagtg cttcagaggg tgatcattaa acagccggtt ctgacagtca
26641 atagtacacc tatggatgtc ctcgagtcgg aaactatgcg ctggccgagc acatcaccag

26701 ccagggatct ggagcgaact tccgcgttac acacttgaac gacatcgtec cccggttgcc
26761 acccatggac tttggattca gccagccaag tccagaatac tggatcacca gtggcaccgg
26821 agccagtgtc acggcgtcgg atattgaact catcgaggga atcaattcga cggcggggaa
26881 tgcaggcgaa gcaacggtgg acgttttggc tcaattgtgg tactttttcg caatttcaga
26941 gtgtctgcta tagctggaca gtccgatgaa ataagtgcgg agagaaagtg taaatagtaa
27001 ttagtatata atcaggcaga gaagcagtgg tggtcagaga agaaagagtg agatcccatt
27061 acgtagcaga taaaccacgt gtggaggcgc tgttcctcca cttgcagttg cggccatcaa
27121 tcatcttctt ctcttactt tctgccacca taactctcat cctgccagct gtgcgcatcc
27181 cgggtttcaa caacagtcgc ctccggggcc tccgtggttc tcctttatta ttccatccgc
27241 cggccgacgt ttcaccctca tctgcgccg ccgcaattct ctccggatcg gtcaagtcac
27301 tcgaaccgcc gcccgcatcg acctcaccga cccgaccgtc tgcgatcgcc catccgtgca
27361 gccatgttcc ctgccccgtc actctgattt gacgctcaat tgacttccgt ggcatacata
27421 catccagcag catatgatca cagctcccc gcctcgtgcc agtctcgccg ccgcccgcgt
27481 tgccagccat gctattcttc cgcggcggag ttaacagtcc gcgcgggcgc aaccaagaaa
27541 aagcaatctt ctgtaccatt cttatcatct acgctctata ctctctcttc ttogctggga
27601 gctccgatcc caaacggtcc gagacgtctg ctgaggatgc ggtgaaggga ggtccgcac
27661 gggactcgcg gtcggggaag gcttccccgt ccgcctcggc ttcgcgacag cgtcaggccc
27721 tggacaagga tatggtggtg gctagtatga agggagatga tgtctcatgg ctgtacacat
27781 attttccga gtggcataaa agcatctacg tggcggacga caagaaagcg ccgttgacgg
27841 tggctttgaa taaggggagg gagtccatgg tgtatctgac gtgagtctcc gtgtcttggc
27901 cttggcctat ctgggcacac ttctactgc tacagatata taattgataa ttacgacaac
27961 cttcccgact cagtcctctt catccattcc caaagatacc aatggcaca cgacgaccct
28021 tactacgatg gtgtgcccgt tctccgcaat ttccaggtec cgtatctcca gaaacaaggg
28081 tatgtgaate tgcgctgcgc ctgggtcctg ggctgcccgg cggagattca tcccctcagc
28141 gatacgcatc gcaacgatgt ccacgcaggg gaatacttca agactgggtt catggaactg
28201 ttcccggggg cggagggtccc tgaagaagtg ggggtctcat gttgcgcgca attcgccgtc
28261 acacgctgga agatccgcga gcggcctaag agtgactatg tgcgatatag gaaatggttg
28321 actgagacga cgttacctga cgatctcagc ggacggataa tggaatactc atggcacagt
28381 atgtccctaa atcccacttc agcgccaagg gcaattgact aacagtaaca gtgattttcg
28441 gcaagggccc tgtgtactgc ccgacagcag aagaatgcta ctgcaaggta tttggactat
28501 gcaatttgga ttgtccgact gagggagatt gcgctggtcg ctatgttctc ccgcatact
28561 cctcccttcc aaagggatgg ccgtatgtcg gctggaaggg tcagaagcag gaccctagca
28621 agggactccc ggaggcttga tcggatacct actatctcat agttgagaag atttgttcga
28681 gtataagatg ttgtgactcg ggtatacac tgcctactga gatcaagaaa agggactcac
28741 tacaagcatg atgaatgctg tagtagaatc agagcggcag cacatctggc aaaacagggtg
28801 agcgccagcg gcagcagccg ttctattccc actgggaatt tcctttgaaa cagccagatc
28861 tcaactcctc taccgctgcc acacacagtt cggcacacca aaaaaggcat atggcatata
