

UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN

Escuela de Posgrado

DOCTORADO EN CIENCIAS AMBIENTALES

**BIOPROSPECCIÓN DE BACTERIAS TERMÓFILAS EN
LOS GÉISERES DE CALIENTES, CANDARAVE
(TACNA-PERÚ) PARA PRODUCCIÓN
BIOTECNOLÓGICA DE ENZIMAS
TERMOESTABLES**

TESIS

PRESENTADA POR:

Mgr. ROBERTO CASTELLANOS CABRERA

Para optar el Grado Académico de:

DOCTOR EN CIENCIAS AMBIENTALES

TACNA – PERÚ

2021

UNIVERSIDAD NACIONAL JORGE BASADRE GROHMANN

ESCUELA DE POSGRADO

DOCTORADO EN CIENCIAS AMBIENTALES


**BIOPROSPECCIÓN DE BACTERIAS TERMÓFILAS EN LOS GÉISERES DE
CALIENTES, CANDARAVE (TACNA-PERÚ) PARA PRODUCCIÓN
BIOTECNOLÓGICA DE ENZIMAS TERMOESTABLES**

Tesis sustentada y aprobada el 30 de agosto del 2021; estando el jurado calificador integrado por:

PRESIDENTE:


.....
Dr. EDWIN MARTIN PINO VARGAS


SECRETARIO:


.....
Dr. DALADIER MIGUEL CASTILLO COTRINA

MIEMBRO:


.....
Dr. MARIO RODRIGO ESPARZA MANTILLA

ASESOR:


.....
Dr. MARIO RODRIGO ESPARZA MANTILLA

AGRADECIMIENTOS

Agradezco a mi asesor Dr. Mario Rodrigo Esparza Mantilla, por haber despertado en mi persona, el interés en la investigación bacteriana.

A la Universidad Nacional Jorge Basadre Grohmann, por permitirme realizar esta tesis con el apoyo del Canon Minero, sobre Canon y Regalías Mineras.

Al Laboratorio de Biotecnología Enzimática de la Unidad de Investigación de la Facultad de Ciencias.

A todas las autoridades de nuestra Universidad, que de alguna manera u otra me han apoyado para la realización de la presente tesis doctoral.

DEDICATORIA

A mis queridos padres, Orestes y María Elena.

A mi querida y amada esposa Manuela.

A mis hijos Elberth, Helen, Holbert, Deanny y
Bethsy.

A mis nietos y bisnietos, con amor y ternura.

CONTENIDO

AGRADECIMIENTOS	iii
DEDICATORIA	iv
CONTENIDO	v
RESUMEN	x
ABSTRACT	xi
RESUMO	xii
INTRODUCCIÓN	1
CAPÍTULO I: PLANTEAMIENTO DEL PROBLEMA	6
1.1. Descripción del problema	6
1.1.1. Antecedentes del problema	7
1.1.2. Problema de la investigación	7
1.2. Formulación del problema	8
1.3. Justificación e importancia	8
1.4. Alcances y limitaciones	9
1.5. Objetivos	9
1.5.1. Objetivo general	9
1.5.2. Objetivos específicos	10
1.6. Hipótesis	10
CAPÍTULO II: MARCO TEÓRICO	12
2.1. Antecedentes del estudio	12
2.2. Bases teóricas	17
2.2.1 Bioprospección	17
2.2.2 Géiseres de Calientes-Candarave	19
2.2.3 Bacterias termófilas	20

2.2.4 Enzimas termoestables	22
2.3. Definición de términos	29
CAPITULO III: MARCO FILOSÓFICO	32
CAPÍTULO IV: MARCO METODOLÓGICO	34
4.1. Tipo y diseño de la investigación	34
4.2. Población y muestra	35
4.3. Operacionalización de variables	36
4.4. Técnicas e instrumentos de recolección de datos	37
4.4.1 Ubicación geográfica de cada géiser	37
4.4.2 Determinación de temperaturas y pH del agua termal	39
4.4.3 Cultivo y crecimiento de bacterias termófilas	40
4.4.4 Cultivo axénico de cepas bacterianas aisladas	42
4.4.5 Extracción de ADN genómico de cepas axénicas	43
4.4.6 Obtención de genes ADNr 16S	43
4.4.7 Envío de productos pcr para secuenciación	44
4.5. Procesamiento y análisis de datos.....	44
CAPÍTULO V: RESULTADOS	50
5.1. Géiseres de Calientes – Candarave	50
5.2. Extracción y purificación del ADN genómico	55
5.3. Aislamiento y purificación del gen ADNr 16S	56
5.4. Identificación molecular de bacterias termófilas cultivables	58
5.5. Clasificación taxonómica de especies bacterianas	63
5.6. Construcción del árbol filogenético	67
5.7. Bioprospección de enzimas de interés biotecnológico	72
5.7.1. Bioprospección de enzimas termoestables lipolíticas	74

5.7.2. Bioprospección de enzimas termoestables proteolíticas	78
CAPITULO VI: DISCUSIÓN	102
CONCLUSIONES	110
RECOMENDACIONES	111
REFERENCIAS BIBLIOGRÁFICAS	112
ANEXOS	120
Anexo 1. Secuencias ADNr 16S Forward enviadas por Macrogen Inc.....	120
Anexo 2. Secuencias ADNr 16S Reverse enviadas por Macrogen Inc. ...	136
Anexo 3. Secuencias Consenso de bacterias termófilas cultivables.	151

ÍNDICE DE TABLAS

Tabla 1. Operacionalización de variables del estudio.....	37
Tabla 2. Ubicación geográfica de los géiseres de Calientes-Candarave.....	53
Tabla 3. Control de calidad y concentración del ADN genómico.	55
Tabla 4a Identificación molecular de bacterias usando el gen ADNr 16S...	59
Tabla 4b Identificación molecular de bacterias usando el gen ADNr 16S ...	60
Tabla 5. Clasificación taxonómica de especies de bacterias termófilas	69
Tabla 6. Especies de bacterias termófilas obtenidas.	70
Tabla 7. Enzimas hidrolíticas (Lip, Glu, Pro) por especies bacterianas	80
Tabla 8. Enzimas lipolíticas en especies de bacterias termófilas.....	84
Tabla 9. Enzimas glucolíticas en especies de bacterias termófilas	91
Tabla 10. Enzimas proteolíticas en especies de bacterias termófilas.....	96

ÍNDICE DE FIGURAS

Figura 1. Mapa geográfico mostrando la zona de muestreo.	39
Figura 2. El volcán Yucamani.....	49
Figura 3. Géiser inactivo.....	50
Figura 4. Géiser en actividad con fumarolas presentes.	51
Figura 5. Área de Conservación Regional Vilacota-Maure	53
Figure 6. Extracción de ADN genómico bacteriano.	55
Figura 7. Amplificación y purificación del gen ADNr 16S	56
Figura 8. Árbol filogenético del género Bacillus.....	68
Figura 9. Especies de bacterias termófilas aisladas.	70
Figura 10. Especies de bacterias cultivables identificadas molecularmente	71
Figura 11. Enzimas lipolíticas de especies bacterianas termófilas.	76
Figura 12. Enzimas glucolíticas de especies bacterianas termófilas.	78
Figura 13. Enzimas proteolíticas de especies bacterianas termófilas.	79
Figura 14. Enzimas termoestables por especies bacterianas.....	81

RESUMEN

Se realizó bioprospección de bacterias termófilas cultivables en los géiseres Calientes-Candarave (Tacna, Perú) para la obtención de enzimas termoestables de aplicación biotecnológica. De once géiseres se obtuvo el sedimento de cada uno de ellos, el cual fue diluido y cultivado en medio sólido enriquecido, obteniendo 48 cepas de bacterias termofílicas, las cuales fueron aisladas y purificadas. De cada cepa aislada se extrajo la cantidad necesaria de ADN genómico, que se purificó mediante electroforesis horizontal y se obtuvieron los genes del rDNA 16S mediante la técnica de PCR. El ADN de estos genes purificados se envió a Macrogen para su secuenciación. Las secuencias enviadas por Macrogen fueron procesadas y depuradas usando varios softwares. Estos permitieron realizar la identificación molecular de las cepas bacterianas termofílicas. Cada secuencia de rDNA 16S se procesó utilizando el software BioEdit para obtener secuencias consenso. Utilizando la herramienta de búsqueda de alineación local básica (BLAST), se obtuvo la identificación de las especies bacterianas correspondientes. Luego de realizar el alineamiento múltiple y obtener el árbol filogenético por MEGA X, se estableció la presencia de 9 especies bacterianas pertenecientes al Dominio *Bacterias*, filo *Firmicutes*, clase *Bacilli*, orden *Bacillales*, familia *Bacillaceae* con los géneros *Aeribacillus*, *Anoxybacillus*, *Bacillus* y *Geobacillus*. La bioprospección de cepas bacterianas termófilas obtenidas se realizó utilizando la base de datos Patricbrc, para conocer las enzimas hidrolíticas termoestables en cada especie identificada. Se obtuvieron un total de 744 enzimas hidrolíticas, 298 lipolíticas, 227 glucolíticas y 224 proteolíticas de importancia biotecnológica.

Palabras clave: Bioprospección, bacterias termófilas, enzimas termoestables.

ABSTRACT

Bioprospecting of cultivable thermophilic bacteria was carried out in the Calientes-Candarave geysers (Tacna, Peru) to obtain thermostable enzymes for biotechnological application. From eleven geysers the sediment of each one of them was obtained, which was diluted and cultivated in enriched solid medium, obtaining 48 strains of thermophilic bacteria, which were isolated and purified. The necessary amount of genomic DNA was extracted from each isolated strain, which was purified by horizontal electrophoresis and the 16S rDNA genes were obtained by the PCR technique. The DNA from these purified genes was sent to Macrogen for sequencing. The sequences sent by Macrogen were processed and debugged using various softwares. These allowed to carry out the molecular identification of the thermophilic bacterial strains. Each 16S rDNA sequence was processed using BioEdit software to obtain consensus sequences. Using the basic local alignment search tool (BLAST), the identification of the corresponding bacterial species was obtained. After performing the multiple alignment and obtaining the phylogenetic tree by MEGA X, the presence of 9 bacterial species belonging to the Bacteria Domain, Firmicutes phylum, Bacilli class, Bacillales order, Bacillaceae family with the *Aeribacillus*, *Anoxybacillus*, *Bacillus* and *Geobacillus* genera was established. The bioprospecting of thermophilic bacterial strains obtained was carried out using the Patricbrc database, to know the thermostable hydrolytic enzymes in each identified species. A total of 744 hydrolytic, 298 lipolytic, 227 glycolytic and 224 proteolytic enzymes of biotechnological importance were obtained.

Keywords: Bioprospecting, thermophilic bacteria, thermostable enzymes.

RESUMO

A bioprospecção de bactérias termofílicas cultiváveis foi realizada nos gêiseres Calientes-Candarave (Tacna, Peru) para a obtenção de enzimas termoestáveis para aplicação biotecnológica. O sedimento de cada uma delas foi obtido de onze gêiseres, os quais foram diluídos e cultivados em meio sólido enriquecido, obtendo-se 48 cepas de bactérias termofílicas, as quais foram isoladas e purificadas. A quantidade necessária de DNA genômico foi extraída de cada cepa isolada, a qual foi purificada por eletroforese horizontal e os genes 16S rDNA obtidos pela técnica de PCR. O DNA desses genes purificados foi enviado à MacroGen para sequenciamento. As sequências enviadas pela MacroGen foram processadas e depuradas em diversos softwares. Estes permitiram realizar a identificação molecular das cepas bacterianas termofílicas. Cada sequência de rDNA 16S foi processada usando o software BioEdit para obter sequências de consenso. Utilizando a ferramenta básica de busca por alinhamento local (BLAST), foi obtida a identificação das espécies bacterianas correspondentes. Após realizar o alinhamento múltiplo e obter a árvore filogenética por MEGA X, foi estabelecida a presença de 9 espécies bacterianas pertencentes ao Domínio Bactéria, filo Firmicutes, classe Bacilli, ordem Bacillales, família Bacillaceae com os gêneros *Aeribacillus*, *Anoxybacillus*, *Bacillus* e *Geobacillus*. A bioprospecção das cepas bacterianas termofílicas obtidas foi realizada por meio do banco de dados Patricbrc, para conhecer as enzimas hidrolíticas termoestáveis em cada espécie identificada. Foi obtido um total de 744 enzimas hidrolíticas, 298 lipolíticas, 227 glicolíticas e 224 proteolíticas de importância biotecnológica.

Palavras-chave: Bioprospecção, bactérias termofílicas, enzimas termoestáveis.

INTRODUCCIÓN

La biodiversidad en la biósfera del Perú es muy extensa, y abarca las cuatro regiones naturales (costa, sierra, selva y nuestro mar peruano). En estas regiones existe una diversidad de animales, vegetales y microorganismos. Aparte de estos ecosistemas y su biodiversidad, existen, microecosistemas constituidos en los géiseres de Calientes-Candarave, una zona volcánica, formado por los volcanes Yucamani Mayor y Yucamani Menor. En estos géiseres se encuentran pequeños ecosistemas, los cuales están formados por los dominios Bacteria, Arquea y Eucaria. El estudio comprendió la bioprospección de bacterias termófilas cultivables presentes en los géiseres de Calientes-Candarave (Tacna-Perú) para la obtención de enzimas termoestables a través de una base de metadatos de libre acceso y de mayor capacidad.

La bioprospección o prospección de la biodiversidad, se define como la búsqueda sistemática, clasificación e investigación de nuevas fuentes de compuestos químicos, genes, proteínas, enzimas, antibióticos, y productos que poseen un alto valor económico, y que se localizan en los componentes de la amplia diversidad biológica.

La bioprospección es una práctica muy antigua, durante los últimos veinte años ha adquirido una importancia extraordinaria, debido a los avances tecnológicos que se han dado en la farmacéutica, biotecnología y agricultura. Lo anterior ha permitido descifrar gran parte de la fisiología de los seres vivos, con el objetivo de explorar la potencialidad de las llamadas sustancias bioactivas y la opción de encontrar en ellas, la solución a importantes problemas de salud (Duarte Torres, 2011).

La bioprospección hace uso de las técnicas moleculares empleadas en la biotecnología del ADN recombinante, para beneficio de la humanidad, a través de la actividad de industrias químicas, agrícolas, farmacéuticas, etc. Esta herramienta permite un proceso industrial eficiente y de bajo costo, que caracteriza a bacterias cuyos genes codifican enzimas termoestables para degradar una cantidad de sustratos industriales que tienen uso directo, tanto en la alimentación como en el aspecto industrial (Suárez, 1997).

Existen varios tipos de bioprospección, como la bioprospección de microorganismos de ambientes extremos, bioprospección de microorganismos marinos, bioprospección de microorganismos de suelos, etc. Nosotros bioprospectamos bacterias termófilas, cultivándolas en medios nutritivos y extrayendo su ADN genómico, para luego, obteniendo sus genes ADNr 16S, luego de secuenciarlos, las secuencias procesadas y depuradas, nos permitió su identificación. Con esta identificación, utilizando la base de metadatos, se ha logrado obtener que especies identificadas producen enzimas termoestables de interés biotecnológico.

Los géiseres de Calientes-Candarave, nos ofrecen una gran oportunidad para aprovechar sus recursos naturales que se encuentran en sus aguas y sedimentos termales, constituido por los tres dominios de microorganismos existentes, bacterias, arqueas y eucarias. La biodiversidad existente está presente en cada uno de estos recipientes térmicos, constituyendo una fuente inagotable de bacterias, arqueas y microalgas termófilas.

Una biocenosis (comunidad biótica, comunidad biológica, comunidad ecológica o simplemente comunidad) viene a ser el conjunto de organismos de todas las especies que coexisten en un espacio definido llamado biotopo, que ofrece las condiciones ambientales necesarias para su supervivencia. Por ello, la existencia de microorganismos termófilos constituye una microbiocenosis. El

término biocenosis fue acuñado por Karl Mobius en 1877, quien subrayaba la necesidad de enfocar la atención no en el individuo sino en el conjunto de individuos.

Un ecosistema, según la definición original de Tansley (1935), está formado por la biocenosis junto con su ambiente físico o biotopo. Se trata de una unidad compuesta de organismos interdependientes que comparten el mismo hábitat. Los ecosistemas suelen formar una serie de cadenas que muestran la interdependencia de los organismos dentro del sistema (Armenteras y Col. 2016).

Lokendra et al. (2011) hace de conocimiento que los volcanes en agonía ya no arrojan lava sino agua hirviente con muchas sustancias químicas distintas; cuando se entibia este caldo, es capaz de albergar bacterias o formar moléculas precursoras de vida; se cree que la vida tuvo origen en estas condiciones, cuando no había aún mucho oxígeno atmosférico (Lokendra Kumar, 2011).

El autor de “El planeta viviente” afirmó que cuando hay agua caliente, calentada por la lava, en los géiseres; allí crecen organismos un poco más evolucionados, como las algas cianófitas, que resisten al calor pero que además aprovechan la luz solar. Colonizan todos los alrededores ingiriendo las sustancias químicas del agua; cuando los hilos de agua proveniente del géiser se enfrían, aparecen las moscas de la salmuera, que ingieren algas y depositan huevos (Attenborough, 2021).

En la bioprospección hace uso de las técnicas moleculares empleadas en biotecnología para beneficio de la humanidad, a través de la actividad de la industria química, farmacéutica, agrícola, entre otras. Se define como la búsqueda dirigida de microorganismos con capacidades económicas útiles, como la producción de nuevos fármacos (antibióticos), enzimas, nutrientes, etc.

Es una herramienta científica que ha contribuido al progreso social y científico de la especie humana.

Obtener enzimas termoestables recombinantes es una necesidad urgente para el tratamiento de residuos sólidos orgánicos tales como celulosas, polisacáridos pécticos, proteínas, almidones y sustratos similares, para convertirlos en productos como glucosa, ácidos galacturónicos, aminoácidos y un buen número de otros productos interesantes para las industrias de alimentos, farmacéuticas, papel, etc.

Por ello, la bioprospección de bacterias termófilas para la producción de enzimas termoestables hidrolíticas es importante, puesto que estas enzimas se obtendrían por inducción por sustrato, comprobaríamos su actividad y especificidad enzimática a temperaturas y pH óptimos. Posteriormente, para su aplicación industrial se emplearía la técnica del ADN recombinante para producir enzimas termostables, que serían producidas en biorreactores a temperatura ambiente.

El lugar es apropiado para realizar este tipo de investigaciones, es accesible, queda a 40 km de la ciudad de Candarave, se puede llegar fácilmente en una camioneta de doble tracción y se puede realizar la toma de muestras, tanto de agua, sedimentos, tapetes, formaciones rocosas y vegetales superficiales, para realizar investigaciones hasta determinar los tiempos de evolución de estos microorganismos.

Actualmente, varias especies microbianas se utilizan en procesos industriales, y el número de productos microbianos sigue creciendo, ya que los microorganismos permiten la producción de altas concentraciones de moléculas con características químicas específicas y pequeños impactos ambientales. Además, la capacidad de modificar genéticamente los microorganismos, tanto

para aumentar la producción como para producir nuevas moléculas de interés, contribuye en gran medida a su aplicación industrial.

A pesar de todo el éxito en el uso beneficioso de los microorganismos para la humanidad, la diversidad y las interdependencias de esos microorganismos sigue siendo un campo por explorar. Teniendo en cuenta que la mayoría de los microorganismos no pueden crecer fácilmente en un laboratorio (< 1%) y que hay aproximadamente 10^{30} células microbianas en la Tierra (un número que supera todas las estrellas conocidas en el universo en nueve órdenes de magnitud), el potencial biotecnológico de los microorganismos es claramente poco utilizado.

Esto ha dificultado la comprensión de los ecosistemas microbianos y ha ido más allá de nuestra capacidad para descubrir y utilizar nuevas funciones beneficiosas de microorganismos, como nuevas enzimas para reacciones químicas y bioquímicas de interés, procesos para aumentar la biorremediación, biomarcadores para el diagnóstico de enfermedades y objetivos terapéuticos.

Por lo tanto, las tecnologías de bioprospección tradicionales y avanzadas se han desarrollado y empleado para aumentar el conocimiento sobre la diversidad microbiológica de los procesos funcionales en los ecosistemas microbianos, y también para la identificación y uso de nuevos productos biotecnológicos.

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA

1.1. DESCRIPCIÓN DEL PROBLEMA

Candarave es una provincia andina situada a 4,100 metros sobre el nivel del mar, en cuyo valle se puede observar la ciudad desde lo alto de la carretera, también se observan los sembrados de alfalfa, ganado vacuno pastando, al igual de los corderos, que se han adaptado al clima frío de la zona. Hacia afueras de la ciudad también puede observarse los auquénidos peruanos, observan siempre al volcán Yucamani y se alimentan del ichu producido a lo largo del valle de los géiseres.

A 45 kilómetros de distancia se encuentra el Valle de los Géiseres, en el cual se localizan más de 85 géiseres de diferentes tamaños que se forman en las profundidades de la falda del volcán Yucamani. Estos géiseres, emiten ruidos y burbujeos de agua termal que llega a muy elevadas temperaturas, pero lo interesante de estos géiseres es que en su interior se encuentra vida latente, constituido por los dominios Bacteria, Arquea y Eucaria.

Esta situación hace interesante realizar un estudio bioprospectivo para identificar molecularmente los tipos de bacterias que existen y determinar mediante el uso de la base de datos las enzimas termoestables, las cuales pueden tener muchas aplicaciones biotecnológicas, dada a que contendrían genes que pueden codificar para biocatalizadores, antibióticos y otras sustancias que deberíamos estudiar.

El problema es que hasta ahora no se ha realizado ninguna prospección con estos microorganismos, siendo el presente estudio el primero que intenta

bioprospectar para obtener bacterias termófilas cultivables que puedan ser inducidas por sustratos para producir enzimas termoestables o termozimas, que mediante la aplicación de la tecnología del ADN recombinante o insertando los genes nativos en otras bacterias mesófilas de rápido crecimiento, como la *Escherichia coli*, producir biotecnológicamente estas enzimas termoestables de manera industrial.

1.1.1. Antecedentes del problema

En la provincia de Candarave (Tacna-Perú) existe una zona volcánica que forma un valle a lo largo del cual, se han formado unos 85 géiseres. De estos géiseres, emanan diariamente aguas subterráneas calientes, las cuales van a desembocar en el río Calientes. Han pasado miles de años, sin que estos géiseres hayan sido aprovechados geotécnica o biológicamente. No se han realizado estudios bioprospectivos que permitan conocer las especies de bacterias termófilas. presentes en los géiseres.

1.1.2. Problema de la investigación

El problema de investigación ha sido abordado, describiendo el lugar donde se encuentra el objeto de la investigación, que son los géiseres de Calientes-Candarave, cuyo ambiente extremo de aguas termales que llegan hasta 92 °C de temperatura, pH variable de 6-8, con minerales pesados de concentración letal para organismos mesófilos, es un ambiente natural para microorganismos termófilos, como bacterias, arqueas y microalgas, cuya importancia, radica en su constitución genómica (tener genes que codifican a diferentes enzimas, antibióticos y metabolitos que son de amplio interés para la biotecnología industrial, alimenticia y farmacéutica.

La solución del problema fue bioprospectar las bacterias termófilas existentes en los géiseres de Calientes-Candarave, cuya identificación molecular permitirá obtener, a través del uso de la base de metadatos, las enzimas termoestables de importancia industrial.

1.2. FORMULACIÓN DEL PROBLEMA

¿Es posible que la biospección de bacterias termófilas existentes en los géiseres de Calientes-Candarave, determinen su identificación molecular inmediata y que a través del uso de la metabase de datos existente, nos permita conocer que enzimas termoestables producen?

1.3. JUSTIFICACIÓN E IMPORTANCIA

La diversidad microbiana en los géiseres de Calientes-Candarave ha sido poco explorada en comparación con plantas y animales superiores que habitan en esta zona. Las condiciones extremas presentes en la región abren una gran oportunidad para realizar investigaciones científicas importantes en el campo de la biotecnología microbiana, posibilitando aportes significativos hacia el desarrollo de antibióticos, toxinas, péptidos activos, etc.

La búsqueda de nuevos compuestos con actividad antibacteriana producidos por bacterias procedentes de los géiseres, constituyen un campo importante, ya que se pueden encontrar poblaciones únicas de microorganismos capaces de producir nuevas sustancias antagonistas, incluyendo antibióticos de importancia y aplicación en diversas áreas como producción de alimentos, minería, procesamiento de basura, biorremediación ambiental, productos de interés para la agricultura y medicina.

1.4. ALCANCES Y LIMITACIONES

Los alcances en este estudio permitirá la identificación de especies de bacterias termófilas y que utilizando la base de metadatos existentes determinar que enzimas hidrolíticas lipolíticas, glucolíticas y proteolíticas contienen en su genoma. Esto permitirá conocer que los genes de dichas enzimas podrían generarse utilizando la tecnología del ADN recombinante para convertirlas biotecnológicamente en enzimas termoestables.

Las limitaciones de índole natural, debido a que el muestreo de aguas termales se debe realizar in situ dado a la sensibilidad de las bacterias que deben ser obtenidas y conservadas en termos adecuados para su mantenimiento hasta el momento que se procese el agua y el sedimento, y el inconveniente podría estar limitado a las lluvias y crecidas de los ríos existentes.

Esto está relacionado con el área geográfica y la época o período en que se realizan los muestreos. En cuanto a los métodos o técnicas empleadas no existe inconveniente alguno; asimismo se cuenta con financiamiento, tiempo disponible y recursos que serán utilizados en el presente estudio. El tipo, cantidad y calidad de los datos y de la información obtenida fueron obtenidas sin ningún inconveniente.

1.5. OBJETIVOS

1.5.1. Objetivo general

Realizar la bioprospección de bacterias termófilas en los géiseres de Calientes-Candarave (Tacna-Perú), para la producción biotecnológica de enzimas termoestables.

1.5.2. Objetivos específicos

- a) Realizar bioprospección de bacterias termófilas en los géiseres de Calientes-Candarave para obtener cepas bacterianas.
- b) Aislar cepas bacterianas cultivables de sedimentos y obtener su ADN genómico extrayendo el gen ADNr 16S para su secuenciación.
- c) Identificar cepas bacterianas utilizando secuencias de genes ADNr 16S mediante la reacción en cadena de la polimerasa.
- d) Obtener las enzimas hidrolíticas termoestables mediante el uso de la base de metadatos .

1.6. HIPÓTESIS

1.6.1. Hipótesis general

La bioprospección de bacterias termófilas de los géiseres de Calientes-Candarave permite su identificación molecular, para obtener las exoenzimas hidrolíticas termoestables, mediante el uso de la base de metadatos existentes.

1.6.2. Hipótesis específicas

- a) La bioprospección de bacterias termófilas en los géiseres de Calientes-Candarave permite conocer la existencia de las especies de bacterias presentes.
- b) El aislamiento y purificación del ADN genómico se obtiene de las cepas bacterianas termófilas cultivadas en medios nutritivos.

c) La obtención de genes de ADNr 16S para la identificación molecular de cada especie bacteriana mediante la técnica del PCR es posible.

d) Las enzimas hidrolíticas termoestables de cada especie bacteriana es posible a través de la base de metadatos.

CAPÍTULO II

MARCO TEÓRICO

2.1. ANTECEDENTES DEL ESTUDIO

2.1.1. Antecedentes a nivel nacional

Se desconoce de estudios bioprospectivos que se hayan desarrollado en microorganismos que habitan condiciones extremas, sobre todo de bacterias termófilas encontradas en géiseres volcánicos cuyas temperaturas de sus aguas termales llegan hasta 92 °C. De tal manera que, las zonas altoandinas de Calientes, Candarave son ecosistemas muy importantes para proporcionar los sitios de extremos fríos y calientes para explorar la diversidad microbiana existente.

2.1.2. Antecedentes a nivel internacional

Las comunidades microbianas en aguas termales volcánicas a gran altura han sido ampliamente estudiadas en la mayor parte del mundo. La vida microbiana no parece estar limitada a ambientes específicos. Durante las últimas décadas, ha quedado claro que las comunidades microbianas se pueden encontrar en las condiciones más diversas, incluidos los extremos de temperatura, presión, salinidad y pH.

Estos microorganismos, llamados extremófilos, producen biocatalizadores que son funcionales en condiciones extremas. En consecuencia, las propiedades únicas de estos biocatalizadores han resultado en varias aplicaciones novedosas de enzimas en procesos industriales. De un

estudio reciente, surgieron importantes enfoques para ampliar la gama de aplicaciones de las extremozimas.

Tanto el descubrimiento de nuevas especies extremófilas como la determinación de secuencias genómicas proporcionan una ruta a nuevas enzimas, con la posibilidad de que éstas conduzcan a nuevas aplicaciones. Sólo una pequeña fracción de los microorganismos en la tierra han sido explotados.

Las nuevas técnicas en el cultivo y la producción de extremófilos, como también las relacionadas con la clonación y expresión de sus genes en huéspedes heterólogos, aumentarán el número de transformaciones impulsadas por enzimas en aplicaciones químicas, alimentarias, farmacéuticas y otras aplicaciones industriales. De igual importancia, las evoluciones dirigidas proporcionan enfoques para mejorar la estabilidad de las enzimas y modificar la especificidad de maneras que pueden no existir en el mundo natural (Kumar Lokendra, 2011).

El uso de organismos lignocelulolíticos solos a altas temperaturas podría reducir potencialmente el requerimiento de celulasa. Además, dichos cultivos facilitan la producción de etanol a alta temperatura y ofrecen la posibilidad de recuperación de etanol in situ. Sin embargo, se requiere más investigación sobre las vías metabólicas, la regulación de la formación del producto final y la construcción de microorganismos termófilos/termotolerantes genéticamente modificados con alta tolerancia al etanol para la utilización óptima de dichos microbios en fermentaciones industriales (Arora R, 2015).

Las endosporas inactivas de las bacterias termófilas (termosporas) se pueden detectar en sedimentos marinos fríos después de la incubación a alta temperatura. Las termosporas en el lecho marino frío pueden explicarse por un historial de dispersión originado en los hábitats de los reservorios de petróleo de biosfera profunda, donde la migración hacia arriba de los fluidos de petróleo en

las filtraciones de hidrocarburos transporta células viables hacia el océano suprayacente (Chakraborty et al., 2018).

El calentamiento de cada sedimento a 50 °C seguido de la secuenciación del gen 16S rRNA reveló varias termosporas con una distribución cosmopolita en toda el área de estudio, así como termosporas que estaban más restringidas geográficamente. Entre las termosporas que tienen una distribución más limitada, 12 OTUs de ocho linajes diferentes se detectaron repetidamente en sedimentos que contienen hidrocarburos termogénicos (Chakraborty et al., 2018).

Los termófilos son los organismos que están adaptados para vivir a altas temperaturas. La supervivencia de estos organismos a altas temperaturas es posible debido a la termoestabilidad de sus enzimas. Muchas enzimas termoestables, como la ADN polimerasa Taq, aldolasa, amilasa, lipasa, proteasa, celulasa, ARN polimerasa, etc. encuentran una serie de aplicaciones comerciales debido a su termoestabilidad. Por lo tanto, el aislamiento de bacterias termófilas a partir de fuentes naturales y su identificación son muy importantes para el descubrimiento de nuevas enzimas industriales.

La deconstrucción de la biomasa es un proceso fundamental para la fabricación de productos objetivo utilizando células microbianas y sus enzimas. Pero las enzimas que poseen un papel importante en la descomposición de la biomasa permanecen relativamente inexploradas. Los microorganismos termofílicos son de especial interés como fuente de nuevas enzimas termoestables. Muchos microorganismos termófilos poseen propiedades adecuadas para uso biotecnológico y comercial. Existe, de hecho, una demanda considerable de una nueva generación de enzimas estables que sean capaces de soportar condiciones severas en procesos industriales al reemplazar o complementar los procesos químicos tradicionales.

Una bioprospección de microorganismos termófilos de aguas termales del Himalaya occidental para la producción de xilanas, una enzima con una actividad significativa a altas temperaturas, por sus aplicaciones industriales, como el bioblanqueo de pulpa y la conversión de materiales lignocelulósicos en sustratos fermentables para la producción de biocombustibles económicos y ambientalmente atractivos.

En este estudio se logró seleccionar microbios productores de xilanas termofílica (EC 3.2.1.3) de diversos sitios termales de Himalayas occidentales. Entre los aislados, se seleccionaron 13 para la producción de enzimas interesadas (xilanas). Sólo un solo aislado, aproximadamente TP28, mostró una actividad máxima de xilanas. Después de optimizar todos los parámetros, este aislado productor de xilanas mostró un aumento de 10 veces en la actividad enzimática (Chauhan et al. 2015)

La metodología de superficie de respuesta se utilizó para generar el modelo de proceso para obtener condiciones óptimas para fuentes de nitrógeno seleccionadas, fuentes de carbono e iones metálicos para la producción máxima de xilanas y mostró un aumento del 10 % en la actividad enzimática. La naturaleza termofílica y alcalina del TP28 es de gran importancia para convertir productos de desecho agrícolas lignocelulósicos en azúcares fermentables para la producción de biocombustibles (Chauhan et al. 2015)

Las comunidades microbianas en aguas termales en elevaciones altas se han estudiado ampliamente en todo el mundo. En este sentido, las regiones del Himalaya de la India son ecosistemas valiosos para proporcionar los sitios "fríos" y "calientes" extremos para explorar la diversidad microbiana. Se aislaron un total de 140 bacterias termófilas de 12 muestras recolectadas de aguas termales de Manikaran y Yumthang del Himalaya indio (Sahay, 2017).

Los aislamientos bacterianos se estudiaron para determinar el perfil filogenético, las propiedades de crecimiento en condiciones variables y las fuentes potenciales de enzimas hidrolíticas termoestables extracelulares como la proteasa, amilasa, xilanasas y celulasas. Basándose en la producción de hidrolasas extracelulares, se seleccionaron e identificaron 51 aislamientos de Manikaran y de Yumthang utilizando la secuenciación del gen ARNr 16S que incluyó 37 especies distintas de 14 géneros diferentes, a saber, *Anoxybacillus*, *Bacillus*, *Brevibacillus*, *Brevundimonas*, *Burkholderia*, *Geobacillus*, *Paenibacillus*, *Planococcus*, *Pseudomonas*, *Rhodanobacter*, *Thermoactinomyces*, *Thermobacillus*, *Thermonema* y *Thiobacillus* (Sahay, 2017).

La mayor parte de la bioprospección potencial está relacionada actualmente con el estudio de los extremófilos y su uso potencial en procesos industriales. Las celulasas microbianas recientes encuentran aplicaciones en diversas industrias y constituyen un grupo importante de enzimas industriales. Se ha realizado una cantidad considerable de trabajo sobre las celulasas microbianas, especialmente con el resurgimiento del interés en la producción de etanol de biomasa empleando celulasas y el uso de celulasas en la industria textil y del papel (Acharya, 2012).

El método más eficiente de hidrólisis de biomasa lignocelulósica es a través de la sacarificación enzimática utilizando celulasas. La revisión analiza el conocimiento actual sobre los microorganismos termofílicos productores de celulasas, sus adaptaciones fisiológicas y el control de la expresión génica de celulasas (Acharya, 2012).

Thankappan et al (2018) afirmaron que el aprovechamiento de los biocatalizadores del nicho de aguas termales de ambientes extremos para la conversión de biomasa es significativo y prometedor debido a las características

especiales de las extremozimas atribuidas por la biogeoquímica intrigante y las condiciones extremas de estos entornos (Thankappan, 2018).

En el presente estudio, 38 aislamientos bacterianos obtenidos de aguas termales de Manikaran (95 °C), Kalath (50 °C) y Vasist (65 °C) de Himachal Pradesh se analizaron para detectar glicosil hidrolasas mediante técnica de enriquecimiento in situ utilizando biomasa lignocelulósica (LCB). Sobre la base de su potencial hidrolítico, se seleccionaron 5 aislamientos y fueron *Bacillus tequilensis* (VCB1, VCB2 y VSDB4) y *Bacillus licheniformis* (KBFB2 y KBFB3) (Thankappan, 2018).

2.2. BASES TEÓRICAS

2.2.1. Bioprospección

La bioprospección consiste en la identificación, evaluación y explotación sistemática de la diversidad de la vida en un lugar determinado, y su objetivo es buscar recursos genéticos para fines comerciales. Respecto a los microorganismos, comprende principalmente estrategias para explotar partes cultivables y no cultivables de la biodiversidad microbiana. Otro método que se ha utilizado en la investigación es la prospección in silico de secuencias genómicas de microorganismos que ya están disponibles en las bases de datos. Todo eso tiene la intención de identificar microorganismos, genes, enzimas y/o vías metabólicas para otras aplicaciones biotecnológicas en la industria o en la investigación en sí.

2.2.1.1. Bioprospección in silico

Moreira de Almeida et al (2015) afirma que la prospección in silico se basa en el descubrimiento de nuevos genes y vías metabólicas a través del análisis de secuencias genómicas en bases de datos. La disminución de los costos de la

secuenciación ha permitido generar una gran cantidad de información sobre genomas microbianos. La comunidad científica ha puesto a disposición cada vez más genomas completos de bacterias, hongos y arqueas. La prospección de esos genomas a través de métodos genómicos comparativos y de análisis evolutivo, con la ayuda de herramientas bioinformáticas, ha permitido la rápida identificación de nuevos genes y enzimas con una amplia gama de funciones y aplicaciones (Moreira de Almeida, 2015).

La identificación y validación de nuevas enzimas relacionadas con la degradación de la biomasa lignocelulósica o el metabolismo de los azúcares en bacterias y levaduras son buenos ejemplos de dónde se puede aplicar esta técnica. Se ha empleado la combinación de diferentes estrategias de bioprospección para identificar nuevos bioproductos. La bioprospección in silico, por ejemplo, puede usarse para diseñar sondas moleculares, que luego se emplean en estrategias para la prospección de microorganismos cultivables o bibliotecas metagenómicas (Moreira de Almeida et al. 2015).

2.2.1.2. Bioprospección de bacterias termófilas no cultivables

Moreira de Almeida et al (2015) indica que la bioprospección de bacterias termófilas no cultivadas se ha realizado principalmente mediante el uso de metagenómica. Con esta técnica, el material genético de los microorganismos de un lugar determinado se colecta, aísla y amplifica en bibliotecas de ADN. Las bibliotecas metagenómicas no solo permiten la identificación de los principales grupos microbianos en el géiser muestreado, sino también su caracterización genética (secuencia de ADN en el grupo bacteriano) y la prospección funcional. Se han construido bibliotecas metagenómicas de los entornos más distintos y se ha obtenido y utilizado de manera efectiva información para el microbiota de sistemas complejos, como los intestinos de insectos o los océanos, para aplicaciones biotecnológicas (Moreira de Almeida, 2015)

2.2.1.3. Bioprospección de bacterias termófilas cultivables

La bioprospección de bacterias termófilas cultivables requiere del cultivo y selección de microorganismos para una aplicación de interés, basada en la biodiversidad microbiana que existe en un determinado hábitat o que se almacena en una colección. Las pruebas miniaturizadas y los sistemas de detección en línea se pueden usar para explorar un gran número de microorganismos que producen compuestos bioactivos y para identificar moléculas de interés. Además, el cultivo de microorganismos que de otra manera no se cultivarían se mejoró con el aumento de los conocimientos de fisiología microbiana, bioquímica y ecología, mediante el empleo de técnicas de fenotipado a gran escala, que permiten el análisis simultáneo de varios rasgos y facilitar la optimización de los medios de cultivo. En el presente estudio nos limitaremos a realizar la bioprospección de bacterias termófilas cultivables de los géiseres de Calientes-Candarave.

2.2.2. Los géiseres de Calientes-Candarave

El campo geotérmico de Calientes está ubicado en la Cordillera Occidental de los Andes en el sur de Perú, en la región de Tacna a una altitud de 4400 msnm dentro de los sistemas volcánicos con presencia de varias estructuras tectónicas. En Calientes podemos distinguir dos cadenas volcánicas alineadas de NS, en el extremo oeste del valle hay un complejo volcánico Yucamane-Calientes-Yucamane Chico, mientras que en el extremo este hay centros volcánicos alineados a NESW que han presentado una actividad que posiblemente hace más de 4 millones de años. El campo de Calientes se caracteriza por presentar una gran cantidad de manifestaciones geotérmicas como fuentes termales. Las aguas termales en Calientes tienen temperaturas de

descarga en el rango de 50 °C a 87 °C, pH casi neutro de 6,90 a 8,34 y conductividad entre 1160 a 5920 $\mu\text{S}/\text{cm}$ (Cruz Pauccara, 2010).

Llegando por la trocha desde el pueblo de Candarave, grandes columnas de vapor se asoman desde la parte más verde del valle, donde abundan comunidades de arbustos y hierbas, gracias a la tibia agua de pequeños riachuelos, que bajan desde el cerro que está a la derecha. Coordenadas GPS: Latitud: -17.117789 Longitud: -70.144606 y 4300 msnm. Calientes está conformado por más de ochenta fuentes termales y géiseres de todos los tamaños y formas, donde lo más bullangueros se complementan, armónicamente, con las más apacibles, en todo un concierto de sonido y color, que va desde el verde o azul hasta las más hermosas tonalidades turquesas. Un paisaje simplemente fantástico, que más parece parte de un paisaje prehistórico que parte de la geografía moderna (Cruz Pauccara, 2010).

2.2.3. Bacterias termófilas

Las bacterias termófilas son microorganismos cuyo tamaño varía entre 2 a 5 μm (micrómetros) de diámetro por 5 a 10 μm de longitud. Estas bacterias se encuentran en las profundidades de las tierras volcánicas, en forma de esporas, en donde han permanecido por millones de años desde su origen en nuestro planeta Tierra.

Al tener contacto con el agua de las filtraciones andinas o lluvias, que son filtradas a través de las hendiduras de la zona volcánica, tiene contacto con los vapores que se originan en las capas profundas de la Tierra produciendo un cambio brusco en la temperatura del agua lo que determina un flujo de vapores que emergen hacia el exterior a través de los géiseres. Este proceso ha producido la remoción de los suelos profundos, originando con el agua caliente

la destrucción de las esporas bacterianas, permitiendo el desarrollo de las bacterias en estos ambientes extremos.

Un termófilo es un organismo, un tipo de extremófilo , que prospera a temperaturas relativamente altas, entre 41 y 122 °C. Muchos termófilos son arqueas, aunque pueden ser bacterias. Se sugiere que las eubacterias termófilas estuvieron entre las primeras bacterias. Los termófilos se encuentran en varias regiones de la Tierra calentadas geotérmicamente , como las fuentes termales como las del Parque Nacional Yellowstone y los respiraderos hidrotermales de aguas profundas , así como materia vegetal en descomposición, como turberas y compost.

Los termófilos pueden sobrevivir a altas temperaturas, mientras que otras bacterias o arqueas se dañarían y, a veces, morirían si se exponen a las mismas temperaturas. Las enzimas de los termófilos funcionan a altas temperaturas. Algunas de estas enzimas se utilizan en biología molecular , por ejemplo, la polimerasa taq utilizada en la PCR . "termófilo" se deriva del griego : θερμότητα (thermotita), que significa calor , y del griego : φιλία (philia), amor.

Los termófilos se pueden clasificar de varias maneras. Una clasificación de estos organismos según sus temperaturas de crecimiento óptimas: Simplemente termófilos: 50–64 °C; Termófilos extremos 65–79 °C; Hipertermófilos de 80 °C y más, pero no < 50 °C.

En una clasificación relacionada, los termófilos se clasifican de la siguiente manera: Los termófilos obligatorios (también llamados termófilos extremos) requieren temperaturas tan altas para el crecimiento, mientras que los termófilos facultativos (también llamados termófilos moderados) pueden prosperar a altas temperaturas, pero también a temperaturas más bajas (por

debajo de 50 °C). Los hipertermófilos son termófilos particularmente extremos para los cuales las temperaturas óptimas están por encima de 80 °C.

