

Problemática a resolver:

En los últimos años de la guerra fría, miembros de la Agencia Nacional de Defensa (CONTROL) detectaron y destruyeron un laboratorio secreto de KAOS que contaba con equipamiento apto para tareas vinculadas con biología molecular y bioprocesos. La infraestructura y equipamiento del laboratorio sugería que en el mismo se trabajaba en el desarrollo y producción de nuevas herramientas para ser empleadas en una potencial guerra biológica a nivel global.

Antes de la destrucción final del laboratorio, las fuerzas de tareas recuperaron muestras biológicas desconocidas provenientes de uno de los fermentadores. El análisis inicial basado en microscopía electrónica sugirió que el material correspondía a un sistema viral desconocido hasta el momento. Por otra parte, los primeros estudios bioquímicos mostraron que el ente biológico estaba formado mayoritariamente por proteínas y ácidos nucleicos, y en menor proporción por lípidos. En base a estos resultados se realizó una purificación exhaustiva de ADN, el cual fue enviado al centro de secuenciación masiva de CONTROL para obtener la información nucleotídica. El ensamblado genómico final tiene 16117 nucleótidos que fueron depositados provisionalmente en la base de datos de CONTROL bajo el registro XX894275.

Ustedes trabajan en el centro de análisis bioinformático de CONTROL y tienen a su cargo la caracterización *in silico* de nuevas secuencias genómicas, por lo cual deben hacerse cargo de dilucidar la identidad del nuevo genoma secuenciado.

La información obtenida será publicada en el Top Secret CONTROL Journal of Biotechnology del mes próximo y será distribuida a todos los miembros de la agencia para su conocimiento. Por este motivo se recomienda que los resultados del estudio bioinformático sean presentados de la forma más concreta y clara posibles.

Algunas de las preguntas que usted debe contestar son:

- 1- Analice la secuencia nucleotídica y determine la existencia de ORFs.
- 2- Seleccione uno de los ORFs encontrados, tradúzcalo y obtenga la secuencia aminoacídica codificada en el mismo.
- 3- Utilice la secuencia aminoacídica obtenida en el paso anterior para caracterizarla contestando las siguientes preguntas:
 - a- ¿Con que secuencias de la base de datos tiene mayor similitud?, identifique proteínas a las que se parece y cuál sería el posible organismo de origen.
 - b- La proteína analizada, ¿presenta algún patrón o motivo característico?
 - c- ¿Cuál es la estructura secundaria de la misma?
 - d- ¿Cuál podría ser la función de la proteína en el organismo incógnita?
- 4- Teniendo en cuenta los resultados obtenidos proponga una hipótesis sobre la identidad del nuevo genoma secuenciado. Finalmente, mencione los análisis bioinformáticos que llevaría a cabo para dilucidarla.

El registro provisional es:

LOCUS XX894275 16117 bp DNA CIRCULAR 03-JUN-2015
DEFINITION Bioterrorism Agent 1, complete genome