28921 acataacata ctactggcat cttcatgagc gaatgctgtc attgacaact cctgatgtat
28981 tataacaaca ctatttccat tccatgtgaa gtaataacta ccacaaaaat atactactcc
29041 atctctcaga tcttgcatca catctatta tactcaaaaa aaatattgat ccctcgctca
29101 accccttatg cctcaggcat gtgatattgc aaggttgctt tatgctcagg caggtattac
29161 tataacggtc tagatcgccc atggaatcaa ttaccgccc atatcgaagc ccatcagatg
29221 cagacgcac ccctaacaag gcataagacc catataaaca gagacgacgt cgtgaatttc
29281 gaagctacaa tgaacttgat cgataactcg atatctgctt atcgtcagaa tacccttttc

29341 cgggacgtta atctggttca atggctgaaa caaaatgacg gaataactag ctaaccatag
29401 tgagttacat gtagactatg attccataga tatatgcaa ggcttgagga ataatggagc
29461 aggactgtgg gaattttattg cttactattc actactagtt agtatatagc acggtttcag
29521 tatggttggtg tgtatagtta gatgggaagt tgggtggttg tctgtgttgt gccgttcgat
29581 ctagtagatt gcattagaac agatgttgat agggacagga ttatgattgt gacctcgtat
29641 agcacaaagg atatatgcag ggtagtggat gattctaccc cggtgatac taggcgttgg
29701 atgcgagctg cctgaatgag tgggctcagg agcaatttgt atcccaccgt catctataag
29761 taaggaaatg aaacagcaac gcagtcagtt gtccataggct tttgatgagt tttgagcgag
29821 tctatgtcta caaaaaccgg aacggaaaag aaaccacccc tgacaaagga aaaaactttt
29881 attatctcgg ctcatgagta acagtatact caacatcctt cttaataatc atcttcgtgc
29941 tctccgaact agccacgcta tccccgtccg ttcgactact gccatcgttt gcagtagtgc
30001 tgagcatgcc accgtatcgt tcgggcacaa atgtaggttg ggaatgtcgg gggttgttca
30061 tgccgtggcc cgcctcttcc tgattatcca ttgactgcaa aacgtagctc tgggactgct
30121 tcgatacatc ccgagcgtat ttgcccgaact tggatctagc gctgctgggt ccaagggatga
30181 ctccccctgc gttgccgtag gtgacgctga tggcgcgcat gaagggggcg aggcagaagg
30241 aggtgcacac gatgagggcg taattcatgc ccatttggtt ccagacgtag ccgtcgatta
30301 tgtctagtgt cgggttcgat gagtaggcat atttttgcag atggtaaaga tgtagagatg
30361 tgaatacgag gagactaatg taagcaccgg gttaggtctg ttgtgtttgt gtagggcggg
30421 actgcttacg gtaaacgaac agcgaaagag gaaagaacga cgagttttcg agttgttggc
30481 atctgcaggc ttttgacgag aaggaccgag aagaggaaga gcatgagctc aatgatgctg
30541 tccattacac cgatgtattc ccatttttgg aactggtgag gtgcggtcag tgtaactgca
30601 aggacgcgga gcgatgggct ttataccaat ccagtacagt ggtcaatggg tcttgcccat
30661 ggtctgttca attcgcagtt cacaaggatg atcagaactg aagacagaac ccaacagatg
30721 cttaatccca gcaccccgta aaggacgaaa tgatggatgc gttggggagt caaccgtgag
30781 agcatcccaa gcacacagca cttggagaga aagaccgtga tcaagaacaa agtatctgta
30841 gcagctgttg tctgcatgga tataaccccg gttaaattgt gctctcgtcc tgaatccaaa
30901 caacgtaccc gctgaacctt cagcactga tctgaagaaa gaagccggat agacgttccg
30961 aatccttttg cagagtttga catgacgagt gaagtttgaa taatagcgaa agcctattga
31021 agacgcagtg attaggatcg catttgtatt tgtatcactg gagatatagg gttaaagctt
31081 accgtagcaa cgaataatag atagtcattc cgtccaaacc ccggactcag tattgttctg
31141 acatacacgc ggataatcaa gcaaatcaat gtcattatca gggccaaagt cccaacaaca
31201 atcacaatgc cggcatgatg actctcgtca atgaccacaa ggggcttccc agatgccgtc