2.2.4. Enzimas termoestables

La producción de enzimas termoestables a partir de microorganismos termofílicos presenta ciertas ventajas económicas y técnicas, ya que permite su producción a gran escala con un rendimiento predecible, existe una disponibilidad continua y no es necesario eliminar al microorganismo y existe una gran versatilidad de microorganismos de los que se puede obtener un gran número de enzimas diferentes. Gracias a la ingeniería genética se pueden desarrollar mutaciones en el metabolismo de los microorganismos para que estos produzcan mayor cantidad de enzima termoestable.

El primer paso para la producción de enzimas termoestables es la elección de la cepa bacteriana adecuada, lo cual dependerá de las características de la enzima termoestable a producir. Una vez elegida la cepa bacteriana, se le deja crecer en condiciones adecuadas de pH, temperatura y aireación. A continuación, se debe extraer la enzima, lo cual depende de si la enzima es extracelular, intracelular o periplasmática.

De manera que si es extracelular no hay que romper ninguna membrana, pero si es intracelular hay que romper tanto la membrana externa como la interna, y si es periplasmática sólo la membrana externa. Para estos dos últimos casos existen diferentes métodos de rotura celular tanto químicos (con disolventes orgánicos, detergente, choque osmótico) como físicos (sonicación, cizalla líquida o sólida, congelación y descongelación).

Una vez extraída la enzima, esta se aísla, es decir se eliminan todos los ácidos nucleicos que han sido liberados al medio tras la rotura celular, y partículas sólidas como fragmentos de membrana y células parcialmente rotas. Después se concentra y enriquece y por último se purifica.

La producción de enzimas microbianas a nivel industrial se lleva a cabo por dos métodos: a) Fermentación en superficie (método Koji): solo para obtener enzimas extracelulares, b) Cultivo sumergido: se lleva a cabo en grandes fermentadores o biorreactores que controlan perfectamente las condiciones de cultivo. En la actualidad se producen gran cantidad de enzimas microbianas utilizadas en múltiples áreas: industria alimentaria, textil, farmacéutica y papelera.

2.2.5. Identificación molecular de bacterias termófilas

La comparación de las secuencias de los ácidos ribonucleicos permite identificar relaciones filogenéticas entre procariotas. Esta situación ha tenido una gran repercusión en la taxonomía bacteriana, generando el sistema de clasificación existente, permitiendo la identificación inmediata y precisa de las cepas bacterianas.

La amplificación del gen ADNr 16S, y su posterior secuenciación, parte evidentemente del ADN extraído de un cultivo puro bacteriano. Teniendo en cuenta su potencialidad, a medida que los recursos tecnológicos aumenten y el costo se haga más competitivo, la identificación bacteriana basada en el ADNr 16S encontrará probablemente una aplicación más rápida en el laboratorio de microbiología.

Desde su descubrimiento el ácido ribonucleico ribosomal ha sido la estructura molecular más empleada en los estudios de filogenia y taxonomía

bacteriana, cuya utilidad como cronómetro molecular fue propuesto por Carl Woese. De la misma manera, Woese reestructuró la clasificación de los procariontes en dos reinos: Eubacteria y Archaeobacteria, cuya divergencia estuvo muy acentuada como la encontrada entre los procariontes y los eucariontes. Asimismo, se logró clasificar a ambos reinos en divisiones y subdivisiones (Woese, 1987).

También Woese acunó el término dominio para reemplazar al reino como categoría taxonómica de rango superior, y clasificó a los organismos celulares en tres dominios: Bacteria, Archaea y Eukarya, el último de los cuales engloba a todos los seres eucariotas (Woese, 1990).

Desde entonces, el análisis de los ARNr 16S se ha utilizado ampliamente para establecer las relaciones filogenéticas dentro del mundo procariota, causando un profundo impacto en nuestra visión de la evolución y, como consecuencia, en la clasificación e identificación bacteriana. De hecho, las ediciones vigentes de los dos tratados fundamentales de bacteriología, el Manual de Bacteriología Sistemática y Los Procariotas basan su estructuración del mundo procariota en relaciones filogenéticas establecidas con esta macromolécula.

Los ARNr 16S pueden caracterizarse en términos de secuencia parcial, mediante el método de catalogación de oligonucleótidos, utilizado en los estudios pioneros de Woese. Siguiendo esta técnica, el ARNr 16S marcado in vivo, y purificado, se trata con la enzima ribonucleasa T1. Los fragmentos generados se separan, determinándose posteriormente la secuencia de todos aquellos que incluyan al menos seis nucleótidos (nt).

A continuación, las secuencias de la colección de fragmentos correspondientes a diferentes bacterias se alinean y comparan, utilizando

programas informáticos, para calcular finalmente los coeficientes de asociación. Como se verá más adelante, la secuenciación del gen que codifica el ARNr 16S ha sustituido en la actualidad a la secuenciación de catálogos de oligonucleótidos.

Los ribosomas son orgánulos complejos, altamente especializados, que utilizan los organismos para el complicado proceso de síntesis de proteínas. El ribosoma bacteriano tiene un coeficiente de sedimentación de 70S (expresado en unidades Svedberg), y puede disociarse en dos subunidades, la subunidad grande (50S) y la subunidad pequeña (30S). Cada subunidad es un complejo ribonucleoproteico constituido por proteínas ribosómicas y moléculas de ARNr específicas. La subunidad 30S contiene el ARNr 16S y 21 proteínas diferentes (numeradas desde S1-S21, mientras que la subunidad 50S contiene los ARNr 5S y 23S junto con unas 34 proteínas (L1-L34).

En bacterias, los genes que codifican los ARN ribosomales están organizados en operones (conjunto de genes que se transcriben a partir de la misma región promotora). Cada operón ribosómico incluye genes para los ARNr 23S (*rri*), 16S (*rrs*) y 5S (*rrf*), separados por regiones espaciadoras o intergénicas (IG), y contiene además genes para uno o más ARN de transferencia (ARNt). El producto de la transcripción del operón a partir de dos promotores, P1 y P2, situados en la región anterior a *rrs*, será procesado por el enzima ARNasa III mediante cortes en sitios específicos que separan las tres clases de ARNr, el/los ARNt y las dos IG.

El ARNr 16S es un polirribonucleótido de aproximadamente 1,500 nucleótidos, codificado por el gen *rrs*, denominado ADN ribosomal 16S, y de cuya secuencia puede obtenerse información filogenética y taxonómica. El ARNr 16S se pliega en una estructura secundaria, caracterizada por la presencia de segmentos de doble cadena, alternando con regiones de cadena sencilla (Neefs, 1990).

Se trata de secuencias específicas cortas que aparecen en todos (o en la mayor parte de) los miembros de un determinado grupo filogenético, y nunca (o sólo raramente) están presentes en otros grupos, incluidos los más próximos. Por ello, el oligonucleótido firma pueden utilizarse para ubicar a cada bacteria dentro de su propio grupo.

El número de copias del operón ribosómico por genoma bacteriano varía considerablemente, de 1 a 15, siendo relativamente constante a nivel de especie, género e incluso familia (Klappenbach, 2001).

2.2.5.1. Características del ARNr 16S como herramienta taxonómica

Aunque existen cronómetros moleculares alternativos al ARNr 16S, hasta el momento ninguno ha conseguido desplazarle. De hecho, esta macromolécula presenta una serie de características, en base a las cuales fue considerado por Woese como cronómetro molecular definitivo (Woese, C.R 1987).

1. Se trata de una molécula muy antigua, presente en todas las bacterias actuales. Constituye, por tanto, una diana universal para su identificación.

2. Su estructura y función han permanecido constantes durante un tiempo muy prolongado, de modo que las alteraciones en la secuencia reflejan probablemente cambios aleatorios.

3. Los cambios ocurren de manera suficientemente lenta, como para aportar información acerca de todos los procariotas y, junto con las variaciones en los ARNr 18S, a lo largo de toda la escala evolutiva. Los ARNr SSU contienen, sin embargo, suficiente variabilidad para diferenciar no sólo los organismos más alejados, sino también los más próximos.

4. El tamaño relativamente largo de los ARNr 16S (1,500 nt) minimiza las fluctuaciones estadísticas.

5. La conservación en estructura secundaria puede servir de ayuda en las comparaciones, aportando una base para el alineamiento preciso.

6. Dado que resulta relativamente fácil secuenciar los ADNr 16S existen bases de datos amplias, en continuo crecimiento.

Una vez determinada la secuencia de nucleótidos y establecidas las comparaciones, será el grado de similitud entre las secuencias de los ADNr 16S de dos bacterias lo que indique su relación evolutiva. Además, el análisis comparativo de secuencias permite construir árboles filogenéticos, que reflejan gráficamente la genealogía molecular de la bacteria, mostrando su posición evolutiva en el contexto de los organismos comparados.

Hay que tener en cuenta, no obstante, que es la comparación de genomas completos, y no la comparación de los ADNr 16S, la que aporta una indicación exacta de las relaciones evolutivas. En su ausencia, la especie bacteriana se define, en taxonomía, como el conjunto de cepas que comparten una similitud del 70 % o más, en experimentos de reasociación ADN-ADN.

(Stackebrandt, 1994) demostraron que cepas con este nivel de relación presentan típicamente una identidad del 97 % o más entre sus genes ARNr 16S. Así, cepas con menos del 97 % de identidad en las secuencias de sus ADNr 16S es improbable que lleguen a estar relacionadas a nivel de especie. Sin embargo, existen cepas que comparten una similitud inferior al 50 % en experimentos de reasociación, y son por tanto clasificadas en especies diferentes, pero presentan una identidad del 99-100 % a nivel de ADNr 16S.

Por ello, en taxonomía, actualmente se recomienda la identificación polifásica, que utiliza criterios fenotípicos junto con datos de secuenciación (Stackebrandt, 1994).

2.2.5.2. Estudio bioprospectivo de bacterias termófilas

La bioprospección (del griego βίο: vida, y del latín prospectio: exploración), como su nombre lo indica, es el estudio de la naturaleza dedicado al hallazgo de organismos y sustancias con posibles usos para beneficio del ser humano que pueden tener un valor comercial significativo en sectores como el industrial, alimentario, cosmético y farmacéutico, entre otros.

Se entiende entonces como la búsqueda sistemática, clasificación e investigación de nuevas fuentes de compuestos químicos, genes, proteínas, microorganismos y otros productos con valor económico actual o potencial, que forman parte de la biodiversidad.

El estudio se enfoca en aquellas especies y componentes que presentan características físicas y biológicas particulares, capaces de producir sustancias especiales y de adaptarse a condiciones extremas para su supervivencia. Es especialmente en ecosistemas frágiles, polos y fondos marinos donde se encuentran organismos de este tipo.

En la bioprospección se hace uso de las técnicas moleculares empleadas en biotecnología para beneficio de la humanidad, a través de la actividad de la industria química, farmacéutica, agrícola, entre otras. Se define como la búsqueda dirigida de microorganismos con capacidades económicas útiles, como la producción de nuevos fármacos (antibióticos), enzimas, nutrientes, etc. Es una herramienta científica que ha contribuido al progreso social y científico de la especie humana.

Actualmente se pueden utilizar base de datos o metadatos concernientes a las enzimas obtenidas en microorganismos similares o de la misma especie, realizados en otras partes del mundo, esto nos lleva a una bioprospección in silico, o también podemos emplear resultados obtenidos con la misma especie tratadas en otros estudios científicos publicados.

2.3. DEFINICIÓN DE TÉRMINOS

ADN: Acido nucleico constituido por una secuencia particular y estable de

Amplificación génica: Producción de réplicas o copias de un gen particular en la célula mediante la reacción en cadena de la polimerasa (PCR, en inglés)

Análisis filogenético: A partir de secuencias moleculares, construir un árbol filogenético que refleje sus relaciones. Desde un punto de vista de usuario final, se puede ver como una caja negra: “entran secuencias y salen árboles”.
Objetivo: Comprender los distintos métodos de análisis filogenético. Saber manejar algunas herramientas para realizar análisis filogenéticos

Árbol filogenético: Gráfico que representa las relaciones filogenéticas entre los diferentes taxones tal como son entendidas por un investigador particular. Es una hipótesis sobre las relaciones filogenéticas de un taxón.

Bioprospección: Actividad de exploración y colecta de muestras de recursos fitogenéticos de una especie cultivada.

Caracterización molecular: Detección de variantes en las secuencias de ADN, que no son afectados por el ambiente, cuando se comparan distintos fenotipos, poblaciones o especies.

Cladograma: Esquema dicotómico que muestra una hipótesis sobre las relaciones filogenéticas de varios taxones. Se construye en los supuestos del análisis cladista. No refleja el grado de divergencia. Cada nodo está definido por una o varias sinapomorfías. Las cladogramas son dirigidas ya que de otro modo no se pueden identificar los clados (si no está dirigido se está en presencia de una red o red de Wagner). Las redes preceden a las cladogramas cuando la polarización del árbol se realiza a posteriori. Hipótesis de relaciones de parentesco entre grupos monofiléticos. Adopta la forma de un árbol con ramificaciones dicotómicas o nodos. Al final de cada rama se sitúan los taxones terminales. Cada nodo debe estar justificado al menos con una sinapomorfía.

Evolución: Cambio en la constitución genética de una población que deriva en la formación de nuevas especies a través del tiempo con la participación de factores evolutivos como: la selección natural, mutación, migración y el aislamiento geográfico.

Filogenia (=filogénesis, genealogía): Historia del desarrollo evolutivo de las especies. Curso histórico de la descendencia de los seres vivos. Denominamos así al patrón de relaciones de parentesco de los organismos o, dicho de otra forma, al estudio de las relaciones entre los linajes de los organismos bajo una perspectiva histórica.

Gen: La unidad estructural, funcional y de recombinación hereditaria. Secuencia de nucleótidos en el ADN que da lugar a la formación de sustancias proteicas y no proteicas.

Genoma: El acervo total de genes de un organismo. Identifica con un código alfanumérico, y que lo distingue del resto en un banco de germoplasma.

Marcador genético: Cualquier atributo que puede ser identificable,

Marcador molecular: Una secuencia específica de nucleótidos que permite localizar un gen de interés con mayor exactitud, en ausencia del efecto ambiental en la expresión de esos genes.

Mutación: Alteración o cambio natural o inducido en un gen, cromosoma o genoma de un individuo y/o especie, que es fijable y heredable.

Raíz: En un árbol filogenético es el nodo que da origen a todos los terminales. El nodo más reciente a todos los clados es la raíz del árbol

Similitud: Utilizamos aquí este término en el sentido de la semejanza morfológica entre individuos o taxones. Constituye uno de los dos criterios de agrupación.

Sistemática filogenética: Procede de Hennig (1966). Está basada en la clasificación mediante las relaciones filogenéticas entre taxones monofiléticos. Se busca la secuencia de transformaciones de estados plesiomorfos a apomorfos. Escuela sistemática, fundada por el entomólogo alemán Willi Hennig, que propugna la agrupación de taxones exclusivamente a partir del parentesco filogenético.

Taxón natural: Es un grupo de organismos que existe en la naturaleza. En el ámbito de la Sistemática Filogenética se entiende a una especie particular o a cualquier grupo de organismos monofilético.

Taxonomía filogenética: la rama de la cladística concerniente con la representación verbal (más que con la reconstrucción o estimación) de las relaciones filogenéticas. Fue desarrollada por Kevin de Queiroz a principios de los 90.

CAPÍTULO III

MARCO FILOSÓFICO

El estudio de las relaciones entre las especies ha formado parte en los estudios de taxonomía clásica desde el siglo XVI, tal como puede extraerse de la evolución de los trabajos de algunos autores como Cesalpino, Bauhin, Ray, Tournefort y, por supuesto, Linneo. Sin embargo, no fue hasta la segunda mitad del siglo XX cuando, tras el planteamiento de la revolucionaria Teoría de la Evolución, se formalizó el interés en esclarecer las relaciones entre las especies. Así, de la mano del botánico Walter Zimmermann (1930, 1934) y del zoólogo Willi Hennig (1966) se desarrollaron las bases metodológicas para la reconstrucción de relaciones de parentesco entre las especies.

Históricamente, estas relaciones filogenéticas se inferían a partir del análisis de datos generalmente morfológicos y en menores medidas anatómicas, químicas y cromosómicas. No fue hasta el desarrollo de la técnica molecular de la PCR (Polymerase Chain Reaction; Kleppe et al., 1971, Sanger & Coulson, 1975, Saiki et al., 1985, Mullis & Faloona, 1987), cuando se pudo incorporar el estudio comparado de regiones de ADN al esclarecimiento de las relaciones evolutivas entre los organismos.

Como en el ambiente se encuentran una serie de microorganismos, plantas y animales, considero conveniente tratar el aspecto de la identificación molecular de las cepas bacterianas existentes en microcuerpos de aguas termales en donde se encuentran una gran variedad de microorganismos termófilos. A estos microorganismos termófilos conocidos con el nombre de bacterias, es que nos vamos a dedicar durante estos meses en que se llevarán a cabo los estudios moleculares de estos microorganismos. Debe considerarse que se pretende establecer las comunidades microbianas (bacterias) que se encuentran en dichos cuerpos de agua.

Se pretende establecer nuevos conocimientos de especies de bacterias termófilas existentes en dichos microambientes termales, estableciendo la comunidad bacteriana presente. Las identificaciones de las especies de bacterias presentes en éstos géiseres permitirán conocer su genoma, especialmente sus genes productores de

enzimas hidrolíticas, cuyo estudio y aplicación biotecnológica permitirá establecer nuevas perspectivas en el uso de la tecnología del ADN recombinante para la producción de enzimas recombinantes industriales cuya actividad catalítica se desarrolle a más de 60 °C.

CAPÍTULO IV

MARCO METODOLÓGICO

4.1. TIPO Y DISEÑO DE LA INVESTIGACIÓN

El tipo de investigación fue descriptiva, debido a que analiza una comunidad bacteriana con el fin de establecer su naturaleza, comportamiento y características; no existiendo manipulación de variables, ya que éstas se observan y describen tal como se presentan en su ambiente natural (géiseres). La metodología fue esencialmente descriptiva, sin embargo, utilizó algunos elementos cualitativos y cuantitativos. Esta investigación se realizó en el Laboratorio de Biotecnología Enzimática de la Facultad de Ciencias de la Universidad Nacional Jorge Basadre Grohmann, Tacna – Perú.

El diseño de la investigación obedece a los objetivos específicos establecidos en el presente trabajo. Es una estructura donde las variables y los sujetos se encuentran organizados con el fin de obtener datos y así responder las preguntas de investigación. El diseño de la investigación es no experimental tipo transeccional, porque va a describir a una o más cepas bacterianas, identificarlas molecularmente y buscar en la base de datos existentes, como el BacDive, establecer sus características, propiedades, metabolismo, las especies de bacterias termófilas cuyos estudios se han realizado en otras partes del mundo y que se encuentran concentrados en diversas plataformas de almacenaje de datos.

De esta manera, obtenemos lo que buscamos, una información que nos diga que metabolitos consumen y que producen, ello nos indicarán la presencia de enzimas hidrolíticas que producen para alimentarse.

El presente estudio se estructuró de acuerdo a los objetivos específicos que se consideran en relación al objetivo general establecido. La primera corresponde a la obtención de muestras de agua y sedimentos para bioprospectar la búsqueda de microorganismos cultivables en el laboratorio, estableciendo el tipo de microorganismo y sus características principales por observación microscópica.

La segunda tiene que ver con la extracción de ADN genómico para luego obtener el gen ADNr 16S, cuya secuencia, como marcador molecular, nos permitirá establecer su anotación como una especie bacteriana, la cual fue comparada con otras secuencias de ADNr 16S del GenBank para su identificación taxonómica.

La bioprospección termina en la selección de especies de bacterias termofílicas que tienen la capacidad de producir termoenzimas como amilasas, lipasas, esterases, pectinasas, proteinasas y celulasas, las cuales podrán ser empleadas para la producción biotecnológica de estas enzimas termoestables.

4.2. POBLACIÓN Y MUESTRA

Se ha realizado un muestreo por conveniencia, de un tamaño poblacional de 85 géiseres, se ha tomado once (11) como muestra, teniendo en consideración que de sus aguas y sedimentos se van a obtener muestras repetidas para encontrar cepas bacterianas que puedan brindarnos una cantidad adecuada de especies cultivables. En este muestreo no probabilístico es muy importante definir con claridad los criterios de inclusión y exclusión, y cumplirlos rigurosamente.

La población bacteriana termófila estuvo constituida por las comunidades existentes en cada uno de los microecosistemas termales en el valle de los

géiseres de Calientes-Candarave (Tacna, Perú). La muestra constituida por las cepas bacterianas obtenidas en cada géiser mediante filtración y cultivo.

El filtrado fue realizado cuando la cantidad de bacterias fue muy pequeña, como en casos de los géiseres con aguas relativamente cristalinas. En esta técnica se necesitan al menos 1000 ml de agua que atraviesen una membrana delgada de un filtro, con poros tan pequeños que no permitan el paso de bacterias, de esta forma éstas son retenidas en la superficie del filtro. En el caso de los sedimentos, haciendo diluciones y repicados.

4.3. OPERACIONALIZACIÓN DE VARIABLES

La operacionalización de las variables se realizó, partiendo del título establecido “Bioprospección de bacterias termófilas en los géiseres de Calientes-Candarave (Tacna, Perú) para producción biotecnológica de enzimas termoestables”. Se describió el marco metodológico, considerando el tipo y diseño de investigación establecida, así como la cantidad de géiseres en los cuales se realizarán los muestreos de agua y sedimentos, para realizar la bioprospección de bacterias termófilas cultivables, luego de ser identificadas molecularmente, fueron sometidas a una inducción por sustrato para obtener las enzimas termoestables.

Tabla 1*Operacionalización de variables del estudio*

Variable	Concepto	Indicadores	Indice	Instrumentos
Independientes				
Géiseres	Pozos de aguas termales subterráneas	Temperatura pH	°C 0-14	Multiparámetro multiparámetro
Cepas bacterias termófilas	Aeróbicas 65-70 °C. Crecimiento rápido	Temperatura Absorbancia Secuencias 16S	°C 0,000 Gen 16S	Incubadoras Espectrofotómetro Programas BioEdit, BLAST, MEGA X,
Dependiente				
Enzimas termoestables	Proteínas solubles catalizadoras	Enz. Lipolíticas Enz. Glucolíticas Enz. Proteolíticas	EC 3.1. EC 3.2. EC 3.4.	Base de metadatos BacDive Patricbrc

4.4. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

Las técnicas para recolección de datos fueron de naturaleza descriptiva. Cepas bacterianas se encuentran presentes sobre la superficie de los sedimentos y en el agua termal. El método usado para la obtención de muestras de agua y sedimento se describe mas adelante. Las cepas bacterianas son posibles de cultivar en medios nutritivos líquidos o sólidos. De cada una de las cepas bacterianas se obtuvo el ADN genómico, el cual permitió su identificación molecular. A estas cepas bacterianas se indujo con sustratos específicos en mayor contenido en el medio de cultivo para producir enzimas termoestables, y se ha medido su actividad enzimática. Entre los métodos empleados fueron:

4.4.1. Ubicación geográfica de cada géiser

La ubicación geográfica de los géiseres de Calientes para el muestreo biológico se realizó utilizando un GPS, el cual midió los datos en unidades UTM. Se ubicaron geográficamente once géiseres (Figura 1). El estudio se realizó con las cepas bacterianas que fueron aisladas de las aguas y sedimentos termales

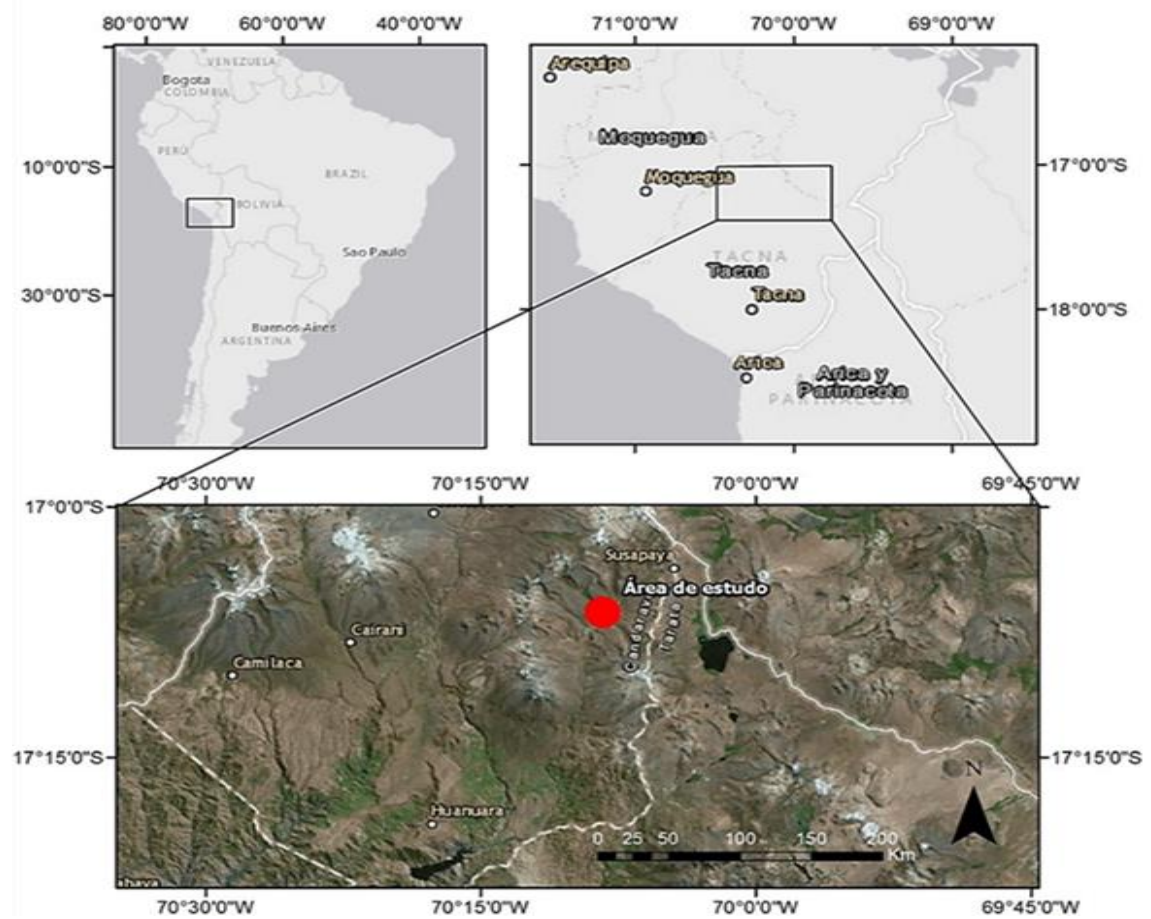
de los géiseres, situados aproximadamente a 4300 de altitud entre las coordenadas geográficas 11°06'25" L.S. y 76°10'58" L.O. El número de géiseres muestreados fue de once (11) a razón que constituye aproximadamente del 8 y 13 % de los 85 géiseres que se encuentran ubicados a lo largo del valle termal y que actualmente algunos de ellos se están secando.

Las muestras de agua y sedimentos se recolectaron antes de la temporada de lluvias de bombas manuales de 11 sitios diferentes de géiseres termales en la región del denominado Valle de los Géiseres de Calientes, Candarave (Tacna, Perú). Las muestras de agua se recolectaron en botellas de almacenamiento diferentes. Se usó muestra de agua para el aislamiento de bacterias, por lo que se mantuvo sobre hielo. Las muestras de sedimento también fueron obtenidas paralelamente con las muestras de agua. Estas muestras se colocaron en frasco de boca ancha, previamente esterilizados y se conservaron a 4 °C para su uso posterior.

Los sitios seleccionados fueron G1 (17°05.720' S; 70°07.549' W), G2 (17°05.883' S ; 70°07.639' W), G3 (17°05.861' S; 70°07.652' W), G4 (17°05.855' S; 70°07.659' W), G5 (17°06.162' S; 70°07.988' W), G6 (17°06.152' S; 70°07.995' W), G7 (17°06.328' S; 70°08.251' W), G8 (17°06.653' S; 70°08.364' W), G9 (17°07.014' S; 70°08.703' W), G10 (17°07.015' S; 70°08.666' W) y G11 (17°07.104' S; 70°08.749' W) (Figura 1).

Figura 1

Mapa geográfico de ubicación geográfica de la zona de muestreo



4.4.2. Determinación de temperatura y pH del agua termal

La obtención de muestras de agua y sedimentos, fue realizado utilizando un sistema de colección consistente de una manguera de resistencia térmica de una pulgada de diámetro de seis metros de largo, conteniendo en su interior otra manguera de un cuarto de pulgada resistente a los 150 psi, por la cual se inyectaba aire a presión a través de una bomba de aire portátil alimentado con una batería de 12 V de corriente continua.

La parte terminal consistió de un depósito cribado para el filtrado del sedimento que era removido por la fuerza de la presión del aire. Por el lado

superior se obtuvo el agua conjuntamente con los sedimentos pequeños, en un depósito de dos litros de capacidad. Sólo se obtuvo aproximadamente un litro, que luego de asentarse el sedimento en el fondo del depósito, inmediatamente era filtrado utilizando un equipo de filtración al vacío cuyas membranas fueron guardadas en frascos pequeños estériles de 100 ml de capacidad.

El sedimento asentado permaneció en los recipientes grandes hasta su procesamiento inmediato en el Laboratorio de Biotecnología Enzimática, ubicado en la Ciudad Universitaria de la Universidad Nacional Jorge Basadre Grohmann de Tacna. La temperatura y el pH del agua fue determinada por un multiparámetro, que contiene un sensor termopolar tipo K. La temperatura se midió en grados celsius (°C) y el pH en su escala conocida de cero (0) a catorce (14).

4.4.3. Cultivo y crecimiento de bacterias termófilas

De un total de once muestras de agua y sedimentos correspondientes a cada uno de los géiseres muestreados, se consideró conveniente trabajar solamente con las muestras de sedimentos de cada géiser, puesto que rendían mayor cantidad de muestras biológicas.

Se pesó 10 g de cada muestra de sedimento en un matraz Erlenmeyer de 250 ml y se agregó 90 ml de solución salina estéril al 0,85 % y se colocó en la incubadora orbital a 70 °C agitando los frascos por 15 minutos a 150 rpm. Luego, se transfirieron alícuotas de 1 ml de cada muestra diluida a tubos que contenían 9 ml del mismo diluyente, hasta obtener la dilución deseada. Seguidamente, se sembraron alícuotas de 10 µl de cada dilución por triplicado en los medios nutritivos sólidos a 65 °C en la incubadora orbital por 48 horas.

Pre-enriquecimiento de las muestras

Se preparó 90 ml de caldo LB Lennox utilizando como solvente agua de géiser filtrada, para ello se filtraron las muestras de agua (1 L) con una membrana de 0,2 μm , luego se agregó 10 g de muestra de sedimento a los frascos que contenían el caldo LB y fueron incubadas por 48 horas a 70 °C. de cultivo específicos para obtener por inducción las enzimas amilasas, celulasas, lipasas y pectinasas (Brock, 1978).

Los medios de cultivo para producción de amilasas utilizaron almidón como sustrato, para celulasas se empleó carboximetilcelulosa (CMC), para lipasas se utilizó tributirina y para pectinasas fue utilizado pectina.

El conteo de unidades formadoras de colonias (UFC) se efectuó por triplicado para cada punto de muestreo. La siembra de bacterias se realizó en agar nutritivo (peptona 5 g, extracto de carne 3 g, extracto de levadura 1 g, glucosa 5 g, agua destilada 1000 ml, agar 15 g, pH 6,8) (Wollum, 1982); los hongos, en papa dextrosa agar (extracto de papa 4,0 g, glucosa 20 g agar 15 g) y los microorganismos solubilizadores de fosfato (MSF) en medio Pikovskaya (glucosa 10 g, $\text{Ca}_3(\text{PO}_4)$ 5 g, $(\text{NH}_4)_2\text{SO}_4$ 0,5 g, NaCl 0,2 g, $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 0,1 g, KCl 0,2 g, extracto de levadura 0,5 g, $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ 0,002 g, $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 0,002 g, purpura de bromocresol 0,1 g, agar 15 g, agua destilada 1000 ml, el fosfato $\text{Ca}_3(\text{PO}_4)$ fue autoclavado por separado) (Patiño & Sánchez de Prager, 2012).

Las UFC fueron incubadas a 28 °C y contadas al segundo, tercero y quinto día, para bacterias, MSF y hongos respectivamente.

4.4.4. Cultivo axénico de cepas bacterianas aisladas

Once muestras se obtuvieron directamente del fondo de los géiseres, primero se obtuvieron las muestras de agua, que se colocaron en termos de 2 litros de capacidad, previamente estériles. Luego, se obtuvieron las muestras de sedimentos, entre 20 a 50 gramos, los cuales fueron colocados en bolsas de plástico previamente esterilizadas y codificadas.

Las muestras de agua fueron filtradas al vacío a través de una membrana de 0,22 μm . Cada membrana fue colocada después de la filtración en un tubo Falcon de 15 ml conteniendo 10 ml de medio de cultivo enriquecido, colocados en incubación a 70 °C en una estufa orbital durante 48 horas. Luego se realizó el repicado en placas Petri con agar conteniendo medio enriquecido protegido con papel kraff para mantener su humedad y evitar el secado del medio de cultivo.

Del mismo modo, en un matraz Erlenmeyer de 250 ml, se realizó la dilución de 10 g de sedimento en 90 ml de medio de cultivo enriquecido, colocados en incubación a 70 °C en una incubadora por rotación a 150 rpm por 48 horas, procediendo como se indicó en Material y Métodos.

De las once muestras de sedimento tratado, se han obtenido un total de 48 cepas bacterianas termofílicas, las cuales fueron aisladas y preparadas para extraer el ADN genómico, y luego obtener los genes ADNr 16S por reacción en cadena de la polimerasa. De las 48 cepas bacterianas obtenidas se ha realizado el procesamiento para la obtención del ADNr 16S de cada muestra de ADN genómico y luego de su tratamiento, purificación y concentración, los ADNr 16S fueron enviados a MacroGen Inc. para su secuenciación correspondiente.

4.4.5. Extracción de ADN genómico de las cepas axénicas

Se procedió a cultivar cada cepa en medio enriquecido sólido de cada muestra bacteriana a la temperatura de 70 °C para evitar la contaminación con bacterias mesófilas y obtener aproximadamente 250 mg de cada cepa bacteriana, colocando éstas en microtubos codificados y utilizando el protocolo especificado por ZymoBIOMICS que permite la toma de muestras de 250 mg de sedimento que fueron agregados en los tubos de lisis. Luego, se adicionó 750 µl de solución de lisis al tubo y se tapó bien. Se siguió el protocolo indicado, obteniéndose ADN genómico de cada cepa bacteriana. El ADN genómico se extrajo y se purificó de acuerdo con Sambrook y Russell (1954).

4.4.5.1. Obtención de genes *ADNr 16S*

El ADN genómico extraído de las especies bacterianas seleccionadas se utilizó luego para amplificar el gen 16S rRNA. Luego, el producto amplificado se verificó usando electroforesis en gel de agarosa y, como se muestra en la figura, se amplificó una banda de alrededor de 1500 pb, que especifica el gen de ARN 16S, de todas las muestras de ADN bacteriano (Figura 7.).

El ADN genómico de cada muestra se extrajo y se purificó de acuerdo con Green y Sambrook (2012). La secuenciación del *ADNr 16S* del aislado y la amplificación del gen diana se realizaron utilizando el cebador bacteriano universal 1492R (5'-TAC GGY TAC CTT GTT ACG ACT T-3 ') y el cebador específico para el dominio bacteria 27F (5'-AGA GTT TGA TCM TGG CTC AG-3 '). La amplificación del ADN se llevó a cabo en las siguientes condiciones: desnaturalización a 94 °C, 5 min, seguida de 30 ciclos de 94 °C por 30 s, 52 °C, 30 s, 72 °C por 1,5 min y extensión final a 72 °C por 10 min. Los productos de PCR amplificados de aislados bacterianos se analizaron por electroforesis con gel de agarosa al 1 %.

El gen 16S rRNA separado en gel de agarosa se escindió y purificó de acuerdo con las instrucciones del fabricante. La pureza del gen de ARNr 16S purificado en gel se verificó en gel de agarosa usando electroforesis y, como se muestra en la Figura 7, solo se obtuvo una banda aguda de 1500 pb de las tres muestras purificadas en gel. Esto fue seguido por la cuantificación del producto de PCR purificado. Se encontró que la cantidad total de ADN de cada muestra amplificada estaba dentro del rango de ADN requerido para la secuenciación (155 ng de ADN basado en 10 ng por 100 pb).

El producto PCR se purificó utilizando el kit de purificación PCR rápida QIA (Qiagen). Los productos PCR purificados fueron secuenciados por Macrogen, Korea del Norte, Utilizando un analizador genético MiSeq.

4.4.5.2. Envío de productos PCR para secuenciación

Los productos de PCR purificados se enviaron a Macrogen Inc de la República de Korea del Sur, vía aérea desde Lima a Korea. La solicitud de servicio se realiza vía e-mail. Macrogen al cabo de una semana envió los resultados de la secuenciación vía e-mail. Esta información se encuentra considerado en el Anexo correspondiente. Las secuencias remitidas por Macrogen Inc. son procesadas utilizando los métodos bioinformáticos.

4.5. PROCESAMIENTO Y ANÁLISIS DE DATOS

4.5.1. Validez y confiabilidad de instrumentos

La validez y confiabilidad de los instrumentos que fueron utilizados para aplicar los métodos en el presente estudio, contaron con la calibración y mantenimiento de las empresas que adjudicaron estos equipos. Asimismo se ha

utilizado la base de metadatos Patricbrc reconocida mundialmente y actualizada al año 2019.

4.5.2. Técnicas de análisis de datos

Las técnicas de análisis de datos corresponden al uso de los resultados obtenidos durante el muestreo y actividades en el laboratorio de Biotecnología Enzimática ubicado en la Ciudad Universitaria de la Universidad Nacional Jorge Basadre Grohmann.

4.5.3. Dinámica para realización de la investigación

La dinámica secuencial para la ejecución del estudio, consistió en obtener de cada géiser seleccionado, muestras de agua y sedimento en las cantidades indicadas anteriormente. La muestra de 2 litros de agua contenida en un termo, fué filtrada al vacío utilizando una membrana de 0,22 μm . También de cada géiser se obtuvo una muestra de 4 litros de agua en una galonera estéril, para utilizarla en la preparación de medios de cultivo, previa filtración al vacío y esterilización correspondiente. Los sedimentos obtenidos de cada géiser muestreado y colocados en bolsas de plástico previamente autoclavadas y codificadas, se utilizaron para obtener las cepas bacterianas.

La estrategia de la investigación consistió en obtener de cada géiser muestreado, 10 gramos de sedimento, los cuales colocados en matraces de 250 ml fueron diluidos con 90 ml de caldo de cultivo nutritivo con el objeto de desprender la mayor cantidad de microorganismos de las superficies de los sedimentos. Estos matraces (11) fueron colocados en una incubadora orbital durante 48 horas a 70 °C y con agitación de 150 rpm.

Luego del tiempo transcurrido, de cada matraz se midieron 10 ml del sobrenadante que fueron diluidos en 90 ml de solución salina isotónica al 0,9 % de NaCl, inmediatamente se tomaron 10 ml del primer diluido (1/10) y fueron colocados en matraces conteniendo 90 ml de solución isotónica (1/100) y así, sucesivamente, hasta obtener una absorbancia aproximada de 0.600 a 600 nm de longitud de onda.

Con cada dilución obtenida de cada géiser, se procedió a colocar en tres placas Petri, 100 µl de cada dilución, realizando un extendido total en cada placa conteniendo agar nutritivo. Luego todas las placas fueron incubadas a 70 °C durante 24 horas para permitir el crecimiento y desarrollo de las cepas bacterianas que hubiera.

Luego de las 48 horas de incubación, de cada colonia formada se obtendrán por picadura y se sembrarán en cajas Petri conteniendo medio de agar nutritivo, esperando otras 24 y 48 horas hasta que expresen una colonia uniforme en color y forma. Teniendo la seguridad de haber obtenido una colonia purificada, se procedió a su cultivo en medios de cultivo líquido conteniendo los sustratos indicados anteriormente, para realizar la inducción de enzimas por sustrato.

Es así, como se lograrán cepas bacterianas, las cuales serán sometidas a pruebas de actividad enzimática, tanto cualitativamente como cuantitativamente. Todos los productos químicos utilizados para los experimentos fueron de grado analítico. Para la preparación de medios de cultivo, los productos químicos utilizados fueron Bacto peptona, extracto de carne, triptona, almidón, extracto de levadura y agar que fueron adquiridos de laboratorio acreditados.

Otras sales inorgánicas tales como, cloruro de sodio, sulfato de manganeso, sulfato ferroso, cloruro de magnesio se adquirieron de Merck

Chemicals Ltd. Otros reactivos como glicerol, HCl, sulfato de amonio, fueron adquiridos de Fluka (Suiza). Los reactivos para amplificación por PCR se adquirieron de Sigma Chemical Company (St. Louis, MO EE.UU.).

La presente investigación requirió de los servicios de secuenciación de Macrogen Inc. Korea. Además del apoyo del Laboratorio de Biotecnología Enzimática implementado con los Fondos del Canon y Regalías Mineras, correspondientes a la Universidad Nacional Jorge Basadre Grohmann, a través del Vicerrectorado de Investigación.

Además de las técnicas y métodos utilizados en la presente tesis doctoral, se ha hecho uso de los metadatos existentes. El término de metadatos, se utilizó desde el año 1960 para describir un conjunto de datos, obtenidos durante los procesos de muestreos, de todo tipo, resultados de análisis físicos, químicos, biológicos, matemáticos, etc.

En la actualidad se está dando más importancia por que existe el mundo virtual. La primera palabra fue la de dato sobre dato, pero a partir de 2004 fue evolucionando hasta convertirse en los que hoy conocemos como metadatos.

La etimología de este término consta de dos palabras, una griega y otra latina. Por un lado la palabra griega “meta”, que significa después de o más allá de, y por otro lado el vocablo latino “datum”, que significa dato.

Así mismo, se forma la expresión metadatos como más allá de los datos. En base a esto, metadatos son un conjunto de datos que describen el contenido informativo de un recurso, de archivos o de información de los mismos. Es decir, es información que describe otros datos. No obstante, la palabra metadatos no tiene una definición exclusiva, ya que son varias las expresiones con las que se

conoce, tales como informaciones sobre datos, informaciones sobre informaciones o datos sobre informaciones.

A pesar de que la palabra metadatos se emplea en la actualidad, generalmente para el ámbito de la informática, el concepto es anterior a Internet, si bien el interés que ha suscitado ha hecho que se utilice en este entorno.

El término metadatos es principalmente muy usado en el ámbito de la informática, enfocado a las empresas debido a la adquisición de una relevancia única en los últimos tiempos como consecuencia de las grandes cantidades de información que se manejan actualmente en las compañías.

Por ello se emplean los metadatos como ayuda para gestionar los datos o información con que trabajan para una mayor efectividad y rendimiento.

En este estudio se ha utilizado dos software que contiene una gran cantidad de datos e informaciones sobre bacterias identificadas y estudiadas por muchos autores. La ventaja que se obtienen utilizando los metadatos es el ahorro de tiempo y de dinero en la obtención de resultados inmediatos.