ACCESSION XX894275
 KEYWORDS .
 SOURCE Unknown virus
 ORGANISM Unknown virus
 Unclassified; dsDNA
 REFERENCE 1 (bases 1 to 16117)
 AUTHORS CONTROL (National Defense Agency)
 TITLE Complete genome sequence of an unknown viral system recovered
 from a KAOS secret laboratory involved in the development of tools for
 the biological war.
 JOURNAL Unpublished
 COMMENT PROVISIONAL REFSEQ: This record has not yet been subject to
 final review
 COMPLETENESS: full length
 FEATURES Location/Qualifiers
 source 1..16117
 /organism="Unknown virus"
 /mol_type="genomic DNA"
 /isolate="Bioterrorism Agent 1"
 BASE COUNT 4740 a 3335 c 3297 g 4745 t
 ORIGIN
 1 accaaacgtg ttttatcggc aggcgcagaa gcggcggcca ttttaaataa ataaacgata
 61 acgccgttgg tgggtgtggg catgtagaat gtaatgttgt tatcttggtc gccatcgggt
 121 tgggtataaat aaatggtata ccgtcgacga cgttcccgat ccgcaaatgg cacgtatact
 181 cgccgacgca ggagttctgg atacagacgg cgaccgggca ggccgcgcac ctatcgccgc
 241 agccgcagtc gttctgccac acgccgtacg ggatacagaa gacgtcacta ttagacataa
 301 gttttacaat tatacatttt ttataagttt tacaattata catttttttt actctatttt
 361 ttttctttta tttaaatcct attttgggcg agtggcggcg gttgagatct tgtcatcgct
 421 ttaactcgct ttacgagtag aattctacgt gtaaaacaca atcaagagat gatgtcattt
 481 gtttttcaaa actgaattta agaaatgatg tcatttggtt ttcaaaactg aactggcttt
 541 acgagtataa ttctacttgt aaaacataat caagggatga tgtcatttgt ttttcaaaac
 601 tgaactcgct ttacgagtag aattctacgt gtaaaacaca atcaagagat gatgtcattt
 661 gtttttcaaa actgaattta agaaatgatg tcatttggtt ttcaaaactg aactggcttt
 721 acgagtataa ttctacttgt aaaacataat caagggatga tgtcatttgt ttttcaaaac
 781 tgaactggct ttacgagtag aattctacgt gtaaaacaca atcaaaagat gatgtcattt
 841 gtttttcaaa actgaatgat gtcatttggt tttcaaaact gaactggctt tacgagtaga
 901 attctacgtg taaaacacaa tcaaaaaatg atgtcatttg tagaatgatg tcatttggtt
 961 ttcaaaatta aactcgcttt acgagtagaa ttctacgtgt aacgcatgat caaaaaatga
 1021 tgtcatcttt tacacatgat tataaacgtg tttatgtatg attcatttgt ttttcaaaac
 1081 tgaactcgct ttacgagtaa aattctactt gtaacgcaag atcggtggtg gatgtcattt
 1141 atttgtgcaa agctgatgtc attttttgca cacgattata aacacaatca aataatgact
 1201 catttgtttt tcaaaactga actcgcttta cgagtaaaat tctacttgta acgcgagatc
 1261 ggtggatgct gtcataatgg cccaaatttt tatataaagg ggccactttg ctatggggaa
 1321 ttaaaattcg tccgagctcg ataggtcgct cctgtagtgt aacacgaaca ggttgatcga
 1381 attggccgac gcgggaacgc tgaatgtcaa tattttatct aatctgttgc gcggataata
 1441 taaattgata ttcatTTTTT ttgtttcagt tgatcgtcgg cggaacgcc gagtggcagg
 1501 tggactcgct tatgaaaatc cacacttaca atcagttgaa agaggcgttt cgcgattact
 1561 gtccacacac gttcgttatc tcgtgcgaca acgcgtttcg cgttatgcgc gtttactacg
 1621 acgaaaacaa gggcaagctg ttggcgtttt gcaacattaa catgaacaat cgcgtgttac
 1681 aattttatTT tgtgataaaa atggatttgt attcgtacaa aaaatgctac aacaatcaca
 1741 tattttcaac gtgtcgcaac cgctgttcca actataacac gtttggttgc cggggcgtaa
 1801 aaggcgtgca catggacaag attaatgtga tcaagttaa acgcaacggc agttcgtacg
 1861 aagaaaaatc taatgctttg gacaagtttt tgagtaacgc caacagggtt cacatgcaaa
 1921 cgcccgatgat tgaaggcacg tacatgcgtt tccgccgtgc gcaacggtgc cgcaacaact
 1981 gcgtcagcga cgacacgcgt ccgtttacat ttgaaaagtT tgatgcagat ttcgaggtcg
 2041 tggacccggc cacgttgagt aataacattg cgcccgatgat ggctgtctac gacattgaaa
 2101 cacattcgga cggacacaat agtccaaac cggacgtgga cgttatcatg tgcacgggt
 2161 tatcggtata caaggacgat cactttgaaa aaatttgttt tgtgtaccac aaagaacccg
 2221 tggaaattcc gcacacaaac gacgacacgc acgtagtcgt gtttttcaac gagattgaca