31261 atgacccgag ccacccagag acgtcagcaa gcgaccgtga gtggcctatc gtcccatgca
31321 ctcccgaaat tcaagagatc gaattaagat cccctaattg agggctctct gaatataagg
31381 accagtacat aaatgaacat gcatgcaggt tttaggtgca ttccgaagcg aacgcaataa
31441 gaccatatct tcgggctcaa gaccatggtg gcaaggctcag ttaggttatc aaggatctgg
31501 cccagtcctt cgaggtttgg tgaaattgag atctattatt ggcggcagta atagcacaaa
31561 gggctcagct agagcgtatg gaaggatttt tgggtccagat tcggaaagtg gccaaaatgt
31621 gttgagagga acgaagggcg ctttgttcat tacggctgat gttcggaacc taaagcccct
31681 gtagtggttc gagtaactaa attcgaccct atacaaatat aagttaataa cggaacagga
31741 agcacttgca gggatgtata ctaccttagg atttagtata attgcaacac gcaatcacat
31801 cgcacccgtc acaaaccatg gcagactaag tacccttagc tcgactagac tgccactacc
31861 ctgtgctagt gtttagaccg agtgttccgt cgaggcgagc cggacatggt tcttgggtgac
31921 cagtgc aaat ggaaaacagt agcctcgcgg agataacagc tcacgaatcg ctataagatt

31981 gcagcctggc ttttaaaaca tagtatcgac tgggtgtgata ctagacagga gcgggcatcc
32041 ttctggatcc accaaggagc gctccgggga atgtgtccgc agcggcgctt ccgttgatcg
32101 ccggggaagt tgcccggatc agttagggct cgtgaaaaga gcgctggtaa gctcccggtg
32161 ttccgggtggg ccgcggcggg atcgtctggc gaacaagatc cggcggggcc atcgccgcga
32221 tatatgaaat atatagttgg ctccagtagt aataaaatcc tgccatttga gtttcgtagc
32281 attatcagtg gctgaattaa accatattgt agagaaaaag atttcgatat taaccccact
32341 gtgggtagag gccaaggat taagatcctg gccgtccgct atacatgaat cgccagttaa
32401 cataggttcg aaatccgttg cgccggcgaa aatggtgcag gcacttcgga cactgttgca
32461 gccgataggc tgcggtggag accgccaatg tagacctgtt attgcgctag gatcatcgaa
32521 atggcaagag gtggcgtgtc catgcagcag aaaaaatgac tcaattgatg aacagccaaa
32581 tcaaatttcg cgcttgacca gggccacggg cggcacataa gttttccgcc catacaacac
32641 atagtagatg atggaaaaaa tcaacatgcc accatacatg agcgagctcc agttcatgtt
32701 ttcggggtcc acagtcgagt aaagggggaa catgacgaaa acgaactccg ggatgaggaa
32761 agccagggcc ccgaggttta cccaaagtcc ccatcgggccg agtgaccacc tgtgtggggg
32821 caacggttcc ccgcgaattc ttttccatga gacacaggaa acagtgacga tgtaggaatt
32881 caacaatgct ccgacagtta gggatgatca ggcgtctagg gcgacagtcg agccgatgtt
32941 gatgagcgat agcagcatcg tgagcacgag ggacactaga acagcattga gggggatgtt
33001 ccaatttggg ttgacctgcg tgcgcatcag tatttagcat agtagagctg aaatcaaata
33061 aaaggtggga acccttacat gaccgaaaaa ggtagccagt gggacgcccc catcacgggc
33121 gaatgaccaa agttgacgag acgcagtggc gacctcagt atgacactgg caataagcga
33181 catggtgaca attgcagtca ttgcgtttgt agctgcgtaa cttttggtgg cattgaaaaa
33241 gacctgaatg aaagggtaac cagtagcgga gtcgaggacc tcatctacat tacccaaggt
33301 gaaacacaaa gtgaccataa caacaaaggc caggaccccg tttagggagc acgctgccat
33361 gatagccttg ggaaggggtg tagaagcgtc tcgaacttcc tcggctgtgt gccgcattag
33421 catttctaaa tgatcgagcc agagctttta ctacacatg tggacagtac agtcgtagcc
33481 aatcatggca gtgatagcag taatcagacc gacctcgtt gcggtaccat tgctattcca
33541 atcaccaccg ttcgtaaact gggtgaaaac cgttttcgca tccttgatgg gtgccagcac