En la presente tesis, se ha realizado primeramente la obtención de datos primarios, temperatura, pH, coordenadas geográficas de los géiseres, pero también se han obtenido agua y sedimentos, de los cuales se ha logrado evidenciar microorganismos existentes, como las bacterias termófilas. Éstas, se encuentran en mayor cantidad en los sedimentos, en fondo de los géiseres constituyendo comunidades en estos microecosistemas.

Este trabajo ha demostrado que luego de identificar las bacterias termófilas, utilizando un método rápido (antes se requería de mucho tiempo para determinar la identificación a nivel de género) para obtener su genoma (ADN genómico) y obteniendo este ADN genómico se puede obtener un marcador

molecular como lo es el ARNr 16S, cuyas secuencias permiten rápidamente establecer las especies en 99,5 % de nivel de confiabilidad y el mismo porcentaje de similitud.

Lo cierto, es que una vez que hemos logrado determinar la especie bacteriana, inmediatamente, utilizando los metadatos, podemos obtener una buena cantidad de información, sobre todo de bacterias que codifican enzimas hidrolíticas.

Para ello se han utilizado las bases de metadatos de BacDive y Patricbrc, para la determinación de enzimas hidrolíticas que se encuentran presentes en cada una de las bacterias identificadas. Esto en base a que mucho investigadores de todo el mundo ya han realizado sus estudios diferentes cuyos datos han sido ingresados a la Base de Metadatos como contribución al desarrollo científico.

De estos metadatos se valen ahora los investigadores, para obtener resultados ya expuestos anteriormente. Esto nos sirve de guía o ayuda para programar el aislamiento, purificación y liofilización de enzimas que las bacterias termófilas puedan producir.

CAPÍTULO V

RESULTADOS

5.1. GÉISERES DE CALIENTES-CANDARAVE

El área de trabajo constituido por 85 fuentes termales que forman parte del conocido Valle de los Géiseres de Calientes-Candarave, ubicados a 4,300 msnm, a lo largo del río Calientes, cuyas temperaturas fluctúan entre 55 y 92 °C, pertenece al área de Conservación Regional Vilacota-Maure de Tacna (Figuras 1-5).

Figura 2

El volcán Yucamani se encuentra situado a 20 kms de Candarave



Un géiser es una forma de fuente termal, que erupla columnas de vapor y agua caliente, materias sólidas, líquidas y gaseosas por aberturas o grietas de la corteza terrestre. Los géiseres se originan cuando se vaporiza una gran cantidad de agua subterránea, empujando violentamente el líquido a la

superficie. La existencia del fenómeno es rara, y estos suelen estar en lugares donde existe actividad volcánica; para que se dé el efecto de calentamiento del agua subterránea de tal manera que hierva, lo usual es que haya magma en la vecindad.

Figura 3

Géiser inactivo cuyas aguas en horas de la mañana estuvieron brotando como si se alimentaran para mantenerse calientes y pacíficas



Con el tiempo, la actividad en estos lugares puede ser constante o intermitente; lamentablemente en algunos lugares, ya sea por causas naturales o la actividad humana, estos han dejado de hacer sus típicas descargas de agua y vapor. A pesar de las altas temperaturas que llegan a alcanzar, existe vida dentro de ellos.

Los géiseres constituyen verdaderos microsistemas, dado que presentan los elementos bióticos que estarían conformados por la presencia de bacterias, arqueas, microalgas que han sido encontradas en dichos cuerpos de agua. Los

elementos abióticos constituidos por el agua termal que llega a temperaturas entre los 55° a 92 °C, con un pH variable y presencia de muchas cortezas calcáreas presentes en su alrededor.

Figura 4

Géiser en actividad con fumarolas presentes



Cada géiser constituye un microclima, especial, cuya microfauna, dado que existe un ciclo de nutrientes en el microsistema, esta microfauna es capaz de digerir casi cualquier sustancia orgánica y algunas sustancias inorgánicas, presentes en dicho ambiente. Estos microorganismos, aunque son de crecimiento y vida muy corta, pueden ser analizados por tener potencialidades que otros microbios no los tienen.

De los 85 géiseres existentes, se consideraron 11 géiseres para la realización del muestreo de las comunidades bacterianas existentes en estos microecosistemas. En el mes de julio del año 2018, se trabajó en once

estaciones situadas en biotopos de fondos rocosos y con poca sedimentación (Tabla 1) de las cuales solamente se consideraron once muestras.

En el presente trabajo se ha realizado once muestreos, consistentes en la obtención de agua y sedimentos presentes en dichos géiseres, además de su ubicación geográfica y obtención de parámetros físicos, químicos y biológicos.

Asimismo, la Tabla 2 muestra las coordenadas geográficas de cada uno de los géiseres en los cuales se ha realizado la toma de muestras de agua y sedimentos, cuyas características se especifican por su apariencia del color y el área que aparentan aproximadamente. También se puede apreciar las temperaturas y pH de cada géiser escogido. Las temperaturas fluctúan entre 65 y 86 grados centígrados (°C), con un promedio aritmético de 75,4 °C y una desviación estándar de 7,6 °C. Este promedio califica a estos géiseres de tipo termófilo, aunque en otro géiser se presentó una temperatura de 92 °C.

Tabla 2

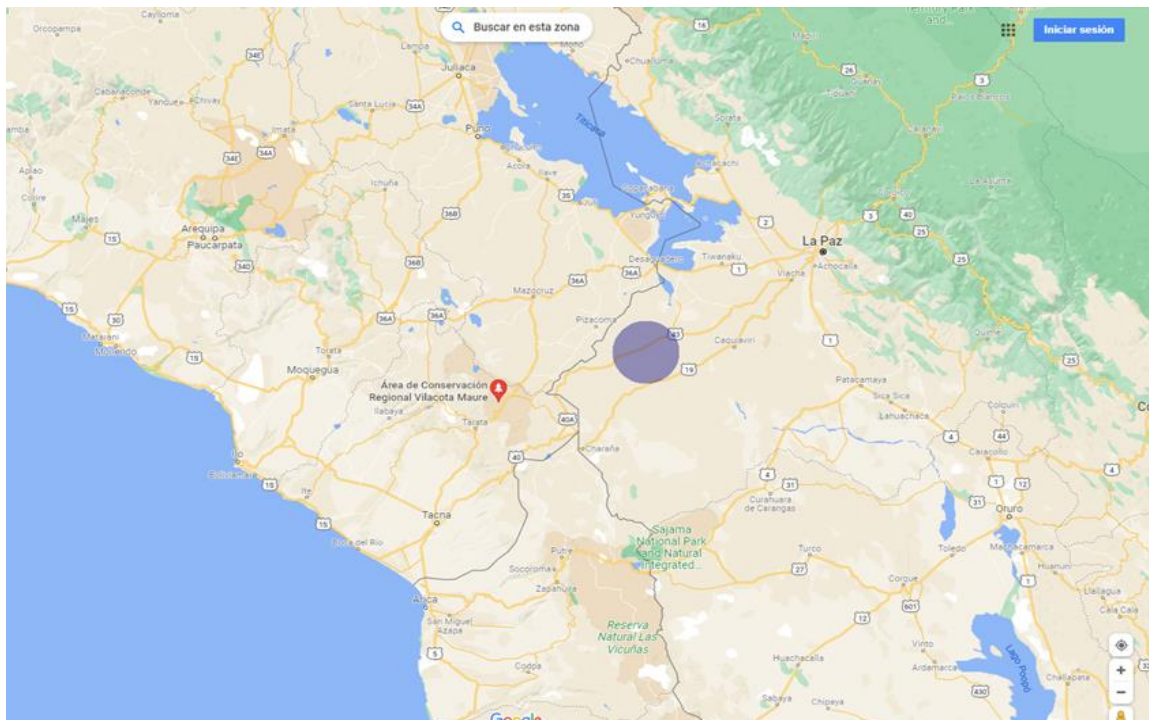
Ubicación geográfica de los Géiseres de Calientes-Candarave

Géiser	Coordenadas Geográficas	Área m ²	Sedimento	pH	Temp °C
G1	17°05.720' S ; 70°07.549' W	15	S. ralo-biofilms	7,70	86
G2	17°05.883' S ; 70°07.639' W	135	S. marrón	7,35	75
G3	17°05.861' S ; 70°07.652' W	4	S. gris	6,35	65
G4	17°05.855' S ; 70°07.659' W	12	S. marrón	7,17	79.5
G5	17°06.162' S ; 70°07.988' W	180	S. ralo	7,27	75
G6	17°06.152' S ; 70°07.995' W	30	S. rojo-verde	7,52	69
G7	17°06.328' S ; 70°08.251' W	48	S. marrón-rojo	7,28	80
G8	17°06.653' S ; 70°08.364' W	18	S. gris	7,29	86
G9	17°07.014' S ; 70°08.703' W	6	S. gris	6,98	80
G10	17°07.015' S ; 70°08.666' W	28	S. marrón	7,57	65
G11	17°07.104' S ; 70°08.749' W	15	S. marrón	7,05	69
	Promedio			7,23	75,4
	Desviación			0,36	7,6

Nota. **S**: Sedimento ralo, gris, marrón, rojo y verde; predominancia marrón y gris

Del mismo modo, el pH fluctuó entre 6,35 y 7,70 con un promedio aritmético de 7,23 y una desviación estándar de 0,36. estos valores establecen que el pH de los géiseres es casi neutro (7,00). Este aspecto de neutralización del agua permite la existencia de los tres dominios de la naturaleza en los géiseres analizados, por la existencia de microorganismos constituidos por bacterias, arqueas y microalgas. Los diferentes colores de los fondos son debido al crecimiento de diferentes clases de bacterias termófilas que utilizan la cantidad de elementos químicos que están presentes en dichos fondos. Así, se encuentran sedimentos de color gris, marrón, rojo y en algunas oportunidades ausentes de color.

Figura 5
Área de Conservacion Regional Vilacota-Maure



Nota. Google Earth Pro.

5.2. EXTRACCIÓN Y PURIFICACIÓN DEL ADN GENÓMICO

El ADN genómico extraído de las 48 cepas bacterianas seleccionadas fue cuantificado en un espectrofotómetro UV-VIS midiendo la absorbancia a 260 y 280 nanómetros de longitud de onda, calculando la relación 260/280. La concentración de ADN extraído se encontró entre 18,640 ng/μl y 132,768 ng/μl. Se encontró que la relación 260/280 fue de 1,756 – 2,069, que está de acuerdo con el valor que debería tener un ADN puro. En la Tabla 3 podemos observar el control de calidad del ADN genómico y la concentración obtenida en 12 tubos conteniendo ADN genómico.

Tabla 3

Control de calidad y concentración del ADN genómico

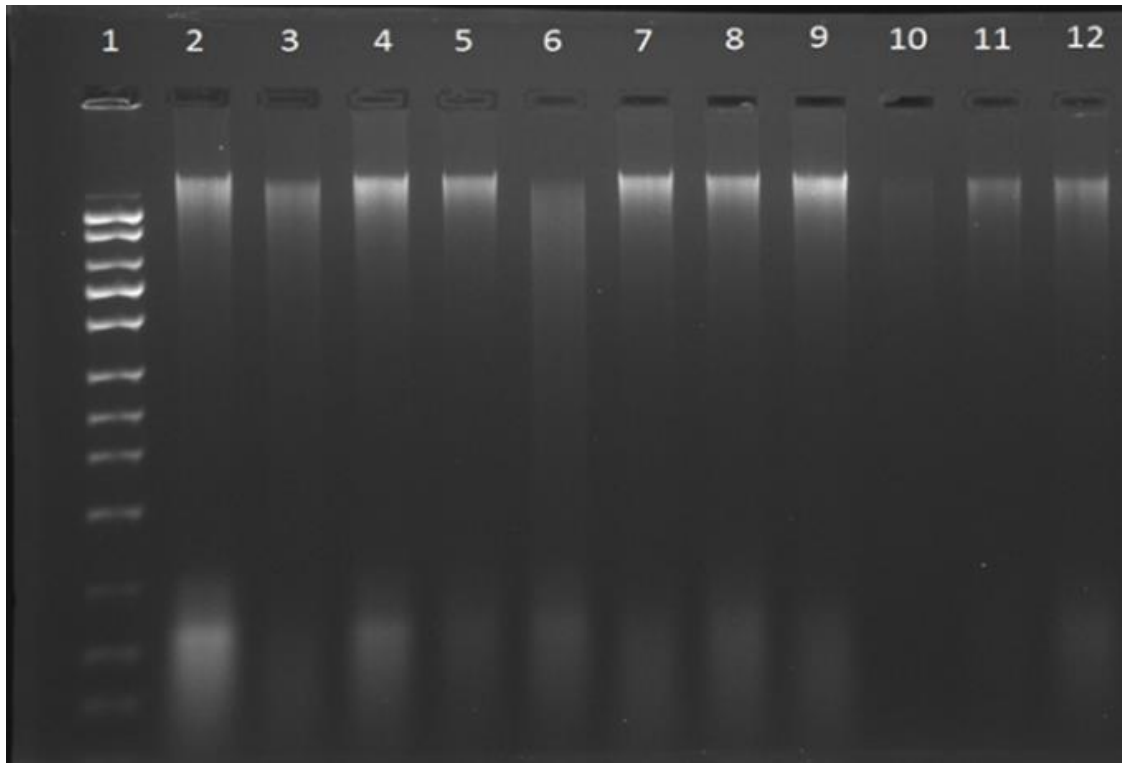
N°	Código	Abs 260	Abs 280	Abs 320	Abs 260	Abs 280	ADNg 260/280	ADNg (ng/ul)
2	A01	0,206	0,133	0,062	0,133	0,065	2,046	132,768
3	B01	0,106	0,076	0,045	0,054	0,027	2,032	54,375
4	C01	0,138	0,092	0,047	0,083	0,041	2,005	82,937
5	D01	0,125	0,089	0,053	0,066	0,032	2,035	65,619
6	E01	0,131	0,088	0,044	0,078	0,040	1,978	78,499
7	F01	0,121	0,081	0,043	0,070	0,034	2,067	69,691
8	G01	0,148	0,094	0,043	0,095	0,046	2,069	94,657
9	H01	0,160	0,120	0,079	0,074	0,037	2,012	74,199
10	A02	0,085	0,074	0,062	0,019	0,010	1,917	18,640
11	B02	0,087	0,067	0,048	0,034	0,016	2,061	33,946
12	C02	0,119	0,088	0,047	0,065	0,037	1,756	65,004
13	D02	0,139	0,091	0,043	0,089	0,044	2,023	88,535

El ADN genómico fue observado después de haber realizado la electroforesis en gel de agarosa, de acuerdo al protocolo señalado por Sambrook y Russell (1996).

5.3. AISLAMIENTO Y PURIFICACIÓN DEL GEN ADNr 16S

Figura 6

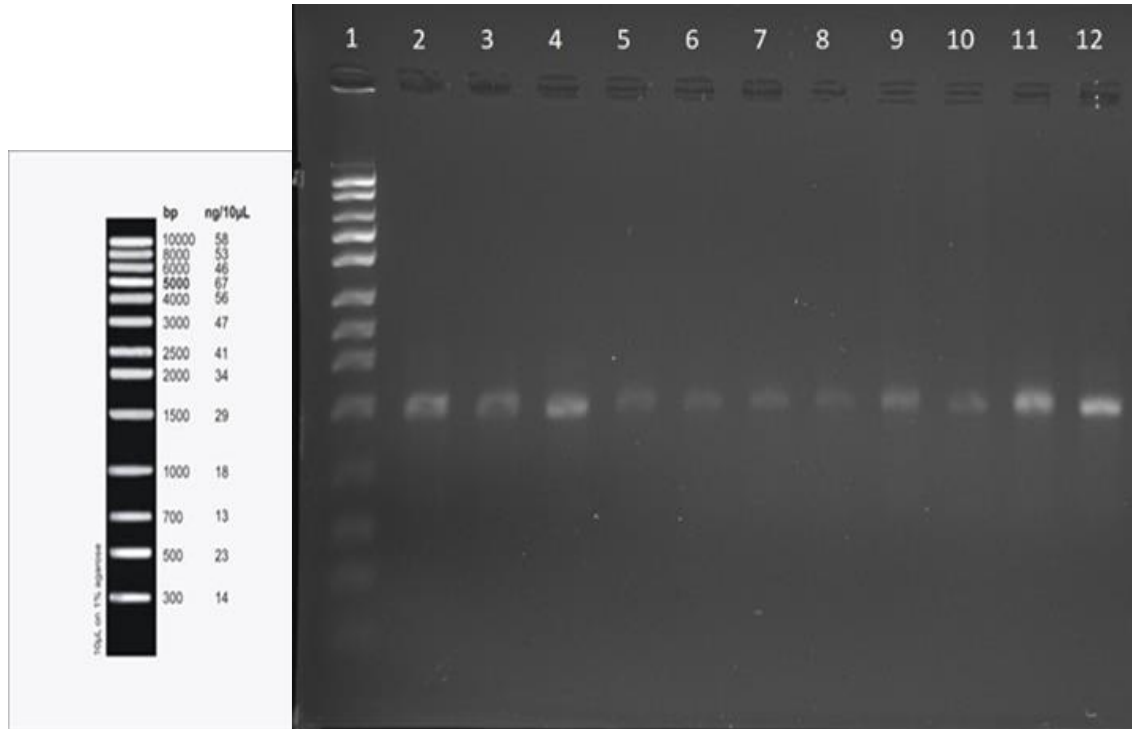
Extracción de ADN genómico bacteriano



Los resultados de las muestras de genes ADNr 16S que fueron aceptables para su secuenciación en la plataforma de Illumina MiSeq. Como resultado de la secuenciación realizada por Macrogen Inc, se nos ha remitido las secuencias de cada una de las 48 muestras de ADNr 16S, en calidad de 48 secuencias directas (forward) y 48 secuencias inversas (reverse) las que han sido procesadas en el Laboratorio de Biotecnología Enzimática de la UNJBG, utilizando el software BioEdit para depurar los nucleótidos repetidos y obtener las secuencias consenso.

Figura 7

Amplificación y purificación del gen ADNr 16S



Las secuencias consenso han sido procesadas introduciendo cada secuencia consenso en el Software Blast, programado para nucleótidos del ARNr 16S y cuyos resultados se encuentran especificados en la Tabla 4. En total se obtuvieron 48 especies de bacterias termófilas, que mediante el uso de la base de metadatos BacDiive, se ha podido establecer su taxonomía y otras características, como algunas características de cada cepa bacteriana. Bac Dive es la base de metadatos de diversidad bacteriana más grande del mundo para información bacteriana estandarizada, es de libre acceso.

Las secuencias forward y reverse de los amplicones obtenidos, es decir, las secuencias contigs o de consenso, se ha procedido a determinar en el programa BLAST las identificaciones correspondientes a las especies obtenidas (OTUs).

Las secuencias obtenidas fueron limpiadas de nucleótidos no deseables utilizando el programa BioEdit. Las secuencias deducidas se sometieron al algoritmo BLAST del Centro Nacional de Biotecnología, Bethesda, MD, EE. UU., para recuperar las secuencias homólogas en GenBank.

Los arboles filogenéticos se construyeron realizando un análisis de matriz de distancia utilizando el sistema NT. La búsqueda en la base de datos y las comparaciones se realizaron con la base de datos BLAST.

Las secuencias enviadas por MacroGen tienen sus extremos con nucleótidos no codificables que deben ser depurados. Para ello se pueden utilizar varios softwares que se encuentran disponibles en el Internet, en diferentes páginas web. Nosotros hemos procedido con el Software Bioedit, que es muy amigable en su empleo. También, puede emplearse Mega X, o Chromas.

De la tabla, los productos pcr (reacción en cadena de la polimerasa) se ha procesado cada gen ADNr 16S, el cual ha sido secuenciado en forma directa (Forward 27F) y en forma inversa (Reverse 1492R). Para cada uno de estos genes se ha utilizado el software BioEdit, primero para realizar la limpieza de nucleótidos repetidos en cada una de las dos cadenas (directa e inversa). En el mismo software se ha obtenido el gen consenso, realizando el alineamiento de dos pares de bases.

5.4. IDENTIFICACIÓN MOLECULAR DE BACTERIAS TERMÓFILAS

Posteriormente, después de haber obtenido la secuencia consenso, dicha secuencia ha sido ingresada en el software BLASTn, para obtener por comparación con otras secuencias casi idénticas, el nombre científico de la bacteria que corresponde a dicho gen por comparación al 99 o 100% de identidad.

Tabla 4a*Identificación molecular de bacterias usando el gen ADNr 16S*

Especies bacterianas	Accesión	N	% ID	Código
Aeribacillua composti	NR159152.1	1370	98,76	C02
Aeribacillus composti	NR159152.1	1364	98,83	A07
Aeribacillus composti	NR159152.1	1363	99,12	P03
Aeribacillus composti	NR159152.1	1377	99,06	P06
Aeribacillus composti	NR159152.1	1368	98,54	P09
Aeribacillus composti	NR159152.1	1359	98,97	C04
Aeribacillus composti	NR159152.1	1362	96,77	C07
Aeribacillus composti	NR159152.1	1422	99,42	C09
Aeribacillus composti	NR144578.1	1423	98,94	L04
Aeribacillus composti	NR144578.1	1373	98,76	L07
Aeribacillus composti	NR114089.1	1358	98,97	L08
Aeribacillus composti	NR113285.1	1387	99,64	L12
Aeribacillus composti	NR026515.1	1426	96,46	A09
Aeribacillus composti	NR026515.1	1410	98,51	C10
Anoxybacillua kamchatkensis	NR115109.1	1333	98,27	A06
Anoxybacillus kamchatkensis	NR115109.1	960	98,21	A05
Anoxybacillus kamchatkensis	NR115109.1	1337	99,03	A12
Anoxybacillus kamchatkensis	NR115109.1	1121	98,66	P10
Anoxybacillus kamchatkensis	NR115109.1	1359	99,19	L03
Bacillus coagulans	NR041523.1	945	99,89	P01
Bacillus coagulans	NR041523.1	1009	99,21	C12
Bacillus coagulans	NR041523.1	850	99,88	L01
Bacillus haynesii	NR157609.1	914	99,89	A04
Bacillus haynesii	NR113974.1	1374	99,21	L05

En total se procesaron 96 secuencias, obteniendo 48 secuencias consenso. En el Anexo 3 pueden observarse las secuencias consenso obtenidas, cuyos nombres científicos obtenidos pueden visualizarse en la Tabla 4a y 4b.

Tabla 4b*Identificación molecular de bacterias usando el gen ADNr 16S*

Especies bacterianas	Accesión	N	% ID	Código
Bacillus hisashii	NR114578.1	1365	99,27	P02
Bacillus hisashii	NR144578.1	1423	98,94	P04
Bacillus hisashii	NR144578.1	1373	98,76	P07
Bacillus hisashii	NR144578.1	1348	99,11	C01
Bacillus hisashii	NR144578.1	1375	98,76	C06
Bacillus hisashii	NR159152.1	1377	99,06	L06
Bacillus hisashii	NR118996.1	1384	99,71	L11
Bacillus licheniformis	NR118996.1	1385	99,78	A02
Bacillus licheniformis	NR118996.1	1379	99,27	A08
Bacillus licheniformis	NR118996.1	1384	99,71	P11
Bacillus licheniformis	NR118996.1	1365	99,27	C03
Bacillus licheniformis	NR118996.1	661	99,85	C11
Bacillus smithii	NR112634.1	1382	99,28	A10
Bacillus subtilis	NR112686.1	1405	99,43	A11
Bacillus subtilis	NR102783.1	1339	99,71	C05
Bacillus subtilis	NR113285.1	1387	99,64	P12
Geobacillus kaustophilus	NR115285.1	1266	99,61	A01
Geobacillus kaustophilus	NR115285.1	1335	99,41	A03
Geobacillus kaustophilus	NR115285.1	1360	98,68	C08
Geobacillus kaustophilus	NR115285.1	1356	99,41	L02
Geobacillus kaustophilus	NR159152.1	1368	98,54	L09
Geobacillus kaustophilus	NR115109.1	1121	98,66	L10
Geobacillus kaustophilus	NR114089.1	1358	98,97	P08
Pseudoxanthomonas taiwanensis	NR113974.1	1374	99,21	P05

En la Tabla 4a podemos apreciar las especies de bacterias identificadas, más su número de accesión, la cantidad de nucleótidos presentes en el gen, el porcentaje de identidad de cada especie y el código correspondiente. En esta tabla se indican cuatro especies de bacterias termófilas, con sus respectivos número de accesión, número de nucleótidos de cada consenso, el porcentaje de identidad entre más de 100 especies, lo cual permite tomar la especie que tiene mayor porcentaje de identidad. Al final se encuentra el código correspondiente a cada especie identificada y ordenada.

En la Tabla 4b podemos apreciar las especies de bacterias identificadas, más su número de accesión, la cantidad de nucleótidos presentes en el gen, el porcentaje de identidad de cada especie y el código correspondiente. En esta tabla se indican cinco especies de bacterias termófilas, con sus respectivos número de accesión, número de nucleótidos de cada consenso, el porcentaje de identidad entre más de 100 especies, lo cual permite tomar la especie que tiene mayor porcentaje de identidad. Al final se encuentra el código correspondiente a cada especie identificada y ordenada.

De las tablas se ha logrado la identificación molecular utilizando el gen ADNr 16S de cada una de las bacterias termófilas. El resultado indica que existen 14 especies de *Aeribacillus composti*, 5 especies de *Anoxybacillos kamchatkensis*, 3 especies de *Bacillus coagulans*, 2 especies de *Bacillus haynesii*, 7 especies de *Bacillus hisashii*, 5 especies de *Bacillus licheniformis*, 1 especie de *Bacillus smithii*, 3 especies de *Bacillus subtilis*, 7 especies de *Geobacillus kaustophilus* y una especie de *Pseudoxanthomonas taiwanensis*.

De todas las secuencias indicadas, para el cultivo de especies de bacterias consideradas para realizar el screening de selección de bacterias productoras de enzimas, se han escogido aquellas que tengan mayor porcentaje de similitud. Siendo las siguientes especies de bacterias consideradas:

El producto de PCR se purificó utilizando el kit de purificación de PCR rápida QIA (Qiagen). Los productos de PCR purificados fueron secuenciados por MacroGen, Korea del Sur, utilizando un analizador genético de Illumina MiSeq. La secuencia deducida se sometió al algoritmo BLAST del Centro Nacional de Biotecnología, Bethesda, MD, EE. UU., Para recuperar las secuencias homólogas en GenBank.

La identificación molecular de las bacterias termófilas fue realizada en el Laboratorio de Biotecnología Enzimática, utilizando los programas libres, como el BioEdit, para realizar la limpieza de cada uno de los amplicones forward y reverse, a fin de obtener la secuencia contig o de consenso. Utilizando el programa BLAST, se determinó la anotación de la secuencia contig con 100 secuencias del Banco de Datos de GenBank. Determinando la especie a un porcentaje de similitud mayor del 99 %. Todas las secuencias (48) fueron procesadas.

El alineamiento múltiple de secuencias y de las obtenidas en la base de datos se realizó con el programa CLUSTALX versión 2.1 (Jeanmougin et al. 1998), el alineamiento fue editado manualmente con el programa BIOEDIT versión 7.0.5.3. El árbol filogenético fue construido con el programa MEGA versión 6.0 (Tamura et al., 2013) empleando los métodos neighbor-joining (Saitou y Nei 1987) y minimum-evolution (Rzhetsky y Nei, 1993). El análisis de bootstrap de 500 replicaciones fue usado para evaluar la estabilidad relativa de las ramas del árbol filogenético (Felsenstein, 1985)

Concretamente, se trata de organismos diminutos, generalmente unicelulares, llamados termófilos. El término termófilo se aplica a organismos vivos que pueden soportar condiciones extremas de temperatura relativamente altas, por encima de los 50 °C. Es un subtipo de vida extremófila. Muchos termófilos pertenecen al dominio Archaea. Estos organismos tienen una tasa de crecimiento rápida, pero de corta duración.

Son hipertermófilos aquellos organismos que habitan a altas temperaturas, que normalmente llegan al punto de ebullición. Por lo común crecen bien y se reproducen a temperaturas mayores a 65 °C, un calor letal para la mayoría de los seres vivos. Los hipertermófilos son procariontes, como algunas bacterias, principalmente arqueas, en las que además imperan otras

condiciones extremófilas. No obstante, las observaciones probaron que es posible la vida en temperaturas elevadas, incluso algunas bacterias prefieren temperaturas superiores al punto de ebullición del agua. Docenas de esas bacterias son conocidas hoy en día. En el Laboratorio las bacterias termofilas prefieren temperaturas entre 55 y 70 C°

5.5. CLASIFICACIÓN TAXONÓMICA DE ESPECIES BACTERIANAS

La base de metadatos bacterianos BacDive (<http://bacdive.dsmz.de>) se ha convertido en un recurso completo de datos estructurados sobre taxonomía, morfología, fisiología, cultivo, aislamiento y datos moleculares de procariontes. Con su versión actual (7/2018), la base de datos ofrece información para 63 669 cepas de bacterias y arqueas, incluidas 12, 715 cepas de tipo. Durante los últimos desarrollos de BacDive, se priorizó el enriquecimiento de la información sobre las cepas existentes. Esto ha resultado en un aumento del 146% del contenido de la base de datos en los últimos tres años.

Se integraron conjuntos de datos especialmente ricos de 4782 descripciones de especies anotadas manualmente en la Revista Internacional de Microbiología Sistemática y Evolutiva que arrojó datos fenotípicos estandarizados para 5468 cepas tipo. Otra mejora importante del contenido se logró mediante la movilización de los resultados de la prueba del 8977 Analytical Profile Index (API®) que constituyen datos fisiológicos para la identificación de 5237 cepas.

BacDive ofrece una colección de datos API® única con respecto al tamaño y la diversidad. Además, se integraron datos sobre perfiles de ácidos grasos y pruebas de susceptibilidad a antibióticos. Una interfaz gráfica de usuario revisada y nuevas herramientas de búsqueda como el buscador de pruebas

API®, el TAXplorer o la búsqueda de fuentes de aislamiento microbiano mejoran significativamente la experiencia del usuario.

Se integraron datos sobre perfiles de ácidos grasos y pruebas de susceptibilidad a antibióticos. Una interfaz gráfica de usuario revisada y nuevas herramientas de búsqueda como el buscador de pruebas API®, el TAXplorer o la búsqueda de fuentes de aislamiento microbiano mejoran significativamente la experiencia del usuario. Se integraron datos sobre perfiles de ácidos grasos y pruebas de susceptibilidad a antibióticos. Una interfaz gráfica de usuario revisada y nuevas herramientas de búsqueda como el buscador de pruebas API®, el TAXplorer o la búsqueda de fuentes de aislamiento microbiano mejoran significativamente la experiencia del usuario.

Con la base de datos de BacDive (The Bacterial Diversity Metadatabase) se ha logrado determinar la clasificación taxonómica de cada bacteria identificada molecularmente utilizando el gen ADNr 16S. Esta base de datos también contiene información correspondientes a: Nombre y clasificación taxonómica, Morfología, Condiciones de cultivo y crecimiento, Fisiología y metabolismo, Aislamiento, muestreo e información ambiental e Información de secuencia.

Esta base de datos, es alimentada con los resultados de las investigaciones que se dan continuamente por otros investigadores, para ello, cuenta con un formulario que se debe llenar cuando uno ha realizado una nueva investigación. En esta oportunidad, solamente hemos considerado el nombre y clasificación taxonómica de cada una de las especies identificadas. De acuerdo con la Base de Datos BacDive se ha logrado obtener las clasificaciones taxonómicas de cada una de las especies bacterianas cultivables:

1. *Aeribacillus composti* cepa LBE-1 (Anexo 1)

Nombre y clasificación taxonómica:

Dominio	Bacterias
Filo	Firmicutes
Clase	Bacilos
Pedido	Caryophanales
Familia	Bacillaceae
Género	<i>Aeribacillus</i>
Especies	<i>Aeribacillus composti</i>
Nombre científico	<i>Aeribacillus composti</i> Finore et al. 2017

2. *Anoxybacillus kamchatkensis* cepa LBE-2 (Anexo 2)

Nombre y clasificación taxonómica:

Dominio	Bacterias
Filo	Firmicutes
Clase	Bacilos
Pedido	Caryophanales
Familia	Bacillaceae
Género	<i>Anoxybacillus</i>
Especies	<i>Anoxybacillus kamchatkensis</i>
Nombre científico	<i>Anoxybacillus kamchatkensis</i> Kevbrin et al. 2006

3. *Bacillus coagulans* cepa LBE-3 (Anexo 3)

Nombre y clasificación taxonómica:

Dominio	Bacterias
Filo	Firmicutes
Clase	Bacilos
Orden	Caryophanales
Familia	Bacillaceae
Género	<i>Bacillus</i>
Especie	<i>Bacillus coagulans</i>
Nombre científico	<i>Bacillus coagulans</i> Hammer 1915

4. *Bacillus haynesii* cepa LBE-4 (Anexo 4)

Nombre y clasificación taxonómica:

Dominio	Bacterias
Filo	Firmicutes
Clase	Bacilos
Orden	Caryophanales
Familia	Bacillaceae
Género	<i>Bacillus</i>
Especies	<i>Bacillus haynesii</i>
Nombre científico	<i>Bacillus haynesii</i> Dunlap y col. 2017

5. *Bacillus hisashii* cepa LBE-5 (Anexo 5)

Nombre y clasificación taxonómica:

Dominio	Bacterias
Filo	Firmicutes
Clase	Bacilos
Orden	Caryophanales
Familia	Bacillaceae
Género	Bacillus
Especies	<i>Bacillus hisashii</i>
Nombre científico	<i>Bacillus hisashii</i> Nishida et al. 2015

6. *Bacillus licheniformis* cepa LBE-6 (Anexo 6)

Nombre y clasificación taxonómica:

Dominio	Bacterias
Filo	Firmicutes
Clase	Bacilos
Orden	Caryophanales
Familia	Bacillaceae
Género	Bacillus
Especies	<i>Bacillus licheniformes</i>
Nombre científico	<i>Bacillus licheniformes</i> (Weigmann 1898) (Chester 1901)

7. *Bacillus smithii* cepa LBE-7 (Anexo 7)

Nombre y clasificación taxonómica:

Dominio	Bacterias
Filo	Firmicutes
Clase	Bacilos
Pedido	Caryophanales
Familia	Bacillaceae
Género	Bacillus
Especies	<i>Bacillus smithii</i>
Nombre científico	<i>Bacillus smithii</i> Nakamura et al. 1988

8. *Bacillus subtilis* cepa LBE-8 (Anexo 8)

Nombre y clasificación taxonómica:

Dominio	Bacterias
Filo	Firmicutes
Clase	Bacilos
Orden	Caryophanales
Familia	Bacillaceae
Género	Bacillus
Especies	<i>Bacillus subtilis</i>
Nombre científico	<i>Bacillus subtilis</i> Ehrenberg 1835 (Cohn 1872)

9. *Bacillus kaustophilus* cepa LBE-9 (Anexo 9)

Nombre y clasificación taxonómica:

Dominio	Bacterias
Filo	Firmicutes
Clase	Bacilos
Orden	Caryophanales
Familia	Bacillaceae
Género	Geobacillus
Especies	<i>Geobacillus kaustophilus</i>
Nombre científico	<i>Geobacillus kaustophilus</i> (Nazina et al. 2001)

5.6. CONSTRUCCIÓN DEL ÁRBOL FILOGENÉTICO

A continuación, se da a conocer las nueve secuencias de las especies identificadas mediante el Software BLAST para realizar el alineamiento múltiple de las secuencias y así construir el árbol filogenético que establecerá las relaciones de las especies de bacterias identificadas molecularmente.

El análisis bioinformático ha determinado el procesamiento de las secuencias del gen ADNr 16S obtenidos del ADN genómico de cada una de las 48 cepas de bacterias termófilas obtenidas en 11 géiseres de Calientes, Candarave-Tacna, Perú.

Figura 8

Árbol filogenético que muestra la posición de las cepas de los géneros de *Anoxybacillus*, *Geobacillus*, *Aeribacillus* y *Bacillus*, lo cual nos indica la estrecha relación que existe de la Familia Bacillaceae y por ende, del filo Firmicutes que pertenece al Dominio Bacteria.

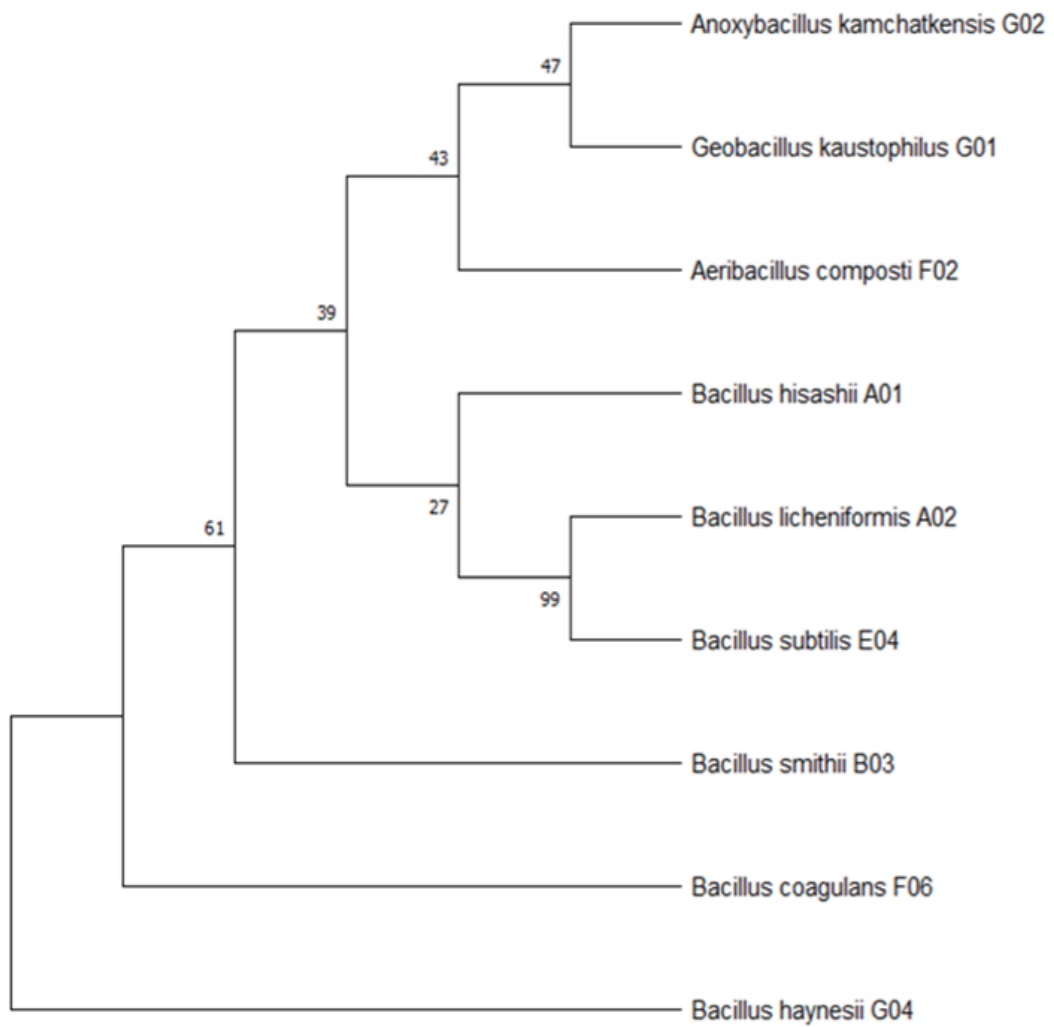


Tabla 5

Clasificación taxonómica de especies de bacterias termófilas

Dominio:	BACTERIA
Filo:	FIRMICUTES
Clase:	BACILLI
Orden:	BACILLALES
Familia:	BACILACEAE
Género:	BACILLUS
Especie:	<i>Bacillus licheniformis</i> <i>Bacillus haynesii</i> <i>Bacillus smithii</i> <i>Bacillus subtilis</i> <i>Bacillus coagulans</i> <i>Bacillus hisashii</i>
Género:	AERIBACILLUS
Especie:	<i>Aeribacillus composti</i>
Género:	ANOXYBACILLUS
Especie:	<i>Anoxybacillus kamchatkensis</i>
Género:	GEOBACILLUS
Especie:	<i>Geobacillus kaustophilus</i>

La comunidad bacteriana termófila de once géiseres de Calientes Candarave representados taxonómicamente en cuatro géneros: *Bacillus* con seis especies distintas, *Aeribacillus*, *Anoxybacillus* y *Geobacillus* con una especie. La Tabla 5 contiene las especies de bacterias termófilas seleccionadas para la obtención de enzimas hidrolíticas o hidrolasas. Estas enzimas catalizan la hidrólisis de varios enlaces.