2281	tgataatgtc	atTTTTTgat	TTTTTaaagt	ttgttaaccc	cgacgtgatt	ctcgactaca
2341	acggcgacgt	gtttgatttg	ccgtacatac	gtgCGcgctt	aaagggtaac	aaaccacagt
2401	tgcgCCgtta	tgatttgccc	cctttgcaag	ccaatacaaa	gttgtttatc	accaagattg
2461	gtaataaaaac	ggatacgtat	tatttttaatt	attacattca	cattgatttg	tacaaatatt
2521	ttggagtaga	cgccaacaat	cgCGacgtgg	aaaatTTTca	actaaacacg	ctcagtaaat
2581	attatTTTggg	cgatgccaag	gtcGattttg	attggcaaac	gatggTgcgc	atgtacaaca
2641	acaaacaatt	gggcaccatt	attgaataca	acgtgcagga	ttgcttgttg	ccgattaaat
2701	tgTTTTTaaa	attaaaactg	aacgatttca	tgtactcgca	atgtTTaatg	tatcgCctat
2761	gcaccgacga	TTTTatctgc	aacatttcgc	acttgattag	tacgacgttt	TTTcaccttg
2821	cTTtgactaa	cacgcgtaca	aactttgtcta	cagggcTTga	agaacaagat	gTTTactTTT
2881	ttgacaaaaa	tgattttgga	aacatgtctg	gcaacgcggg	catgtcgcgt	ttgcaacgca
2941	agcttacgcc	tcttaaagac	gtgcccGCCa	acgcgatttg	TTTgggcgcc	atcgccgaaa
3001	cggTcaaaata	tgaaggcggc	aaagtgttgc	agccccgtgc	cggcgTgtac	gagtacgcgt
3061	TTTcgTtaga	TTTcaattcg	ttgtatctca	ccattatgat	tgacatttgc	gcttgtctca
3121	caaacccta	TTTgtgcgaa	gacggcaacg	tgtaccttaa	ccaagacaag	gaagcgataa
3181	acgtcaagtt	gttGctcaaa	ctgcttaa	agcgcagtga	tctaaagaaa	tgccgtgaca
3241	atcaaacCGa	atccgaattt	ttgtatgact	tgtacgacca	aatgcaaaac	ctaagcaagc
3301	gtactgCcaa	cagtattttat	ggttattacg	gtattTTTTg	caaagcgttg	gccaattaca
3361	ttacaagagt	gggCCgcgaa	aaactaaccg	ccgcaattag	tataattgaa	ggtctcagtg
3421	aagacCCCGa	attgttacgt	gaatttggtt	tgacgacgct	caagtTTaat	gtgctTTtacg
3481	gagacacaga	ttcaacgTTT	gtgttgCCca	cgTTtaatcg	cgatgaaatt	cccgaagagc
3541	aacgcattggc	cgCGcttacg	cgcatTTgtg	cgTccgtgga	ggcgcgcggt	aacgcacaat
3601	tcaccaacCG	ttacaaaaatg	gcgtttgaaa	acttgatgag	cgTacttGta	ttgttaaaaa
3661	agaaaaagta	ttgttacatt	aacaacaacg	gcaaaattgt	gtTTaaagga	tggttggtca
3721	aaaaggacat	gccagtgttt	atgcgtgtag	cgTTtagatc	agccatggag	cacgtgttgc
3781	gtCGccagga	cctgcCGgga	tgTTTaaaca	gTTtgcaaaa	cgacatgctt	atgtattTTa
3841	acgcgtTTtg	cggcgacaag	cctTTa	attacagTTT	tagcatgacg	tacaatgacg
3901	gCGccggcaa	agcaaacaa	cggtgcCGc	aggaaccgca	agaagcact	atgccgatta
3961	acgcgcgcgt	agttaccgtc	gcgcgcatt	gtcgcgaaat	cctggTTagc	aaaggcacgg
4021	atTTTgtgcc	aggcaacggc	gaccgaattc	catatgtgct	gcttgacatt	gagggtaaaa
4081	ccactgataa	aagttaccct	ttgcgtTTgt	ttgatgtaac	cacgaaacgc	attagctggc