33601 ccaaggggt ataatcacia tgaaaaatcc ggcaacgtgg aggatcaaga tgaaggcttc
33661 aacgagaggg agcttcttgg ctaggaaagt gttgaaaaga acggagaaaa gggatgatggc
33721 aattgtaatg agagtaccgt gccaccgctc agaaacataa tcagggtaat tcaggataac
33781 cagtccttga atgatggtac ctgctaggaa agcaattgac acaatcgccg attcccagcc
33841 tgtggcacag agccatccta gatgcgaccg ggcttattag gctcagccaa ttttcagcgg
33901 gaagcaacac ctaccagtga tataactcaa gaacttctgg tacttcctag gtgcaaaactc
33961 tgatacccaa tggatattggc caccagcagt ggggtgcccta tgaatcgatc agctgggtga
34021 taatttttagg atgaagtgcg tgcacttaca tggaagccat ttccgccaga ctggcgaaca
34081 gcagcaacac tccaaccgaa acaacgatga aaccccagat aaatccagcc gtgccaccat
34141 caaccagagg cgaactcatc atactgcccc aaccttaata tcgagggcaa accagggaga
34201 acgtatgact tacgttagca cgacttccca tgaagcaatg agggtagatc cgaaggcaac
34261 aattgagatg aacctaaaat ttctctgagg agaccattag aaagagactt tttaccgagc
34321 tggcagataa cgaaccctga gaacttgctc ctttcccaaa gcctgcatat ccctttggtc
34381 ggccatagtg ccgatatact tctgggcgac atgacctttg gcaagccctt gatcagcgaa
34441 tacagtggaa atattatttc cacccttggg ggttggtgagt tcaacattat ccggcatttc
34501 ttcggtgtcc tctgtctctc ctctaggttt tgattcttgc aaagaacaaa cacgaatttc
34561 aaggcgaggg acaagtagct atatttcaat ccagtccttc agggccctgc gtcgtgcatc

34621 tggattagcc aacccacccat ttactgacga gttcccggaa gggaaactct gcccttcac
34681 tcagtcacgc acagggggta gcgggggcct agaagccacc aatgtggatg agaagcctca
34741 gtagccaaat acagctggct ctaattcaac ttgcgctggc tgcgataaga ggagcttcat
34801 cgaatcctgg gtgttccact atcaggctcg acccacctgg cgctttcaag ctaaagcggtt
34861 ggcgtgggct aagtggacta gtgttaatct aacgccctcg ggggagaacc gcgccatacc
34921 atcccgcgtc ggttgtaata gtacttcgtc cggttataga cttttctccg actgtttcag
34981 cttattaag ctttcgggta agaagccatg tctcggcgaa gttggggatg tcgttttgaa
35041 aagctacgag gactgatcgg agcccgatca ttgccttact acgacagtat atcttgttag
35101 ccgttgcttt cactggcgtc tgccgctcgg tagattaacg agggatctcg aatatagtgt
35161 accatagata gtgtaccata gacacttaag ccttcctttc tttcagcaac agtaaaactt
35221 ggtgtttgcg gatagaggat acttaagcaa tgacagctcc taaacggcac agtctaggca
35281 ttactgttca gttcgaaggt ccctagaacc tgtagtggca gctatcgggc cactcagtgg
35341 tatgtctatt agtcgaggaa gcaaatacaa tggacaaata gaatcaaata accgtacatt
35401 ctcccacgtg accgtcaaca agcgtctcac ctgccatgaa ttactaatcc caacgacgtc
35461 atcgtctgat tcccgcgcat atcaacactt ctatacagct tccagaagat ccatctgtct
35521 tgaaagttgc cctattcgaa cagttcgtgc tcttgaggga aacaaaaatg gcggaaacat
35581 atcccattgg tacacgcaaa gaatgcgaac gtctagtcaa ggactggggc ttcaagcatg
35641 tctttacttg ggcggacgga aggtgagtaa caggaaacca tggttgtcgt gctatgtacc
35701 tcttagtatc ctgtttcata ttataatttt ccgcttagca attattgtaa ttaagtcctt
35761 gttgatgctg atactgtcct agcaatgccc actatccacc aactcccat ggcggtgtca
35821 ccacccatct aatacggcag gggactctgg caatcaccta ccagatgac aatgctaagt