Algunas de estas enzimas plantean problemas porque tienen una especificidad muy amplia y no es fácil decidir si dos preparaciones descritas por diferentes autores son iguales o si deben incluirse en entradas diferentes. Si bien el nombre sistemático siempre incluye 'hidrolasa', el nombre común está, en la mayoría de los casos, formado por el nombre del sustrato con el sufijo -asa. Se entiende que el nombre del sustrato con este sufijo, y ningún otro indicador, significa enzima hidrolítica. Cabe señalar que las enzimas en EC 3.4 (peptidasas) tienen nombres recomendados en lugar de nombres comunes

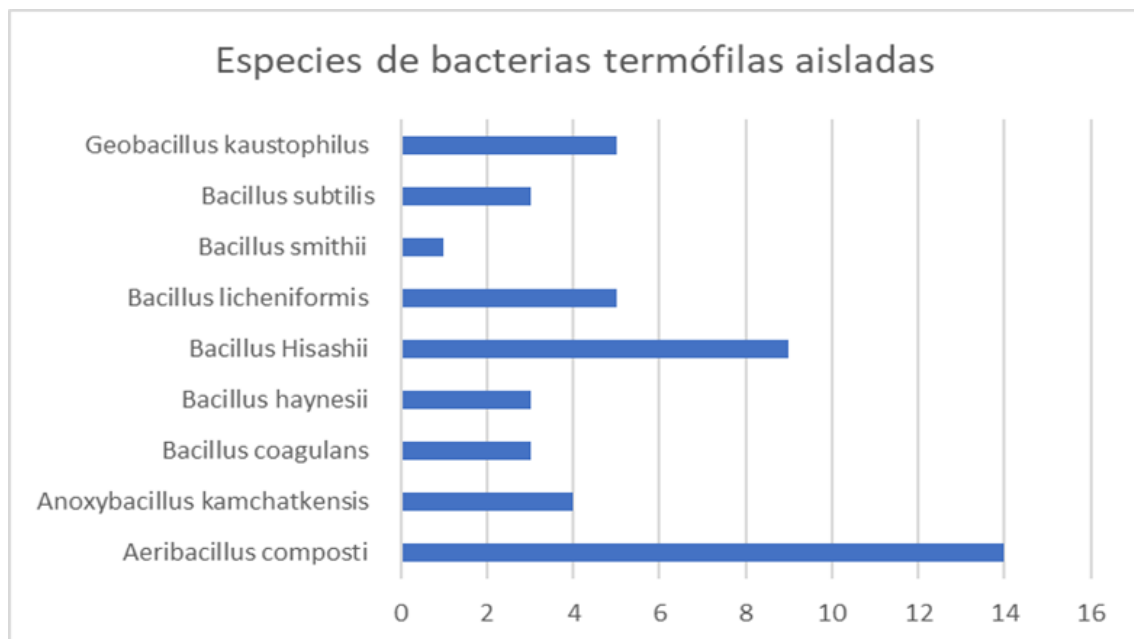
Tabla 6

Especies de bacterias termófilas obtenidas

ESPECIES DE BACTERIAS TERMÓFILAS CULTIVABLES	% Ident	n
<i>Aeribacillus composti</i> strain N.8	99,54	14
<i>Anoxybacillus kamchatkensis</i> strain KG8	99,19	4
<i>Bacillus coagulans</i> strain NBRC 12583	99,26	3
<i>Bacillus haynesii</i> strain NRRL B-41327	99,90	3
<i>Bacillus hisashii</i> strain N-11	99,37	9
<i>Bacillus licheniformis</i> strain DSM 13	99,78	5
<i>Bacillus smithii</i> strain NBRC 15311	99,28	1
<i>Bacillus subtilis</i> strain 168	99,41	3
<i>Geobacillus kaustophilus</i> strain BGSC 90A1	99,41	5

Figura 9

Especies de bacterias termófilas aisladas

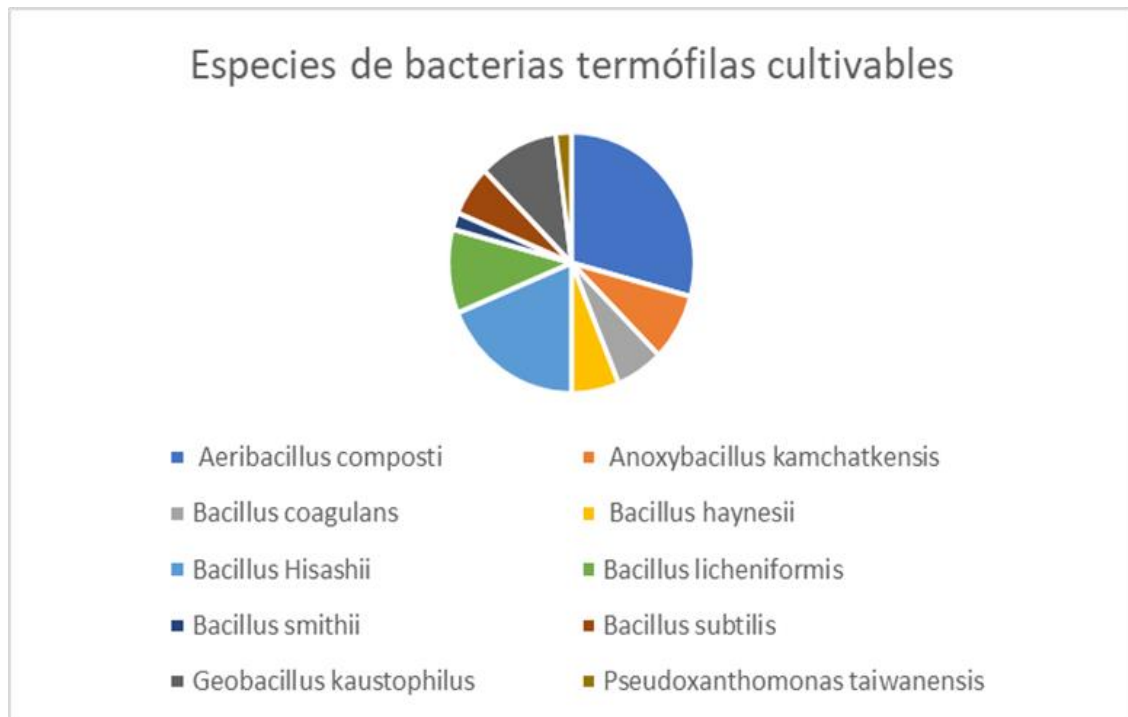


La Figura 9 nos muestra las barras que indican el número de especies de bacterias termófilas resultantes después de haber realizado el ordenamiento correspondiente, figurando sólo la especie que contaba con el mayor porcentaje de identidad, cuya variación estuvo considerado entre 99,19 y 99,90 % (Tabla 6). Se puede apreciar que *Aeribacillus composti* representó la mayor frecuencia

(14), mientras que la especie de *Bacillus smithii* representó la menor frecuencia (1).

Figura 10

Especies de bacterias cultivables identificadas molecularmente



Asimismo, podemos señalar que *Bacillus hisashii*, seguidos de *Bacillus licheniformis* y *Geobacillus kaustophilus* representaron frecuencias de 9 y 5 respectivamente.

Con el uso del metadato Patricbrc se han obtenido una buena cantidad de información referente a las enzimas hidrolíticas producidas por las bacterias obtenidas. Solamente se han separado las enzimas oligoméricas (es decir, que tienen una sola cade polipeptídica de aminoácidos, esto permite la posibilidad de ser clonada en el plásmido o vector en otra bacteria, como *E. coli* para su multiplicación en grande, a nivel industrial.

Identificadas las cepas bacterianas se han obtenido las enzimas hidrolíticas clasificadas en EC 3, considerando a las enzimas lipolíticas, glucolíticas y proteolíticas. Se han obtenido 56 enzimas lipolíticas, representadas por las más importantes, como carboxilesterasas, metilesterasas, fosfoesterasa, acetiesterasas, fosfatasa alcalina, lisofosfolipasas, fitasas, lipasa, monoacilipasas, palmitoil coA hidrolasa, fosfatasa ácida, gluconolactonasa, y fosfatasas.

La sigla LBE1, representa al Laboratorio de Biotecnología Enzimática y los números 1, 2, 3, 4, 5, 6, 7, 8, y 9, representan a las especies: *Aeribacillus composti*, *Anoxybacillus kamchatkensis*, *Bacillus coagulans*, *Bacillus haynesii*, *Bacillus hisashii*, *Bacillus licheniformes*, *Bacillus smithii*, *Bacillus subtilis*, y *Geobacillus kaustophilus*.

5.7. BIOPROSPECCIÓN DE ENZIMAS DE INTERÉS BIOTECNOLÓGICO

Obtenida la información del aislamiento de cepas bacterianas termófilas con un elevado grado de identificación entre el 99,2 y 99,9 %, es posible proceder a la bioprospección de enzimas termoestables de interés biotecnológico mediante el uso de la metabase que actualmente se encuentra a disposición gratuita para todo investigador que desea explotarla.

He tenido la oportunidad de utilizar la base de datos denominada PATRIC 3.6.10 que me ha permitido obtener una gran cantidad de información sobre las enzimas que pertenecen a la clasificación de Hidrolasas (EC 3.1.1.1) que al igual de la base de datos KEGG ENZYME: 3.1.1.1 me da información referencial.

Patric 3.6.10 me ha permitido obtener las enzimas hidrolasas de importancia biotecnológica, que se encuentran en el genoma de muchas especies de bacterias de diferentes procedencias y que han sido estudiados por

muchos investigadores del mundo y que han ido aportando poco a poco hasta lograr ahora una gran base de datos que nos permiten establecer las principales enzimas hidrolíticas que se encuentran como genes que codifican enzimas termoestables en las nueve especies de bacterias que han sido identificadas molecularmente en este estudio.

En función de su acción catalítica específica, las enzimas se clasifican en 6 grandes grupos o clases:

- Clase 1: OXIDORREDUCTASAS
- Clase 2: TRANSFERASAS
- Clase 3: HIDROLASAS
- Clase 4: LIASAS
- Clase 5: ISOMERASAS
- Clase 6: LIGASAS
- Clase 7: RIBOZIMAS

Se entiende por hidrólisis enzimática, la hidrólisis que se produce mediante un grupo de enzimas llamadas hidrolasas. Las enzimas que se utilizan en aplicaciones biotecnológicas son las de la clase 3, es decir hidrolasas, que comprenden actualmente 13 subclases, como se indican abajo:

EC 3.1 Actuando sobre enlaces éster

EC 3.2 Glucosilasas

EC 3.3 Actuando sobre enlaces éter

EC 3.4 Actúa sobre enlaces peptídicos (peptidasas)

EC 3.5 Actuando sobre enlaces carbono-nitrógeno

EC 3.6 Actúa sobre anhídridos de ácido

EC 3.7 Actuando sobre enlaces carbono-carbono

EC 3.8 Actuando sobre enlaces halogenuros

EC 3.9 Actuando sobre enlaces fósforo-nitrógeno

EC 3.10 Actuando sobre enlaces azufre-nitrógeno

EC 3.11 Actuando sobre enlaces carbono-fósforo

EC 3.12 Actuando sobre enlaces azufre-azufre

EC 3.13 Actuando sobre enlaces carbono-sulfuro

5.7.1. Bioprospección de enzimas termoestables lipolíticas

Las enzimas termoestables que pertenecen a la clase EC 3.1 que actúan sobre los enlaces éster, las denominamos como lipolíticas. Bajo el nombre de esterasas se comprenden una amplia cantidad de importantes enzimas que producen la ruptura del enlace éster formado por los ácidos inorgánicos (fosfórico, sulfúrico) u orgánicos diversos con diferentes alcoholes o tioles. Pese a las numerosas investigaciones existe todavía cierta incoherencia, dada la escasa especificidad de acción para muchas esterasas.

Se nombran mediante el nombre del sustrato seguido de la palabra hidrolasa, y cuando la enzima es específica para separar un grupo en particular, éste puede utilizarse como prefijo. En algunos casos este grupo puede ser transferido por la enzima a otras moléculas y se considera la hidrólisis misma como una transferencia del grupo al agua.

A veces suele utilizarse el nombre común de la enzima, por lo que muchas veces el sufijo –asa, nos indicará generalmente que se trata de una hidrolasa. Por ahora no ha sido posible encontrar una nomenclatura sistemática para todas las péptido-hidrolasas, por lo que hoy en día, para las nuevas enzimas, su nombre se forma según: fuente de péptido hidrolasa seguido, si es necesario, de una letra para distinguirla de otra enzima.

En la Tabla 8 puede apreciarse de manera general las enzimas que se encuentran presentes en cada una de las nueve especies de bacterias termófilas, así se han obtenido enzimas lipolíticas (que degradan los triglicéridos o grasas, aceites, además de degradar las esterasas. La tabla, además indica la

presencia de dichas enzimas en cada una de las especies obtenidas e identificadas como LBE-1 hasta LBE-9. Podemos apreciar el número de enzimas lipolíticas por cada una de las bacterias identificadas molecularmente, además de indicar el número y clase de enzima que están presentes en cada especie bacteriana. En la descripción de la figura 11 podemos observar la cantidad de enzimas por especie, indicado por cada una de las barras.

Las enzimas lipolíticas bacterianas han recibido un gran protagonismo debido a las ventajas que presentan: poseen diversas actividades catalíticas, se pueden obtener en gran cantidad, son económicas, se producen de forma regular, son estables y su proceso de producción es más factible y seguro, debido a que los microorganismos son fáciles de cultivar por sus simples exigencias nutritivas y su tiempo de generación más reducido (Olempska-Beer et al., 2006).

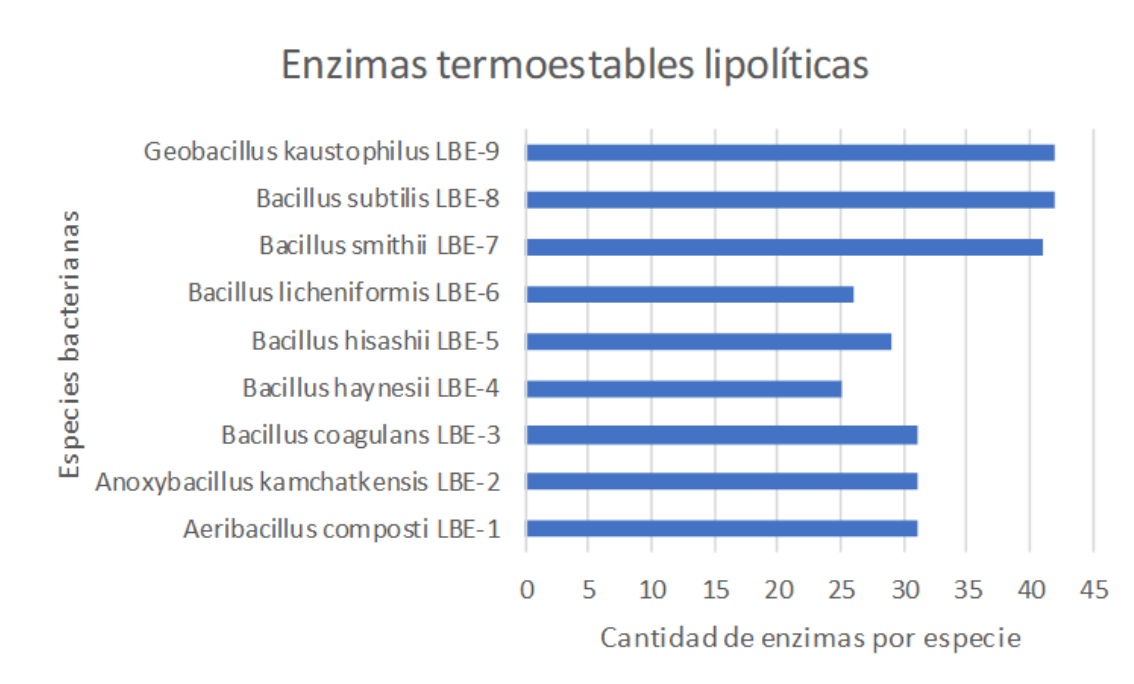
Las enzimas lipolíticas son biocatalizadores que llevan a cabo la síntesis, reacción de hidrólisis o transesterificación de la grasa. Las enzimas lipolíticas que se encuentran en la naturaleza se han utilizado para la producción de alimentos durante millones de años, pero en las últimas décadas han adquirido una gran importancia; debido a las múltiples aplicaciones tecnológicas en las que son capaces de actuar ya sus ventajas propias frente a los catalizadores convencionales no biológicos. Una de las principales fuentes de estos catalizadores es bacteriana, debido a que son fáciles y rápidos de cultivar, entre otras ventajas, y gracias a la biología molecular se pueden obtener en gran cantidad (Navarro-González y Periago, M.J., 2012).

Las fitasas son una clase especial de fosfatasas que catalizan la hidrólisis secuencial del fitato. La incapacidad de las plantas para utilizar el fósforo a partir de los fitatos del suelo es debido a la baja actividad de fitasas en sus raíces. Muchos de ellos pueden solubilizar el fósforo a partir de los fitatos, mediante la

liberación de fitasas. Este proceso permite la movilización del fósforo hacia las plantas y un mejor aprovechamiento de este nutriente. Una alternativa es mejorar las rizobacterias en cuanto a su capacidad de solubilizar los fitatos del suelo, mediante la transformación genética. En este trabajo el gen *phyL* de *Bacillus licheniformis* fue clonado en el vector de liberación suicida pJMT6 (vector derivado del sistema pUT/mini Tn5). La construcción recombinante que contiene un marcador de selección no antibiótico, fue transformada en *Escherichia coli* CC118/ pir. Un clon transformante (F16) fue seleccionado y posteriormente caracterizado (Fernández, M.T. et al., 2009).

Figura 11

Enzimas lipolíticas de especies bacterianas termófilas



La Figura 11 muestra las barras azules que representan la cantidad de enzimas por cada especie de bacteria termófila, siendo mayor en *Geobacillus kaustophilus*, seguido de *Bacillus subtilis* y *Bacillus smithii*.

En la Figura 11 se observa la distribución de enzimas lipolíticas termoestables presentes en las nueve especies de bacterias termófilas identificadas molecularmente usando el ADNr 16S. Como puede observarse, la enzima acetil esterasa se encuentra en todas las especies bacterianas. La carboxilesterasa en ocho especies, seguido de la monoglicérido lipasa, lisofosfolipasa, 3-fitasa, lipasa precursora, palmitoil hidrolasa, y las demás que se encuentran posteriormente.

5.7.2. Bioprospección de enzimas termoestables glucolíticas

Las enzimas termoestables que pertenecen a la clase EC 3.2 que actúan sobre los enlaces éster, las denominamos como glucosidasas. Las glucosidasas (también conocidas como glucósido hidrolasas) catalizan la hidrólisis de enlaces glucosídicos para generar glúcidos menores. Son enzimas extremadamente comunes con papeles importantes en la naturaleza como en la degradación de biomasa, como celulosa y hemicelulosa, en la defensa contra las bacterias, en mecanismos de patogénesis y en el normal funcionamiento celular. Junto a las glucotransferasas, las glucosidasas forman la mayor maquinaria catalítica para la síntesis y rotura de enlaces glucosídicos.

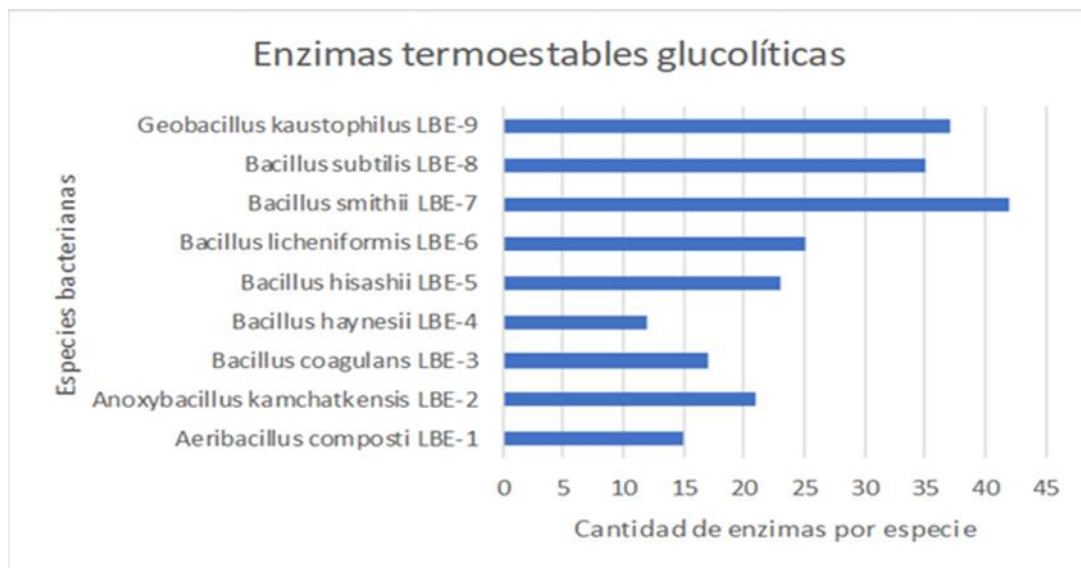
También se han obtenido 53 enzimas glucolíticas, como alfa-glucosidasas, beta-glucosidasas, betagalactosidasas, pullulanasa, levanasa, poligalactorunasa, chitinasas, alfa-amilasas, alfa-celulasas, alfa-xilanasas, alfa-galactosidasa, chitosanasa, etc.

La Figura 12 muestra la presencia de enzimas glucolíticas termoestables en todas las especies bacterianas aisladas y purificadas de los géiseres de Calientes-Candarave, mostrando una abundancia de la alfa-galactosidasa en 8 especies, seguida de pullulanasa (7 especies), oligo 1,6 glucosidasa, neopullulanasa y alfa-glucosidasa en 6 especies. Luego, las enzimas xilosidasa,

acetilglucosaminidasa, beta-galactosidasa y beta glucanasa; es decir, la celulasa, están presentes en 5 especies de bacterias termófilas. Contando las demás que se encuentran en 4, 3, 2, y 1 especie.

Figura 12

Enzimas glucolíticas de especies bacterianas termófilas



Debe tenerse presente que cada barra indica la presencia de cada enzima en el número de especies que se indica. En la Figura 12 puede observarse que *Bacillus smithii*, presentan mayor cantidad de enzimas glucolíticas, seguidor de las bacterias *Geobacillus kaustophilus* y *Bacillus subtilis*.

5.7.3. Bioprospección de enzimas termoestables proteolíticas

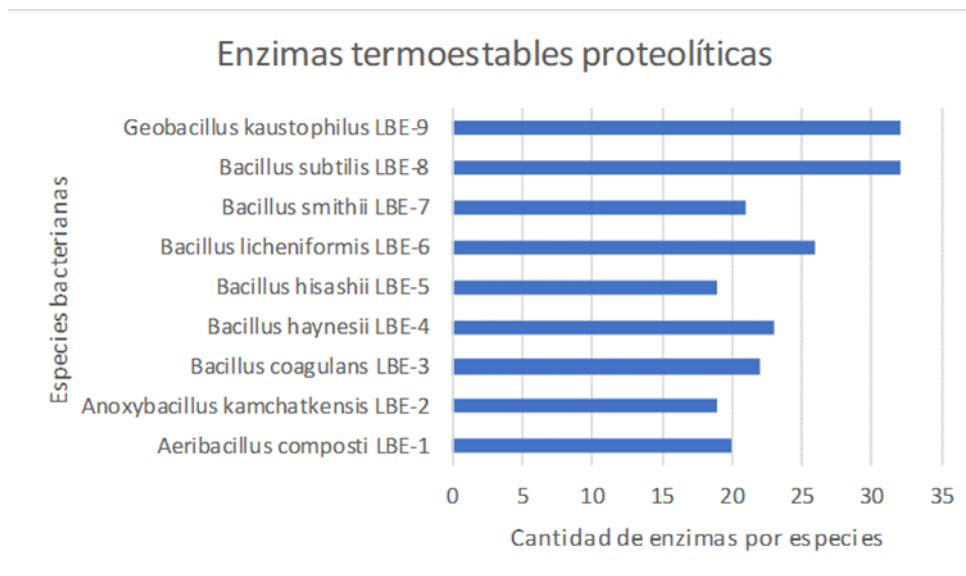
Las enzimas termoestables que pertenecen a la clase EC 3.4 que actúan sobre los enlaces peptídicos, las denominamos como proteolíticas o proteasas. Las proteasas también conocidas como peptidasas, son capaces de hidrolizar enlaces peptídicos de otras proteínas. Se encuentran distribuidas ampliamente en animales, plantas, y microorganismos y tienen gran importancia comercial, debido a que ocupan el mayor porcentaje de ventas en el mercado mundial de

enzimas. Sus aplicaciones son principalmente en las áreas médicas y farmacéuticas, así como en la industria alimentaria y de detergentes.

Su clasificación involucra dos grupos: exoproteasas y endoproteasas, dentro de estas últimas se encuentran las proteasas ácidas, mejor conocidas como aspartil proteasas, las cuales son de gran interés, principalmente en la industria de los alimentos y las bebidas. Aquellas de origen fúngico presentan gran demanda debido a las características que poseen, lo que ha dado como consecuencia el desarrollo de distintos métodos para aumentar su producción.

Figura 13

Enzimas proteolíticas de especies bacterianas termófilas



El objetivo de esta revisión bibliográfica, es dar a conocer las generalidades de las proteasas de origen fúngico, específicamente las de tipo aspartil, así como su importancia y aplicaciones biotecnológicas en la industria médica, farmacéutica, de las bebidas y principalmente en la alimenticia, además de hacer mención de los avances en los métodos para aumentar su producción.

Se han obtenido un total de 37 enzimas proteolíticas, que degradan proteínas constituidas por cadenas peptídicas, como aminopeptidadasas, proteasas, carboxipeptidasas, termolisina, subtilisina, serina proteasas, etc.

Las enzimas termoestables lipolíticas, glucolípidas y proteolíticas, se han obtenido utilizando la base de metadatos Patricbrc (<https://user.patricbrc.org/>), obteniendo la data existente por cada una de las cepas de especies de bacterias termófilas aisladas e identificadas utilizando las secuencias 16 S que han sido procesadas en el Laboratorio de Biotecnología utilizando los softwares BioEdit, Blast y Mega X.

Tabla 7.

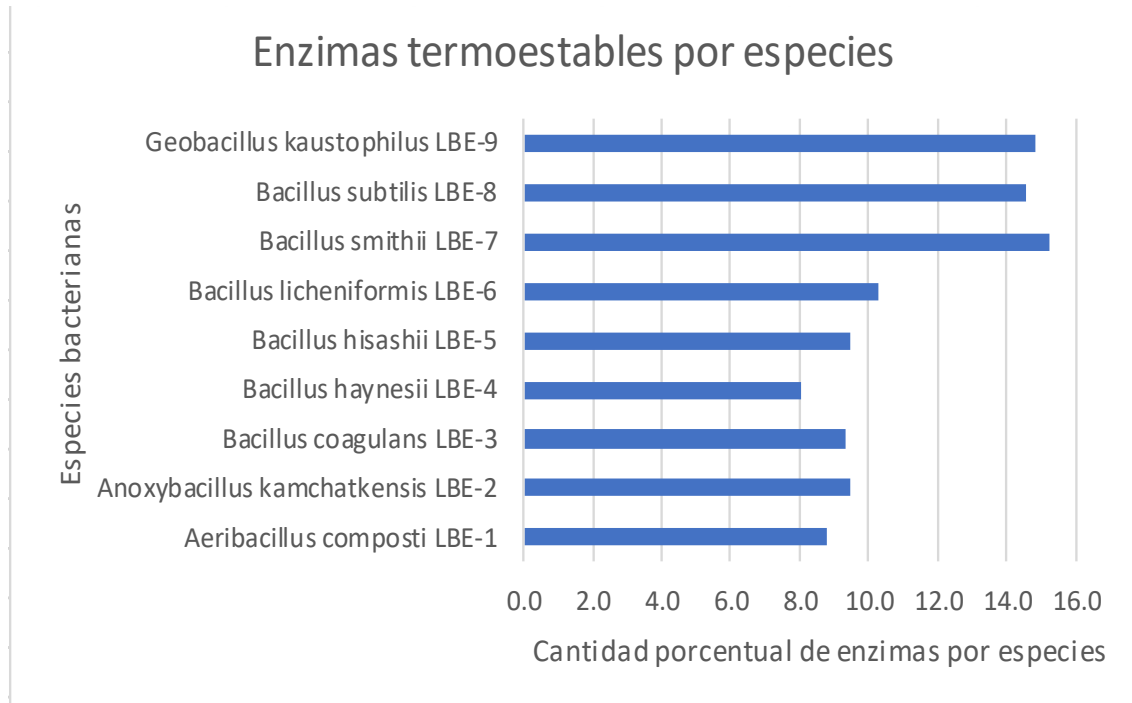
Enzimas hidrolíticas (Lip, Glu, Pro) por especies de bacterias

Especies de bacterias cultivables	Lip	Glu	Pro	Total	%
<i>Aeribacillus composti</i> LBE-1	31	15	20	66	9,3
<i>Anoxybacillus kamchatkensis</i> LBE-2	31	21	19	71	8,9
<i>Bacillus coagulans</i> LBE-3	31	17	22	70	10,3
<i>Bacillus haynesii</i> LBE-4	25	12	23	60	10,7
<i>Bacillus hisashii</i> LBE-5	29	23	19	61	8,9
<i>Bacillus licheniformis</i> LBE-6	26	25	26	77	12,1
<i>Bacillus smithii</i> LBE-7	41	42	21	114	9,8
<i>Bacillus subtilis</i> LBE-8	42	35	32	109	15,0
<i>Geobacillus kaustophilus</i> LBE-9	42	37	32	111	15,0
Total	298	227	224	744	100
%	39,8	30,3	29,9	100	

En la tabla 7 se puede apreciar la cantidad de enzimas lipolíticas (Lip) presentes por especies bacterianas, pudiéndose observar que la especie que presenta mayor cantidad de enzimas fue *Bacillus smithii* (114), seguido de *Geobacillus kaustophilus* (111) y *Bacillus subtilis* (109) y en menor cantidad, *Bacillus haynesii* (60), *Bacillus hasashii* (60), y *Aeribacillus composti* (66). También podemos observar que las lipo enzimas estuvieron presentes en mayor porcentaje (39,8 %) y en porcentajes de 30 %, las Glucoenzimas y Proteoenzimas.

Figura 14

Enzimas termoestables por especies bacterianas termófilas



En la Figura 14 se muestra en la abscisa las enzimas peptidasas presentes en cada una de las especies de cepas bacterianas termófilas. Puede apreciarse que la abscisa representa el número de especies bacterianas que contienen genes que codifican a dichas enzimas. La enzima carboxil-terminal proteasa se encuentra presente en la mayor parte de especies bacterianas, seguido de lipoproteína peptidasa y serina proteasa. Las aminopeptidasas que liberan los aminoácidos leucina, valina, fenilalanina y tirosina, están presentes en 5 especies de bacterias.

Las enzimas termoestable carboxipeptidasa, la señal peptidasa 1 y la bacilo peptidasa F se hallan presentes en 4 especies bacterianas y las demás enzimas se encuentran presentes en todas las especies bacterianas.

La relación de enzimas lipolíticas (esterasas y lipasas) por cada cepa de especie bacteriana, representado por la extensión LBE (Laboratorio de Biotecnología Enzimática), que son nueve especies bacterianas termófilas encontradas en las aguas termales.

Así, en función de las especies identificadas utilizando el gen ARNr 16S se han utilizado los datos correspondientes a las enzimas hidrolíticas encontradas en otros estudios realizados y obtenidos de dichas especies, utilizando los genomas completos e identificando los genes correspondientes, que están disponibles públicamente en PATRIC y las especies de bacterias termófilas cultivables que se han identificado, en un número de 9 especies, se ha encontrado un total de 744 enzimas hidrolíticas, siendo 298 lipolíticas, 227 glucolíticas y 224 proteolíticas.

Asimismo, podemos indicar que en cada una de las especies bacterianas, las enzimas lipolíticas estuvieron presentes entre 25 y 42 enzimas; las enzimas glucolíticas, entre 12 y 42 enzimas, y las proteolíticas entre 19 y 32 enzimas.

Esto constituye un buen indicador para la realización, de aislamientos de las enzimas que sean elegidas para estudios mucho más contundentes, es decir, de enzimas de importancia biotecnológica. En realidad, esto ahorra tiempo y dinero, que hasta ahora, los investigadores, solamente alcanzan a estudiar pocas enzimas, dado que la inducción para cada enzima representa la utilización de sustratos adecuados y ensayados anteriormente. La extracción de enzimas y su purificación, también representa una pérdida de tiempo, porque cada enzima, tiene su metodología propia.

Como resultado de una selección de microorganismos termofílicos, se aislaron 48 cepas de bacterias termófilas cultivables formadoras de esporas, de los sedimentos de los géiseres de Calientes-Candarave, Tacna-Perú. De las 48

cepas axénicas se obtuvieron 10 especies de bacterias termófilas, de las cuales 47 pertenecieron al filo Firmicutes y una al filo proteobacteria. El filo Firmicutes presentó una constancia de la Familia *Bacillaceae* con cuatros géneros distintos, *Aeribacillus*, *Anoxybacillus*, *Bacillus* y *Geobacillus*, con una estrecha relación entre ellos.

Asimismo, la abundancia de cada uno de los géneros presentes fue de 14 especies para *Aeribacillus composti*, 9 especies para *Bacillus hisashii*, 5 para *Geobacillus kaustophilus*, 5 para *Bacillus licheniformis*, 4 para *Anoxybacillus kamchakensis*, 3 especies para *Bacillus subtilis*, 3 especies para *Bacillus coagulans*, 3 especies para *Bacillus haynesii* y uno para *Bacillus smithii*. Se debe indicar que el porcentaje de similitud existente se encuentra por encima del 99 por ciento, esto es importante porque nos asegura con mayor certeza que en la búsqueda de enzimas hidrolíticas existentes en cada una de las especies pueden considerarse como probablemente cierta de acuerdo a la base de datos existentes sobre el particular.

Utilizando la base de datos de Bac Dive, se ha determinado la organización taxonómica de cada una de las especies identificadas. Asimismo, utilizando la base de datos de Patricbrc, se ha obtenido de cada de las especies, una gran cantidad de información, respecto solamente a las enzimas hidrolíticas que son motivo de nuestro estudio. Este proceso no ha ahorrado mucho tiempo y también en el aspecto económico, en el gasto de muchos medios de cultivo que se tendría que utilizar durante el desarrollo del estudio.

Con los datos encontrados, es más rápido, determinar distintas estrategias de estudio, en la obtención en físico, de la enzima seleccionada, además de su purificación y liofilización.

Existe un total de 9 especies distintas de cepas de bacterias termófilas y lo que se ha realizado después, es identificar por medio de la Data Base, que enzimas hidrolíticas podrían encontrarse en cada especie determinada.

En un ambiente por encima de los 70 °C, se considera que los termófilos no pueden existir, debido al crecimiento rápido de las bacterias termófilas formadoras de esporas, algunas de las cuales tienen un tiempo de generación tan rápido como 10 min, pero en el laboratorio, disminuyendo la temperatura a 65 °C duran más tiempo con vida, convirtiéndose en esporas cuando sus condiciones no sean las necesarias para su sobrevivencia.

Tabla 8.

Enzimas lipolíticas en especies de bacterias termófilas.

1. *Aeribacillus composti*

6-phosphogluconolactonase (EC 3.1.1.31)
Lysophospholipase (EC 3.1.1.5)
Monoglyceride lipase (EC 3.1.1.23)
Putative acetyl esterase YjcH (EC 3.1.1.-)
Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
Exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
Ribonuclease Z (EC 3.1.26.11)
Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)
GTP pyrophosphokinase (EC 2.7.6.5)
D-aminoacyl-tRNA deacylase (EC 3.1.1.96)
Endonuclease IV (EC 3.1.21.2)
Alpha-ribazole-5'-phosphate phosphatase (EC 3.1.3.73)
Phosphoserine phosphatase RsbX (EC 3.1.3.3)
Inositol-1-monophosphatase (EC 3.1.3.25)
Alkaline phosphatase (EC 3.1.3.1)
5'-nucleotidase SurE (EC 3.1.3.5)
Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)
Inorganic pyrophosphatase PpaX (EC 3.1.3.18)
Hydroxyacylglutathione hydrolase (EC 3.1.2.6)
Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)
Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)
Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)
Ribonuclease HIII (EC 3.1.26.4)
Ribonuclease HII (EC 3.1.26.4)
Chemotaxis response regulator protein-glutamate methyltransferase CheB (EC 3.1.1.61)
Ribonuclease III (EC 3.1.26.3)
Carboxylesterase (EC 3.1.1.1)

Manganese-dependent protein-tyrosine phosphatase (EC 3.1.3.48)
Histidinol-phosphatase (EC 3.1.3.15)
Carboxylesterase (EC 3.1.1.1)
5-methylcytosine-specific restriction enzyme A (EC 3.1.21.-)
Ribonuclease M5 (EC 3.1.26.8)
Peptidyl-tRNA hydrolase (EC 3.1.1.29)
Acyl-CoA hydrolase (EC 3.1.2.20)
2',3'-cyclic-nucleotide 2'-phosphodiesterase (EC 3.1.4.16)
Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)
Ribonuclease P protein component (EC 3.1.26.5)
D-glycero-D-manno-heptose 1,7-bisphosphate phosphatase (EC 3.1.1.-)
Alpha-ribazole-5'-phosphate phosphatase (EC 3.1.3.73)
2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase (EC 3.1.3.87)

2. *Anoxybacillus kamchatkensis*

Carboxylesterase (EC 3.1.1.1)
5'-nucleotidase (EC 3.1.3.5)
Histidinol-phosphatase (EC 3.1.3.15)
Endonuclease IV (EC 3.1.21.2)
Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
Exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
Ribonuclease Z (EC 3.1.26.11)
Histidinol-phosphatase (EC 3.1.3.15)
Lysophospholipase (EC 3.1.1.5); Monoglyceride lipase (EC 3.1.1.23)
D-aminoacyl-tRNA deacylase (EC 3.1.1.96)
Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)
GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase II
Ribonuclease III (EC 3.1.26.3)
Chemotaxis response regulator protein-glutamate methyltransferase CheB (EC 3.1.1.61)
Ribonuclease HII (EC 3.1.26.4)
Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
Phage exonuclease (EC 3.1.11.3)
2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase (EC 3.1.3.87)
Ribonuclease P protein component (EC 3.1.26.5)
Pimeloyl-[acyl-carrier protein] methyl ester esterase BioH (EC 3.1.1.85)
Inositol-1-monophosphatase (EC 3.1.3.25)
Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)
Peptidyl-tRNA hydrolase (EC 3.1.1.29)
Ribonuclease M5 (EC 3.1.26.8)
Acyl-CoA hydrolase (EC 3.1.2.20)
Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)
Putative acetyl esterase YjcH (EC 3.1.1.-)
Phage serine/threonine protein phosphatase NinI (EC 3.1.3.16)
Phage exonuclease (EC 3.1.11.3)
Phage serine/threonine protein phosphatase NinI (EC 3.1.3.16)

3. *Bacillus coagulans*

1,4-dihydroxy-2-naphthoyl-CoA hydrolase EC 3.1.2.28
2',3'-cyclic-nucleotide 2'-phosphodiesterase EC 3.1.4.16
3'-nucleotidase EC 3.1.3.6
6-phosphogluconolactonase EC 3.1.1.31

Acid phosphatase EC 3.1.3.2
 Acyl-CoA hydrolase EC 3.1.2.20
 Alkaline phosphatase EC 3.1.3.1
 Carboxylesterase EC 3.1.1.1
 Chemotaxis response regulator protein-glutamate methyltransferase EC 3.1.1.61
 D-aminoacyl-tRNA deacylase EC 3.1.1.96
 Endonuclease IV EC 3.1.21.2
 Exodeoxyribonuclease III EC 3.1.11.2
 Exodeoxyribonuclease VII large subunit EC 3.1.11.6
 Exodeoxyribonuclease VII small subunit EC 3.1.11.6
 Fructose-1,6-bisphosphatase, GlpX type EC 3.1.3.11
 Glycerophosphoryl diester phosphodiesterase EC 3.1.4.46
 Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase EC 3.1.7.2
 Inositol-1-monophosphatase EC 3.1.3.25
 Lysophospholipase EC 3.1.1.5
 Monoglyceride lipase EC 3.1.1.23
 Manganese-dependent protein-tyrosine phosphatase EC 3.1.3.48
 Peptidyl-tRNA hydrolase EC 3.1.1.29
 Probable 2-phosphosulfolactate phosphatase EC 3.1.3.71
 Putative acetyl esterase YjcH EC 3.1.
 Ribonuclease HII EC EC 3.1.26.4 3.1.26.4
 Ribonuclease HIII EC 3.1.26.3
 Ribonuclease III EC 3.1.26.8
 Ribonuclease M5 EC 3.1.26.5
 Ribonuclease P protein component EC 3.1.26.11
 Ribonuclease Z EC 3.1.3.23
 Sugar phosphatase YidA EC 3.1.21.3

4. *Bacillus haynesii*

1,4-dihydroxy-2-naphthoyl-CoA hydrolase (EC 3.1.2.28)
 2',3'-cyclic nucleotide 3'-phosphodiesterase (EC 3.1.4.37)
 2',3'-cyclic-nucleotide 2'-phosphodiesterase (EC 3.1.4.16)
 5'-nucleotidase (EC 3.1.3.5)
 2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase (EC 3.1.3.87)
 (2)3-hydroxyisobutyryl-CoA hydrolase (EC 3.1.2.4)
 3-phytase (EC 3.1.3.8)
 6-phosphogluconolactonase (EC 3.1.1.31)
 Acyl-CoA hydrolase (EC 3.1.2.20)
 Alkaline phosphatase (EC 3.1.3.1)
 Carboxylesterase (EC 3.1.1.1) (3)
 Cephalosporin-C deacetylase (EC 3.1.1.41)
 Chemotaxis response regulator protein-glutamate methyltransferase CheB (EC 3.1.1.61)
 D-aminoacyl-tRNA deacylase (EC 3.1.1.96)
 Endonuclease IV (EC 3.1.21.2)
 Endonuclease V (EC 3.1.21.7)
 Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
 Fructose-1,6-bisphosphatase, Bacillus type (EC 3.1.3.11)
 Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)
 Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
 Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)
 Histidinol-phosphatase (EC 3.1.3.15)
 Inositol-1-monophosphatase (EC 3.1.3.25)
 Lipase precursor (EC 3.1.1.3)

Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)
Lysophospholipase (EC 3.1.1.5); Monoglyceride lipase (EC 3.1.1.23)
Manganese-dependent protein-tyrosine phosphatase (EC 3.1.3.48)
Palmitoyl-CoA hydrolase precursor (EC 3.1.2.2)
Peptidyl-tRNA hydrolase (EC 3.1.1.29)
Phage exonuclease (EC 3.1.11.3)
Phosphoserine phosphatase RsbX (EC 3.1.3.3)
Pseudouridine-5' phosphatase (EC 3.1.3.-)
Putative acetyl esterase YjcH (EC 3.1.1.-)
Pyrimidine 5'-nucleotidase YjjG (EC 3.1.3.5)
Ribonuclease HII (EC 3.1.26.4)
Ribonuclease HIII (EC 3.1.26.4)
Ribonuclease III (EC 3.1.26.3)
Ribonuclease M5 (EC 3.1.26.8)
Ribonuclease P protein component (EC 3.1.26.5)
Ribonuclease Z (EC 3.1.26.11)
Sphingomyelinase (EC 3.1.4.12)
Sugar phosphatase YidA (EC 3.1.3.23)

5. *Bacillus hisashii*

1,4-dihydroxy-2-naphthoyl-CoA hydrolase (EC 3.1.2.28)
5'-nucleotidase (EC 3.1.3.5) (2)
6-phosphogluconolactonase (EC 3.1.1.31)
Acyl-CoA hydrolase (EC 3.1.2.20)
Alkaline phosphatase (EC 3.1.3.1)
Carboxylesterase (EC 3.1.1.1)
Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)
D-aminoacyl-tRNA deacylase (EC 3.1.1.96)
Endonuclease IV (EC 3.1.21.2)
Exodeoxyribonuclease III (EC 3.1.11.2)
Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)
Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)
Inositol-1-monophosphatase (EC 3.1.3.25)
Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)
Lysophospholipase (EC 3.1.1.5); Monoglyceride lipase (EC 3.1.1.23)
Manganese-dependent protein-tyrosine phosphatase (EC 3.1.3.48)
Peptidyl-tRNA hydrolase (EC 3.1.1.29)
Phosphoglycolate phosphatase (EC 3.1.3.18)
Probable 2-phosphosulfolactate phosphatase (EC 3.1.3.71)
Putative acetyl esterase YjcH (EC 3.1.1.-)
Ribonuclease HII (EC 3.1.26.4)
Ribonuclease III (EC 3.1.26.3)
Ribonuclease M5 (EC 3.1.26.8)
Ribonuclease Z (EC 3.1.26.11)
Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)

6. *Bacillus licheniformis*

1,4-dihydroxy-2-naphthoyl-CoA hydrolase (EC 3.1.2.28)
2',3'-cyclic nucleotide 3'-phosphodiesterase (EC 3.1.4.37)
2',3'-cyclic-nucleotide 2'-phosphodiesterase (EC 3.1.4.16)

5'-nucleotidase (EC 3.1.3.5)
 2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase (EC 3.1.3.87)
 3-hydroxyisobutyryl-CoA hydrolase (EC 3.1.2.4)
 3-phytase (EC 3.1.3.8)
 6-phosphogluconolactonase (EC 3.1.1.31)
 Acyl-CoA hydrolase (EC 3.1.2.20)
 Alkaline phosphatase (EC 3.1.3.1)
 Carboxylesterase (EC 3.1.1.1) (3)
 Cephalosporin-C deacetylase (EC 3.1.1.41)
 Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)
 D-aminoacyl-tRNA deacylase (EC 3.1.1.96)
 Endonuclease IV (EC 3.1.21.2)
 Endonuclease V (EC 3.1.21.7) (2)
 Exodeoxyribonuclease VII large subunit (EC 3.1.11.6) (2)
 Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)
 Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46) (4)
 Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)
 Histidinol-phosphatase (EC 3.1.3.15)
 Inositol-1-monophosphatase (EC 3.1.3.25)
 Lipase precursor (EC 3.1.1.3)
 Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)
 Lysophospholipase (EC 3.1.1.5); Monoglyceride lipase (EC 3.1.1.23)
 Manganese-dependent protein-tyrosine phosphatase (EC 3.1.3.48)
 Palmitoyl-CoA hydrolase precursor (EC 3.1.2.2) (2)
 Peptidyl-tRNA hydrolase (EC 3.1.1.29)
 Phage exonuclease (EC 3.1.11.3)
 Phosphoserine phosphatase RsbX (EC 3.1.3.3)
 Pseudouridine-5' phosphatase (EC 3.1.3.-)
 Putative acetyl esterase YjcH (EC 3.1.1.-)
 Pyrimidine 5'-nucleotidase YjjG (EC 3.1.3.5)
 Ribonuclease HII (EC 3.1.26.4)
 Ribonuclease HIII (EC 3.1.26.4)
 Ribonuclease III (EC 3.1.26.3)
 Ribonuclease M5 (EC 3.1.26.8)
 Ribonuclease P protein component (EC 3.1.26.5)
 Ribonuclease Z (EC 3.1.26.11)
 Sphingomyelinase (EC 3.1.4.12)
 Sugar phosphatase YidA (EC 3.1.3.23)

7. *Bacillus smithii*

1,4-dihydroxy-2-naphthoyl-CoA hydrolase (EC 3.1.2.28)
 2',3'-cyclic-nucleotide 2'-phosphodiesterase (EC 3.1.4.16)
 2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase (EC 3.1.3.87)
 Acyl-CoA hydrolase (EC 3.1.2.20)
 Alkaline phosphatase (EC 3.1.3.1)
 Carboxylesterase (EC 3.1.1.1)
 Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)
 D-aminoacyl-tRNA deacylase (EC 3.1.1.96)
 Endonuclease IV (EC 3.1.21.2)
 Exodeoxyribonuclease III (EC 3.1.11.2)
 Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
 Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)
 Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)

Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)
 Histidinol-phosphatase (EC 3.1.3.15)
 Inositol-1-monophosphatase (EC 3.1.3.25)
 Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)
 Lysophospholipase (EC 3.1.1.5); Monoglyceride lipase (EC 3.1.1.23)
 Manganese-dependent protein-tyrosine phosphatase (EC 3.1.3.48)
 Peptidyl-tRNA hydrolase (EC 3.1.1.29)
 Putative acetyl esterase YjcH (EC 3.1.1.-)
 Ribonuclease HII (EC 3.1.26.4)
 Ribonuclease HIII (EC 3.1.26.4)
 Ribonuclease III (EC 3.1.26.3)
 Ribonuclease M5 (EC 3.1.26.8)
 Ribonuclease P protein component (EC 3.1.26.5)
 Ribonuclease Z (EC 3.1.26.11)

8. *Bacillus subtilis*

1,4-dihydroxy-2-naphthoyl-CoA hydrolase (EC 3.1.2.28)
 2',3'-cyclic-nucleotide 2'-phosphodiesterase (EC 3.1.4.16)
 5'-nucleotidase (EC 3.1.3.5)
 2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase (EC 3.1.3.87)
 3-phytase (EC 3.1.3.8)
 6-phosphogluconolactonase (EC 3.1.1.31)
 Acyl-CoA hydrolase (EC 3.1.2.20)
 Alkaline phosphatase (EC 3.1.3.1)
 Carboxylesterase (EC 3.1.1.1)
 Cephalosporin-C deacetylase (EC 3.1.1.41)
 Chemotaxis response regulator protein-glutamate methylesterase CheB (EC 3.1.1.61)
 D-aminoacyl-tRNA deacylase (EC 3.1.1.96)
 Endonuclease IV (EC 3.1.21.2)
 Endonuclease V (EC 3.1.21.7)
 Exodeoxyribonuclease III (EC 3.1.11.2)
 Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
 Exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
 Fructose-1,6-bisphosphatase, *Bacillus* type (EC 3.1.3.11)
 Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)
 Gluconolactonase (EC 3.1.1.17)
 Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
 Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)
 Histidinol-phosphatase (EC 3.1.3.15)
 Inositol-1-monophosphatase (EC 3.1.3.25)
 Lipase precursor (EC 3.1.1.3)
 Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)
 Lysophospholipase (EC 3.1.1.5); Monoglyceride lipase (EC 3.1.1.23)
 Manganese-dependent protein-tyrosine phosphatase (EC 3.1.3.48)
 Peptidyl-tRNA hydrolase (EC 3.1.1.29)
 Phage exonuclease (EC 3.1.11.3)
 Phosphoserine phosphatase RsbX (EC 3.1.3.3)
 Probable 2-phosphosulfolactate phosphatase (EC 3.1.3.71)
 Pseudouridine-5' phosphatase (EC 3.1.3.-)
 Putative acetyl esterase YjcH (EC 3.1.1.-)
 Pyrimidine 5'-nucleotidase YjgG (EC 3.1.3.5)
 Ribonuclease HII (EC 3.1.26.4)
 Ribonuclease HIII (EC 3.1.26.4)

Ribonuclease III (EC 3.1.26.3)
Ribonuclease M5 (EC 3.1.26.8)
Ribonuclease P protein component (EC 3.1.26.5)
Ribonuclease Z (EC 3.1.26.11)
Sugar phosphatase YidA (EC 3.1.3.23)

9. *Geobacillus kaustophilus*

1,4-dihydroxy-2-naphthoyl-CoA hydrolase (EC 3.1.2.28)
2',3'-cyclic-nucleotide 2'-phosphodiesterase (EC 3.1.4.16)
3'-nucleotidase (EC 3.1.3.6)
2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase (EC 3.1.3.87)
5'-nucleotidase (EC 3.1.3.5)
Alkaline phosphatase (EC 3.1.3.1)
Carboxylesterase (EC 3.1.1.1)
Chemotaxis response regulator protein-glutamate methyltransferase CheB (EC 3.1.1.61)
D-aminoacyl-tRNA deacylase (EC 3.1.1.96)
Endonuclease IV (EC 3.1.21.2)
Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
Exodeoxyribonuclease VII small subunit (EC 3.1.11.6)
Fructose-1,6-bisphosphatase, GlpX type (EC 3.1.3.11)
Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)
Guanosine-3',5'-bis(diphosphate) 3'-pyrophosphohydrolase (EC 3.1.7.2)
Histidinol-phosphatase (EC 3.1.3.15)
Inositol-1-monophosphatase (EC 3.1.3.25)
Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)
Peptidyl-tRNA hydrolase (EC 3.1.1.29)
Phage exonuclease (EC 3.1.11.3)
Pimeloyl-[acyl-carrier protein] methyl ester esterase BioH (EC 3.1.1.85)
Probable 2-phosphosulfolactate phosphatase (EC 3.1.3.71)
Putative acetyl esterase YjcH (EC 3.1.1.-)
Ribonuclease HII (EC 3.1.26.4)
Ribonuclease HIII (EC 3.1.26.4)
Ribonuclease III (EC 3.1.26.3)
Ribonuclease M5 (EC 3.1.26.8)
Ribonuclease P protein component (EC 3.1.26.5)
Ribonuclease Z (EC 3.1.26.11)
Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)

Nora. Elaboración propia.