4141	TTaaacacat	gaccattTTa	aatacgtTTa	tgaacgagct	gttggaatt	TTTggtgacc
4201	agcacaaaaga	tgtgattgcc	gaatgttacg	aggTTattat	agcaacatac	atgcaaaacc
4261	aagtgtacga	caaaaaacgc	gctTTatttg	ttccaattgc	ctacaaacat	aaagtgcaaa
4321	acaaacaagt	tacaaaaaac	tacaaaggt	ctaaatgcaa	agcGCCaagc	gacacggagg
4381	acagcgaaga	agaagaagaa	gatgaggacg	ttccaccCGa	agtcaaatgt	gcaaacacaa
4441	cgTTTaaatt	ttgtTTgtat	aaagtTaaat	aaaacaaatt	tgTTTataat	taaatggTTT
4501	ttatTTTccg	tgCGcgatta	gctgcGcaat	caaaatcagt	ggacgacaaa	tacggatcct
4561	gagacaacat	ggccacaatt	aggtatacaa	gaaaaaccaa	aagtattgca	tattgaggac
4621	tcataTTTtat	tatttGttac	acttagcact	taaaactagg	tacattTaaa	TTaaaatcat
4681	TTTTattaaa	tgacatatct	aaattTaca	aaactTTatc	gtagggTcta	tagtgtTTTt
4741	caaaagctTT	acgaaattca	gcacacaaag	ttgtTTcgta	aaatTTTtga	taattTctt
4801	tgcgtaacaa	tgcatgcaaa	aacttatcca	aaaatggaac	agccaattcg	atggctTTat
4861	ctactTTtagt	ttcgtcaatg	ggTTtgCGc	ccggtcCGca	TTTactTTc	aaaataatac
4921	cgatCGcttc	caatggacta	ttgtTcaaat	ccaaacattt	tagattgtgt	tcgtgtatcg
4981	aatccgattt	taagattTcc	ttgtagtaca	cgtaaccgtc	cctaggatta	cgTTTataca
5041	tgagaatgtg	cgataaaaaat	aaacgaaccg	gTTTTgtaag	atcttcgaaa	tacgctTTTt
5101	cctgtgggta	TTTcttGTTt	ttggcatgaa	agtatatcga	accattgaat	tgcatcgact
5161	ctaaaaattc	atgatccgta	tacactacac	agaatctgtt	gcgaacgccc	ctgtcgtaat
5221	cgctaattgtg	caatggTTTg	ttgttgacca	ccaacaattt	gtaattggct	tcgtattTTT
5281	gactaccctg	atatttgcgg	cagacactgt	tactTTTgct	cgaatcggcg	gtgctTTTga
5341	aaaaagaatc	gttacattct	ttgagttcgt	taatgacgta	caattgcgaa	atcaattTgt
5401	tggcctccat	ttcgtcagtt	tctTTTTTgg	acaaggtata	TTTgtcccg	tcgcgtTTTat
5461	gtactacaat	aatggattct	agcagatcga	aaaagctaga	TTTgcccga	ccgggttcgc
5521	cgttcaaaata	tataacaacat	TTTTcgtagt	ctgtcgggat	gcctaagcta	gctccaaaaat
5581	gcatcattaa	caatgaattt	TTTacattaa	aatttGtgaa	caatctaaaa	tacaaataac
5641	cacgtacgac	ttgtttcaca	aatagcgggtg	aatatgtTTT	gacatcgatg	cgcgacatta
5701	ttacacgcat	atagaaacga	gtcaaccatt	tggccaaatc	gtccgacggt	ctggctacaa
5761	TTaatTTgtc	ccaccacaca	ctatattTgc	gaagtatcac	tattgtattg	gcataattTT
5821	tgtaaaaatg	atcaaaaatac	tgTtcattac	tattaccact	aatataatca	ctattgccat
5881	cgtcattgca	attattTTTt	ggtgtctTTg	caaaatcgtc	atagtcataa	ttgtcagcgg