35881 tgcataatgg ggaggtgaaa aaggagacgt ttggggttgg tgcgagactg gatgttcag
35941 ctggcaagtt acacgaggtg tggattggcg aggatgggtg cgaatatgtt attggtgaat
36001 aggtaggagc ggatgtgggt tacggcttgt cccttacgga ttatgatcac ttgagcgatg
36061 tgttgtatta tgagtttgct ctattgtcta atctatagac attaaactgt atcttgtacg
36121 gcgtccagca aattttccat atgtccaatc cttttcatct atgttttagt tgtgaaatat
36181 aacaatcatt attcgtgctc atctagttag taaatcgaa atcaagtctg caaggcaatc
36241 aattagacat gttaccagtc aatcttgtag aaggaaaact caccgcttgt ttccaaggct
36301 cttattggct gccagatatt cctccctctc ttcgctggac atggcattcc aacgggcagc
36361 gttgtacctc acagcctcat tagcattgta gtcattggtat cggagaattc ggtgagccat
36421 acttggtgcg ccagatgtaa taaaacttgg ctccaatgaa caaagccatg ctgaacacgg
36481 acaaggcgat cagaacctta ttaccacccc ggtagtaagg cttgtctttt tcttggtaga
36541 tctgcaaggg gtggtcagag atgagtaata atgcaaagtt cgtgtttcac tcacctgact
36601 ggcaatgata ctagatgctt ggaccatcat gttgtatact gcacttgcaa cagtgcgtgt
36661 tcggaccgat ccgcggttcc tagacgtcat ggccacaatg attgcgtgga cgtaaggcgc
36721 tccaatcaac agagcattga taatccaccg cgcccatcca ctggctccat caggcagaac
36781 ttcgagcgcg atgagcaaga ccaagcacca aaattcgctg acaacaccga ggagtagacg
36841 ctggttgaag cgttctgaga tccatgtcca gaatagtagg ttgatgatga agatgacata
36901 tgcggggatg gtcaggaggt tcgtttcgaa ggtgccgaac ccaagcgact tcaactccag
36961 ggtaatgtag ttggtagcag gtgtgttggg aatcatccat gtcaggccaa gcagatagat
37021 gggccacatg tcgtagtcgc acagagcatg ccaaagcatc ttcggcggtta ccgcctgacg
37081 gttgtgcatg cttcccttgc tgggacgtc acggagcacc cggttgacta ggatcttctc
37141 ctgcgcgtcg ttgaaccatc cgtctttgcc gcggaagccg cctttggtct gtgtgggaga
37201 tgcaggcatc cagaatgcgg cgatgattcc gatgattcct gtgatgaggc cctcaaaagc

37261 gaacagatag cgccatcctc cagtataacc gttagaaccg tgaatgtgga gaaaaccgaa
37321 agccagaaaa gcgccaacga tagatgtggc ttcgtatgaa acccagaagt aactcaaccg
37381 ttgcggcaat tcttttagatt tgtagaagaa ggacaggaaa aggatagtgt cggcaatgct
37441 agtcgttggt aggattgttc cccaagtagg aaggtgtagg aagacttaca atcctccctc
37501 caggaggccg agaagcgctc tgcaggcgaa atatccggct ctgttcttga gaaaggcctg
37561 gaatgctgcg acgaaactcc aggcgaccat ttgaatggga atccagcggg ctgggcccgc
37621 cttcttagac agtaactgcg aaggaagctc agaaaagagg aaagtgaggt agaaaattgt
37681 ctggccggta ttgtaatcgt tcgtggcatc tccgaggtcc ccaagcatgc tgtccgagtt
37741 tgcttgaaca atattgcctc tgtcaagctg aagtgcgaaa aacgtaacac atgcgaaggt
37801 acaaactcgg atgtcaatct attgcgctta gctatTTTTT aacagaaagg cttggatgaa
37861 tagtaattac cttcctaate agttgttttt cttcttcttc cgtccattca aagtatggat
37921 cccagcgatg cgcgctctca tattcgggaa taggcttgta aaacctgggg tcttccgtcg
37981 tatcaaatgc gcagtcagtc acctttgcga ctctgtgttc cgagtcgctc agtgaggatg
38041 atgatgccga tgatgtcgag acgtcagggg tcttcttggt gtccctcgctc gcggagaccg
38101 tgtccgcggg gacttccttg acgcgctctg tggtcacctc gcctgtatag caaggatgatg
38161 ggaggcagaa taacagagac ggaccggcct cgggtggttg agctcaagaa tcaacaagcg