Tabla 9

Enzimas glucolíticas en especies de bacterias termófilas

1. *Aeribacillus composti*

Alpha-galactosidase (EC 3.2.1.22)
Neopullulanase (EC 3.2.1.135)
5'-methylthioadenosine nucleosidase (EC 3.2.2.16) @ S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)
beta-N-acetylglucosaminidase (EC 3.2.1.52)
Maltose-6'-phosphate glucosidase (EC 3.2.1.122)
Uracil-DNA glycosylase, family 1 (EC 3.2.2.27)
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)
6-phospho-beta-glucosidase (EC 3.2.1.86)
6-phospho-beta-glucosidase (EC 3.2.1.86)
Membrane-bound lytic murein transglycosylase D precursor (EC 3.2.1.-)
A/G-specific adenine glycosylase (EC 3.2.2.-)
DNA-3-methyladenine glycosylase II (EC 3.2.2.21)
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)
Maltose-6'-phosphate glucosidase (EC 3.2.1.122)

2. *Anoxybacillus kamchatkensis*

Glucan 1,4-alpha-maltohexaosidase (EC 3.2.1.98)
alpha-amylase (EC 3.2.1.1) @ Pullulanase (EC 3.2.1.41)
beta-fructofuranosidase (EC 3.2.1.26)
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)
Oligo-1,6-glucosidase (EC 3.2.1.10)
Neopullulanase (EC 3.2.1.135)
Alpha-glucosidase (EC 3.2.1.20)
5'-methylthioadenosine nucleosidase (EC 3.2.2.16) @ S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)
Xylan 1,4-beta-xylosidase (EC 3.2.1.37)
beta-galactosidase (EC 3.2.1.23)
alpha-galactosidase (EC 3.2.1.22)
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)
Pullulanase (EC 3.2.1.41)
Oligo-1,6-glucosidase (EC 3.2.1.10)
Uracil-DNA glycosylase, family 1 (EC 3.2.2.27)
Sucrose-6-phosphate hydrolase (EC 3.2.1.26)
Sucrose-6-phosphate hydrolase (EC 3.2.1.26)
DNA-3-methyladenine glycosylase II (EC 3.2.2.21)
A/G-specific adenine glycosylase (EC 3.2.2.-)
Alpha-glucosidase (EC 3.2.1.20)
Phage lysozyme R (EC 3.2.1.17)

3. *Bacillus coagulans*

5'-methylthioadenosine nucleosidase EC 3.2.2.9
S-adenosylhomocysteine nucleosidase EC 3.2.1.86
6-phospho-beta-glucosidase EC 3.2.2.-
A/G-specific adenine glycosylase EC 3.2.1.22

Alpha-galactosidase EC 3.2.1.23
Beta-galactosidase EC 3.2.2.21
DNA-3-methyladenine glycosylase II EC 3.2.2.23
Formamidopyrimidine-DNA glycosylase EC 3.2.1.98
Glucan 1,4-alpha-maltohexaosidase EC 3.2.1.20
Maltodextrin glucosidase EC 3.2.1.122
Maltose-6'-phosphate glucosidase EC 3.2.1.135
Neopullulanase EC 3.2.1.10
Oligo-1,6-glucosidase EC 3.2.1.-
Spore peptidoglycan hydrolase (N-acetylglucosaminidase) EC 3.2.1.93
Trehalose-6-phosphate hydrolase EC 3.2.2.27
Uracil-DNA glycosylase, family 1 EC 3.2.2.16

4. *Bacillus haynesii*

2,6-beta-fructan 6-levanbiohydrolase (EC 3.2.1.64)
5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)
6-phospho-beta-glucosidase (EC 3.2.1.86)
A/G-specific adenine glycosylase (EC 3.2.2.-)
Alpha-galactosidase (EC 3.2.1.22)
Alpha-glucosidase (EC 3.2.1.20) (2)
alpha-L-arabinofuranosidase (EC 3.2.1.55)
alpha-xylosidase (EC 3.2.1.177)
Arabinan endo-1,5-alpha-L-arabinosidase (EC 3.2.1.99) (2)
Arabinogalactan endo-1,4-beta-galactanase (EC 3.2.1.89)
Beta-1,4-glucanase (cellulase) (EC 3.2.1.4) (2)
Beta-fructofuranosidase (EC 3.2.1.26)
Beta-galactosidase (EC 3.2.1.23) (3)
Beta-glucosidase (EC 3.2.1.21) (4)
Beta-N-acetylglucosaminidase (EC 3.2.1.52)
Cellulose 1,4-beta-cellobiosidase (reducing end) (EC 3.2.1.176)
Chitinase (EC 3.2.1.14)
Chitodextrinase (EC 3.2.1.14)
DNA-3-methyladenine glycosylase II (EC 3.2.2.21) (2)
Endo-beta-1,3-1,4 glucanase (licheninase) (EC 3.2.1.73)
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)
Glucan 1,4-alpha-maltohexaosidase (EC 3.2.1.98)
Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)
Levanase (EC 3.2.1.65)
Maltose-6'-phosphate glucosidase (EC 3.2.1.122)
Mannan endo-1,4-beta-mannosidase (EC 3.2.1.78)
Mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase (EC 3.2.1.96)
Neopullulanase (EC 3.2.1.135)
Oligo-1,6-glucosidase (EC 3.2.1.10)
Polygalacturonase (EC 3.2.1.15)
Pullulanase (EC 3.2.1.41)
Sucrose-6-phosphate hydrolase (EC 3.2.1.26)
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)
Unsaturated rhamnogalacturonyl hydrolase (EC 3.2.1.172) (2)
Uracil-DNA glycosylase, family 1 (EC 3.2.2.27)
Xylan 1,4-beta-xylosidase (EC 3.2.1.37) (2)

5. *Bacillus hisashii*

5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)
6-phospho-beta-glucosidase (EC 3.2.1.86) (6)
A/G-specific adenine glycosylase (EC 3.2.2.-)
alpha-galactosidase (EC 3.2.1.22)
Alpha-glucosidase (EC 3.2.1.20)
Beta-fructofuranosidase (EC 3.2.1.26)
Beta-galactosidase (EC 3.2.1.23)
Beta-glucosidase (EC 3.2.1.21)
DNA-3-methyladenine glycosylase (EC 3.2.2.20)
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)(2)
Levanase (EC 3.2.1.65)
Maltodextrin glucosidase (EC 3.2.1.20)
Maltose-6'-phosphate glucosidase (EC 3.2.1.122) (2)
Mannosylglycerate hydrolase (EC 3.2.1.170)
Oligo-1,6-glucosidase (EC 3.2.1.10)
Pullulanase (EC 3.2.1.41)
Trehalose-6-phosphate hydrolase (EC 3.2.1.93) (2)
Uracil-DNA glycosylase, family 1 (EC 3.2.2.27)
Xylan 1,4-beta-xylosidase (EC 3.2.1.37)

6. *Bacillus licheniformis*

2,6-beta-fructan 6-levanbiohydrolase (EC 3.2.1.64)
5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)
6-phospho-beta-glucosidase (EC 3.2.1.86)
A/G-specific adenine glycosylase (EC 3.2.2.-)
alpha-galactosidase (EC 3.2.1.22)
Alpha-glucosidase (EC 3.2.1.20) (2)
Alpha-L-arabinofuranosidase (EC 3.2.1.55)
Alpha-xylosidase (EC 3.2.1.177)
Arabinan endo-1,5-alpha-L-arabinosidase (EC 3.2.1.99)
Arabinogalactan endo-1,4-beta-galactanase (EC 3.2.1.89)
Beta-1,4-glucanase (cellulase) (EC 3.2.1.4)(2)
Beta-fructofuranosidase (EC 3.2.1.26)
Beta-galactosidase (EC 3.2.1.23) (4)
Beta-glucosidase (EC 3.2.1.21) (2)
Beta-N-acetylglucosaminidase (EC 3.2.1.52)
Cellulose 1,4-beta-cellobiosidase (reducing end) (EC 3.2.1.176)
Chitinase (EC 3.2.1.14)
Chitodextrinase (EC 3.2.1.14)
Cyclo(L-leucyl-L-leucyl) synthase (EC 2.3.2.22)
DNA-3-methyladenine glycosylase II (EC 3.2.2.21)
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)
Gamma-glutamyltranspeptidase (EC 2.3.2.2)
Gellan 1,4-alpha-maltohexaosidase (EC 3.2.1.98)
Inosine-uridine preferring nucleoside hydrolase (EC 3.2.2.1)
Isochorismatase (EC 3.3.2.1) [bacillibactin] siderophore
Levanase (EC 3.2.1.65)
L-O-lysylphosphatidylglycerol synthase (EC 2.3.2.3)
Maltose-6'-phosphate glucosidase (EC 3.2.1.122)

Mannan endo-1,4-beta-mannosidase (EC 3.2.1.78)
Mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase (EC 3.2.1.96)
Neopullulanase (EC 3.2.1.135)
Oligo-1,6-glucosidase (EC 3.2.1.10)
Polygalacturonase (EC 3.2.1.15)
Pullulanase (EC 3.2.1.41)
Sucrose-6-phosphate hydrolase (EC 3.2.1.26)
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)
Unsaturated rhamnogalacturonyl hydrolase (EC 3.2.1.172) (2)
Uracil-DNA glycosylase, family 1 (EC 3.2.2.27)
Xylan 1,4-beta-xylosidase (EC 3.2.1.37)

7. *Bacillus smithii*

5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)
6-phospho-beta-glucosidase (EC 3.2.1.86)
A/G-specific adenine glycosylase (EC 3.2.2.-)
Alpha-galactosidase (EC 3.2.1.22)
Alpha-glucosidase (EC 3.2.1.20)
DNA-3-methyladenine glycosylase II (EC 3.2.2.21) (2)
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)
Maltose-6'-phosphate glucosidase (EC 3.2.1.122)
Pullulanase (EC 3.2.1.41)
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)
Uracil-DNA glycosylase, family 1 (EC 3.2.2.27)

8. *Bacillus subtilis*

2,6-beta-fructan 6-levanbiohydrolase (EC 3.2.1.64)
5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)
6-phospho-beta-glucosidase (EC 3.2.1.86) (5)
A/G-specific adenine glycosylase (EC 3.2.2.-)
Alpha-amylase (EC 3.2.1.1)
Alpha-galactosidase (EC 3.2.1.22)
Alpha-galacturonidase (EC 3.2.1.67)
Alpha-glucosidase (EC 3.2.1.20)
Alpha-L-arabinofuranosidase (EC 3.2.1.55) (2)
Arabinan endo-1,5-alpha-L-arabinosidase (EC 3.2.1.99) (2)
Arabinogalactan endo-1,4-beta-galactanase (EC 3.2.1.89)
Arabinoxylan arabinofuranohydrolase (EC 3.2.1.55)
Beta-1,4-glucanase (cellulase) (EC 3.2.1.4)
Beta-galactosidase (EC 3.2.1.23) (2)
Beta-N-acetylglucosaminidase (EC 3.2.1.52)
Chitosanase precursor (EC 3.2.1.132)
DNA-3-methyladenine glycosylase II (EC 3.2.2.21) (2)
Endo-1,4-beta-xylosidase (EC 3.2.1.8)
Endo-beta-1,3-1,4 glucanase (licheninase) (EC 3.2.1.73)
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)
Glucuronoarabinoxylan endo-1,4-beta-xylosidase (EC 3.2.1.136)
Levanase (EC 3.2.1.65)
Maltose-6'-phosphate glucosidase (EC 3.2.1.122)
Mannan endo-1,4-beta-mannosidase (EC 3.2.1.78)

Mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase (EC 3.2.1.96)
Neopullulanase (EC 3.2.1.135)
Oligo-1,6-glucosidase (EC 3.2.1.10) (2)
Pullulanase (EC 3.2.1.41)
Sucrose-6-phosphate hydrolase (EC 3.2.1.26)
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)
Unsaturated rhamnogalacturonyl hydrolase (EC 3.2.1.172) (2)
Uracil-DNA glycosylase, family 1 (EC 3.2.2.27)
Xylan 1,4-beta-xylosidase (EC 3.2.1.37)

9. *Geobacillus kaustophilus*

5'-methylthioadenosine nucleosidase (EC 3.2.2.16)
S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)
6-phospho-beta-glucosidase (EC 3.2.1.86) (5)
A/G-specific adenine glycosylase (EC 3.2.2.-)
alpha-amylase (EC 3.2.1.1) @ Pullulanase (EC 3.2.1.41)
alpha-galactosidase (EC 3.2.1.22)
Alpha-glucosidase (EC 3.2.1.20)
alpha-L-arabinofuranosidase (EC 3.2.1.55)
Arabinogalactan endo-1,4-beta-galactanase (EC 3.2.1.89)
beta-galactosidase (EC 3.2.1.23)
beta-N-acetylglucosaminidase (EC 3.2.1.52)
DNA-3-methyladenine glycosylase II (EC 3.2.2.21)
Endo-1,4-beta-xylanase (EC 3.2.1.8) (2)
Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)
Glucan 1,4-alpha-maltohexaosidase (EC 3.2.1.98)
Neopullulanase (EC 3.2.1.135) (2)
Pullulanase (EC 3.2.1.41)
Spore peptidoglycan hydrolase (N-acetylglucosaminidase) (EC 3.2.1.-)
Trehalase (EC 3.2.1.28) @ Cytoplasmic trehalase (EC 3.2.1.28)
Trehalose-6-phosphate hydrolase (EC 3.2.1.93)
Uracil-DNA glycosylase, family 1 (EC 3.2.2.27)
Xylan 1,4-beta-xylosidase (EC 3.2.1.37)
Xylan alpha-1,2-glucuronosidase (EC 3.2.1.131)

Tabla 10

Enzimas proteolíticas en especies de bacterias termófilas

1. *Aeribacillus composti*

Signal peptidase I (EC 3.4.21.89)
Glutamyl aminopeptidase (EC 3.4.11.7); Deblocking aminopeptidase
Spore germination endopeptidase Gpr (EC 3.4.24.78)
Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.24)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
Carboxyl-terminal protease (EC 3.4.21.102)
Gamma-D-glutamyl-meso-diaminopimelate peptidase (EC 3.4.19.11)
Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)
Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)
Signal peptidase SipW (EC 3.4.21.89), required for TasA secretion
ATP-dependent Clp protease proteolytic subunit ClpP (EC 3.4.21.92)
Carboxyl-terminal protease (EC 3.4.21.102)
Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)
ATP-dependent protease La (EC 3.4.21.53) LonB Type I
ATP-dependent protease La (EC 3.4.21.53) Type I
Lipoprotein signal peptidase (EC 3.4.23.36)
Serine protease (EC 3.4.21.-)
Thermostable carboxypeptidase 1 (EC 3.4.17.19)
Signal peptidase I (EC 3.4.21.89)
ATP-dependent protease subunit HslV (EC 3.4.25.2)
Cytosol aminopeptidase PepA (EC 3.4.11.1)
Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)
Methionine aminopeptidase (EC 3.4.11.18)
Leader peptidase (Prepilin peptidase) (EC 3.4.23.43) / N-methyltransferase (EC 2.1.1.-)
SOS-response repressor and protease LexA (EC 3.4.21.88)
Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
Serine protease, DegP/HtrA, do-like (EC 3.4.21.-)
Microsomal dipeptidase (EC 3.4.13.19)
Gamma-glutamyltranspeptidase (EC 2.3.2.2) @ Glutathione hydrolase (EC 3.4.19.13)
Serine protease, DegP/HtrA, do-like (EC 3.4.21.-)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)

2. *Anoxybacillus kamchatkensis*

Carboxyl-terminal protease (EC 3.4.21.102)
ATP-dependent Clp protease proteolytic subunit ClpP (EC 3.4.21.92)
Thermostable carboxypeptidase 1 (EC 3.4.17.19)
Hydrogenase maturation protease (EC 3.4.24.-)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
Spore germination endopeptidase Gpr (EC 3.4.24.78)
Gamma-D-glutamyl-L-diamino acid endopeptidase I (EC 3.4.19.11)
ATP-dependent protease La (EC 3.4.21.53) Type I
ATP-dependent protease La (EC 3.4.21.53) LonB Type I

Proline dipeptidase (EC 3.4.13.9)
 Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)
 Glutamyl aminopeptidase (EC 3.4.11.7); Deblocking aminopeptidase
 Leader peptidase (Prepilin peptidase) (EC 3.4.23.43) / N-methyltransferase (EC 2.1.1.-)
 Lipoprotein signal peptidase (EC 3.4.23.36)
 ATP-dependent protease subunit HslV (EC 3.4.25.2)
 Signal peptidase I (EC 3.4.21.89)
 SOS-response repressor and protease LexA (EC 3.4.21.88)
 Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.-.-)
 Serine protease, DegP/HtrA, do-like (EC 3.4.21.-)
 Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.24)
 SOS-response repressor and protease LexA (EC 3.4.21.88)
 Microsomal dipeptidase (EC 3.4.13.19)
 Methionine aminopeptidase (EC 3.4.11.18)
 D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)

3. *Bacillus coagulans*

Aminopeptidase C EC 3.4.22.40
 Aminopeptidase S (Leu, Val, Phe, Tyr preference) EC 3.4.11.24
 ATP-dependent Clp protease proteolytic subunit ClpP EC 3.4.21.92
 ATP-dependent protease La LonB Type I EC 3.4.21.53
 ATP-dependent protease La Type II EC 3.4.21.53
 ATP-dependent protease subunit HslV EC 3.4.25.2
 Carboxyl-terminal protease EC 3.4.21.102
 Cytosol aminopeptidase PepA EC 3.4.11.1
 D-alanyl-D-alanine carboxypeptidase EC 3.4.16.4
 Glutathione hydrolase EC 3.4.19.13
 Leader peptidase (Prepilin peptidase) EC 3.4.23.43
 Lipoprotein signal peptidase EC 3.4.23.36
 Methionine aminopeptidase EC 3.4.11.18
 Muramoyltetrapeptide carboxypeptidase EC 3.4.17.13
 Proline dipeptidase EC 3.4.13.9
 Pyrrolidone-carboxylate peptidase EC 3.4.19.3
 Serine protease, DegP/HtrA, do-like EC 3.4.21.-
 Signal peptidase I EC 3.4.21.89
 SOS-response repressor and protease LexA EC 3.4.21.88
 Spore germination endopeptidase Gpr EC 3.4.24.78
 Thermostable carboxypeptidase 1 EC 3.4.17.19
 Tripeptide aminopeptidase EC 3.4.11.4

4. *Bacillus haynesii*

Acylamino-acid-releasing enzyme (EC 3.4.19.1)
 Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.24)
 Arginyl aminopeptidase (EC 3.4.11.6) @ Leucyl aminopeptidase (EC 3.4.11.10)
 ATP-dependent Clp protease proteolytic subunit ClpP (EC 3.4.21.92) (2)
 ATP-dependent protease La (EC 3.4.21.53) LonB Type I (2)
 ATP-dependent protease subunit HslV (EC 3.4.25.2)
 Bacillopeptidase F precursor (EC 3.4.21.-)
 Bacillopeptidase F precursor (EC 3.4.21.-)
 Carboxyl-terminal protease (EC 3.4.21.102) (2)
 Cytosol aminopeptidase PepA (EC 3.4.11.1)
 D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4) (7)

D-aminopeptidase dipeptide-binding protein DppA (EC 3.4.11.-)
 Deblocking aminopeptidase (EC 3.4.11.-)
 Gamma-D-glutamyl-L-diamino acid endopeptidase I (EC 3.4.19.11)
 Gamma-D-glutamyl-meso-diaminopimelate peptidase (EC 3.4.19.11) (3)
 Gamma-DL-glutamyl hydrolase PgsS (EC 3.4.19.-), catalyzes PGA release
 Gamma-glutamyltranspeptidase (EC 2.3.2.2) (2)
 Glutathione hydrolase (EC 3.4.19.13) (2)
 Glutamyl aminopeptidase (EC 3.4.11.7)
 Glutamyl endopeptidase precursor (EC 3.4.21.19)
 Leader peptidase (Prepilin peptidase) (EC 3.4.23.43)
 Lipoprotein signal peptidase (EC 3.4.23.36)
 Major intracellular serine protease precursor (EC 3.4.21.-)
 Methionine aminopeptidase (EC 3.4.11.18) (2)
 Microsomal dipeptidase (EC 3.4.13.19)
 Muramoyltetrapeptide carboxypeptidase (EC 3.4.17.13)
 Proline dipeptidase (EC 3.4.13.9)
 Pyrrolidone-carboxylate peptidase (EC 3.4.19.3)
 Serine protease, DegP/HtrA, do-like (EC 3.4.21.-) (3)
 Signal peptidase I (EC 3.4.21.89) (3)
 Signal peptidase SipW (EC 3.4.21.89), required for TasA secretion
 SOS-response repressor and protease LexA (EC 3.4.21.88)
 Spore germination endopeptidase Gpr (EC 3.4.24.78)
 Subtilisin E (EC 3.4.21.62)
 Thermostable carboxypeptidase 1 (EC 3.4.17.19)
 Tripeptide aminopeptidase (EC 3.4.11.4)

5. *Bacillus hisashii*

Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.24)
 ATP-dependent Clp protease proteolytic subunit ClpP (EC 3.4.21.92)
 ATP-dependent protease La (EC 3.4.21.53) LonB Type I
 ATP-dependent protease La (EC 3.4.21.53) Type II
 ATP-dependent protease subunit HslV (EC 3.4.25.2)
 Carboxyl-terminal protease (EC 3.4.21.102)
 D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
 Deblocking aminopeptidase (EC 3.4.11.-)
 Gamma-glutamyltranspeptidase (EC 2.3.2.2)
 Glutathione hydrolase (EC 3.4.19.13)
 Glutamyl aminopeptidase (EC 3.4.11.7)
 Leader peptidase (Prepilin peptidase) (EC 3.4.23.43)
 Lipoprotein signal peptidase (EC 3.4.23.36)
 Methionine aminopeptidase (EC 3.4.11.18)
 Microsomal dipeptidase precursor (EC 3.4.13.19)
 Proline dipeptidase (EC 3.4.13.9)
 Serine protease, DegP/HtrA, do-like (EC 3.4.21.-)
 Signal peptidase I (EC 3.4.21.89)(2)
 SOS-response repressor and protease LexA (EC 3.4.21.88)
 Spore germination endopeptidase Gpr (EC 3.4.24.78)
 Thermostable carboxypeptidase 1 (EC 3.4.17.19)
 Tripeptide aminopeptidase (EC 3.4.11.4)

6. *Bacillus licheniformis*

Acylamino-acid-releasing enzyme (EC 3.4.19.1)

Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.24)
 Arginyl aminopeptidase (EC 3.4.11.6) @ Leucyl aminopeptidase (EC 3.4.11.10)
 ATP-dependent Clp protease proteolytic subunit ClpP (EC 3.4.21.92)
 ATP-dependent protease La (EC 3.4.21.53) LonB Type I
 ATP-dependent protease La (EC 3.4.21.53) Type I
 ATP-dependent protease subunit HslV (EC 3.4.25.2)
 Bacillopeptidase F precursor (EC 3.4.21.-)
 Carboxyl-terminal protease (EC 3.4.21.102)
 Cytosol aminopeptidase PepA (EC 3.4.11.1)
 D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
 D-aminopeptidase dipeptide-binding protein DppA (EC 3.4.11.-)
 Deblocking aminopeptidase (EC 3.4.11.-)
 Gamma-D-glutamyl-L-diamino acid endopeptidase I (EC 3.4.19.11)
 Gamma-D-glutamyl-meso-diaminopimelate peptidase (EC 3.4.19.11)
 Gamma-DL-glutamyl hydrolase PgsS (EC 3.4.19.-), catalyzes PGA release
 Gamma-glutamyltranspeptidase (EC 2.3.2.2)
 Glutathione hydrolase (EC 3.4.19.13)
 Glutamyl aminopeptidase (EC 3.4.11.7)
 Glutamyl endopeptidase precursor (EC 3.4.21.19)
 Glutathione hydrolase (EC 3.4.19.13)
 Leader peptidase (Prepilin peptidase) (EC 3.4.23.43)
 Lipoprotein signal peptidase (EC 3.4.23.36)
 Major intracellular serine protease precursor (EC 3.4.21.-)
 Methionine aminopeptidase (EC 3.4.11.18)
 Microsomal dipeptidase (EC 3.4.13.19)
 Muramoyltetrapeptide carboxypeptidase (EC 3.4.17.13)
 Proline dipeptidase (EC 3.4.13.9)
 Pyrrolidone-carboxylate peptidase (EC 3.4.19.3)
 Serine protease, DegP/HtrA, do-like (EC 3.4.21.-)
 Signal peptidase SipW (EC 3.4.21.89) (4)
 SOS-response repressor and protease LexA (EC 3.4.21.88)
 Spore germination endopeptidase Gpr (EC 3.4.24.78)
 Subtilisin E (EC 3.4.21.62)
 Thermostable carboxypeptidase 1 (EC 3.4.17.19)
 Tripeptide aminopeptidase (EC 3.4.11.4)

7. *Bacillus smithii*

Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.24)
 ATP-dependent Clp protease proteolytic subunit ClpP (EC 3.4.21.92)
 ATP-dependent protease La (EC 3.4.21.53) LonB Type I (2)
 ATP-dependent protease subunit HslV (EC 3.4.25.2)
 Carboxyl-terminal protease (EC 3.4.21.102)
 Cytosol aminopeptidase PepA (EC 3.4.11.1)
 D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4) (4)
 Deblocking aminopeptidase (EC 3.4.11.-)
 Gamma-D-glutamyl-L-diamino acid endopeptidase I (EC 3.4.19.11)
 Glutamyl aminopeptidase (EC 3.4.11.7)
 Leader peptidase (Prepilin peptidase) (EC 3.4.23.43)
 Lipoprotein signal peptidase (EC 3.4.23.36)
 Methionine aminopeptidase (EC 3.4.11.18) (2)
 Microsomal dipeptidase (EC 3.4.13.19)
 Muramoyltetrapeptide carboxypeptidase (EC 3.4.17.13)
 Proline dipeptidase (EC 3.4.13.9)

Pyrrrolidone-carboxylate peptidase (EC 3.4.19.3)
Serine protease, DegP/HtrA, do-like (EC 3.4.21.-)
Signal peptidase I (EC 3.4.21.89) (2)
SOS-response repressor and protease LexA (EC 3.4.21.88)
Spore germination endopeptidase Gpr (EC 3.4.24.78)
Thermostable carboxypeptidase 1 (EC 3.4.17.19)
Tripeptide aminopeptidase (EC 3.4.11.4)

8. *Bacillus subtilis*

Acylamino-acid-releasing enzyme (EC 3.4.19.1)
Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.24)
Arginyl aminopeptidase (EC 3.4.11.6)
Leucyl aminopeptidase (EC 3.4.11.10)
ATP-dependent Clp protease proteolytic subunit ClpP (EC 3.4.21.92)
ATP-dependent protease La (EC 3.4.21.53) LonB Type I
ATP-dependent protease La (EC 3.4.21.53) Type I
ATP-dependent protease subunit HslV (EC 3.4.25.2)
Bacillolysin, extracellular neutral metalloprotease (EC 3.4.24.28)
Bacillopeptidase F precursor (EC 3.4.21.-)
Carboxyl-terminal protease (EC 3.4.21.102)
Carboxyl-terminal protease (EC 3.4.21.102)
Cytosol aminopeptidase PepA (EC 3.4.11.1)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
D-aminopeptidase dipeptide-binding protein DppA (EC 3.4.11.-)
Deblocking aminopeptidase (EC 3.4.11.-)
Gamma-D-glutamyl-L-diamino acid endopeptidase I (EC 3.4.19.11) (3)
Gamma-DL-glutamyl hydrolase PgsS (EC 3.4.19.-)
Gamma-glutamyltranspeptidase (EC 2.3.2.2)
Glutathione hydrolase (EC 3.4.19.13)
Glutamyl aminopeptidase (EC 3.4.11.7)
Glutamyl endopeptidase precursor (EC 3.4.21.19)
Leader peptidase (Prepilin peptidase) (EC 3.4.23.43)
Lipoprotein signal peptidase (EC 3.4.23.36)
Major intracellular serine protease precursor (EC 3.4.21.-)
Methionine aminopeptidase (EC 3.4.11.18) (2)
Muramoyltetrapeptide carboxypeptidase (EC 3.4.17.13)
Peptidase E (EC 3.4.11.2)
Proline dipeptidase (EC 3.4.13.9)
Pyrrrolidone-carboxylate peptidase (EC 3.4.19.3)
Serine protease, DegP/HtrA, do-like (EC 3.4.21.-)
Signal peptidase I (EC 3.4.21.89)
Signal peptidase SipW (EC 3.4.21.89)
SOS-response repressor and protease LexA (EC 3.4.21.88)
Spore germination endopeptidase Gpr (EC 3.4.24.78)
Subtilisin E (EC 3.4.21.62)
Thermostable carboxypeptidase 1 (EC 3.4.17.19)
Tripeptide aminopeptidase (EC 3.4.11.4)

9. *Geobacillus kaustophilus*

Acylamino-acid-releasing enzyme (EC 3.4.19.1)
Aminopeptidase S (Leu, Val, Phe, Tyr preference) (EC 3.4.11.24)
ATP-dependent Clp protease proteolytic subunit ClpP (EC 3.4.21.92)

ATP-dependent protease La (EC 3.4.21.53) LonB Type I
ATP-dependent protease subunit HslV (EC 3.4.25.2)
Carboxyl-terminal protease (EC 3.4.21.102)
Cytosol aminopeptidase PepA (EC 3.4.11.1)
D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)
D-aminopeptidase (EC 3.4.11.19)
Gamma-D-glutamyl-L-diamino acid endopeptidase I (EC 3.4.19.11)
Gamma-glutamyltranspeptidase (EC 2.3.2.2)
Glutathione hydrolase (EC 3.4.19.13)
Glutamyl aminopeptidase (EC 3.4.11.7)
Leader peptidase (Prepilin peptidase) (EC 3.4.23.43)
Lipoprotein signal peptidase (EC 3.4.23.36)
Methionine aminopeptidase (EC 3.4.11.18)
Microsomal dipeptidase (EC 3.4.13.19)
Proline dipeptidase (EC 3.4.13.9)
Pyrrolidone-carboxylate peptidase (EC 3.4.19.3)
Serine protease, DegP/HtrA, do-like (EC 3.4.21.-)
Signal peptidase I (EC 3.4.21.89) (2)
SOS-response repressor and protease LexA (EC 3.4.21.88)
Spore germination endopeptidase Gpr (EC 3.4.24.78)
Thermolysin (EC 3.4.24.27)
Thermostable carboxypeptidase 1 (EC 3.4.17.19)
Tripeptide aminopeptidase (EC 3.4.11.4)

CAPÍTULO VI

DISCUSIÓN

La estrategia para realizar el presente estudio, consistió primero en utilizar el gen ADNr 16S para identificar las bacterias que estuvieron presentes en los lugares de muestreo. Identificando molecularmente las bacterias en un rango del 99,2 a 99,6 de porcentaje de similitud nucleotídica, al comparar nuestras secuencias consenso, con las secuencias presentes en la Base de Datos GenBank de BLAST- NCBI, se pudo obtener las identificaciones a nivel de especies.

Segundo, utilizando la Base de metadatos de BacDive, se estableció la clasificación taxonómica de cada una de las especies, es decir la posición que tienen dentro de la escala taxonómica como dominio, filo, clase, orden, familia, género y especie.

Tercero, al obtener un número de nueve especies de bacterias, clasificada, como se indicó en Resultados, estos nombres científicos se utilizaron para establecer a través del uso de la base de metadatos de Patricbrc, las enzimas hidrolíticas presentes en cada una de la especies identificadas.

Ahora conocemos que todo genoma tiene información respecto a elementos genéticos de cada especie de bacteria empleada, por lo que existen bases de datos o metadatos que registran cientos de miles de secuencias de cada genoma.

Las bases de datos presentes, en BacDive y Patricbrc, de disposición libre, aunque existen muchos más, nos ha llevado a utilizar el análisis de genes que codifican a enzimas hidrolíticas correspondientes a bacterias termófilas cultivables con un nivel de certeza única, porque los datos obtenidos

corresponden a las mismas especies que hemos obtenido en los lugares de muestreo, aunque no se descarta que existan algunos errores en los análisis.

Asimismo, se ha contrastado con los estudios realizados anteriormente en nuestro laboratorio de varios tesis que han publicado sus tesis en la Revista Ciencia & Desarrollo de la Universidad Nacional Jorge Basadre Grohmann, de Tacna, Perú.

Desde 2014, se han realizado trabajos relacionados con la producción de enzimas hidrolíticas termoestables, teniendo una experiencia en la producción de enzimas celulasas, que degradan la celulosa (Zapata & Castellanos, 2014; Salazar et al., 2015; Ninaja et al., 2018).

Asimismo, se ha logrado, se ha logrado obtener enzimas que degradan al almidón, obteniendo enzimas amilasas (enzima hidrolítica que cataliza la escisión o desdoblamiento de un enlace mediante hidrólisis. Las enzimas proteolítica, hidroliza, con distinto grado de especificidad, las proteínas nativas o fragmentos peptídicos.

Las enzimas termoestables —tales como polimerasas, ligasas, nucleasas, transcriptasas inversas, polinucleótido quinasa y enzimas de restricción— desempeñan un papel importante en la biotecnología. El uso, cada vez más frecuente, de técnicas de ADN requiere la identificación y producción de nuevas enzimas termoestables.

Por otro lado, se considera que las aguas termales de pH casi neutro y baja concentración de sal ofrecen un tipo de sistema de cultivo puro para termófilos moderados. En nuestro caso, las bacterias termófilas de los géiseres se encuentran entre un rango de 55 a 92 °C de temperatura del agua termal. Lo que hace posible, que en un futuro cercano queramos obtener enzimas

termoestables, para ser utilizadas en la industria alimenticia o farmacéutica, tenemos que realizarlo a través de una producción en biorreactores de mayor capacidad.

Dado que Perú es un país que en sus Andes se presentan muchos volcanes, hay muchas fuentes termales distribuidas, y se han realizado muchas observaciones sobre las bacterias termofílicas. Pero nuestro objetivo principal fue aplicar una bioprospección que permita en corto tiempo realizar un muestreo de aguas y sedimentos termales, para conseguir el ADN genómico mediante incubación y crecimiento de las colonias de bacterias posibles y sin identificar.

En bacterias termófilas, el hecho de haberlas cultivado y observar las características de sus colonias, no son suficientes para identificarlas al grado de especies. Por ello, es necesario el ADN genómico, para obtener los genes ADNr 16S que por reacción en cadena de la polimerasa, empleando los cebadores 16S (Forward 27F y Reverse 1492R) puedan permitir por comparación con la Base de Datos del GenBank, su anotación como especie. Esto se ha realizado en nuestro caso, pudiendo obtener 48 cepas bacterianas aisladas de los once géiseres muestreados y que ha permitido su clasificación taxonómica, mostrada en Resultados.

Las especies identificadas, han permitido, mediante las bases de datos existentes en diferentes páginas webs colgadas en el Internet, tales como el BacDive, Patric, Brenda, etc., determinar las diferentes enzimas hidrolíticas pertenecientes a la clasificación E.C. 3. de acuerdo a la Lista de Enzimas Clase 3 – Hidrolasas del Comité de Nomenclatura de la Unión Internacional de Bioquímica y Biología Molecular (NC-IUBMB) Versión LTEX preparada por Andrew McDonald, Escuela de Bioquímica e Inmunología, Trinity College Dublin, Irlanda.2019 IUBMB (Anexo 3).

De acuerdo a esta clasificación se han obtenido los nombres de las enzimas obtenidas de la base de metadatos de Patricbrc que contiene todas las enzimas hidrolíticas de microorganismos bacterianos, entre ellos de las bacterias termófilas obtenidas.

De la base de metadatos se han obtenido las enzimas presentes en el genoma bacteriano de las especies *Aeribacillus composti*, *Anoxybacillus kamchatkensis*, *Bacillus subtilis*, *B. coagulans*, *B. smithii*, *B. Hisashi*, *B. Haynesii*, *B. licheniformes* y *Geobacillus kaustophilus*, todos pertenecientes al filo Firmicutes.

Cada especie indicada, contiene las enzimas hidrolíticas para lípidos, polisacáridos y proteínas, nombradas como enzimas que actúan sobre enlaces ésteres (EC 3.1), enzimas glucosilasas (EC 3.2) y enzimas que actúan sobre enlaces peptídicos (peptidasas) (EC 3.4) la mayor parte actúan sobre sustratos comerciales, como aceites, grasas, almidones, proteínas, que son de importancia biotecnológica.

Utilizando la Base de Metadatos de Patricbrc, permite obtener las enzimas hidrolíticas correspondientes a las especies de bacterias mencionadas. Por un lado, la precisión de encontrar las enzimas hidrolíticas en esta base de datos, es mayor cuando se trata a nivel de géneros.

Ferrer et al en el año 2014, aisló una bacteria termófila productora de enzima amilasa, de los géiseres de Calientes, Candarave (Tacna-Perú), la cual por análisis de la secuencia del gen ARNr 16S presentó un 99 % de identidad con *Bacillus licheniformis*. Para determinar la capacidad amilolítica se evaluó en medio sólido presentando un halo de hidrólisis de 29,5 mm de diámetro a las 72 horas. En la evaluación de la producción de amilasa se obtuvo una concentración

máxima de 1410,57µg/mL de azúcares reductores a inicios de la fase estacionaria a 54,6 horas de incubación a 60 °C y pH 7.