5941	cgggtgacgac	gacattgcta	tcgatggcag	cgactgtact	gtcttccatc	gttagaattt
6001	ttaacaggat	atttgcagac	gaaaattcat	atagtaacat	gtcaatatcg	atcgtgtcca
6061	atttggttga	caattgatcg	attagctgta	accgacgctc	gtacacgata	ggagcatatt
6121	tttcaattat	aacgctcgtt	gtttgagcgc	ccaatatattt	agtgtatggt	tcgggcgcat
6181	agtgcaaaaca	caaatacaat	tctacaagtt	gagcattgtc	acaaaacaac	tcaaagatta
6241	atgtcaattt	gaaagctttt	atgttgacat	tcatttgagc	gacgcacgaa	caagaagtct
6301	gtgtcgtacc	cgcttcttta	cattcgacgc	aacgtaaatt	cttaatgagg	tctgacattt
6361	tagtatcggt	caaataaata	ccgtatatga	ttaattcgtt	tggtgaagag	ttccagattt
6421	cgtgaaaata	ctggttcaat	ttggaatggt	ccgcgttttt	gcattgacga	caattgtcga
6481	acgaattaat	tatcgacatg	ttagttttga	tgacttttat	gtcgcgacac	actttggcca
6541	cgtgataagt	tttaaaaatt	tccatttcat	acttggcatt	gttcagcatg	taatcgatag
6601	tttccttggg	taaaaaattca	ttttcatcga	tatatctaaa	cggattgact	aagcaattgc
6661	caattatgaa	cggacaacta	ttgtgataat	gattgatgaa	cacattgaaa	acaccctggt
6721	cggcaaaaaca	caaatacttc	caattattaa	atttgattgt	ggacaatttt	acgcttggtc
6781	ctgtttcggg	tactttgaac	aattcatcgt	cttttttcac	gggaacatag	tgtttgccgt
6841	tgaacacgta	aaagctgcct	tgagactcga	gttttttcaa	aaaacccaaa	cacaacacat
6901	taggcggtaa	tttacaagtc	ataacccttt	cgtatgtgta	tgcccatcgt	tcggtgaaac
6961	tcatgtcgtc	attgagtga	tttaaatata	tacaatagtg	tatggcataa	taatatccga
7021	gtagcacgca	aggattttcc	attgaaaaaa	aatgtaaact	atcacaaaac	ttcttgtaca
7081	ctcttgtcga	caattgaaca	tagggatcac	atcgtgctct	cactttcacc	aatgattcgg
7141	catcgttctt	gtagagcgtc	aaacataaca	gttccaagta	gagtttgatg	tcggtctcac
7201	aaaactcgaa	tcgttgtcga	ctgaccattt	tccagacaac	aatgattaaa	taatcaaaat
7261	tgaaatagtt	actttcgtc	aaataacgaa	tcaaaacgtc	accgtccaca	ttttcttgat
7321	tttgagccc	ttcgatcatg	cttgttttga	ttttgtccaa	actgatatcg	atttcgtttt
7381	tgatgagttc	atagttttca	ctgctcgtg	ttatgttttt	gattatggtt	ggtgagaatt
7441	tttgactttt	cacagtttca	tagttactaa	aagtcttgtc	gtcaaacacg	cgcatactac
7501	gcaaatcgat	ctgaacgatg	tcgctaaact	ttggcgccgt	aacacattct	tgcataggga
7561	tatcgtcacg	tatatattca	aacagatcct	tgctcgaata	taccagtttg	ggtaggattt
7621	tgcatacacc	cttgctgccc	tccgacatgc	gtatcgtgaa	cagtgtatcg	ttggaatcgt
7681	taaaaatcga	atgtccgttg	acaaacagcg	tgtttcgatt	gcacgtcatg	cacaattcga
7741	cattcaaaaa	atgttccgga	taccaaacga	atagattaac	attgccgagg	ctgcgattgt