38221 actcagggga catcgggcct ttttgacgtg tccagcaccg gtagacgact gtcgagactc
38281 atcacgcaat tttctgttaa agtgtcagca cgatggacta gttgttagca cgagagctgt
38341 gggtagctat acgtagccaa ttgtcgtatg agtgggtaaa gttcaggcaa cctagaacgt
38401 tgagcgcacg gatgtgcgac aagaaatggt tggcggcatg ggggccaacg tcaactacgg
38461 ccgagaaaact gcttaagggc tgatatgcct gattcaatgg ccgcataatc tgatatggct
38521 ttactttccg aggtgcttgt tagtgctaga gacttgaccg atcttgagca aaagtcccg
38581 caggggatgt aatcacgcaa ggggagcata actgacgtta tgaggaacaa gtaaggtaa
38641 cggacaagtg ccgggactga gattagcggc tgctccagat ctccgcagct gatgcaaagg
38701 gtccaataat agctggcagt ttgtatcagg ttccgcagac atgcgatgtc accatcgcat
38761 tggcttatga tgatatgctt ggggcttgac atcgcaacat tgtcttttct gctgattggc
38821 aaagatacgc ggcttatagt gggggaatct gatacgaata gtgttttcat caagtttttg
38881 aagatcagga tgagcatttt ggagggtgga gatgcggcgg acatttgacg attcatcgcg
38941 acgagataac agtaccagta ccacctcaca atctactaca cagcagagtc agtaaagaat
39001 attacacaga taatttcgac ccagttcaaa acacattaag actttgtggg ataaagcctt
39061 aggcggggga tgaggggtgga ttgtaagtcg ccttgccggc aagcttaatt ggggcgagtt
39121 gcgcgaagca acagcactat cagccatgc ttggcccgtc tggtagcag caaccatttt
39181 ctctccctc tttcccatgc acttcattga ctgcattaaa tgcagcaacg acctcgacgc
39241 gaagagtgcg agagcccatg ttggacttaa gtggaattta atgatttata tccgcagatt
39301 atccaaaaat agtcaagggc atctttccag ccattcacaa actgtgctgg gctgcagcca
39361 tggctcgggt cgaagtgaac cgataattga taaccatagc actgacgtct taactttctt
39421 gtattatcat tgcaccattg acatatcgga tcggcgagtc atgctggctt acactacttt
39481 cccgtggctc gcgtaactac ttgtccgagt tccaagcata tccgactgca ggcacctcct
39541 tgatacctga cccccgaaa atggcgtcaa tccatacccc gcaggaattg cgtgacaaaa
39601 gacggatatt cggcatgggt agcagagtat atatgggtgc agggttctgt gcaggatgca
39661 ttctggaaat caccacaaga acagcacagc tgatcatagtt gcacctgcag acatctaaca
39721 atatggcccc atctctgggt gaggacactc ctctgtcgct ccgggcggca cctaagtcca
39781 cgcttaagac cgatgctggt tttaacaagg agaacgtgat tggctatggg gaggcctaca
39841 agcatgagaa cgagctcaag ggcactgaga agcagccgcc agcttcgttt cccaactatc

39901 ttctgtgtctg ggacaatgaa acagaaaggt agcttacact actcatatat ttatctggct
39961 tcaactactga cagcttagat accctcccct ccagcccttc gaacactatg accacggaaa
40021 ggatgccgat ccgaccttcc ctgacctcct acccaaggac aaagtcgaag tagatgacat
40081 tacacctacc attggcaccg aggtccgcgg tgtccagctg agccagctga gcaaggaagg
40141 aaaagaccag ctggctttgt acgtagctca gcggaagggt gtcggtacga tcccccatct
40201 gtaagacgag atgacctttc cattaacaaa ccctaccagc attccgcgac caagacttgc
40261 cacatctgcc tatcgagaaa gccctagaat tcggcgggcta cttcggcaga caccacatcc
40321 accagacctc tggagcaccc aagggcttcc cgagatcca cctcgtgcac cgagggtgccg
40381 atgacaggag tggagctgaa ttctgttccc aacgcaccaa caccatcacc tggcactcgg
40441 acgtgacctt cgaagagcag cccctggca cgaccttcc gtacattctg gacggaccat
40501 caagcgggtg tgacactctc ttgcgcgaca tggcacaggc atacagacgt ctatcgctg