El extracto amilolítico tuvo un grado de purificación de 2,3 veces con sulfato de amonio, actuando a una temperatura óptima de 58 °C con una actividad de 10,7 U/mL ya un pH óptimo de 6,7 con una actividad de 8,3 U/mL, además la velocidad máxima (V_m) fue de 229,5µg de almidón hidrolizado por ml de enzima por minuto y el valor de la constante de Michaelis (K_m) fue de 8,7 mg de almidón por ml enzima (Ferrer et al., 2014).

Dieciocho bacterias termófilas, codificadas y morfológicamente diferentes, fueron seleccionadas por su habilidad para producir enzimas celulolíticas. Entre ellas se seleccionó CB-13 por su alto potencial para producir celulasa. El análisis de su gen ARNr 16S reveló su estrecha relación con *Geobacillus thermoparaffinivorans*. Adicionalmente, se intentó separar y purificar la β -glucosidasa mediante IEXC. El respectivo análisis de sus fracciones resultantes por SDS-PAGE, definió el peso molecular de la enzima con un valor aproximado de 92,6 kDa. Finalmente, la caracterización funcional de las celulastas totales, mostró el pH de 7,0 y la temperatura de 70 °C como óptimos para la actividad (Ramírez & Castellanos, 2017).

En este trabajo se estudia las proteasas alcalinas producidas por dos bacterias termófilas designadas como BP-2 y BP-4 que fueron aisladas de sedimentos de los géiseres de Candarave, mostrando capacidad proteolítica en medio sólido. Del análisis de las secuencias del gen ARNr165 de cada bacteria se determinó que la bacteria BP-2 tiene 99 % de identidad con la especie *Bacillus licheniformis* y la bacteria BP-4 tiene un 99 % de identidad con la especie *Geobacillus thermoparaffinivorans* (Naquiche et al., 2015).

En otra investigación se tuvo como objetivo la separación de tinta de papel impreso de oficina utilizando un microbio termófilo amilolítico. Se evaluaron los parámetros óptimos de pH, temperatura, concentración de enzima y de papel impreso para una máxima separación de tinta por la actividad amilolítica de *Geobacillus thermoparaffinivorans*, procedente de los géiseres de Tacna, Perú (Espinal et al., 2020).

Las bacterias termófilas de los géiseres de Calientes-Candarave, también producen enzimas pectinasas, que degradan las pectinas y que permiten obtener una serie de productos, como pectinas, heteropolisacáridos, componentes principales de la capa media de la pared celular de la piel de los cítricos (naranjas, limones y mandarinas), manzanas, melocotones y otros vegetales. Se aislaron veinte cultivos bacterianos gram positivos, de los cuales 12 fueron productores de pectinasas, siendo el cultivo LBE-P4 el de mayor actividad enzimática con 0,154 U/ml y 0,069 mg/ml de proteínas totales, que fue identificado molecularmente como *Geobacillus kaustophilus* LBE-P4 (Salazar et al., 2020).

En esta investigación, se estudió la cinética de la enzima celulasa producida por la bacteria *Bacillus sp* GCB-13, aislada de los géiseres de Candarave. La bacteria alcanzó una fase logarítmica a partir de las 24 horas de incubación. La producción enzimática fue desarrollada usando CMC como sustrato. El extracto enzimático crudo obtenido fue aplicado sobre sustratos lignocelulósicos naturales como tallos y hojas del geranio, y sobre sustratos sintéticos como papel bond y carboximetilcelulosa (CMC), para evaluar su potencial de degradación. (Salazar et al., 2019).

Los almidones de papa (*Solanum tuberosum*), camote (*Ipomoea batatas*) y yuca (*Manihot esculenta*) fueron hidrolizados por la enzima amilasa obtenida por cultivo líquido de la bacteria termófila, designada como *Bacillus licheniformis*

BA-3 y aislada de los géiseres de Calientes, Candarave, Tacna. Se obtuvo el extracto crudo, el cual fue filtrado, centrifugado y precipitado con sulfato de amonio al 60 % y dializado por 24 horas en refrigeración a -5 °C. Asimismo, la bacteria termófila utilizada presentó el 99 % de identidad mediante la secuenciación del gen ARNr 16S. Se concluyó que existe mayor afinidad de la enzima producida por *Bacillus licheniformis* BA-3 con el almidón de *Ipomoea batatas* por ser el menor valor de Km obtenido (Zúñiga et al., 2019).

Se recolectó muestras de agua y sedimento de ocho fuentes termales de Calientes (Candarave-Tacna) las cuales fueron incubadas a 50°C por 48 horas en medio de cultivo carboximetilcelulosa (CMC). La identificación molecular de cada una de estas cepas corresponde, en el mismo orden, a *Bacillus sp* y *Paenibacillus sp*. Con respecto a la producción de proteínas totales de la cepa GCB-13, no se presentó diferencias significativas a 45°C, 50°C y 55°C; caso contrario a la producción de azúcares reductores donde sí se determinó una mayor productividad a 45 °C y 50 °C. Por otro lado, la cepa GCC-13 no reveló diferencias significativas en la producción de proteínas totales y azúcares reductores; en consecuencia, para este caso en particular, se concluye que existe un rango más amplio de acción (Ninaja et al., 2018).

En la presente investigación se aisló una bacteria termófila productora de enzima amilasa, de los géiseres de Calientes, Candarave (Tacna —Perú), la cual por análisis de la secuencia del gen ARNr 16S presentó un 99 % de identidad con *Bacillus licheniformis*. El extracto amilolítico tuvo un grado de purificación de 2,3 veces con sulfato de amonio, actuando a una temperatura óptima de 58 °C con una actividad de 10,7 U/ml ya un pH óptimo de 6,7 con una actividad de 8,3 U/mL, además la velocidad máxima (V_m) fue de 229,5 pg de almidón hidrohizado por ml de enzima por minuto y el valor de la constante de Michaelis (K_m) fue de 8,7 mg de almidón por mide enzima (Ferrer, C. et al., 2015).

Se investigó la capacidad de decoloración de las bacterias termófilas con actividades tipo lacasa sobre tintes sintéticos industriales, especialmente, tintas de impresión que se liberan en los efluentes industriales en el momento del destintado del papel reciclado. Por tal motivo, se seleccionaron los cultivos bacterianos termófilos *Geobacillus thermoleovorans* CB-13, *Geobacillus thermoparaffinivorans* CB-25, *Geobacillus thermoleovorans* FT-06 y *Geobacillus kaustophilus* TM-12, que presentaron mayor actividad enzimática tipo lacasa en medio sólido para los cuatro sustratos evaluados (2-MP; 2,6-DMP; ABTS y SRG) (Salazar & Castellanos, 2018).

CONCLUSIONES

1. La bioprospección de las bacterias termofílicas cultivables en los once géiseres de Calientes-Candarave ha permitido obtener un total de 48 cepas de especies de bacterias termofílicas, clasificadas taxonómicamente en el dominio Bacteria, filo Firmicutes, clase Bacillia, Orden *Bacillales*, Familia *Bacillaceae*, Géneros *Aeribacillus*, *Anoxybacillus*, *Bacillus* y *Geobacillus*.
2. Se obtuvieron un total de nueve especies de bacterias termofílicas: *Aeribacillus composti*, *Anoxybacillus kamchatkensis*, *Bacillus licheniformis*, *Bacillus hisashii*, *Bacillus haynesii*, *Bacillus smithii*, *Bacillus coagulans*, *Bacillus subtilis* y *Geobacillus kaustophilus*.
3. Todas las especies obtenidas contenían genes que codificaban para la expresión de enzimas hidrolíticas como lipolíticas (298), glucolíticas (227) y proteolíticas (224), con un total de 744 enzimas termoestables.
4. Todas las enzimas hidrolíticas mencionadas podrían ser producidas biotecnológicamente mediante la inserción de los genes que codifican a enzimas en plásmidos o vectores de la bacteria *Escherichia coli*, para su producción comercial porque son monoméricas, es decir de una sola cadena polipeptídica.
5. Se ha logrado demostrar la necesidad de los metadatos existentes de libre utilización, los cuales permiten un ahorro económico y de tiempo para realizar bioprospecciones de cual naturaleza, basado en los estudios realizados por otros investigadores y que se encuentran a disposición del usuario.

RECOMENDACIONES

1. El presente estudio ha determinado que existe la necesidad de aplicar la tecnología del ADN recombinante para procurar explotar el gran tesoro microbiano que tenemos en la zona protegida de Maure-Vilacota de la provincia de Candarave donde se encuentran los géiseres de Calientes-Candarave que debe ser protegidos por alguna ordenanza de la Municipalidad de Candarave.
2. Asimismo, es necesario, invitar a los estudiantes de microbiología a que profundicen sus estudios en estos microorganismos termófilos, dado que cada uno de ellos, constituye una gran fuente de sustancias, metabolitos, antibióticos y enzimas, que podrían ser empleados en el futuro para la solución de diversos problemas farmacéuticos, alimentarios, industriales y hasta en la biorremediación de los suelos y pasivos mineros de la Región de Tacna.
3. Para fortalecer las fuentes de metadatos existentes en el mundo, es necesario primero que el Instituto General de Investigación de la Universidad Nacional Jorge Basadre Grohmann exija a cada investigador que es beneficiado con los fondos del Canon, sobrecanon y regalías, almacene su data obtenida de cada investigación, respetando la posesión de los mismos.

REFERENCIAS BIBLIOGRÁFICAS

- Acharya, S. (2012). Bioprospecting Thermophiles for Cellulase Production: A Review. *Brazilian Journal of Microbiology*, 844-856.
- Armenteras, D.G.T. (2016). Revisión del concepto de ecosistema como "unidad de la naturaleza" 80 años después de su formulación. *Ecosistemas*, 83-89.
- Arora, R.B.S. (2015). Bioprospecting thermophilic thermotolerant microbes for production of lignocellulosic ethanol: A future perspective. *Renew Sust Energy Rev*, 699-717.
- Attenborough, David. (2021). Una vida en nuestro planeta: Mi testimonio y una visión para el futuro. Primera edición. *Editorial Planeta, S. A.*
- Castillo, D., Castellanos, R., Esparza, M. (2012). Caracterización metagenómica de poblaciones microbianas biolixiviantes aisladas desde los géiseres de Calientes (Candarave – Tacna) para aplicaciones en minería: del genoma a las aplicaciones industriales. Universidad Nacional Jorge Basadre Grohmann. *CONCYTEC*.
- Castellanos, R., Esparza, M., Zúñiga, A., Ferrer, C., Zapata, H. Naquiche, A., Ninaja, F., Salazar, I y Calizaya, J. (2013). Biotecnología de microorganismos termófilos de la región Tacna-Perú como plataforma base de producción de enzimas termófilas para aplicaciones industriales. Universidad Nacional Jorge Basadre Grohmann. *CONCYTEC*.
- Castellanos, R., Zapata, H.B., Ferrer, I.C., Naquiche, A.J., Salazar, J.I., Zúñiga, A.Z., Pilco, N., Esparza, M. (2014). Producción biotecnológica de enzimas industriales a partir de microorganismos termófilos aislados de los Géiseres de Calientes

- Candarave-Tacna. Instituto General de Investigación. Universidad Nacional Jorge Basadre Grohmann.
- Castellanos, R., Zapata, H.B., Naquiche, A.J., Salazar, J.I., Zúñiga, A.Z., Yllanes, M.V, Salazar, F., Ramírez, H., Samillán, S., Castillo, D. (2015). Tratamiento biotecnológico con enzimas termoestables bacterianas para el destintado del papel periódico y de oficina en su proceso de reciclado. *UNJBG-FONDECYT*.
- Castellanos-Cabrera, R. (2021) *Análisis filogenético molecular de la comunidad bacteriana termófila del valle de los géiseres de Calientes, Candarave –Tacna (Perú)*. Tesis de Grado Doctoral no publicada. Universidad Nacional Mayor de San Marcos. Lima, Perú.
- Chauhan, S., Seth, C.A. & Seth, A. (2015). Bioprospecting thermophilic microorganisms from hot springs of western Himalayas for xylanase production and its statistical optimization by using response surface methodology. *Journal of Pure and Applied Microbiology*. 9. 1417-1428.
- Cruz Pauccara, V. (2010). *Geochemical characterization of thermal waters in the Calientes Geothermal Field, Tacna, South of Peru*. Proceedings World Geothermal Congress 2010, Bali, Indonesia. Bochum: International Geothermal Association, 1-7.
- Cruz, V., Vargas, V. & Matsuda, K. (2010) Geochemical Characterization of Thermal Waters in the Calientes Geothermal Field, Tacna, South of Peru. *Proceedings World Geothermal Congress 2010*. Bali, Indonesia, 25-29
- Duarte-Torres, O.(2011). La bioprospección en Colombia. EXPEDITIO, (7). Recuperado a partir de:
<https://revistas.utadeo.edu.co/index.php/EXP/article/view/732>

Espinal, D., Zúñiga, A., Castillo, D. & Castellanos, R. (2020). Separación de tinta de papel impreso por actividad amilolítica microbiana. *Ciencia & Desarrollo*, N° 27, pp. 121-132 (2020). DOI: <https://doi.org/10.33326/26176033.2020.27.1002>

Felsenstein, J. (1985). Phylogenies and the Comparative Method. *The American Naturalist*. 125 (1): 1.

Fernández, M.T.; Rodríguez, H.; Gonzalez, T. y Goire, I. (2009) Construcción de un vector para la integración cromosomal de un gen de fitasa de *Bacillus licheniformis*. *Rev. peru biol.* vol.16, n.1, pp.109-114. Disponible en: http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1727-99332009000100014&lng=es&nrm=iso. ISSN 1727-9933.

Ferrer, C., Zúñiga, A., Salazar, I., Naquiche, A., Zapata, H. & Castellanos, R. (2015) Caracterización de la enzima amilasa de la bacteria termófila *Bacillus licheniformis* BA-3 aislada de los géiseres de Candarave (Tacna-Perú). *Ciencia & Desarrollo* N° 20: 65-68 / ISSN 2304-8891.

Hennig, W. (1966). Phylogenetic Systematics. *University of Illinois Press, Urbana*. Pp.263.

Klappenbach, J.A, Saxman, P.R, Cole, J.R, Schmidt, T.M. (2001) rrndb: the ribosomal RNA operon copy number database. *Nucl Acids Res* 29:181-4.

Mullis, K.B. & F.A. Faloon. (1987). Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology* 155: 335-350.

- Finore, I., Gioiello, A., Leone, L., Orlando, P., Romano, I., Nicolaus, B. & Poli, A. (2017) *Aeribacillus composti* sp. nov., un bacilo termófilo aislado del compost de orujo de la almazara. *IJSEM* 67: 4830–4835 (DOI 10.1099 / ijsem.0.002391
- Instituto Leibniz DSMZ. Gleim, D., Kracht, M. & Weiss, N. (1980). Procariotas Nomenclatura actualizada - compilación de todos los nombres de Bacteria y Archaea, válidamente publicado según el Código Bacteriológico desde el 1. de enero de 1980, y cambios de nomenclatura válidamente publicados.
- Kevbrin, V.V., Zengler, K., Lysenko, A.M. & Wiegel, J. (2005). *Anoxybacillus kamchatkensis* sp. Nov., Una nueva bacteria aeróbica facultativa termófila con un pH óptimo amplio del valle del Geysir, Kamchatka. *Extremófilos* 9: 391-398.
- Klappenbach, J. A. (2001). rrndb: the Ribosomal RNA Operon Copy Number Database. *Nucleic Acids Research*, 181-184.
- Lokendra-Kumar, G. A. (2011). Extremophiles: A Novel Source of Thermostable Agarase from *Bacillus* sp.BI-3, a Thermophilic Bacterium isolated from Hot Spring,. *Biothechnology*, 121-135.
- Madigan, M. T., Martinko, J. M., Bender, K. S., Buckley, D. H., Stahl, D. A., & Brock, T. (2014). *Brock Biology of Microorganisms* (14° Edition). Boston: Pearson.
- Moreira de Almeida, J. R., García-Collares, D. & Dias-Barbosa, P.F. (2015). Microbial bioprospecting. *Agroenergy. Ministry of Agriculture, Livestock and Food Suppl. Brasil*.
- Navarro-González, I. y Periago, M.J. (2012). *Enzimas lipolíticas bacterianas: propiedades, clasificación, estructura, aplicaciones tecnológicas y aspectos legales*. Departamento de Tecnología de los Alimentos, Nutrición y

Bromatología. Área de Nutrición y Bromatología. Facultad de Veterinaria. Universidad de Murcia. Campus de Espinardo 30100; Murcia, España.

Naquiche, A.J., Zúñiga, A.Z., Ferrer, C.I., Salazar, I.J., Zapata, H.B. & Castellanos, R. (2015). Caracterización de dos bacterias termófilas (BP-2 y BP-4) con capacidad proteolítica aislados en los géiseres de Candarave. Tacna – Perú. *Ciencia & Desarrollo*. (Tacna) ISSN 2304-8891; 2015; 19:49-53.

Neefs, J. M. (1990). Compilation of small ribosomal subunit RNA sequences. *Nucleic acids research*, 2237-2317.

Ninaja, F., Castillo, D. & Castellanos, R. (2018) Evaluación de enzimas celulolíticas bacterianas aisladas de las fuentes termales de Calientes (Candarave-Tacna). *Ciencia & Desarrollo*, Vol. 17, 23 (2) 55 – 61.

Ramírez, H. & Castellanos, R. (2017) Producción de celulasas bacterianas termoestables utilizando papel periódico reciclado como sustrato. *Revista Ciencia y Tecnología para el Desarrollo-UJCM*; 3(6):8-21.
<http://dx.doi.org/10.37260/rctd.v3i6.88.g72>

Reimer, L.C., Vetcinova, A., Carbasse, J.S., Söhngen, C., Gleim, D., Ebeling, C., Overmann, J. (2019). BacDive in 2019: Bacterial phenotypic data for High-throughput biodiversity analysis. *Nucleic Acids Res.* 8;47(D1):D631-D636. doi: 10.1093/nar/gky879.

Sahay, H. N. (2017). Hot springs of Indian Himalayas: potential sources of microbial diversity and thermostable hydrolytic enzymes. *3 Biotech*, 1-11.

Saitou, N. & Nei, N. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, 4: 406-425.

Samillán, S. & Castellanos, R. (2016). *Biodegradación de aceites residuales por lipasas termoestables de bacterias termófilas*. Tesis de Titulación. Facultad de Ciencias. UNJBG. Tacna, Perú.

Salazar, F.V. & Castellanos, R. (2018). Biodecoloración de tintes sintéticos industriales por bacterias termófilas con actividad enzimática tipo lacasa. *Revista Ciencia y Tecnología para el Desarrollo-UJCM*; 4(7):9-21.
DOI: <http://dx.doi.org/10.37260/rctd.v4i7.97.g81>

Salazar, I., Zúñiga, A., Naquiche, A. & Castellanos, R. (2019) Caracterización enzimática de una celulasa producida por *Bacillus* sp. (GCB-13) para la degradación de sustratos lignocelulósicos. *Ciencia & Desarrollo*, Vol. 18, 25 (2) 61 – 67. ISSN 2304-8891 SSN 2617-6033. DOI:
<https://doi.org/10.33326/26176033.2019.25.865>

Salazar, I., Ramírez, H., Yllanes, M., Zúñiga, A., Naquiche, A., Castillo, D. & Castellanos, R. (2020) Selección e identificación de una nueva bacteria productora de pectinasa a partir de fuentes geotermales. *Ciencia & Desarrollo N° 27*, pp. 143-149. ISSN-L 2304-8891 DOI:
<https://doi.org/10.33326/26176033.2020.27.1004>

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., Arnheim, N. (1985) Enzymatic amplification of beta-globin genomic. *Dec 20;230(4732):1350–1354*.

Sanger, F. & Coulson, A.R.(1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology* 94: 441-446.

- Stackebrandt, E, Frederiksen, W, Garrity, G.M, Grimont, P.A, Kampfer, P, Maiden, M.C, et al. (2002) Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* 52:1043-7.
- Stackebrandt, E, Goebel, B.M. (1994) Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* 44:846-9.
- Stackebrandt, E.A.(1994). Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. *International Journal of Systematic Bacteriology*, 846-849.
- Pocaterra, F., Beltrán-Zambrano, R., López-Rodríguez, F. y Romero, F. (1997) Biodiversidad, Bioprospección y Bioseguridad. *Instituto de Estudios Ecologistas del Tercer Mundo. Ediciones ABYA-YALA. Quito-Ecuador.*
- Thankappan, S., Kandasamy, S., Joshi, B. et al. (2018) Bioprospecting thermophilic glycosyl hydrolases, from hot springs of Himachal Pradesh, for biomass valorization. *AMB Expr* (2018) 8: 168.
- Zimmermann, W. (1930). *Die Phylogenie der Pflanzen*. G. Fischer, Jena, Germany.
- Zimmermann, W. (1934). Research on phylogeny of species and of single characters. *American Naturalist* 68: 381-384.
- Woese, C.R, Stackebrandt, E, Macke, T.J, Fox, G.E. (1985) A phylogenetic definition of the major eubacterial taxa. *Syst Appl Microbiol* 6:143-51.
- Woese C.R. (1987) Bacterial evolution. *Microbiol Rev.*51:221-7.

Woese, C. K. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci USA*, 4576-4579

Zúñiga, A., Ferrer, C., Salazar, I., Naquiche, A. & Castellanos, R.(2019) Estudio cinético de la hidrólisis del almidón de *Solanum tuberosum*, *Ipomoea batatas* y *Manihot esculenta* con amilasa termoestable de *Bacillus licheniformis* BA-3 aislada de los géiseres de Candarave-Tacna. *Ciencia & Desarrollo*, Vol. 18, 24 (1) 18 - 26 (2019)ISSN 2304-8891 Versión impresaISSN 2617-6033 Versión electrónica DOI: 10.33326/26176033.2019.24.781

Anexo 1

Secuencias ADNr 16S Forward enviadas por Macrogen Inc

>A01_27F

```
TTGCATGCGGCGTGCTATACATGCAGTCGAGCGAACCAATAAGAAGCTTGCTTTTTGTTGGTTAGCGGCGGACGTT
GTGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGATTAT
CTTTCCGCCTGGAGAGATAAGGAAAGATGGCTTTTGTCACTACTTACAGATGGGCCCGCGGCGCATTAGCTAGTTG
GTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGTATCGGCCACACTGGGACTGAGAC
ACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCG
CGTGAGCGAAGAAGGTCTTCCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGGAGGGAATGCCGGTAC
CTTGACGGTACCTGACGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTT
GTCCGGATTTATTGGGCGTAAAGCGCGCGCAGGCGGTCTTTAAGTCTGATGTGAAATCTTGGCGCTCAACCGCA
AGCGGTCAATTGGAAACTGGGGACTTGAGTGCAGAAGAGGAAAGCGGAATCCACGTGTAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACCAAGTGGCGAAGGCGGCTTTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGG
AGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACAGATGAGTGCTAAGTGTGGAGGGTTTCCGCCCTTC
AGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGTGCAGAACTGAAACTCAAAGGAATTGACGG
GGGCCCGCACAAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCTCC
TGACGCCCTAGAGACAGGGTCTTCCCTTCCGGGACAGGATGACAGGTGGTGCATGTTGTCGTCAGCTCGTGTC
GTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACCCCTGGGTTCTAGTTGCCAGCATTACAGTTGGGCACTCTAG
AGCGACTGCCGGCGACAAGTCCGGAGGAAAGTGGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTAC
ACACTGGCTACAATGGAAGGGTACAAAGGGGCGAGCCAAACCGCCAAACGCAAAAGCGAATCCCAAAAAACCTTTCT
CAGTTTTCGATTGCAAGCGTGCAAACTCGCTGGTGAACCCGGAATCGCTTGTATACGCGGGATAACCAACGCC
CGCGTGAATACTTTCCGGGGCTTTT
```

>B01_27F

```
ATGCATGCGCATCTATACATGCAGTCGAGCGGACCGACGGGAGCTTGCTCCCTTAGGTCAGCGGCGGACGGGT
GAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATGCTTGATT
GAACCGCATGTTCAATATAAAAGGTGGCTTTCAGCTACCCTTACAGATGGACCCGCGGCGCATTAGCTAGTTG
GTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGTATCGGCCACACTGGGACTGAGAC
ACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCG
CGTGAGTGATGAAGTTTTTCCGGATCGTAAACTCTGTTGTTAGGGAAGAACAAGTACCGTTTGAATAGGGCGGTAC
CTTGACGGTACCTAACCGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTT
GTCCGGAATTTATTGGGCGTAAAGCGCGCGCAGGCGGTTTTTAAGTCTGATGTGAAAGCCCCCGGCTCAACCGGG
GAGGGTCAATTGGAACTGGGGAACCTTGAGTGCAGAAGAGGAGAGTGGAATCCACGTGTAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACCAAGTGGCGAAGGCGACTCTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGG
AGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACAGATGAGTGCTAAGTGTGGAGGGTTTCCGCCCTTT
AGTGCTGCAGCAAACGCATTAAGCACTCCGCCTGGGGGAGTACGGTGCAGAACTGAAACTCAAAGGAATTGACG
GGGGCCCGCACAAAGCGGTGGAACCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCC
TCTGACAACCTTAGAGATAGGGCTTCCCTTCCGGGAGCAGGTGACAGGTGGTTGATGTTGTCGTCAGCTCGT
GTCGTGAGATGTTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTTGATCTTAGTTGCCAGCATTACAGTTGGGCACT
TAAGGTGACTGCCGGTGACAAACCGGAAGAAAGGGTGGGGATGAACTTCAAATCATCATGGCCCTTATGAAACCG
GGGCTACCACACTTGCTACAATGGGGCAAAAAAAGGGAGCCAAAACCCCGGGGGTAAACCAATCCCCAAAAT
TGTTTTCAATTTGAAACCAATCTCCAACCTCACAGCTGAAACTGAAACCTTAAATACCGAATAACCAAAACC
CGCGGGGAAAAAATTTCCGGGGCTCTGTAACCCACCCCTCCACCACACAGTTTTTTAACCCTAA
```

>C01_27F.

```
GGCAGTGCGCTGCCTATACATGCAGTCGAGCGAACCAATAAGAAGCTTGCTTTTTGTTGGTTAGCGGCGAACGGG
TGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGATTATC
TTTCCGCCTGGAGAGATAAGGAAAGATGGCTTTTGTCACTACTTACAGATGGGCCCGCGGCGCATTAGCTAGTTGG
TGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGTATCGGCCACACTGGGACTGAGACA
CGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCG
GTGAGCGAAGAAGGTCTTCCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGGAGGGAATGCCGGTACC
TTGACGGTACCTGACGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTTG
TCCGGATTTATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTTAAGTCTGATGTGAAATCTTGGCGCTCAACCGCAA
GCGGTCAATTGGAACTGGGGGACTTGAGTGCAGAAGAGGAAAGCGGAATCCACGTGTAGCGGTGAAATGCGTAG
AGATGTGGAGGAACACCAAGTGGCGAAGGCGGCTTTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGG
GCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACAGATGAGTGCTAAGTGTGGAGGGTTTCCGCCCTTC
AGTGCTGCAGCTAACGCATTAAGCACTCCGCCCGGGGAGTACGGTGCAGAACTGAAACTCAAAGGAATTGACG
GGGGCCCGCACAAAGCGGTGGAACCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCT
CCTGACCCCTTGAGAGACAGGGTCTTCTTCCGGGACAGGATGACAGGTGGTTGATGTTGTCGTCAGCTCGT
GTCGTGAGATGTTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTGGGTTCTAGTTGCCAGCCATTCAATTTGGGCA
CTTAAGAGCGACTGCCGGCGACAAGTCCGAAGAAAGGTGGGGGATGAACTTCAAATCATCCAGGCCCTTATAA
CCCGGGCTACACACCTGGTTCCAATGGGAGGGTACAAAGGGCGCCAAACGCGCCACCCCAAAACCGATTCCCA
GAAAAACATTCTTCAATTTGAATTTGAAGGGGGCAATCCCCCGCTTAAACCGGAATCCCTTTTTATTCCGGAT
AAACAAGGCCCGGTGAAAAATTTCCGGGCTGTTTACCCTCGGTTTCCCTTCC
```

>D01_27F.

GTGCAGTGCGCATGCTATACATGCAGTCGAGCGGACCAAAATCGGAGCTTGCTCTGGTTTGGTCAGCGGGGACGG
GTGAGTAACACGTGGGCAACCTGCCCGCAAGACCGGGATAACTCCGGGAAACCGGAGCTAATACCGGATAACACC
GAAGACCGCATGGTCTTTGGTTGAAAGGCGGCCTTTGGCTGTCACCTGCGGATGGGCCCGCGGCGCAATAGCTAG
TTGGTGAGGTAACGGCTCACCAAGCGGACGATGCGTAGCCCGCCTGAGAGGGTGACCGGCCACACTGGGACTGA
GACACGGCCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGGCGAAAGCCTGACGGAGCGACG
CCGCGTGAGCGAAGAAGGCCTTCGGGTCTGTAAGCTCTGTTGTGAGGGACGAAGGAGCGCCGTTCAAGAGGGC
GGCGCGGTGACGGTACCTCACGAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGC
GAGCGTTGTCGGAAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAAGCCCACGGCTC
AACCGTGGAGGGTCAATTGAAACTGGGGGACTTGAGTGCAGGAGAGGAGAGCGGAATTCACGTGTAGCGGTGA
AATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCTGCAACTGACGCTGAGGCGCGAAA
GCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACAGATGAGTGCTAAGTGTAGAGGGGTC
ACACCCTTTAGTGCTGCAGCTAACGCGATAAGCACTCCGCCCGGGGAGTACGGCCGCAAGGCTGAAACTCAAAG
GAATTTGACGGGGCCGCAACAGCGGTGGAACCAAGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGTC
TTGACATCCCCTGACAACCAAGAAGATTGGGCGTTCCCCCTTCGGGGGACAGGGTGACAGGTGGTTGCATGGT
TGTCGTGAGCTCGTGTCTGAGAAAGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTCGCCTCTTGTGCCAGCA
CCAAAGTGGCACTCTTAAAGGGACTGCCGGCGACAATTCGAAGAAAAGGTGGGGGATAACCTCCAATCATCCA
GGCCCCCTTATAACCGGGGGCTACAACCTGTTTACAATGGGGCGGTACAAGGGGTACAAGGGGTGAAACCCCGGGG
AGGAAAACCCAAAACCCCTTCAATTTTTGAAATGAAGGGGCGAACTTCCCCCGCTGAAAACCGGAATCCCTTTT
ATACCGGAATAAATAGCCCGCGGGAAAAAATTTCCCGGGCTTGTTTAAACCCCGGTTT

> E01_27F

GTGGACTGGCGCTGCCTATACATGCAGTCGAGCGGACCGAAGGAGCTTGCTCCTTTAGGTTAGCGGGGACGGG
TGAGTAACACGTGGGCAACCTGCCCTGCAGACTGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACCCG
AAAACCGCAGGGTTTTCCGGTTGAAAGGCGGCTTTTAGCTGTCACCTGCAGGATGGGCCCGCGCATTAGCTAGT
TGGTGAGGTAACGGCTCACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGAG
ACACGGCCGAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAAGTGTGACGGAGCAACGC
CCCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTGAGGGAAGAACAAGTCCGTTTCAACAGGGCGG
TACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGC
GTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAATCTCGCGGCTCAACC
GCGAGCGGCCATTTGGAACCTGGGGAACCTTGAGTGCAGGAAAAGGGGAGCGGAATTCACGTGTAGCGGTGAAATG
CGTAAAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTGCAGCTGAGGCGCGAAAGCGT
GGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACAGATGAGTGCTAAGTGTAGAGGGTATCCAC
CCTTTAGTGTGACGCAAAACGCAATTAAGCACTCCGCTGGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAAT
TGACGGGGACCCCGCAACAGCGGTGGAACCTTGAGTGCAGGAAAAGGGGAGCGGAATTCACGTGTAGCGGTGAAATG
GACATCCCCTGACAACCCCTAGAAGATAGGGCGTTCCCTTTTCGGGGACAGGGTGACAGGTGGTTGCATGGTTGTC
GTACCTCGTGTCTGAAAAGTTTGGGTTAATTTCCCGCCACCGAGCGCAACCCCTTTGACCTTAGTTGCCACCATT
CAATTTGGGCACTTCTAAGGTGACTGCCCGCTAAAATTTCCGAAGAAAAGGTGGGGGATGACGTCCAATCATCC
ATGCCCTTATGAACCCGGGGTAAACCCCTTCTTACTTCAATTTGGGGGGTAACAAGGGGTAACAAGGGGTAACCC
CGGGTTGGAACCAAAACCCCAAAAAACCCCTTCAATTTTGGATATGAGAGGGGGGCAATCCCCCGCGCTGA
AAAACCGGAAAACCCCTTTTAAATTCGGGAAAAACAATGCGCCCGGTGAAAAAATTTCCCGGGTTTTTTTAAATA
CCCC

> F01_27F

GTGGCATGCGCATGCTATACATGCAGTCGAGCGGACCAGAGATCCGGAGCTTGCTCTGAGTTTGGTCAGCGGGC
ACGGGTGAGTAACACGTGGGCAACCTGCCCGCAAGACCGGGGATAACTCCGGGAAACCGGAGCTAATACCGGATA
ACACCGAAGACCGCATGGTCTTTGGTTGAAAGGCGGCCTTTGGCTGTCACCTGCGGATGGGCCCGCGCGCATTAG
GCTAGTTGGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGGCCTGAGAGGGTGACCGGCCACACTGGG
ACTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGGCGAAAGCCTGACGGAG
CGACGCCGCGTGAGCGAAGAAGGCCTTCGGGTCTGTAAGCTCTGTTGTGAGGGACGAAGGAGCGCCGTTCAAG
AGGGCGGCGCGGTGACGGTACCTCACGAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGG
GGGCGAGCGTTTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAAGCCCACG
GCTCAACCGTGGAGGGTCAATTGAAACTGGGGGACTTGAGTGCAGGAGAGGAGAGCGGAATTCACGTGTAGCG
GTGAAATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGCAACTGACGCTGAGGCGC
GAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACAGATGAGTGCTAAGTGTAGAGG
GGTACACCCCTTAGTGTGCTGACGCTAACCGGATAAGCACTCCCGCCGGGGAGTACGGCCGCAAGGCTGAAACTC
AAAGGAATTGACGGGGGCCGCAACAGCGGTGGAACCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCA
GGTCTTGACATCCCCTGACAACCAAGAGATTGGGCGTTCCTCCCTTCGGGGGACAGGGTGACAGGTGGTTGCAT
GGTTGTGCTGAGCTCGTGTGAGATGTTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTGCCTTAGTTGCCA
GCCACGAAGGTGGGCACTCTAAAAGGGACTGCCCGCAACTTCCGAAGAAAAGGTGGGGGATGACGTCCAATCAT
CCAGGCCCTTATGACCTGGGGTACCACCGTCTTACAATGGGGCGGTACAAGGGGCTTCCAAAACCCCGGG
GGGGAACGAAAACCCCAAAAACCCCTTCAATTTTGAATGTCAGGGGGCAACTCTCCCGCTTAAAACCCGAATTC
CTTTTATATCCCT

>G01_27F

GTGACTGGCGCATGCCTATACATGCAGTCGAGCGGACCAGAAATCCGGAGCTTGCTCTGAGTTTGGTCAGCGGGC
GACGGGTGAGTAACACGTGGGCAACCTGCCCGCAAGACCGGGGATAACTCCGGGAAACCGGAGCTAATACCGGAT
AACACCGAAGACCGCAGGTCTTTGGTTGAAAGGCGGCCTTTGGCTGTCACCTGCGGATGGGCCCGCGGCGCATA

GCTAGTTGGT GAGGTAACGGCTACCAAGGCGACGATGCGTAGCCGGCCTGAGAGGGTGACCGGCCACACTGGG
ACTGAGACACGGCCAGACTCCTACGGGAGGCGACGATAGGGAATCTTCCGCAATGGGCGAAAGCCTGACGGAG
CGACGCCCGCTGAGCGAAGAAGGCCTTCGGGTCGTAAGCTCTGTTGTGAGGGACGAAGGAGCGCCGTTCAAG
AGGGCGGCGCGGTGACGGTACCTCACGAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGG
GGGCGAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAAGCCCACG
GCTCAACCGTGGAGGGTCAATTGAAACTGGGGGACTTGAGTGCAGGAGAGGAGAGCGGAATTCACGTGTAGCG
GTGAAATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGCAACTGACGCTGAGGCGC
GAAAGCGTGGGGAGCAAACAGGATTAGATAACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTTAGAGG
GGTACACCCCTTAGTGTGCTGACGCTAACGCGATAAGCACTCCGCCCGGGGAGTACGGCCGCAAGGCTGAAACTC
AAAGGAATTGACGGGGGCCGACAAAGCGGTGAAACCAGGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCA
GGTCTTGACATCCCCTGACAACCCAAGAGATTGGGCGTCCCCCTTCGGGGGACAGGGGTGACAGGTGGTTGCAT
GGTTGTCGTACGCTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTCGCCTCTTGTGGCA
GCACGAAAGTGGGCACTCTTAAAGGGACTGCCCGCGACAAATCCGAAGAAAGTGGGGGTAACCTTCAAATCAT
CAAGCCCCTTTGAACTGGGGTACCACCGTGTACAAATGGGGCGGTACAAGGGGTTGCAACCCCGGGG
GGAGGAATCCCCCAAACCCCTCTTATTTTTGATTGTAAGGGGGCAATCCCCCGCTTAAACCCCG

>H01_27F

GTGCATGCGCATGCTATACATGACGTCGAGCGGACCAGAGATCCGGAGCTTGCTCTGAGTTGGTCAGCGCGGAC
GGGTGAGTAACACGTGGGCAACCTGCCCGCAAGACCGGGATAACTCCGGGAAACCGGAGCTAATACCGGATAAC
ACCGAAGACCGCATGGTCTTTGGTTGAAAGGCGGCCTTTGGCTGTCACTTGCGGATGGGCCCGCGCGCATTAGC
TAGTTGGT GAGGTAACGGCTACCAAGGCGACGATGCGTAGCCGGCCTGAGAGGGTGACCGGCCACACTGGGAC
TGAGACACGGCCAGACTCCTACGGGAGGCGAGCATAGGGAATCTTCGCAATGGGCGAAAGCCTGACGGAGCG
ACGCCGCGTGAGCGAAGAAGGCCTTCGGGTCGTAAGCTCTGTTGTGAGGGACGAAGGAGCGCCGTTCAAGAG
GGCGGCGCGGTGACGGTACCTCACGAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGG
GGCGAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAAGCCCACGG
CTCAACCGTGGAGGGTCAATTGAAACTGGGGGACTTGAGTGCAGGAGAGGAGCGGAATTCACGTGTAGCGG
TGAAATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGCAACTGACGCTGAGGCGCG
AAAGCGTGGGGAGCAAACAGGATTAGATAACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTTAGAGGG
GTCACACCCCTTAGTGTGCTGACGCTAACGCGATAAGCACTCCGCCCGGGGAGTACGGCCGCAAGGCTGAAACTCA
AAGGAATTGACGGGGGCCGACAAAGCGGTGGAACCAAGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAG
GTCTTGACATCCCCTGACAACCCAAGAAGATTGGGCGTCCCCCTTCGGGGGACAGGGGTGACAGGTGGTTGCAT
GGTTGTCGTACGCTCGTGTGAAATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTCGCCTCTTGTGGCAG
CCACAAAGGTGGGCACTCTTAAAGGGACTGCCCGGCCACAAATCCGAAGAAAAGTGGGGGATAAAATCAAATCAT
CAAGGCCCTTATTAACCGGGGGTTAAACACGTGGTCCAATGGGGCGGTTACAAAGGGGTGCCAACACACCC
CGGGGGGAGCGAAATCCCCAAAACCTC

>A02_27F

GTGACTGCGCATGCTATACATGCAAGTGCAGCGGAGCGACGGGAGCTTGCTCCCTTAGGTGACGGCGGGACGG
GTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATGCTTGA
TTGAACCGCATGGTTCAATCATAAAAGTGGCTTTTAGCTACCACTTACAGATGGACCCCGCGGCGCATTAGCTAGT
TGGTGAGGTAACGGCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCCACTGGGACTGAG
ACACGGCCAGACTCCTACGGGAGGCGAGTAGGGAATCTTCGCAATGGACGAAAAGTCTGACGGAGCAACGC
CGCGTGAGTGATGAAGTTTTGGGATCGTAAACTCTGTTGTTAGGGAAGAACAAGTACCCTTCAATAGGGCGGT
ACCTTGACCGTACCTAACAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCG
TTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTTCTTAAGTCTGATGTGAAAGCCCCCGGCTCAACCG
GGGAGGGTCAATTGAAACTGGGGAACCTTGAGTGCAGAAGAGGAGAGTGGAATTCACAGTGTAGCGGTGAAATGC
GTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGACTCTGTTGTTGTAAGTACTGACGCTGAGGCGGAAAAGCGTG
GGGAGCGAAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTTCCCGCC
CTTTAGTGTGCTGACGAAACGCATTAAGCACTCCGCTGGGGGAGTACGGTCCGCAAGACTGAAACTCAAAGGAATT
GACGGGGGCCGACAAAGCGGTGGAGCATGTGTTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACA
TCCTCTGACAACCCTAGAGATAGGGCTTCCCTTCGGGGGAGAGTGACAGGTGGTTGCATGGTTGTCGTACAGT
CGTGTGCTGAGAAGTTTTGGGGTTAAGTCCCGCAACGAGCGCAACCCCTTGATCTTATGTTGCCAGCCATTCAAGTTGG
CACTCCTAAGGTGACTGCCGTGACAAACCGAAGAAAAGGGGTGGGGATGAAATCAAATAATCAAGGCCCTTAT
AAACCGGGGTAAACAATTTGTTTCAATGGGGCGAAAAAAGGGGGCCAAACCGCCGGGGTAAGCAATCCACAC
AAAATATTTTTTCATTTTTGTATACAAATCTGCCACTCCACCCGCGTGGAAGTTGGAACCCCTTTTTAATCCGGAT
ACAACAAGCCCCGGCGTAAAAATATCTCGGGCTGTTACACCGCCGGTCCCACCACAAAGGTGGGTAACCC
AATTG

>B02_27F

CTGCAGTGCAGTATACATGCAAGTGCAGCGGACCAGGAGCTTGCTCCTTTAGGTTAGCGGCGGACGGGT
GAGTAACACGTGGGCAACCTGCCCTGCAGACGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACACCGA
AAACCGCAGGGTTTTCGGTTGAAAGGCGGCTTTTAGCTGTCACTGCAGGATGGGCCCGCGGCGCATTAGCTAGTT
GGTGAGGTAACGGCTACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGAGA
CACGGCCAAACTCCTACGGGAGGCGAGTAGGGAATCTTCGCAATGGACGAAAAGTCTGACGGAGCAACGCC
CCGTGAGCGAAGAGGTCTTCGGATCGTAAAGCTGTTGTGAGGGAAGAACAAGTCCGCTTCAAGCAGGGCGGT
ACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGC
GTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAATCTCGCGGCTCAACC
CGGAGCGGCCATTGAAACTGGGGAACCTTGAGTGCAGGAAAGGGGAGCGGAATTCACGTGTAGCGGTGAAATG

CGTAAAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTACGCTAAGGCGCGAAAGCGT
GGGGAGCGAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTATCCAC
CCTTTAGTGCTGCAGCAAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATT
GACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACA
TCCCCTGACAACCCTAGAGATAGGGCGTTCCCTTTTCGGGGACAGGGTGACAGGTGGTGCATGGTTGTCGTGAGC
TCGTGTGCTGAAATGTTGGGTTAAGTCCCAGCAACGAGCGCAACCCCTTACCTTAGTTGCCAGCATTCAATTTGGGCA
CTCTAAGGTGACTGCCGGCTAAAATTCCGAAGAAAGGGGGGGGATGACCTTCAAATCATCAAGGCCCTTATAAAC
TGGGGCTACCAACTTGCTACAATGGGGGGGTACAAAGGGCACAAAACCCCGGGGGGAGCAAACCCAAAAAAC
CCCCATTTTTGAATTAAGGCGGCACTTCCCGCGGAAAACCGAACCCCTTTTATTCGGGAAAAAAGCCCCGTG
AAAAATTTCCCGGCTTTTACACCCCCCTCACCCCAAAATTTTACCCCCCAATCTGGGGGTACCTTCCGGGAC
CCCCAATG
GA

>C02_27F

CTGGCTGCGCATGCCTATACATGCAGTCGAGCGGACCGAAGGGAGCTTGCTCCTTTAGGTTAGCGGGGACGGG
TGAGTAACACGTGGGCAACCTGCCCTGCAGACTGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACACCG
AAAACCGCATGGTTTTCGGTTGAAAGGCGGCTTTTAGCTGTACTGCAGGATGGGCCCGCGGCGCATTAGCTAGT
TGGTGAGGTAACCGTACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACCGGGACTGAG
ACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGC
CCCGTAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTGAGGGAAGAACAAGTGCCGTTCAACAGGGCGG
TACCTTAGCGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGC
GTTGTCGGGAATTTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAATCTCGCGGCTCAACC
GCGAGCGGCCATTGAAAACCTGGGGAATTTAGTGCAGGAAAGGGGAGCGGAATTCACGTGTAGCGGTGAAATG
CGTAAAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTACGCTAAGGCGCGAAAGCGT
GGGGAGCGAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTATCCAC
CCTTTAGTGTGCAGCAAACGCATTAAGCACTCCCGGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATT
GACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACA
TCCCCTGACAACCCTAGAGATAGGGCGTTCCCTTTTCGGGGACAGGGTGACAGGTGGTGCATGGTTGTCGTACC
TCGTGTCTGAAAATGTTGGGTTAATTTCCCGCACCGAGCGCAACCCCTTACCTTAGTTGCCAGCATTCAATTTGGC
ACTCTAAGGTGACTGCCGGCTAAAATTCGAAGAAAGGTGGGGGATGACGTCCAATCATCAAGCCCCCTTATGACC
TGGGCTACCACGTCTTAATGGGGGGTACAAAGGGCACAAAACCCCGAGTTGAAACGAATCCCAAAAAACCC
CCCCCTATTTGAAATTTAGGGGGGAAATCCCCCCCCATAAAACCCGAAACCCCTTAAATTCGGGAAAAACAAGCCCC
GTGAAAAAATTTCCCGGTTTTTTTACAACCCCGTCCCCACCAAAAAATTTGTTAACCCCAATCCGTTGGGTACCTT
CCTGAGCCCCCATGGCAATGGT

> D02_27F

GTGACTGCGCTGCCTATACATGCAAGTCGAGCGGACCGAAGGAGCTTGCTCCTTTAGGTTAGCGGGCGGACGGGTG
AGTAACACGTGGGCAACCTGCCCTGCAGACTGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACACCGAA
AACCGCAGGGTTTTTCGGTTGAAAGGCGGCTTTTAGCTGTACTGCAGGATGGGCCCGCGGCGCATTAGCTAGTTG
GTGAGGTAACGGCTACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGAGAC
ACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCC
CGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTGAGGGAAGAACAAGTGCCGTTCAACAGGGGCGGTA
CCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCG
TTGTCGGGAATTTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAATCTCGCGGCTCAACCG
CGAGCGGCCATTGAAAACCTGGGGAATTTAGTGCAGGAAAGGGGAGCGGAATTCACGTGTAGCGGTGAAATGC
GTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTACGCTGAGGCGCGAAAGCGTG
GGGAGCGAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTATCCACC
CTTTAGTGTGCAGCAAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATT
GACGGGGACCCGCACAAGCGGTGGAACCATGTGGTTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGA
CATCCCCCTGACAACCCTAGAAGATAGGGCGTTCCCTTTTCGGGGACAGGGTGACAGGTGGGTGCATGGTTGTCGT
CACCTCGTGTCTGAAAATTTGGGTTAAGTCCCAGCAACGAGCGCAACCCCTTACCTTAGTTGCCACCATTCAGTT
GGGCACTCTAAGGTGACTGCCGGCTAAAATTCGGAAGAAAAGTGGGGGATGAACTCAAATCATCCAGCCCCCTA
AAAACCTGGGCTACCCCTGCTTACAAGGGGGGGGTTCAAAGGGGACCAAAAACCCCGGTTGGACGAATCCCCA
AAAACCCCCCTATTTTGAATTCAGGCGGCACTCCCCCGGCTAAAAAACCGAAATCCTTTTTATTCGGGAAAA
ACATACCCCGTAAAAAATTTCCCGGTTTTGTGTAACACCGCGGGTACCCCAAAAAATTTGGTAAACCCCAAT
CCTGGGGTACCTTTCCGGACCCCCCAATGGCAATTT

>E02_27F

CTGGCTGCGCTGCCTATACATGCAGTCGAGCGGACCGACGGGAGCTTGCTCCTTAGGTCAGCGGGCGGACGGGT
GAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATGCTTGATT
GAACCGCATGGTTCAATCATAAAGGTGGCTTTCAGTACCCTTGCAGATGGACCCGCGGCGCATTAGCTAGTTG
GTGAGGTAACGGCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGAC
ACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCC
CGTGAGTGATGAAGTTTTCGGATCGTAAACTCTGTTGTTAGGGAAGAACAAGTACCCTTGAATAGGGCGGTA
CTTGACGGTACCTAACCAAGAAAGCCAGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTT
GTCCGGAATTTGGGCGTAAAGCGCGCGCAGGCGGTTTTCTTAAGTCTGATGTGAAAGCCCCCGGCTCAACCGG
GAGGGTCAATGAAAACCTGGGGAATTTAGTGCAGAAGAGGAGTGGAAATTCACGTGTAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACCAGTGGCGAAGGCGACTCTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGG

AGCGAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTTAGAGGGTTCCGCCCTTT
AGTGCTGCAGCAAACGCATTAAGCACTCCGCCCGGGGAGTACGGTCGCAAGACTGAAACTCAAAGGAATTGACG
GGGGCCCCGCACAAGCGGTGGAACCAGGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCC
TCTGACAACCCTAGAGATAGGGCTTCCCCTTCGGGGGCAGAGTGACAGGTGGTTGCATGGTTGTCGTAGCTCGT
GTCGTGAGATGTTTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTTGATCTTAGTTGCCAGCATTCAGTTGGGCACTC
TTAAGGTGACTGCCGGTGACAAAACCCGAAAGGGTGGGGGATGAACTTCAAATCATCAAGCCCCCTTATGACCT
GGGGCTACACACGTGGCTACAATGGGGCAAAAACAAAAGGCAACAAAACCCCCCAGGCTAAGCCAATCCCCAAA
TCTTTTTCCATTTCCGAATCCAATCTGGCAATTCACCCGCTTAAAAGTGGAAACCCCTTTTAAATCCGGATAACCAA
GCCCCCGGTAAAAATTTCCCGGGCCTGTTAATCCCGCGCTCACACCCCAAATTTGAAAACCCAAATCCTTGAG
GACCTTTGACCCGCCCATGGCAATTTGAGGGGGGTT

> F02_27F

GTACAGTGGCGCGTGCTACACATGCAGTCGAGCGGACGCTGAACGGAGACTTGCTCCTTTAGGTTAGCGGCGG
ACGGGTGAGTAACACGTGGGCAACCTGCCCTGCAGACTGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAA
CACCGAAAACCGCAGGGTTTTCGGTTGAAAGCGGGCTTTTAGCTGCTACTGCAGGATGGGCCCGCGGCATTAG
CTAGTTGGTGAGGTAACGGCTACCAAGGCGACGATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGA
CTGAGACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGC
AACGCCCGGTGAGCGAAGAAGGTCCTCGGATCGTAAAGCTCTGTTGTCAGGGAAGAACAAGTCCCGTTCCGAACG
GGCGGTACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTG
GCAAGCGTTGTCCGGAATTTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAGTCTGATGTGAAATCTCGCGGC
TCAACCCGCGAGCGGCCATTGAAACTGGGGAACCTTGAGTGCAGGAAAGGGGAGCGGAATTCACGTGTAGCGGT
GAAATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGCGGCTCTCTGGCCTGTAAGTACGCTGAGGCGCGA
AAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTTAGAGGGT
ATCCACCCTTTAGTCTGCAGCAAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGAAGGCTGAAACTCAA
GGAATTTGACCGGACCCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGT
TTGACATCCCCTGACAACCCTAGAGATAGGGCTTCCCCTTCGGGGACAGGGTGACAGGTGGTGCATGTTGTC
GTCAGCTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTTGACCTTAGTTGCCAGCATTCACT
TGGGCACTCTAAGGTGACTGCCGGCTAAAAGTCCGAAGAAAGGGGGGGATGACGTCAAATCATCATGCCCTTA
TGACCTGGGCTACCACGTGCCTACAATGGGGTGGTACAAAGGGCACCGAAAACCCCCGAGGGGGAAACGAATCCA
AAAAACCCCTTTCTATTTCCGATTGCAGGGTGAACCTCCCCTGCTTAAAACCCGAATCCCTTATAAATCCGGGATTA
ACAATGCCCCCGGAAAAAATTTCCCGGGTCTGTTACCCCGCCCGTACCCCCCAAATTTTTTAAACCCCGG
AATGG

>_G02_27F

GTGCAGTGCAGCATGCTATACATGCAGTCGAGCGGACGATTCAAGCTTGCTTTTGAATCGTTAGCGGCGGACGGGT
GAGTAACACGTGGGCAACCTGCCCTGTAGACGGGGATAACCCGAGAAATCGGTGCTAATACCGGATAATACGAA
AGGCCGCATGGTCTTTTCGTTGAAAGGCGGGCAGGCTGTGCTACAGGATGGGCCCGCGGCATTAGCTAGTT
GGTGAGGTAACGGCTACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGATCGGCCACACTGGGACTGAGA
CACGCCCAAGACTCCTACGGGAGGCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCC
GCGTGAGCGAAGAAGGCCCTTCGGGTCGTAAGCTCTGTTGTTAGGGAAGAACAAGTACCGCAGTCACTGGCGGTA
CCTTGACGGTACCTAACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGT
TGTCCGGAATTTGGGCGTAAAGCGCGCGCAGGCGGCTTTTAAAGTCTGATGTGAAAGCCCACGGCTCAACCGT
GGAGGTCATTGAAACTGGGGACTTGAGTGCAGAAGGAGAGCGGAATTCACGTGTAGCGGTGAAATGCG
TAGAGATGTGGAGGAACCCAGTGGCGAAGGCGCTCTGTTGTTGTAAGTGCAGGCTGAGGCGCGAAAGCGTGG
GGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTTAGAGGGTATCCACCCT
TTAGTGCTGTAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGCTCGCAAGAGTGAACCTCAAAGGAATTTGAC
GGGGCTCCGCACAAGCGGTGGAGCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCC
CCTGACAACCCAGGAGAAGATCGGGCTTCCCCTTCGGGGACAGGGTGACAGGTGGTGCATGTTGTCGTACG
CTCGTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTGACCTTAGTTGCCAGCATTCACTTGGG
CACTCTAAGGTGACTGCCGGCTAAAATTCGAAGAAAGGTTGGGGATGACTTCAAATCATCAGGCCCTTATAACCG
GGGCTACACCTTGTACAATGGGCCGTACAAAGGGTCCCGAACCCCCAGGGGGAGCCATCCCCAAAAACCGCT
CTCCATTTTGAATGTAGCGCGGCACTCCCCTGCTTAAAACCGAAACCTTTTAAATTCGGAAAAACAAGCCGGG
GAAAAAATTTCCCGGGCTGTGTAACCCG

>H02_27F

GGCAGTGCAGCATGCCTATACATGCAGTCGAGCGGACGAAATCGGAGCTTGCTCTGAGTTTGGTCAGCGCGGAC
GGGTGAGTAACACGTGGGCAACCTGCCCGCAAGACCGGGATAACTCCGGGAAACCGGAGCTAATACCGGATAAC
ACCGAAGACCGCAGGTCTTTGGTTGAAAGGCGGGCTTTGGCTGCTACTTGCAGGATGGGCCCGCGGCATTAGCT
AGTTGGTGAGGTAACGGCTACCAAGGCGACGATGCGTAGCCGGCCTGAGAGGGTGACCGGCCACACTGGGACT
GAGACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTCCGCAATGGGCGAAAGCCTGACGGAGCGA
CGCCCGTGAGCGGTAAGAAGGCTTCGGGTCGTAAGCTCTGTTGTTGAGGGACGAAAGCCTTCAAGAGG
GCGGCGCGGTGACGGTACCTCACGAGGAAGCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGG
GCGAGCGTTGTCCGGAATTTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAAGTCTGATGTGAAAGCCCACGGC
TCAACCGTGGAGGGTCAATTGAAACTGGGGACTTGAGTGCAGGAGAGGAGAGCGGAATTCACGTGTAGCGGT
GAAATGCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTGCGCTGCAACTGACGCTGAGGCGCGA
AAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTTAGAGGGG
TCACACCCTTTAGTCTGCAGCTAACCGGATAAGCACTCCGCCCGGGGAGTACGGCCGAAGGCTGAAACTCAA
AGGAATTTGACGGGGCCCGCACAAGCGGTGGAACCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGG

TCTTGACATCCCCTGACAACCCAAGAAGATTGGGCGTTCCCCCTTCGGGGGACAGGGTGACAGGTGTTGCAGG
GTTGTCGTGAGCTCGTGTGAGAAATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTCGCCTCTTGTGCCAC
CACTAAAGGTGGGGCACTCTAAAGGGACTGCCGCGACAAATCCGAAGAAAAGGTGGGGGATGAAAATCAAAT
AATCAAGGCCCCCTTATAAACTGGGGGTACCAACCTGGTTACAATGGGGCGGGTACAAAGGGTTGGCGAACCC
CCCCGGGGGAAGGAAATCCACAAAAAACCCCTCTTTTTTTGAATTGAAGGGTGAAATCTCCCCCCCCTGAAAC
CCGAAA

>A03_27F

GGCATGCGCATCCTATACATGCAGTCGAGCGGACCAAATCGGAGCTTGCTCTGAGTTTGGTCAGCGGCGGACGGG
TGAGTAACACGTGGGCAACCTGCCCGCAAGACCGGGATAACTCCGGGAAACCGGAGCTAATACCGGATAACACCG
AAGACCGCAGGTCTTTGGTTGAAAGGCGGCCCTTTGGCTGTCACTTCCGGATGGGCCCGCGGCATAGCTAGTTG
GTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGGCTGAGAGGGTGACCGGCCACACTGGGACTGAGAC
ACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTCCGCAATGGGCGAAAGCCTGACGGAGCGACGCCG
CGTAGCGAAGCAAGGCTTCGGGTCTGAAAGCTCTGTTGGGGGACGAGGAGCGCCGTTGGAAGAGGGCGGC
GCGGTGACGGTACCTCACGAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGGGGCGAGC
GTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAAGCCACGGCTCAACC
GTGGAGGGTCAATTGAAACTGGGGGACTTGAGTGCAGGAGAGGAGCGGAATCCACGTGTAGCGGTGAAATG
CGTAGAGATGTGGAGAACACCACTGGCGAAGGCGGCTCTCTGCCCTGCAACTGACGCTGAGGCGCAAAAGCGT
GGGGAGCAACAGGATTAGATACCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGGTCACAC
CCTTTAGTGTGACGCTAACGCGATAAGCACTCCGCTGGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAAT
TGACGGGGCCCGCAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGACCTTACCAGGTCTTGACAT
CCCCTGACAACCAAGAAGATTGGGCGTTCCCCCTTCGGGGGACAGGGTGACAGGTGGGTGATGGTTGTCGT
CAGCTCGTGTGAGAAAGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTCGCCTCTAGTTGCCAGCACCTAAG
GTGGGCACTCTAAAGGGACTGCCGCGACAAATCCGAAGAAAAGGTGGGGGATAACGTCAAATCATCCAGGCC
CCTTATAACCGGGGGCTACAACTGGTTACAATGGGCGGTACAAAGGGGTGCGAACCCACCCAGGGGGGAGCA
AATCCAAAAAACCCCTCTCTTTTTGAATTAAGGGGGCACTCCCCGGCTTGAACCCGAATCCCTTTAATTC
CGGATAAACATGCCCGCTGGAAAAATTCGCCGGCCTGTGTAACCCCGCGGTTACCCCCAAAAGTTTGACCCCC
CAAATTCGTGGGTACCTCAGGGACCCCCCAATGGGAATTGGGGT

>B03_27F

GGCGTGGCATGCCTATACATGCAGTCGAGCGGACTTTCAAGAAGCTTGCTTTTTGAAAGTTAGCGGCGGACGGG
TGAGTAACACGTGGGCAACCTGCCTGCAAGACGGGGATAACTCCGGGAAACCGGGGCTAATACCGGATAATATCT
TCCTTCGCATGAAGGAAGTTGAAAGGCGGCGCAAGCTGCCGCTTGACAGATGGGCCCGCGGCATTAGCTAGTT
GGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACTGAGAGGGTGATCGGCCACACTGGGACTGAGA
CAGGCCACAGCTCCTACGGGAGGCAGTAGGGAATCTCCGCAATGGACGAAAGTACCGGAAACCGG
GCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTCAGGGAAGAACAAGTACCGTTGCAACAGGGCGGT
ACCTTGACGGTACCTGACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCG
TTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGCTCTTAAGTCTGATGTGAAAGCCCACGGCTCAACCG
TGGAGGGTCATTGAAACTGGGAGACTTGAGTGCAGAAGAGGAGCGGAATCCACGTGTAGCGGTGAAATGCG
TAGAGATGTGGAGGAACACCACTGGCGAAGGCGGCTCTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTGG
GGAGCGAACAGGATTAGATACCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGCTTCCACCC
TTTAGTGTGACGCTAACGCATTAAGCACTCCGCTGGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTG
ACGGGGCCCGCACAAAGCGGTGGAACCAAGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACA
TCCTTCGTACCTCTAGAGGATAGAGGGTTCCCTTCGGGGACGATTGACAGGTGGTTGATGGTTGTCGTCA
GCTCGTGTGAGAAATGTTGGGGTTAAGTCCCGCAACGAGCGCAACCCCTGGACCTTAGTTGCCACCATTCAGTTG
GGCACTCTAAGGTGACTGCCGGTGACAAACCCGAAGAAAAGGTGGGGGATGAACTTCAAATCATCCAGGCCCCCT
TATGAACCCGGGCTCACACCTGGTTCCAATGGGAGGGTTACAAAGGGTCCCGAAACACCCCGGTTGGAGCCAA
TCCCAAAAACATTTTTCTAATTTTTGAATTTGAGGGGGCAATTCCTCCCTGGAAAACCCGAAATCCTTTAATTCG
GGATAAACATGCCCGCGGTAGAAAAATATCCCGGGCCCTGTTAACCCCGCCGTTCCCCACACAAAGTTTGGTAC
CACCAATCCGGGAGGAACCTTACGGGCCCCCAAGGGCGGATTGGGGTTGTTTTGGGGCGGGGGGGCGTTA
AAAGTT

>C03_27F

CTGAAGTGCCTGCCTATACATGCAGTCGAGCGAACCAATAAGAAGCTTGCTTTTTGTTGGTTAGCGGCGGACGG
GTGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGATTAT
CTTCCGCTGGAGAGATAAGGAAAGATGGCTTTTGTCACTTACAGATGGGCCCGCGGCATTAGCTAGTTG
GTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCAACCTGAGAGGGTGATCGGCCACACTGGGACTGAGAC
ACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCG
CGTAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGGAGGGAATGCCGGTAC
CTTGACGGTACTGACGAGAAAGCCACGGCTAACTGCGCAGCAGCCGCGGTAATACGTAGTTGGCAAGCGTT
GTCCGGATTTATTGGGCGTAAAGCGCGCAGGCGGCTCTTTAAGTCTGATGTGAAATCTTCCGGTCAACCGCA
AGCGGTCAATTGAAACTGGGGACTTGAGTGCAGAAGAGGAAAGCGGAATCCACGTGTAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACCACTGGCGAAGGCGGCTTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGG
AGCAACAGGATTAGATACCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTGGGAGGGTTCCGCCCTT
CAGTCTGACGCTAACGCATTAAGCACTCCGCTGGGGGAGTACGGTTCGCAAGACTGAAACTCAAAGGAATTGAC
GGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCT
CCTGACCGCCCTGGAGACAGGGTCTCCCTTCGGGGACAGGATGACAGGTGGTGCATGTTGTCGTGAGCTCGT
GTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTGGTTCTAGTTGCCAGCATTCATTTGGGCACTC

TAGAACGACTGCCGGCCACAATTCCGAAGAAAGGTGGGGGATGAAGTCCAATCATCATGGCCCCCTATAAACTGG
GGCTACAACCTGGTTACAATGGAAGGTTACAAAGGGCGGCAAAACCGGCCACCCAAAACGAATCCCGAAAAACATT
TCTATTTTTGAATGAAGGGGGGACTCCCCCGCGAAAACCGGAAACCCTTAATTCGGGATACACAGCCGCGGGAA
AAAAATCTCGGCCTGTTATACCGCCGTTCCACCAAATTTGTAACCAATTCGTGAGTAACCTAGGACGCCCAT
GGCAGTTGGTGTGTTTCATTGACTGACGGGGGGGGGGGGATTGGGGGAAATTTAAAAA

>D03_27F

GGTTGTACAGTATAACTTATACTTTTTGCTGACAGTCGTAGCGGAGTATTCGTAACCTGCTTTTTGGATCGTTAGCG
GCGGACGGGTGATTAACACGTGGGCAACCTGCCCTGTAGACGGGGATAAACACCGAGAAATCGGTGCTAATACCGG
ATAATACGAAAGGCCGCATGGTCTTTTCGTTGAAAGGCGGCGCAAGCTGTGCTACAGGATGGGCCCGCGGCAT
TAGCTAGTTGGTGAAGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTG
GGACTGAGACACGGCCAGACTCCTACGGGAGGCGACGATAGGGAATCTTCGCAATGGACGAAAGTCTGACGG
AGCAACGCCGCGTGAGCGAAGAAGGCCTTCGGGTCGTAAGCTCTGTTGTTAGGGAAGAACAAGTACCGCAGTCA
CTGGCGGTAGGAGCAACAGGATTACGATAGGAAGCCACGGCTAACCTACGTGCCAGCAGCGCGGTAATACGTAGGT
GGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTCTTTAAGTCTGATGTGAAAGCCCACGG
CTCAACCGTGGAGGGTCAATTGAAACTGGGGACTTGAGTGCAGAGAGGAGAGCGGAATTCACGTGTAGCGG
TGAATGCGTAGAGATGTGGAGGAACACCAAGTGGCGAAGGCGGCTCTCTGGTCTGTAACGTACGCTGAGGCGCG
AAAGCTGGGGAGCAAAACAGGATTAGATACCCTGGTAGTCCACGCGTAAACGATGAGTGCCTAAGTGTAGAGGG
TATCCACCCTTAGTGCTGTAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGCTCGCAAGAGTGAACCTCAA
GGAATTGACGGGGGCCCGACAAGCGTGAGGACCATGCGGTTTAAATTTGAATGACCAGTTAAGACCTTCTCCCGA
GACTCGGAACTAAGGTCTTCCCTCCACATGACAAGTCTACCGTCGTTGCCAGGATTACCATCAGGTCACATGGTGG
AAACTCGGCCTTTAACCAACAA

>E03_27F

GTTATCGGGCGCATGCCCTATACATGCAAGTCGAGCGGACCGAAGGGAGCTTGCTCCTTTAGGTTAGCGGGGAC
GGGTGAGTAACACGTGGGCAACCTGCCCTGCAGACGGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACA
CCGAAAACCGCAGGGTTTTCGGTTGAAAGGCGGCTTTTAGCTGTCACTGCAGGATGGGCCCGCGGCATTAGCT
AGTTGGTGAGGTAACGGCTCACCAAGGCGACAATGCGTAGCCAACTGAGAGGGTGACCGGCCACACTGGGACT
GAGACACGGCCAGACTCCTACGGGAGGCGAGCAGTAGGGAATCTTCGCAATGGACGAAAGTCTGACGGAGCAA
CGCCCCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTCAGGGAAGAACAAGTCCGCTTGAACAGGG
CGGTACCTTGACGTTACCTGACGAGGAAGCCACGGCTAACCTACGTGCCAGCAGCGGTAATACGTAGGTGGC
AAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAGTCTGATGTGAAATCTCGCGGCTC
AACCGCGAGCGGCCATTGAAACTGGGGAACCTTGAGTGCAGGAAAGGGGAGCGGAATTCACGTGTAGCGGTGA
AATGCGTAAAGATGTGGAGGAACACCAAGTGGCGAAGGCGGCTCTCTGGCTGTAACGTACGCTGAGGCGCGAAA
CGGTAGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCGTAAACGATGAGTGCCTAAGTGTAGAGGGTAT
CCACCCTTTAGTGCTGCAGCAAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAG
GAATTGACGGGGACCCGCACAAGCGGTGGAACCATGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGT
CTTGACATCCCCTGACAACCTAGAAGATAGGGCGTTCCCTTTTCGGGGACAGGGTGACAGGTGGTGCATGGTTG
GACTCAGTCTGTGCGTGAATAATGTTGGGTTAAGTCCGCAACCGAGCGCAACCCTTGACCTTATTTGACAGCTT
CATTTGGGCCCTCTAAGGTGACTGCCGGCTAAAAATCCGAAGAAAGGGTGGGGGATGAAGTTCAAATCATCAAGC
CCCCTAAGAACCTGGGGTACCACCTTGCTTCCAATGGGGTGGTTACAAAGGGGAGCCAAAACCCCGGGTGGGA
GCGAATCCAAAAAACCCACCCCTTTTTCGGAATGGAGGGGGGCAATCCCCTCGCTTGAAACCCGAATACCTTTTT
AATTCGGGAAAAAAAGCCCGCGTGGAAAAAAATTTCCCGGTTTTGTTAACCCCCCGCTCCCCCCCCGAGAT
TTTTTACCCCAACCTTGGGGGTACCTTACCAGAAACCCCGCCCAATGGCAAAATTTGGGTTTTGTTTTACCGGTTCCCA
TACAAATTT

>F03_27F

GTGAAGTGCGCATGCCTATACATGCAGTCGAGCGGACCGAAGGGAGCTTGCTCCTTTAGGTTAGCGGCGGACGG
GTGAGTAACACGTGGGCAACCTGCCCTGCAGACGGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACACC
GAAAACCGCAGGGTTTTTCGTTGAAAGGCGGCTTTTAGCTGTCACTGCAGGATGGGCCCGCGGCATTAGCTAG
TTGGTGAGGTAACGGCTCACCAAGGCGACAATGCGTAGCCAACTGAGAGGGTGACCGGCCACACTGGGACTGA
GACACGGCCAGACTCCTACGGGAGGCGAGCAGTAGGGAATCTTCGCAATGGACGAAAGTCTGACGGAGCAACG
CCCCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTCAGGGAAGAACAAGTCCGCTTGAACAGGGCG
GTACCTTGACGTTACCTGACGAGGAAGCCACGGCTAACCTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAG
CGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAGTCTGATGTGAAATCTCGCGGCTCAAC
CGCGAGGGCCATTGAAACTGGGGAACCTTGAGTGCAGGAAAGGGGAGCGGAATTCACGTGTAGCGGTGAAAT
GCGTAAAGATGTGGAGGAACACCAAGTGGCGAAGGCGGCTCTCTGGCTGTAACGTACGCTAAGGCGCGAAAGCG
TGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCGTAAACGATGAGTGCCTAAGTGTAGAGGGTATCCA
CCCTTTAGTGCTGCAGCAAACGCATTAAGCACTCCGCCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAA
TTGACGGGGACCCGCACAAGCGGTGGAACCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTG
ACATCCCCTGACAACCCCTAGAAGATAGGGCGTTCCCTTTTCGGGACAGGGTGACAGGTGGTGCATGGTTGTCGT
CAGCTCGTGTGCGTGAAGGTTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGACCTTAGTTGCCAGCATTCAAGT
GGGACTCTAAGGTGACTGCCGGCTAAAAATCCGAAGAAAGGGTGGGGGATGAACTTCAAATCATCAAGCCCCCT
TAGAACCTGGGCTACCACCGTGGTTACAATGGGGGGGGTACAAAGGGGAGCCAAAACCCCGGGTGGAGCGA
AACCCCAAAAACCTTCTCATTTCTGATTGAGGGGGGCAATCCCCCGGCTGAAAACCGGAATACCTTTTT
AATTCGGGAAAAAACTAGCCCCCGTGAATAATTTCCCGGTTTTGGTAACCCCCCTTCCCCCCCCAAAT
TTGTTAAAACCCCAATTCCTTGGGGTACCTTACCAGGAAACCCCAATGGCAAAATTTGGGGGGGGTTTTTAA
AAAAAGGCC

>G03_27F

GGGGACAGTGCCTTGCCTAACTGCAGTCGAGCGGACCGAAGGGAGCTTGCTCCTTTAGGTTAGCGGCGGACG
GGTGAGTAACACGTGGGCAACCTGCCCTGCAAAACAGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACAC
CGAACACCGCACGGTTTTTCGGTTGAAAGGCGGCTTTTACTGTCACACTGCACGATGGGCCCCGCGGCGCATTAGCTA
GTTGGTGAAGTAACGCTCACCAAGGCGACCATTCGTAGCCACCTGAAAGGGTGACCGGCCACACTGGGACTG
AAACACGGCCAACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAACAACG
CCCCGTGAGCGAAAAAGGTCCTCGATCGTAAAGCTCTTTTGTACAGGAAAAACAAGTGCCTGTTGAAACAGGGCG
GTACCTTGACGGTACCTGACGAGGAAGCCACGGCTAAGTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAG
CGTTGTCCGGAATTTGGGCGTAAAGCGCGCGCAGGCGGTTTCTTAAGTCTGATGTGAAATCTCGCGGCTCAAC
CGCGAGCGGCCATTGAAACTGGGGAATTTGAGTGCAGGAAAGGGGAGCGGAATCCACGTGTAGCGGTGAAAT
GCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTACGCTGAGGCGCGAAAGCG
TGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGTAAAGTGTAGAGGGTATCCA
CCCTTTAGTGTGCAGCAAACGCATTAAGCACTCCGCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAAT
TGAGGGACCCCGTCAACAAGCGGTGAGCTACAAAGGGGCGAAAAACCCCGGTGGAGGGAATCCCAAAAAAC
ATCCCTGACAACCCTAGAGATAGGGCGTTCCCTTTTCGGGAACAAGGTGACAGGTGGTGCATGGTTGTCGTAC
CTCGTGTCTGAAAAATGTTGGTTAATTCGCAACGAGCGCACCTTTGACCTTAGTCCACCATTCAATTTGGG
CCCTCTAAGGTGACTGCCGGCCTAAAAATCCCAAGAAAGTGGGGGATGACGTCAAACCTCCAGCCCCCTTAAAC
TGGGGCAACCCCGTCTACAATGGGGGGTACAAAGGGGCGAAAAACCCCGGTGGAGGGAATCCCAAAAAAC
CCCCCTCATTTTGAATTGAGGGGGGAAACCCCGCTTGAACCCGAAACCCAGTTAATTCGGGAAAAATGGCCCC
GGGGAAACATTCCCGGCCCTTTTACCCTCCGGTCCGGTCAAG

>H03_27F

GTTGACGTGCGCTGCCTATACATGCAGTTCGAGCGAACCAATAAGAAGCTTGCTTTTTGTTGGTTAGCGGCGGACG
GGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGATT
ATCTTTCCGCCTGGAGAGATAAGGAAAGATGGCTTTTGTATCATTACAGATGGGCCCGCGGCGCATTAGCTAGT
TGGTGAAGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACTGAGAGGGTATCGGCCACACTGGGACTGAG
ACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGCTGACGGAGCAACGC
CGCGTAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGGAGGGGAATGCCGGT
ACCTTGACGGTACCTGACGAGAAAGCCACGGCTAAGTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCG
TTGTCCGGAATTTGGGCGTAAAGCGCGCGCAGGCGGTCCTTTAAGTCTGATGTGAAATCTTCGGCTCAACCCG
AAGCGGTCAATGGAAACTGGGGGACTTGTAGTGCAGAAGAGGAAACCGGAATCCACGTGTAGCGGTGAAATGCGT
AGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTTTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTGGG
GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGTAAAGTGTGGAGGGTTCCGCGCT
TCAGTGTGTCAGCTAACGCATTAAGCACTCCGCTGGGGGAGTACGGTTCGCAAGACTGAACTCAAAGGAATTA
CGGGGCCCGCACAAAGCGGTGAGCATGTGTTTTAATTCGAACCAACCGCAAGAACTTACCAGCTTGGAAATG
TCCTGACCGCCCTGGAGACAGGGTCTTCTCTTCGGGGACAGGATGACAGGTGGTGCATGGTTGTCGTAGCTCGT
GTCTGAGATGTTGGGTTAAGTCCCGCAACCACCGCAACCTTGGTTCTAGTTGCCACCATTCAATGGGCACTCT
GAAACAACGTCGGCGGACCATTCGAAGAAAGGTGGGGGAGGAGGTTCAAATAATCATGGCCCTTAGAACCTGG
GCTAACACGTTGTTACAATGGAGGTTCAAAGGGTACAAAGGGGCGAACCCGCCCCACCAAAAGCCGATTCCCGAA
ATTTCTATTTTTGAAGTAAAGCGGCAACTCTCCTCGGTGTAACCACGGAAGCCAGAGAATACCCGATACCCATT
GCCCCGCGGGGAAGCC

>A04_27F

GTGACTGCGCTGCCTATACATGCAAGTCGAGCGGACCGACGGGAGCTTGCTCCCTTAGGTCAGCGGCGGACGGG
TGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATGCTTGAT
TGAACCGCATGTTCCATTCTAAAAAGTTGCTTTTAACTTCAAATTTACGAATGAACCCCCGCCCATTAATTTG
TTAAGTTACCGCCCCACCAAGGCAACAATCCTAACCCAATTTAAAGGTTAACCGCCCCCGGGAAACGAAAAACCC
GCCCCAAAACCCCGGAAGCCACCATTTAGGAATCCTCCCATCGGAACAAAATTTCAACACAAACCCGCCCCG
AATTAATAAAGTTTTTCGAATCTAAAAATTTTTTTTTTAGGAAAAACCAATTCCTTTCAATAAGGCCGTACCCTTACC
GTTCCCTACCCAAAAGCCCCGCCAAATTCCTTCCCACCACCCCGTTATTCTTAGTTGCCAGCCTTTTTCCGAAT
TTATGGGCCTAAAGCGCGCGCAGGCGTTTTCTAATTCGATGTGAAAGCCCGGCTCAACCGGGGAAGGTCATT
GGAAACTGGGGAACCTTGTAGTGCAAAAAAGAAAATGGAATCCACCGTGTAGCGGTGAAATGCGTAAAAATTTGGAGG
AACACCATTGGCAAAGCAACTCTCTGGTCTGTAAGTACCTGAGGCCCGAAAAACCTGGGGAGCGAACAGGATTA
AATACCCTGGTAGTCCACCCCTAAACAATAAATGCTAATTTTTAAAGGTTTCCCCCTTTATTGCTGCACCAAAAC
CATTAAAGCCCTCCCCCGGGGAATACGGTCCCAAGACTGAACTCCAAGGAATTGACGGGGGCCCCCCCAAAACCGT
GGAACCATTTGTTAATTAACCAACCCCAAAAACCTTACCAGGCTTTGAATTCCTTTGAAACCTTAAAAATAGGGCTT
CCCCCTCCGGGCCAAATAACAAGTTTTCCAGGTTTTCCCGCCCCCGGTCCCGAAAATTTGGGTTAATTCCTAAAAA
AGGCAACCTTTTTTTTTTTTTCCCTCTTTTTGGGCCCTTAAAGGAAACCCCTTCAAACCCCAAAAAGGGGG
GGGGAGGGCAAAAAACCAAGCCCCCTTAAACCGGGGCCACACCTTTTTCAATTTGGGAAAAAAAAGGGGGGAA
ACCCCCCCGGGCCCCCCCCCCCATTTTTTTTTTTTTAAAAATTTCCCGCCCCCGGGGAAGGAAAAACCTTTT
TTTACAAAACCCCGGGGAAATTTCCCGGTTTTTAAACCCCGCCCCCAGGGTTAACACACCCCTTAAAAAC
TTCGCCCCCGGGAATTTGGGTCCCGGAA

>B04_27F

GTCATTTACCGCATCCTATCATGCAGTCGAGCGGACAGATGGGAGCTTGCTCCCTGATGTTAGCGGCGGACGGGT
GAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATGCTTGT
GAACCGCATGTTCAAACATAAAAGGTGGCTTCGGCTACCACTTACAGATGACCCCGCGGCGCATTAGCTAGTTG
GTGAGGTAATGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGTAGCGCCACTGGGACTGAGAC

ACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCG
CGTGAGTGATGAAGTTTTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGTTCGAATAGGGCGGTAC
CTTGACGGTACCTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGTGGCAAGCGTT
GTCCGGAATATTGGGCGTAAAGGGCTCGCAGGCGGTTTCTTAAGTCTGATGTGAAAGCCCCCGGCTCAACCGGG
GAGGGTCAATTGGAAACTGGGGAACCTTGAGTGCAGAAGAGGAGAGTGGAAATCCACGTGTAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACCAGTGGCGAAGGCGACTCTCTGGTCTGTAAGTACGCTGAGGAGCGAAAGCGTGGGG
AGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAAAGGGGGTTCCGCCCC
TTAGTGCTGCAGCTAACGCATTAAGCACTCCGCTGGGGGAGTACGGTTCGCAAGACTGAAACTCAAAGGAATTGA
CGGGGGCCCGCACAAAGCGGTGGAACCATGTGGTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTGACAT
CCTCTGACAATCCTAGAAGATAGGACGTCCCTTTCGGGGCAGAGTGACAGGTGGTTGCATGTTTGTGTCAGCT
CGTGTGTCGATGATGTTGGGTTAAGTCCCAGCACGAGCGCAACCCCTTGATCTTAGTTGCCAGCATTGAGTTGGGCA
CTCTAAGGTGACTGCCGGTGACAAACCGGAAGAAAGGTGGGGGATAACGTCAAATCATCAAGGCCCTTATGACC
TGGGGTACAACCTTGCTACAATGGGACAGAACAAGGGGAGCGAAACCCCGGGTTAAGCAATCCCCAAAT
CTTTTTCAATTTGGATCCAAATCTGAAATCAAACGGCTGGAGACGGGAAACTCTTTAAATCCGGAAAAACAAG
GCCGCCGTGAAAAAATTTCCGGCCTGTGTAACCGCCGTTCAACCCAAATTTGTTACACCCAAATCCTGGGGT
ACCTTTAGGACCTCCCATGGACAATTGTGGGGTTTTTTTTCCACAACCTCTTATATAATCCTTAAAAATAATTTTT
GCCCC

>C04_27F

GGCACATGGCGGCAGCCTACACATGCAGTCAACGGCAGCGGGTAGGAAGCTTGCTTCCATGCCCGCGAGTGG
CGGACGGGTGAGGAGCGCATCGGAATCTACTCCGTGCTGGGGGATAACGTAGGGAAACTTACGCTAATACCGCAT
ACGACCTAAGGGTGAAAGTGGGGGACCGCAAGGCCTCACGCGATGGAATGAGCCGATGTCCGATTAGCTAGTTG
GCGGGGTAAGGCCACCAAGGCGACGATCGGTAGCTGGTCTGAGAGGATGATCAGCCACACTGGGACTGAGAC
ACGGCCCAGACTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGCGAAAGCCTGATCCAGCCATGCCG
CGTGGGTGAAGAAGGCCTTCGGGTTGTAAGCCCTTTTTGTTGGGGAAGAAATCCTGTGGTAATACCCGCGGG
GATGACGGTACCCAAAGAATAAGCACCGGCTAAGTCTGTGCGCAGCAGCCGCGGTAATACGAAGGGTCAAGCGT
ACTCGGAATTAAGTGGCGTAAAGCGTAAAGCTGCGTAAAGTGGTGGCTTAAGTCCGTTGTGAAAGCCCTGGGCT
GAATTCGAGTGGATACTGGGTCACTAGAGTGTGGTAGAGGTGGCGGAATTCGGGTGTAGCAGTGAATGCGTA
GAGATCGGGAGGAACACCCGTGGCGAAGGCGGCCACCTGGGCCAACACTGACACTGAGGCACGAAAGCGTGGG
GAGCAAAACAGGATTAGATACCCTGGTAGTCCACGCCCTAACCGATGCGAAGTGGATGTTGGGTTCAATTTGGGACT
CAGTATCGAAGCTAACCGCTTAAAGTTCGCCGCTGGGGGAGTACGGTTCGCAAGACTGAAACTCAAAGGAATTGAC
GGGGGCCCGCACAAAGCGGTGGAATTAGGTGGGTTAATTCGATGCAACGCGAAGAACCCTTACCTGGTCTTGACAT
CCGCGGAATGCCCAGAAAATGGGCGGGTGCCTTCGGGAGCCGCGAGACAGGTGCTGCATGGCTGTGTCAGCT
CGTGTGTCGATGATGTTGGGTTAAGTCCCAGCACGAGCAACCCCTGTCTTAGTTGCCAGCAGTAAATGGTGA
GGAACTAAGGAGAGACCCCGCTGACAAACCGGAAGAAAGGTTGGGGATGACCTCCAGCTCACCGGCCCTTA
CAACCAGGGGTACCACCTTCTACAATGGTGGGGGACAAAAGGCTGCGCACCCCCAGGGGGGAGCCAATCCCA
AAACCCACTCTCTTTCCCGGATTGGGATCTGGCACTCCACCCCTCGAAATCCGAAATCCCTTGTAAATCCAAATAAC
CAATGTTCCGGGGGAAAACATTCCCGGGCCTTGTACACACCGCTTCAACCAGGGAAATGTTGCCAAAAACAAGTT
TAACCCAAAGGGCTGCCGTTGCTACTGGGGTGGTTGTTAGAAGACAGCTCCCTTGGTCCGGGCAAAAAGGGC
AAAAAGCATTAGGTGCAACATAGCGATCATTGCCTTACCAGTACATTAAGGATATAGAAGTGGACGTATAAGGTA
AGCATAATGCAAT