7801	gcggaacg	cacatactcg	ccgatgtcga	tgtcaaattt	gagtttcaaa	tacaaacgcc
7861	atccaaaata	agaaatctca	atattggggc	agtacacgta	gtctcccgcc	gattcagttt
7921	tggtggcata	ttcttccgcg	ttgctcatca	caaattcact	gaacgaaatc	gagtcacgca
7981	cagcttcgta	atatcgttgc	aatataaacg	gtctaatttt	gatggcgaaa	tagtttcctt
8041	gtacacacca	atcgtgactg	tctatacttt	ttacatgatc	tttcgaagac	gacgcagccg
8101	ctccaccact	atcggcatca	acaacaccga	atatttcggt	cacatgcacc	aaattattgt
8161	tggcaatcgc	gctattggtg	ttgcatgatt	tgtaattgca	tcgcgtttgt	ttagtagaaa
8221	ctgcgcacaa	aagtttttcc	agaatttgat	aggatttaat	taaaaacttt	tcctgctggg
8281	cactattttt	gagcactata	gtgtcaacac	accccaaatt	gacgacggtt	tcattgtccg
8341	gttgttgaat	acgattaaaa	atgttctcaa	atattgcgtc	aacactaatt	ggtgcggtgg
8401	ccataagcga	gtccacctgc	cactcggcgt	tcccgccgac	gatcaactga	aacaagaaaa
8461	atgaatatca	atttatatta	tccgcgcaac	agattagata	aaatattgac	catcagcgtt
8521	ccgcgctcgg	ccaattcgat	caacctgttc	gtgttcaact	acagggacga	cctatcgagc
8581	tcgaattcgc	tcgctatttg	gaacataatc	atatcgcctc	atataatttg	taattaaata
8641	tacaattaaa	aactatacaa	ttttttttat	tacaaataat	gatacaattt	ttattattac
8701	atttaatat	gtctactatt	acggttttcta	accatacagt	acaaaaataa	aattacaata
8761	aatacaatta	caaaagtagc	tacatgacca	aacatgaacg	aagtcaattt	ggcggccaat
8821	tcgccttcag	ccatggaagt	gatgtcgttc	agactgggtc	cgacgccgcc	aaacttgggtg
8881	ttctccatgg	tggttatgag	gttgcttttt	tgttgggcaa	taaacgacca	accgctggca
8941	tctttccaac	tgctcgtgata	ggttggtgtg	ccgatgggtc	gaatccaaaa	ttcgacgtcg
9001	tcgtcgattg	ctagttcctt	gtagttgcta	aaatctatgc	attgcgacga	gtccgtgttg
9061	gccacccaac	gcccttcttt	gtaaatgctg	ttggtgtagc	aattactggt	gtgtgccggc
9121	ggattgggtg	acggcatcag	caaaaacgtg	tcgtccgaca	aaaatgttga	agaaacagaa
9181	ttgttcatga	gattgccaat	caaacgctcg	tccaccttgg	ccacggaaac	tatcaggtcg
9241	tgcagcatat	tgtttatctt	gttgatatgc	gcatgcatca	gctcaatggt	cattttcagc
9301	aaatcgtttt	cgtacatcag	ctcctcttga	atatgcatca	ggtcgccttt	ggtggcagtg
9361	tctccttctg	tgtacttggc	tctaacgttg	tggcgccaag	tgggcgcccg	tttcttgact
9421	tggtgctcga	ctttgcgttt	aatgcatctg	ttaaacttgc	aattccacgt	gttttttagaa
9481	agatcatata	tatcattgtc	gatcaaacag	tgttcgcgtg	tcaccgactc	ggggttattt
9541	ttgtcatctt	taatgagcag	acacgcagct	tttacttggc	gcgtggtgaa	cgtagacttt