40561 agttccgtaa gagactgcac gggctgaagg cagttcactc tggcattgag cagggtcaaca
40621 acagtctgaa caaggggggt attgctcgac gggatccaat tacttctgag catccgattg
40681 tcaggactca cccagtaagt accatacact atacatcctt cttagctcga agctaattac
40741 attatctagg taaccggaga gaaagccctc ttctggaacc cccaatgtat gtaccacat
40801 cacatcacc ttatactcta catgaacact aacagaaaaca gtcactcgat acatcgctcg
40861 ctacaagaaa gaagagtctg actaccttct caagttcctc tacgaccaca ttgctctctc
40921 ccaggatctc caaactcggg tcagatggcg tcctggaact gttgtcgttt gggacgtaag
40981 tctagtctcc ctacatactc gaccaaggat agaaaccta cgaatgcaga accgcgttgt
41041 taccattctc gccctcttgc actgggaaga tggtcagaga cgccaccatg cgagaatcac
41101 cccgcaggcg gaagcgccgt atgagacgcc gtacgagggt tagctgcata cgctgacaat
41161 ctacgcta at gaagtgcacg atggcagcgg gactggccta gatatgattt aaactacaga
41221 gcatgaatat actaggagat cattctcaac tgggtgtgcag ttgttgtcca ggacattccc
41281 caatagtaat aatctcctcg attaagcaca tagcaacacc acactcatcg cccctcccct
41341 ttcagtgatc tcaccaccca gcataacaat cttatcgaga agcgactcat aataaccgcc
41401 tactaccag agaaagcagc agatttggtg gacgcagata cagcatcccc aggcggatga
41461 atccta atag ccggtgccat acaagtcccc gtccccgtcc ccgttccctc acccgacccc
41521 atcacatcgt ccatatgaaa cgcctctgac tcgggatgcg tttcaggatc gagaatcctt
41581 tcgggatgat tgtctgcagt agacagggtta tttttgttgt gagtagtagc gggagtagta
41641 gttgtaggat tagttgaatg agatggagga gggacatgtt cgagggtgtt ccatgtaacg
41701 gtgttgtttt cgctggagct ggtggagaat gagcggggag ggtggtgtga ttgggattgg
41761 gatgttgttg gtggcattgt tgatgacggt ggtggatgaa tcagtaacta gtgggtttga
41821 taaagggcaa gttgtagtac tggtagagag agagtgggtc gatagcggta ggtttagagt
41881 gagagggaga gggattgagt atctatgaac tcaggattcg gtcagtataa catacatctc
41941 caacaactaa tagagttatc ttacctgcaa cgaagatatc actagggtag agtgagttca
42001 atttacttca actcttctc ccgtagaaaa aagccaatgt actaatctct cctctctctc
42061 ctccctctct cctccctct ttctcgccat tcaactactc catccgaaac aggaaggaac
42121 aaacaccccc ccaactaccat acttcatcca ttcatataga tagatagata tactattgtt
42181 tcctctcaaa cctccacaac aaccaaccc tccactccc aaacaccctc gccgaagccg
42241 gcaacatcgc attcctctcc tccaacaccc ccaccaattc cctatacgtg tgctcccccg
42301 gaagcctcac ctccctcaat gccaacgact ggtttccttt cccacctcc ggatgcaaga
42361 tctcctcgac attatcaaca tgctggtaaa agaccttcat tgcgagttgt gcagcgggtg
42421 tgtgggatgc gctggctgtg gagtagcgga gggaggggtt gaagggccat atctattcat
42481 ctatttttagt atatagatat atatatatat gttgtaggaa gggacggata gggggagttt

42541 gggggggggg aggggggaag cttacgagta gtccatcttt agctccacag tggatgacga
42601 atctgcggtg gctttcacgt tcaatgagtt cgaggagttg tgcggtgagg attgtttccg
42661 gggggtggga ttcccaggtc tcctctgggg ggttggtgtc ttcgaggat gaggtggtgt
42721 tggcggacag actggttttg aagagtcggt agccgttggc gatggggtcg gggaggccga
42781 gttgggtggt gcaggttgtg caggttagga ttggtggttc tgtttctgtt tctttgtttg
42841 ggtctggttg tgttgatgtt gagtctttgg gagggacc

//