9601	tgttttgagaa	tcatactcac	gccgtctcga	tgaagcgag	tgtccacggt	cacgttaatg
9661	ggattgccct	cagcgtccaa	aatgtatacc	tggcactcgt	cggtgtcgtc	gtggcactca
9721	agcctgctgt	acatttttcga	agtagaaaaca	cgcacatcgcc	acgatttggt	gcacgtgtgg
9781	tacgcaaaagt	gattgttatt	ctgccgtttc	accaactcct	tgcctttgac	ccactggccg
9841	cggccttcgt	tgtcgcgaaa	acactcttcg	ctgtcactgc	cccaacggtc	gattaactct
9901	tcgcccacct	cgcactgctg	cctgatgccc	cacatgagca	aatcttcttt	gcccacagtc
9961	agcgttttca	tggattcttc	gacgcgtgtg	ttgggatcca	gcgagcctcc	gttgtacgca
10021	tacgcctggt	agtaccctt	gtagccaata	atcacgtttt	cgttgtagtc	cgtctccacg
10081	atggtgattt	ccacgtcctt	ttgcagcggt	tccttgggcg	gggtaatgtc	caagttttta
10141	attttgtagc	gacccgtttt	catttgcgcg	ttgcagtgtc	ccgccgcaa	ggcagaatgc
10201	gccgccgcca	aaagcacata	taaaacaata	gcgcctacca	tctcgcttgt	gtgtttctta
10261	tcgaagcctt	ggtatgactg	atttactagt	agcattgagg	catcttatat	accgcaccgc
10321	tatttgGCCA	cgtgacacaa	ggcacgttgt	taaattaata	attctatctt	tttgtcttaa
10381	ttgtaagatt	atttttatct	ggctgttata	aaaacgggat	taagcttaat	cccgttttta
10441	taacagccag	ataaaaaata	tcttacaatt	aagacaaaaa	gatagaatta	ttaatttaac
10501	aacgtgcctt	gtgtcacgtg	gccaaatagc	ggtcgggtat	ataagatgcc	tcaatgggac
10561	cagtgtataa	agatttagag	attgatgttt	gcaacacacc	accttcagaa	tgtatcaacg
10621	atttactgaa	atctgtagat	tcagaagaga	ttagggaata	ttgtaagaag	aagaaatgga
10681	ttatacctga	aattcctacc	aacatagaaa	gggctatgaa	tcaagccagc	atgattatta
10741	atactattct	gatgtttgtc	agtacattag	gtattgttta	tgtcatttat	aaattgtttg
10801	ctcaaaactca	ataaatgtaa	taataaaaaat	tgtatcatta	tttgtaataa	aaaaaattgt
10861	atagttttta	attgtatat	taattacaaa	atatatgagg	cgatatgatt	atgttccaaa
10921	tagcgagcga	ggtacctcgc	aaattttttt	gaaaagcctt	ttacaagttt	tactttgcag
10981	taggttaata	atttttttgt	catacaatgc	aatattttata	tatagcttat	ttttaagaa
11041	tgtcgtagtc	gtttgcaatg	tctattgtct	ttcaatgtaa	ggcacctgaa	agtgtttgga
11101	aatgtacatc	ttcagaaaatc	tccaaggaga	tccatcacat	ccaagtagtg	ctagagtggc
11161	cgtaacagtg	accaagtcat	catgattttc	gcaaatgatc	actaaagctg	aaaccatttt
11221	gaccactttg	gacaaaactt	ttgtggtgag	gaaatctttc	gccacttctt	gtagtctctg
11281	gacttttgtt	gagatttggg	cagtgaacct	gacacaaaaa	gctctaccca	aacctgtgat
11341	gtaatcactc	agccccatag	ctctgcttat	atagacctcc	caccgtacac	gcctaccgcc
11401	catttgcgtc	aatggggcgg	agttgttacg	acattttgga	aagtcccgtt	gattttgggtg
11461	ccaaaacaaa	ctcccattga	cgtcaatggg	gtggagactt	ggaaatcccc	gtgagtcaaa
11521	ccgctatcca	cgcccattga	tgtactgcc	aaaccgcatc	accatggtaa	tagcgatgac
11581	taatacgtag	atgtactgcc	aagtaggaaa	gtcccataag	gtcatgtact	gggcataatg
11641	ccaggcgggc	catttaccgt	cattgacgtc	aatagggggc	gtacttgga	tatgatacac
11701	ttgatgtact	gccaagtggg	cagtttaccg	taaatactcc	accattgac	gtcaatggaa
11761	agtccctatt	ggcgttacta	tgggaacata	cgtcattatt	gacgtcaatg	ggcgggggtc
11821	gttgggCGGT	cagccaggcg	ggccatttac	cgtaagtatt	gtaacgatcg	atcgttacat
11881	aacttacggt	aaatggcccc	cctggctgac	cgcccaacga	ccccgcacca	ttgacgtcaa
11941	taatgacgta	tgttcccata	gtaacgcca	tagggacttt	ccattgacgt	caatgggtgg
12001	agtattttacg	gtaaaactgcc	cacttggcag	tacatcaagt	gtatcatatg	ccaagtacgc
12061	cccctattga	cgtcaatgac	ggtaaatggc	ccgcctggca	ttatgccag	tacatgacct
12121	tatgggactt	tcctacttgg	cagtacatct	acgtattagt	catcgctatt	accatggtga
12181	tgcggtttttg	gcagtacatc	aatgggCGGT	gatagcggtt	tgactcacgg	ggatttccaa
12241	gtctccaccc	cattgacgtc	aatgggagtt	tgttttggca	ccaaaatcaa	cgggactttc
12301	caaaatgtcg	taacaactcc	gccccattga	cgcaaatggg	cggtaggcgt	gtacggtggg
12361	aggtctatat	aagcagagct	atggcaaatg	atggatggtt	cagaaagttt	aatgatgcat
12421	gtaatgctgc	aaaggggattg	gaatggattg	ctaataagat	ttccaaactg	attgaatgga
12481	taaaaaacia	agtacttccc	caagccaaag	aaaaactaga	attttgtagt	aaactcaaac
12541	aacttgatat	actagagaga	caaataacca	ccatgcatat	ctcgaatcca	acacaggaaa
12601	aacgagagca	gttgttcaat	aacgtattgt	ggttggaaaca	aatgtcgcaa	aagtttgccc
12661	cattttatgc	cgttgaatca	aaaagaatca	gggaactcaa	gaacaaaatg	gtaaattata
12721	tgcaatttaa	aagtaaacia	agaactgaac	cagtgtgtgt	attaatccat	ggtacacccg
12781	gttctggtaa	atcattaaca	acatccattg	tgggacgtgc	aattgcagaa	cacttcaatt
12841	cagcagtata	ttcacttcca	ccagatccca	agcactttga	tggttatcag	caacagggaag
12901	ttgtgattat	ggatgatctg	aaccaaatac	cagatggaca	ggatataagc	atgttttgtc
12961	aaatgggtttc	ttcagtggat	ttcttgcttc	caatggctag	tttagataac	aagggcatgt
13021	tattcaccag	taattttgtt	ctagcctcca	caaattctaa	cacactaagc	cccccaacia
13081	tcttgaatcc	tgaagcttta	gtcaggagat	ttggttttga	cctagatata	tgtttgcata
13141	ctacctacac	aaagaatgga	aaactcaatg	caggcatgtc	aaccaagaca	tgcaaagatt
13201	gccatcaacc	atctaatttc	aagaaatgtt	gccccctagt	ctgtggaaaa	gctattagct

13261	tggtagacag	aactaccaac	gttaggtata	gtgtggatca	actgggtcacg	gctattataa
13321	gtgatttcaa	gagcaaaatg	caaattacag	attccctaga	aacactgttt	caataaacat
13381	tgcaaacgac	tacgacattc	tttaaaaaata	agctatatata	aaatattgca	ttgtatgaca
13441	aaaaaattat	taacctactg	caaagtaaaa	cttgtaaaaag	gcttttcaaa	aaaatttgcg
13501	atctagagga	ggaggaggag	gtggtgttgg	cgcagcaacg	tcgccaaactc	ccattgttat
13561	ttctatgcaa	aaccccacac	caacaacggc	ccctcgataa	taaaagacaa	aaataatata
13621	aaatatatgt	ataattaatt	aaattcaaaa	tatatgtata	actactgctt	ctgtgccatg
13681	ttaataccca	agatcgagca	aatctttgat	ttaagccagc	gccaaagggga	gcctgagcac
13741	ccaagcaagg	atagggttgc	tgtcacagta	atgatgtcgc	tgcggtttct	aacaacaatc
13801	accagggcgg	acaaaagtgt	tataatgttc	ccgactaccc	tctccacaac	gggtgcggaa
13861	ccaacccatca	tgtccctgac	ctgccccaca	tagttggata	tctcatcagt	aaatccagca
13921	ccaaaagcgt	tacctaaagct	ctggacgtta	tcggtaaggc	ccatagtcac	ttggttgttc
13981	acgatcttgt	cgccgccagt	gtcaacttgc	aactgaaaca	atatccaaca	tgaacgtcaa
14041	tttatactgc	cctaattggcg	aacacgataa	caatatttct	tttattatgc	cctctaaaac
14101	caacgcggtt	atcgtttatt	tcctgaggaa	ataaacgata	accgcgttgg	ttttagaggg
14161	cataataaaa	gaaatattgt	tatcgtgttc	gccattaggg	cagtataaat	tgacgttcat
14221	gttggtatatt	gtttcagttg	caagttgaca	ctggcggcga	caagatcgtg	aacaaccaag
14281	tgactatggc	tgacggctgg	gttaagaagt	ttactgaggc	agttaacgcc	tttaaggggt
14341	tggactggat	cgctgccaag	ttgtccaagt	tccttgattg	gatcaaatcc	aagatcatcc
14401	cggagctcag	agagagagcg	gagtttggtta	agaatccttag	gcagcttcct	cttctcgagg
14461	cccagatcac	cactttggag	cactccaacc	ctaatacagga	gacccaagaa	cagcttttct
14521	cgaacgtcca	atacctggca	caccactgta	ggaagaacgc	tcgcctctat	gcagcggaag
14581	cccggagagt	tttcgcacta	gagaaacgcg	tccttgagac	aatgcagttc	aagaccaaga
14641	atcgaattga	acctgtttgc	tgtttgatcc	atggaacgcc	gggcactggg	aatcacttg
14701	ccacgactat	tattggcagg	aagatcgctg	agtacgagaa	tagcggggtc	tatagtttgc
14761	cacctgaccc	agaccacttt	gatggctacg	ccgagcaggc	agttgtaatt	atggatgatc
14821	tgcatacagaa	cccagacggc	aaagatatga	gcttgttttg	tcagatgggt	tccaccacc
14881	ccttcgtggt	tccgatggct	gctctcgagg	ataaaggtag	acttttcacc	tccaaatag
14941	tgctggcctc	aacaaatgcc	aaccatatcc	atccagttac	agtcgccgat	ggaaaggccc
15001	ttcagcgccg	cttccacttc	gacacggaca	ttgaattgat	ggatggcttc	gtgaaaaacg
15061	ggaaactaga	tattcagagg	gcaaccgagg	catgtgaaga	ctgttctccg	atcaactttc
15121	agaagtgcatt	gcctctcatt	tgtagtaagg	ctctccagct	ccgtagcaag	aaggggtgatg
15181	gcatgagata	cagcattgat	accatgatca	cagagatgcg	cagggagtca	gcccgcgct
15241	ataatattgg	gaatgttata	gaggcactct	tccaataatt	atacatatat	tttgaattta
15301	attaattata	catatatatt	atattatttt	tgtcttttat	tatcgagggg	ccgttggttg
15361	tgtgggggtt	tgcatagaaa	taacaatggg	agttggcgac	gttgctgcgc	caacaccacc
15421	tcctcctcct	ccagtactgg	aaactattaa	taaaactggt	aaagttatga	atttgtttgt
15481	agacaattttg	gttttgtagc	aatgcaatgt	ttgtaaagaa	atatctacgg	atgaaagatt
15541	tttaaagcca	aaagaatgtt	gcgaatacgc	tatatgcaac	gcgtgctgcg	ttaacatgtg
15601	gaagacggcc	accacgcacg	caaaatgtcc	agcgtgcagg	acatcgtata	aataagcacg
15661	caacgcaaaa	tgggaccacc	tgagttcaaa	cccctgcgca	ttgatgtgag	tgaggaaacc
15721	cctgcaccac	ccgcgatcgc	agatcttcta	gcaagtattg	actccgagga	ggtgcgtgaa
15781	tactgccgca	taaaggggtg	gatagtacag	gagaagatca	cgaaggaaaag	cctcgagcgc
15841	aacgttaacc	gtgcgctcat	cgtgcttcaa	tccgccaccc	tcatagcaac	catctgtggc
15901	gtgatctatg	tggtgtacaa	actgtttgca	ggactccaat	aatatgattt	acactgatcc
15961	cactactggc	gctacgacta	gcacagacgt	cgtccgtcca	caaactattt	aaacaggcta
16021	actccaaaca	tgttcttgac	catcttggtt	gtagtagtaa	ttattgcttt	aataattata
16081	tttgttcaat	ctagcagtaa	tggaacacgc	tcggggg		

//