>D04_27F

GTCTCGGTCCGCATCGTTATCATGCAGTGCAGCGGACCGACGGGAGCTTGCTCCCTTAGGTCAGCGGCGGACGG
GTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATGCTTGA
TTGAACCGCATGGTTCAATCATAAAAGTGGCTTTTAGCTACCATTACAGATGGACCCGCGGCGCATTAGCTAGT
TGGTGAAGTACAGGCTCACCAAGGCGACGATGCGTAGCCGACTGAGAGGGTATCGGCCACTGGGACTGAG
ACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCAACGGGCAACGC
CGCGTGAGTGATGAAGTTTTTCGGATCGTAAACTCTGTTGTTAGGGAAGAACAAGTACCGTTCGAATAGGGCGGT
ACCTTGACGGTACCTAACAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGTGGCAAGCG
TTGTCCGGAATATTGGGCGTAAAGCGCGCGCAGGCGGTTTCTTAAGTCTGATGTGAAAGCCCCCGGCTCAACCG
GGGAGGGTCAATTGGAACCTGGGGAACCTTGAGTGCAGAAGAGGAGAGTGGAAATCCACGTGTAGCGGTGAAATGAC
GTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGACTCTCTGGTCTGTAAGTACGCTGAGGCGGAAAGCGTG
GGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTTCCGCC
CTTTAGTGCTGCAGCAAACGCATTAAGCACTCCGCTGGGGGAGTACGGTTCGCAAGACTGAAACTCAAAGGAATT
GACGGGGGCCCGCACAAAGCGGTGGAACCATGTGGTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTGAC
ATCCTCTGACAACCTTAGAGATAGGGCTTCCCTTTCGGGGCAGAGTGACAGGTGGTGCATGTTTGTGTCAGCT
CGTGTGTCGATGATGTTGGGTTAAGTCCCAGCACGAGCGCAACCCCTTGATCTTAGTTGCCAGCATTGAGTTGGGCA
CTCTAAGGTGACTGCCGGTGACAAACCGGAAGAAAGGTGGGGGATGACGTCAAATCATCAGGCCCTTATAACCG
GGGCTACAACCTGCTTCAATGGGCGAAAACAAGGGAGCCAAAACCCCGGGTAAACAATTTCTCAAATATGTTTC
TCAATTTGAAACCAATCTGGCAATCCACCGCCGGAACCTGAAACCTTTAATCCCGATACACAGCCCGGGAAAT
TCCCGGCTTGAACCCCGTACACCCAAGTTGTTATCCAATCCTGAGTACTTGGTACGCCATGGAATTTGGG
GGGTGTGTTCCCTGTTACCTGTAAAAAAAAGGGGGTGTATGGGTGCGATGCATTGATTAACAATTCAAAGAGAAA
GGAAGTTAAATATATAACATTAGGGA

>E04_27F

GCACTGCGCTGCCTATACATGCAGTTCGAGCGGACAGATGGGAGCTTGCTCCCTGATGTTAGCGGCGGACGGGT
GAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATGGTTGTTT

GAACCGCATGTTCAAACATAAAAGGTGGCTTCGGCTACCACTTACAGATGGACCCGCGGCGCATTAGCTAGTTG
GTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCCACTGGGACTGAGAC
ACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCG
CGTGAGTGATGAAGTTTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGTTTCAATAGGGCGGTAC
CTTGACGGTACCTAACCAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGTGGCAAGCGTT
GTCCGGAATTAATGGGCGTAAAGGGCTCGCAGGCGGTTTTCTAAGTCTGATGTGAAAGCCCCCGGCTCAACCGG
GAGGGTCAATGGAACTGGGGAACCTTGAGTGCAGAAGAGGAGAGTGAATTCCACGTGTAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACCAAGTGGCGAAGGCGACTCTCTGGTCTGTAAGTACGCTGAGGAGCGAAAGCGTGGG
AGCGAACAGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAAAGGGGTTTTCCGCC
TTAGTGTGACGCTAACGCATTAAGCACTCCGCTGGGGGAGTACGGTGCAGACTGAAACTCAAAGGAATTGA
CGGGGGCCCCGACAAGCGGTGGAGCATGTGTTTTAATCGAAGCAACGCGAAGAACCTTACCAGTCTTGACATC
CTCTGACAATCCTAGAGATAGGACGTCCCTTTCGGGGGACAGGTGACAGGTGGTGCATGGTTGTCGTAGCTCGT
GTCGTGAGATGTTGGTTAAGTCCCAGCAGCGCAACCTTGATCTTAGTTGCCAGCATTCAAGTTGGGCACTT
AAGGTGACTCCCGGTGACAAACCCGAAAGAAAGTGGGGGATGACGTCCCAATCATCATATAACCGGGG
TACCCCTGCTTCAATGGACCGAACAAGGGAGCGAAACCCCGGTTAAACAATCCCCAATCTTTTTCCCATTT
GGAACCAATTCTGACTCCACGGCCTGAACTGGAACCTTTAATCCCGGAAAAACAAGCCCCGTGAAAAATTTCC
GGCCTTGTAACCCCGTCAACCCCGAAGTTTTTACCCCAACTCGTAAGTACTTTAAGACCCCATGGGAAAG
GGGGTGGTTAAGATACTTGAAGGAATCCCGGCGGACGTAAGGCGAACATGAGTAAGTTAAGTGAA
GTAATTATACAATTAGATCATAATCGAAGTACATGAA

>F04_27F

GTGGAATGCGGCATGCTATACATGCAAGTCGAGCGGACCGACGGGAGCTTGCTCCCTTAGGTGACGGCGGACG
GGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATGCTTG
ATTGAACCGCATGGTTCAATCATAAAAGGTGGCTTTTAGCTACCACTTACAGATGGACCCGCGGCGCATTAGCTAG
TTGGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCCACTGGGACTGA
GACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACG
CCGCGTGAGTGATGAAGTTTTCCGGATCGTAAACTCTGTTGTTAGGGAAGAACAAGTACCGTTTCAATAGGGCGG
TACCTTGACGGTACCTAACCAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGC
GTTGTCCGGAATTAATGGGCGTAAAGCGCGCGCAGGCGGTTTTCTAAGTCTGATGTGAAAGCCCCCGGCTCAAC
GGGGAGGGTCAATGGAACTGGGGAACCTGAGTGCAGAAAAGGAGAGTGAATTCCACGTGTAGCGGTGAAATGC
GTAGAGATGTGGAGAACACCATTTGGCGAAGGCCACACTCTGGTCTGTAAGTACACTGACACTCAAGCGCGTAAAC
GAGAGCGAACACGATTACATATGCTGTTAGTCCACGCGTTAAACGATTAGAGCTGAGTGGTAAAAAGTCTCCTC
CCCCTTTTTGGCTGCAATCATACCCCTTTTTACCCACGCCCCCGGGGAGAGGACGGCGGGGGGTGATTTAA
AATAAAAAATAAAAAAGACCGGCCGGCCACGATTATTCAGCCAAAATGATGTTGATTAGAATCCGACAGCCCC
CCGAAATATAAACAGGCGGGCGGGGATGGGAACCTTAATTTAGGAAAGGAAACAAGTACCGTTGGGTTGGGGG
GGTGGGGGGGCCCAATTGAAAGGGAAGGATGGATTGGTTTTCTCCCTCCCCCAAGGAGGCTGGTTTTTAAGG
AAAGCTTAAATTTCCCCCAAAACCGGACGGGGGCTTCTTTTGAATTAATAACGTAACCTTTTCCGTTTCGGATT
GAAGGAAAAACCATTTCCCGGGGACCCCGCCCCAAAACCAAAAAAGAAAGGGGGGGGGGGGGGGTGTGGAAT
TCAAATGCTCCCGGATTAAGGCACCCCAATCCCTAATTTTGGAAAGGGGGGGGGGGGGGGGGGGTGTGGG
TTGTGGGGCATGGGTTTTGGAGAACAGAAATTTCCGACACCGGAAACACGCGGAACCCCGGGCGGGGGGGCG
CCTGCAATCCCCTGTAACCCCTTATAAATTTGGTTAAAAATATGTTAAAAAGTTATAAGAAGGAACACTTTTTTAA
TATGAGAAACAAAAGGGGAAAACCGCAGGGGGTTTTGAGAAGTTATACAAAGGAACCCCCCTGGCTGTGGGGCGG
GGCCCCCGGAACCGTTAACTGGCGGAGAAATTTTAAAAAACCCGTTTGGGCCCCCGGGGGGGGGGGGGGG
GCACAAAACCCCGGCAACCCCTTTTTGGAGTTCCGTTTTGCTGAAACCAACAAAGACTCCCTTCGGTG
TCCCTCCCCCTGGCCCCACAAGGAAGCGCGGGGCCCGCAACAGGGGGGGGGCTTTGTGGTGCAAAATCG
GTGACAAAACACTAGAAGAGGGGGATAACCTAACACGAAAATTGGAAAATTT

>G04_27F

GTCAAGGCGCTGCCTATACATGCAGTCGAGCGGACCGACGGGAGCTTGCTCCCTTAGGTGACGGCGGACGGGT
GAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATGCTTGATT
GAACCGCATGGTTCCATTCTAAAAAGTGGCTTTTAACTTCCACTTTCCGATTGACCCCCCGCCAATAACCAAGTTG
TGAAGTTACCGCCCAACCAAGGCGACCAATCCTAACCAACTGAAAAGGTGAACCGCCACCCCGGACCGAAACCC
CGCCCAAAACCCACCGGAAGCCACAATTAGGAATCCTCCCAATGGAACAAAGTCTGAACGAACACCCCCCG
GAATGAAGAAAGTTTTCCGAACCTAAAACCTTTTTGTTAAGGAAAAACCAATTCCGTTTCAATAAGGGCGTACCCT
ACCGTTCCCAACCAAAAGCCCCCGCAAAATACCTTCCCACCACCCCGTTATACCTTAGTTGCCAGCCTTTTTCCC
GAATTAATGGGCCTTAAACCCCGGAGCGTTTTCTTAATTTCTAAGTTAAAAACCCCGCCACCCCGGGAAGGT
CATTTGAAACTGGGGAACCTTAAATCCAAAAAAGAAAATTTGAATTTACCTTTAACCGTTAAATGCCTAAAAATTTGA
AGAACCCCAATTGCCAAAGCCACTTTTTGTCTGTAAGTACCCTTAAGCCCAAGCCTGGGGAACCAACAAGAA
TAAATACCCTGGTAGTCCACCCCTAAACAAGAATGCTAATTTGTTAAAGGTTTTCCCCCTTAAATGCTGCAGCAAA
CCCATTAAAGCATGCTCCCTGGGAGTACGGTGCAGGACTGAAACTCAAAGGAATTGACGGGGCCCGCACAAG
CGGTGGAGCATGTGGTTAATTCAAAGCAACGCAAAAACCTTACCAGTCTTGACATCCCTGACAAACCCCTACAGATA
GGGCTTCCCTTCGGGGCAAAATGACAAGTGGTGCATGGTTGTCGTACCTCCTGTCTAAAATTTTGGGTTAAT
CCCCACCCACCCCAACCTGGATCTTATTTGCCACCAATCAATTTGGGCCCTTAAGGTAAGTACCCTGCGGTGACAAC
CAAAAAAGGTGGGGTACGTTTCTTGGCCCTTTTAACTGGGTTAAACCCGGTCAATTTGGCCGAACAA
AGGGAACCCCGGGGTACCCATTTCCCAACTTTTTTGGAGTTTTTTTTAAAGAAATTTCCAAACCCCGGAAAGT
GAAGCTTGAATTCGGAAACCATGCCGTTTTAATTAACCTGGCTTTTAAACCCGCTGCCACCCAAATTTTAAAC
CAACCGGAGAAACTAGAGGGG

>H04_27F

GTGCAGTGCCTGCTATACATGCAAGTCGAGCGGACCGAAGGGAGCTTGCTCCTTTAGGTTAGCGGGCGGACGGG
TGAGTAACACGTGGGCAACCTGCCCTGCAGACGGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACACCG
AAAACCGCAGGGTTTTTCGTTGAAAGGCGGCTTTTAGCTGTCACTGCAGGATGGGCCCGCGGCATTAGCTAGT
TGGTGAGGTAACGGCTACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGAG
ACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGC
CCCGTAGCGAAGAAGGCTTTCGGATCGTAAAGCTCTGTTGTCAGGGAAGAACAAGTGCCGTTCAACAGGGCGG
TACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGC
GTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAATCTCGCGGCTCAAC
GCGAGCGGCCATTGAAACTGGGAACTTGAAGTGCAGGAAAGGGGAGCGGAATCCACGTGTAGCGGTGAAATG
CGTAAAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTACGCTAAGGCGCGAAAGCGT
GGGGAGCGAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTATCCAC
CCTTTAGTGTGCAGCAAACGCATTAAGCACTCCGCTGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATT
GACGGGGACCGCACAAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACCGGAACCTTACAGGCTTTGACA
TCCCCTGACAACCCTAGAGATAGGGCGTCCCCTTTCCGGGACAGGGTGACAGGTGGTGCATGGTTGTCGTACC
TCGTGTCGTGAAATGTTGGGTTAAGTCCCAGCAACGAGCGCAACCCCTGACCTTAGTTGCCAGCATTCAATTTGGGCA
CTCTAAGGTGACTGCCGGCTAAAATTCGAAGAAAGGGGGGATGACGTCAAATCATCAGGCCCTTAAAACCCGGC
GCTACCCCGTCTTAAAAGGGGTGGTACAAAAGGCGACCAAAACCCCGGTTGACGAAATCCAAAAAACCCCTC
TATTTCAATTTAGGGGGGCAACCCCGGCTAGAAACCCGAATCCCTTTAATTCGGAAAAACAACCCCGG
AAAAAATTTCCCGGTTTTTAAACC

>A05_27F

CTCGACTGCGCTGCCTATACATGCAGTTTCGAGCGGACGATTCAAGCTTGCTTTTGGATCGTTAGCGGGCGGACGGG
TGAGTAACACGTGGGCAACCTGCCCTGTAGACGGGGATAACACCGAAGAAATCGGTGCTAATACCGGATAATACGA
AAGGCCGATGGTCTTTTCGTTGAAAGGCGGCGCAGGCTGTGCTACAGGAGGGGCCCGCGCATTAGCTAGT
TGGTGAGGTAACGGCTACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGATCGGCCACACTGGGACTGAG
ACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGC
CGCGTAGCGAAGAAGGCCTTCGGGTCGTAAGCTCTGTTGTTAGGGAAGAACAAGTACCGCAGTCACTGGCGGT
ACCTTGACGGTACCTAACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCG
TTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTTAAAGTCTGATGTGAAAGGCCACGGCTCAACCG
TGGAGGGTCATTGAAACTGGGGGACTTGAAGTGCAGAAGAGGAGAGCGGAATTCACAGTGTAGCGGTGAAATGC
GTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTG
GGGAGCAAACAGGATTAGATACCCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTATCCACC
CTTTAGTGTGCTAGCTAACGCATTAAGCACTCCGCGGGGAGTACGCTCGCAAGAGTGAAGTCAAAGGAATTG
ACGGGGCCCGCACAAAGCGGTGGAACAGGTGGTTAATTCGAAGCAACCGCAAGAAAGTCTGAGGCTTTGACA
TCCCCTGACAACCCGAGAAGATCGGGCGTCCCCTTTCCGGGGGACAGGGTGACAGGTGGTTGCATGGTTGTCG
TCAGCTCGTGTGAGATGTTTGGGTTAAGTCCCAGCAACGAGCGCAACCCCTCGACCTTATTTGCCACCATCAAT
TGGGCACTCTAAGGTGACTGCCGGCTAAAATTCGAAGAAAGGTTGGGGGATAACCTTCAAATCATGCCCCCT
TATAAACCCTGGGCTACACCCCTTGTCCAAATGGGCGGTACAAAGGGTCCGCAACCCCGGGGAGCAATCC
CCAAAAACCGCTCTCTATTTTTGATTGTAAGGCGGCAATTCCTCTGCTGAAAACCGAAATCCCTTTAATTCGGATA
AACATGGCCGCGGTTAAAAAATTTCCCGGCCCTGTTGTAACCCGCGGGTCCACCCAAAAAGTTTTTGCACACCC
ACACCCGGGAGAACCCTTCCCGG

>B05_27F

CTCAAAATGCGCTGCCTATAATGCAAGTTTCGAGCGGACGATTCAAAAGCTTGCTTTTGGATCGTTAGCGGGCGGACG
GGTGAGTAACACGTGGGCAACCTGCCCTGTAGACGGGGATAACACCGAAGAAATCGGTGCTAATACCGGATAATAC
GAAAGGCCGATGGTCTTTTCGTTGAAAGGCGGCGCAAGCTGTGCTACAGGATGGGCCCGCGGCATTAGCTAGT
GTTGGTGAGGTAACGGCTACCAAGGCGACGATGCTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTG
AGACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAAC
GCCGCGTAGCGAAGAAGGCCTTCGGGTCGTAAGCTCTGTTGTTAGGGAAGAACAAGTACCGCAGTCACTGGCG
GTACCTTGACGGTACCTAACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAG
CGTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTTAAAGTCTGATGTGAAAGGCCACGGCTCAAC
CGTGGAGGGTCATTGAAACTGGGGGACTTGAAGTGCAGAAGAGGGAGAGCGGGAATTCACAGTGTACGGTGAA
ATGCGTAGAGATTGTGGAGGAACACCAGTGGGCGAATGCGGCTCTCTGGTCTGTCACTGACTCTGAGGCGCGGAA
AAGCGTGGGGGAACAAAAAGGGATTTACATTTCCCGGTTGTTCTCCCCCCTAATCTAATGAGTGGTTATA
TGGTTATCAAGGGTGAGAAACTCCCCTTTATTCGTTTAAAGGCCAACTCGAGTTTCCCCCCCCATTTTCCCCCG
GGGGGACTTACTTACCTCGTTAAGAAAGCCGGTTGGGGATCCTTTTCAAGAAAATTAATAACTTGGGGGAAG
GCACCCCGCCAAAGGGGGCTTCGGGGTGGTTAATACCACCCGGCCTTTGTAATTTTTAATTTCCCGAGCGAA
AAACACAACACCGCTCCCTTTGGGAGGCACCCCAATACCCCTCTCAAAGGGGAATTAATAATGTAAGCCCCCGG
CGTTGGGGGACCCCTTTTGGGCGCCCGGCGGAATTAAGAACCCCTAAAATTTTTCTGGGAGGGCCAAAGG
AATAGTGGGGCTTTTTCCCCCCCCCTTTTTTTTTTTTTATGGGTTTTACACCCGGGGGAAAAAAAAGGG
GGTTAACACCCCGGGTACCAAAATTTGGTGAAGAGGGATGGGAAAAATTTGGGATTTTCTTGGTCCCTT
CTTCTCTCCCCCTATAAAAAAGAAAAGGCAACCGAGGTGTGATTTTTTTTTTAAGTGTAAAGGGAAAGATTTGAA
AAAAGGTTCCGGGTCTTCTTTTTAACCCCCACGGCGGGGACAAAAGGGGCCGCGGTTTTTGGATTATTC
TTTTAACCAAGTTAAGACCAAAAATTTTTTTTTCCCCAATTTGTAAGGGGGAAGGGGAAATCTAAAAACCG
GGCAAGAACACCGGGGGCCCCGGGCCCAAAAAAAGAGATAAGAAGGGAAAAAGGTGGGGGGGGGGGGGG
GTTGGGTATTATACAAAATTTCTTAGGGTTTGGGGGGGGGGGCCACCAACCTTTTTTCCCGCGGGGGGG
GGGGGGGGCGGCGCACCAAAATTTCTTCGGCCCCCTGGGGGGGTGGTGGAAATGGAGAAACCCCGTTTT

TGCTCCCCAACACAAAAAATCGGGGGGGGGGGGGCCGCCACGGGCCAAACGGCTCCCAAATTTATTTAGG
TGTTCTGGTGATTTCCGCGTGATGCGGGTTATGGGGATTTAAACAAAAGGAAAACCAGGGGCTGGGAGAATTTGGTA
TTTTTAAGATTTTAAATGGA AAAAGGTAAAACAAAATGGGGGCCTGGGTAGAGTTTAAACGCCCCGCGGGGGG
GGAAGTGGGAGCCCCACAAGCTCTTCAACACCCCTTCCGGACAACAAAACAAAGATGTAGTGGGAATTTGCTC
CGCCCCCCTTTGAGGG

>C05_27F

GTCAAAGGCGCATGCCTATACATGCAGTTCGAGCGGACCGAAGGGAGCTTGCTCCTTTAGGTTAGCGGCGGACGG
GTGAGTAACACGTGGGCAACCTGCCCTGCAGACTGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACACC
GAAAACCGCAGGGTTTTCGGTTGAAAAGGCGGCTTTTAGCTGTCACTGCAGGATGGGCCCGCGGCGCATTAGCTAG
TTGGTGAGGTAACGGCTCACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGA
GACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAAGTCTGACGGAGCAACG
CCCCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTCAGGGAAGAACAAGTCCGCTTGAACAGGGCG
GTACCTTGACGGTACCTGACGAGGAAGCCACGGTACGTCGACGAGCAGCCGCGTAATACGTAGGTGGCAAG
CGTTGTCCGGAATTTGGGCGTAAAGCGCGCGCAGGGCGTTCCCTTAAGTCTGATGTGAAATCTCGCGGCTCAAC
CGCGAGCGGCCATTGAAAACCTGGGGAACCTTGAGTGCAGGAAAAGGGGAGCGGAATTCACGTGTAGCGGTGAAAT
GCGTAAAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTGCAGTAAAGCGCGAAAAGCG
TGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTATCCA
CCCTTTAGTGCTGCAGCAAACGCATTAAGCACTCCGCTGGGGAGTACGGCCGAAGGCTGAAAACCTCAAAGGAAT
TGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTGAC
ATCCCTGACAACCCTAGAGATAGGGCGTCCCTTTTCGGGGACAGGGTGACAGGTGGTGCATGGTTGTGCTGACG
CTCGTGTGTAAGTGTGGTTAAGTCCCGCAACGAGCGCAACCCTTGACCTTAGTGGCAGCATTCAATTTGGCG
ACTCTAAGGTGACTGCCGGCTAAAAATCCGAAGAAAGGGTGGGGGATGACCTTCAAATCATCAAGCCCCCTTAAAA
CCGGGGCTACCCCTGCTTCCAATGGGGGTTACAAAGGGAACCAAAAACCCCGGGGGGAGCGAAACCCAA
AAAACACCTCCTATTTTGAATGAAGGCGGGCACTCCCCCGCTGAAAACCGAAAACCTTTTAATTCGGAAATAAC
ATGCCCGGTGAAAAAATTTCCCGGTTTGTAAACCCCGCTCCACCACAAAATTTGTTAACCCTAAATTCGTGG
GTGACCTTCCGGGACCGCCATGGCAATGGGGTTGGTTGGACAATGCCTGGGCTTTGGTGGGGTGGGTTGGG
TCGGGACCAAGAACAAGAAATAATACCAACCACAACCAAACTTAAACCTAATTA

>D05_27F

GTGACTGCGCTGCCTATACATGCAGTTCGAGCGGACAGATGGGAGCTTGCTCCTGATGTTAGCGGCGGACGGGT
GAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACCGGATGGTTGTTT
GAACCGCATGTTCAAACATAAAAGGTGGCTTCGGCTACCACCTACAGATGGACCCGCGGCGCATTAGCTAGTTG
GTGAGGTAACGGCTCACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGAC
ACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAAGTCTGACGGAGCAACCGCG
CGTGAGTGATGAAGTTTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGTTCGAATAGGGCGGTAC
CTTGACGGTACCTAACCAAGGACGACGGCTAAGTACGTCGACGAGCCGCGTAATACGTAGGTGGCAAGCGTT
GTCCGGAATTTGGGCGTAAAGGGCTCGCAGGCGGTTTTCTTAAGTCTGATGTGAAAGCCCCCGGCTCAACCGGG
GAGGTCAATTGAAAACCTGGGAACTTGAGTGCAGAAGAGGAGAGTGGAAATCCACGTGTAGCGGTGAAATCGTA
GAGATGTGGAGGAACACCAGTGGCGAAGGCGACTCTCTGGTCTGTAAGTGCAGCTGAGGAGCGAAAGCGTGGGG
AGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAAAGGGGTTTTCCGCCCC
TTAGTGCTGCAGCTAACGCATTAAGCACTCCGCTGGGGGATACGGTTCGCAAGACTGAAAACCTCAAAGGAATTTGA
CGGGGCCCCGACAAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCCTTACCAGTCTTGACATC
CTCTGACAATCTTAGAGATAGGACGTCCTTTCGGGGCAGAGTGACAGGTGGTGCATGGTTGTGCTGACTCGT
GTCGTGAGATGTTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGATCTTAGTTGCCAGCATTACAGTTGGGCACTC
TAAGGTGACTGCCGGTGCACAAACCGAAGAAAGGTGGGGGATGACGTTCAATCATCATGCCCTTATGACTGGGGC
TACACACTTGCAACAATGGACAAAACAAAGGGAACCAAAAACAGCCCGTTTTAACCCATTCCCAAATCTTTCTCT
ATTTGGATCCAAATCTGCAACTCCAACGCTGGAATTTGATCCCTTTAATTCGGGTAACCAAGCCCGGGTGA
AAAAATTTCCGGCTGTTAACTCCCCCTCCACCACACAGTTTTTAATCAATTCCCGGAGGTACCTTCTGAAC
GCCCTACTGGCATATTTGGTTTGGGCTCACGCTCAAAGGAAGGGCCAGCGAAGGGGCCCGGGGCATGCGC
TTAAGTTACAGAAGCGAGATACGGATACATAGGCTATGGAATCGAGTAGAGGGCGGGGG

>E05_27F

GTCACTGCGCGTGCCTATACATGCAGTTCGAGCGAACCAATAAGAAGCTTGCTTTTTGTTGGTTAGCGGCGGACGG
GTGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGATTAT
CTTTCGCTGGAGAGATAAGGAAAGATGGCTTTTGTCTCACTTACAGATGGGCCCGCGGCGCATTAGCTAGTTG
GTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCAACCTGAGAGGGTGATCGGCCACACTGGGACTGAGAC
ACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAAGTCTGACGGAGCAACGCCG
CGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGGAGGGAATGCCGGTAC
CTTGACGGTACCTGACGAGAAAAGCCACGGCTAAGTCCAGCAGCCGCGTAATACGTAGGTGGCAAGCGTT
GTCCGGATTTATTGGGCGTAAAGCGCGCAGCGGCTCTTTAAGTCTGATGTGAAATCTTGGCGCTCAACCGCA
AGCGGTCAATTGAAAACCTGGGGACTTGAGTGCAGAAGAGGAAAGCGGAATTCACGTGTAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTTTCTGGTCTGTAAGTGCAGCTGAGGCGCGAAGCGTGGGG
AGCAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTGGGAGGGTTTTCCGCCCTT
CAGTCTGCAGCTAACGCATTAAGCACTCCGCCCGGGGAGTACGGTTCGCAAGACTGAAAACCTCAAAGGAATGAC
GGGGGCCCGCACAAAGCGGTGGAACAGGTGGTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTTGACATC
TCCTGACCGCCCTGGAGACAGGGTCTTCCCTTCCGGGACAGGATGACAGGTGGTTGCATGGTTGTGCTGAGCTCG
TGCTGTGAGATGTTTGGTTAAGTCCCGCAACGAGCGCAACCCTTGGTTCTAGTTGCCAGCATTACAGTTGGGCACT

CTAAACGACTGCCGGCCACAATCCGAAGAAAGGTGGGGGATGAAGTCCAATCATCATGCCCCCTTATAACCGGG
GGTACACCCTGGTTCCAATGGAAGGTTCAAAGGGGCCAAAACGGCGCACCCCAAAGGGAATCCCCAAAACATT
CTCTATTTTGGATGGAAGGGGGCGAATCCCCCGCTGAAACCCCGGAAAACCTTTTAATCCGGAAAAACAGGCCCG
GTGAAAAATATCCCGGGCTTGTAAACCCCGCTTTCCCGCCAGATTTTGTAAACCCCAATTCGGTGAACCTAAAGCC
GCCCATGGCAATGATGGGTTTGGGATTGGAAAGCGGTTATAATTGGCAAATTTTTTAAATAATTTAAGGAATGGTA
AAGACAATGTAATAGAAAGGAAATGAAGAAGAATTAAGGGAAATTAGAAACAAGTAAAAGTAC

>F05_27F

GGGCTGGCGCTGCCTAACTGCAGTCGTGCGGACCTTTTAAAGCTTGCTTTTAAAGGTTAACGGCGGAAGGGTGAG
TAACACGTGGGCCACCTGCCTGTAAGACTGGGATAACACCCGGGAACCGGGGCTAATACCGGAAAGTTTTTCT
CCCCATGGAGGAAAAAGGAAAGGCGGCTTCTGCTGCCACTTACAGATGGGCCCGCGGCGCAATAGCTAGTTGGC
GGGGTAACCGGCCACCAAGGCAACGATGCGTAGCCACCTGAGAGGGTGATCGGCCACATTGGGACTGAGACAC
GGCCAAAACCTCTACCGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCCG
TGAGTGAAAGAAAGCTTCGGGTCGTAACACTCTTTGGCCGGGGAAGAACAAGTCCGTTGCAACAGGGGGGGCGC
CTGACGGTACCCCGCCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAAGTGCAAGCGTTG
TCCCGAATTATGGGGGTAAAGCGCGCGCAGGCGGCTTCTTAAATCTGATGTGAAATCTTGC GGCTCAACCGCAAG
CGGTCAATGGAACTGGGAGGCTTGTAGTGCAAAAAGAGAATGGAATTCACGTGTAGCGGTGAAATGCGTAAA
AATGTGGGAACACCAGTGGCGAAAGCGGCTCTTTGGTCTGTAACCTGACGCTTAAGCGCGAAAGCGTGGGAGC
AAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAATGTTATAAGGTTTCCGCCCTTATGCT
GCAGCTAACGCATAAGCACTCCGCCCGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGGCCG
CACAAACGGTGGAGCAGGTGGTTAATTCTAACAAACGGGAGAACCTTACCAGGTCTTGACATCTCTGACTTCCGT
GAGACAGGGCTTCCCTTCTGGGGACAATGACAAGTGGTGACGTTGTCTCACCTTGTGCGTGAATAATTGGGT
TAATCCCCACCACCCACCTTGTACTTAATTGCACCATTAATTGGGCCCTAAGTGACGGCGGTGACAACCCAAG
AAGGGGGGGGAGAACTCCAAACACCAGCCCTTAAACCGGGGTACCCTGTTCAATGGAGGTACAAGGGTCCGA
AACCCGGGTAGCCATCCCCAAAACATTCCCTTTTTATTAGGGGGCAACCCCCCTAAACCGAAACCTTTATTCCGA
ACAAGCCCGGGGAAAATTTCCGCCTTTAAACCCCCCCCAAAATTTAATCCAAATCTGGGAACCTAAGTCCCAA
TGGGAATGGGATGGGTTTAACTGCCTTGGGTGGGGGGTTTCCGCCAAAGTCAAAA

>G05_27F

CGCAGGTGGCGCTGCCTATACATGCAAGTCGAGCGAACCAATAAGAAGCTTGCTTTTTGTTGGTTAGCGGCGAA
CGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGA
TTATCTTCCGCCTGGAGAGATAAGGAAAGATGGCTTTTGTATCACTTACAGATGGGCCCGCGGCGCATTAGCTA
GTTGGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTG
AGACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAAC
CGCCGCTGAGCGAAGAGGTCTTCGGATCGTAAAGCTTGTGTTAGGGAAGAACAAGTCCGCGCAAGGCGG
GTACCTTGACGGTACCTGACGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGTGGCAAG
CGTTGTCCGGATTTATTGGGCGTAAAGCGCGCGCAGGCGGCTCTTAAAGTCTGATGTGAAATCTTGC GGCTCAACC
GCAAGCGGTCAATGGAACTGGGGGACTTGTAGTGCAAGAGGAAAGCGGAATTCACGTGTAGCGGTGAAATGC
GTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTTTCTGGTCTGTAACCTGACGCTGAGGCGCAAGGCGTG
GGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTGGGAGGGTTTCCGC
CCTTCAGTGTGAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGTGCAGACTGAAACTCAAAGGAAT
TGACGGGGGCCGCAAGCGGTGGAACAGGTGGTTAATTGCAAGCAACGCGAAGAACCTTACCAGGTCTTGA
CATCTCTGACCGCCCTGGAAGACAGGGTCTTCCCTTCGGGGACAGGATGACAGGTGGTTGCTGATGTTGTCGCA
GCTCGTGTGAGAAAGTTGGGTTAAGTCCCGCAACGCAACCCCTTGGGTTCTAGTTGCCAGCCATTTCAGTT
GGGCACTTAGAGCGACTGCCGGCCACAAGTCCGAAGAAAGGGGGGGGAATAACGTTCAAATCATTACAGCCCT
TATGAACCGGGGTACCACCTGTTTCAATGGGAGGGTTCAAAGGGCGGCCAAACGGGCCACCCAAAACCGGAAT
CCCCAAAACATTTCTTATTTTTGGAATGAAGGGCGGAATCCCCCGCCGAAACCGGAAACCCCTTTAATTCCG
GATAACAAAGCCCGGTGAAAAATTTCCGGGCTTTGTAACACCCGCGGCTTCCACCAAGGTTTAAATCCCC
AATCCGTTGGGTAACCTAAGAGCCCCCATGGCAAGTTGTGGGGGGGGTTTTAAATTGGATAGCTGTTTTTCTG
GGGGATTTTTGGGGGGGGTTAAAAAT

>H05_27F

AAGGAGTGCGGCATGCTATACATGCAGTCGAGCGAACCAATAAGAAGCTTGCTTTTTGTTGGTTAGCGGCGGACGT
GGTGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGATT
ATCTTCCGCCTGGAGAGATAAGGAAAGATGGCTTTTGTATCACTTACAGATGGGCCCGCGGCGCATTAGCTAGT
TGGTAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAG
ACACGGCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGC
CGCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGGAGGGAATGCCGGT
ACCTTGACGGTACCTGACGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGTGGCAAGCG
TTGTCCGGATTTATTGGGCGTAAAGCGCGCAGGCGTCTTTAAGTCTGATGTGAAATCTTGC GGCTCAACCGC
AAGCGGTCATTGAAACTGGGGACTTGTAGTGCAAGAGGAAAGCGGAATTCACGTGTAGCGGTGAAATGCGT
AGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTTTCTGGTCTGTAACCTGACGCTGAGGCGCAAGCGTGGG
GAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTGGAGGGTTTCCGCCCTT
CAGTGTGACGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGGTGCAGACTGAACTCAAAGGAATTGACG
GGGGCCGCAACCGGTGGGAGCATGGTTAATTGCAAGCAACCGGAATTCACGTGTAGCGGTGAAATGCGT
CTGACCGCCCTGGAGACAGGGTCTTCCCTTCGGGGACAGGATGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGT
CGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTGGTTCTAGTTGCCAGCATTTCAGTTGGGCACTCTAG
AGGACTGCCGGCGACAAGTCCGAGGAAGGTGGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTAC

ACACTTGCTACAATGGAAGGTACAAAGGGCAACGAAACGGCGAACCCAAAAGCGAATCCCAGAAAACCTTTCTCAG
TTTTCGGATTGGAGGGTGGAAATCTCCCTGCAGGAGAACCAGAAATCGCTAGTATTCCCAGGATACACAGTGCCCCGG
GAAAAAATTTCCCGGCCCTTTGAACC

>A06_27F

GTCGAAGGTGCGGCATGACTTATCATGCAGTCGAGCGGACCGAAGGGAGCTTGCTCCTTTAGGTTAGCGGGCGGACG
GGTGAGTAACACGTGGGCAACCTGCCCTGCAGACGGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACAC
CGAAAACCGCAGGGTTTTCCGGTTGAAAGGCGGCTTTTACTGTCAGGATGGGCCCGCGCCGATTAGCTA
GTTGGTGAAGTAACGCTCACCAAGGCGACAATGCGTAGCCAACTGAGAGGGTGACCGCCACACTGGGACTG
AGACACGGCCAAACTCCTACGGGAGGCAAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAAC
GCCCCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTGACGGGAAGAACAAGTCCGTTTGAACAGGGC
GGTACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAA
CGCTTGTCCGGAATTTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAAGTCTGATGTGAAATCTCGCGGCTCAA
CCGCGAGCGGCCATTGGAAACTGGGGAACCTGAGTGCAGGAAAGGGGAGCGGAATTCACGTGTAGCGGTGAAA
TGCGTAAAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTACGCTGAGGCGCGAAAGC
GTGGGGAGCGAACAGGATTAGATACCTGGTAGTCCACGCCGTAACGATGAGTGTAAAGTGTAGAGGGTATCC
ACCTTTAGTGTGCAGCAACGCATTAAGCACTCCGCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGA
ATTGACGGGGACCGGCACAAGCGGTGGAACCATGTGGTTTAAATTCGAAGCAACCGCAAGACCTTACCAGGCTT
GACATCCCCTGACAACCCTAGAAGATAGGGCGTTCCCTTTCCGGGACAGGGTGACAGGTGGGTGCATGGTTGTC
GTACCTCGTGTCTGAAAAGTTTGGGTTAAGTCCCAGCAACGAGCGCAACCTTGACCTTAAATGCCCACCATTCA
TTTGGGCACTAAGGTGACTGCCCGCTAAAAATCGGAAGAAAGGGGGGGGAAGAAGCACTTCAAATCATCCAGGCC
CCTTACTGACTGGGGTACCCCTTGTCTTACAAGGGGGGGGTACAAGGGCAGCCAAACCCCACTGGAGG
GAAACCCCAACAAACACCTTCAATTTTTGATTTAAGGCTGGCAACTCCCCGGCTGAAAAGACGGAAACCTTTTTTA
TTTTCCGGAATAACATGGCCCCCGGGGATAAAATTTCCCGGGTCTTTTAAACCCCCCCCCCTCACCCCCCA
GAATTTGTGTTAACCCCAAAATCCGTGGGGTGACTTTCCCGAAACCCCCCAATGGCAAATTAGGGAAGGGG

>B06_27F

GGGACTGCGCTTGGCCTAACTGCAGTCGTGCGGACCTTTTAAAGCTTGCTTTTAAAGTTAACGGCGGAAGGGTGA
GTAACACGTGGGCCACCTGCCCGTAAGAATGGGATAACACCCGGGAACCGGGGCTAATACCGGAAAGTTTTTTCC
TCCCACGGAGGAAAAAGGAAAGGCGGCTTCCGCTGCCACTTACAGATGGGCCCGCGGCGCAATAGCTTGTGG
CGGGGTAACCGCCACCAAGGGAACGATGCGTAGCCACCTGAGAAGGTGATCGGCCACATTGGGACTGAAACA
CCGCCAAACTCCTACCGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAATCTGACGGAAACAACGCCCG
TGAGTGAAGAAAGGCTTCTGGTGTGTAAGTCTTTTCCCGGGGAAAAACAAGTCCGTTTGAACAAAGGGGGCGCC
CTGACGGTACCCCGCCACAAAACCCGCTAAATACCTGCCACCAGCCCCGTAATACCTAAGTGGCAAGCGTTG
TCCCGAATTTATGGGGTAAAACGCGCGCAAGCCGCTTCTTAAATCTGATGTGAAATCTTGCGGCTCAACCGCCAG
CGGTCAATTGGAATGGGAGGCTTGAATGCAGAAAAAGAGAATGGAATTTACGTGTAGCGGTGAAATGCGTAAAA
ATGTTGAGGAACACCCTGGCGAAAGCGGCTCTTTTGTCTGTAAGTACGCTTAAAGCGCCAAAACGTGGGGAGCA
AACAGGATTATATACCCTGGTAGTCCACGCCGTAAACGATGAGTGTAAATGTTATAAGGTTTCCGCCCTTAATGCT
GCAGTTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTATCGGCCACACTGGGACTGAGAC
GCCAAAGCGGTGGAGCATGTGGTTAATTCTTAGCAACGCGGAGACCTTACCAGGTCTTGACATCTCTGACCTCCC
TGGAGACAGGGCTTCCCTTCTGGGACAAAATGACAAGTGGTGCAGGGTGTGCTCACCTCTGCTGAAAATTTG
GGTTAATTTCCCCACAAGGGCAACCTTGACCTATTTGCCCCATTAATTTGGGCATCTAAGTGACGGCGGTGACAAC
CCAAGAAAGGTGGGATAAATTCAAAATCAAGCCCTTTAAACTGGGGCACCCCTGTTCAATGGAGGGTAAAAGG
GTCGGAACCCCGGTTAACCTATCCCAAAAATTTCCCTTTGAAATAGGGGGGACCCCCCTGAAAACCGAATCCCT
TTTTTTGAAAAAAGCCCGGGGAAAATTTCCCGCTTTTAAACCCCCCCCAAAATTTAAACCAACCTTGAA
CAAAGCCCCCAGGGAAATGGGGGTGGGGTTGAATCTGATAAGGTGCCCCC

>C06_27F

GTGAGTGCAGCTGCCTATACATGCAGTCGAGCGAACCAATAAGAAGCTTGCTTTTTGTTGGTTAGCGGCGAACGG
GTGAGTAACACGTGGGTAACCTGCCCTGTAAGATCGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGATTAT
CTTTCCGCCTGGAGAGATAAGGAAAGATGGCTTTTGTATCACTTACAGATGGGCCCGCGGCGCATTAGCTAGTTG
GTGAGTTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTATCGGCCACACTGGGACTGAGAC
ACGGCCACAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCG
CGTAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGGAGGGAATGCCGGTAC
CTTGACGGTACCTGACGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCGTT
GTCCGGATTTATGGGCGTAAAGCGCGCGCAGCGGCTCTTTAAGTCTGATGTGAAATCTTGCGGCTCAACCGCA
AGCGGTCAATTGGAATGGGGGACTTGAGTGCAGAAGAGGAAAGCGGAATTCACGTGTAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTTTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGG
AGCAAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGTAAAGTGTGGGAGGGTTTTCCGCCCTT
CAGTGTGCTGAGTAAACGATTAAGCACTCCGCCGGGGAGTACGGTCCGCAAGACTGAAACTCAAAGGAATGAC
GGGGCCCGCACAAAGCGGTGGAACAGGTGGTTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATC
TCCTGACCGCCCTGGAAGACAGGGTCTTCCCTTCCGGGACAGGATGACAGGTGGTTGCATGGTTGCTGTCAGCTC
GTGTCGTGAGATGTTTGGGTTAAGTCCCGCAACGAGCGCAACCTTGGGTTCTAGTTGCCAGCCATTCAGTTGGG
CACTCTAGAGCGACTGCCGGCGACAATTCGAAGAAAAGGTGGGGATGAACTTCAAATCATCAAGGCCCTTATA
GACCTGGGGCTACACCCTGTTACCATGGAAGGGTACAAGGGGAACGAAACGCGCACCCCAAGGCTTACCAGGTCTG
AAAAACATTTCTTATTTTGAATGGAGGGGGGAATTTCCCGGCTGAAAACCGGAAACCTTGTATATCGGAATAAC
AAAGCCCCCGGTAAAAAATTTCCCGGGCTTGGTAAACCCCGCCTCACCCCAAGATTTTAAACCCCAATCCC
TGAGGAACTAAGGGACCCCAATGCAATTTTGGCCTTTTTTTTGTGAAAATGTCGTACGTTGGGGAAAAATTT

TGGATCCGGGTCCTGTGAGGATCAAGAAATAAAGGAAATAAGGAGCAAAAAGGAAAAATTAAGGAGCTAAAACG
GGGT

>D06_27F

CCGACTGGCGCTGCCTATACATGCAGTCGAGCGAACCAATAAAAAGCTTGCTTTTTGTTGGTTAGCGGCGGACGG
GTGAGTAACACGTGGGTAACCTGCCTGTAAGACCGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGATTAT
CTTTCCGCCTGGAGAGATAAGGAAAGATGGCTTTTGCCATCATTACAGATGGGCCCGCGGCATTAGCTAGTTG
GTGAGGTAACGGCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGTATCGGCCACACTGGGACTGAGAC
ACGGCCAGACTCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCG
CGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGGAGGGAATGCCGGTAC
CTTGACGGTACCTGACGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGTGGCAAGCGTT
GTCCGGATTTATTGGGCGTAAAGCGCGCGCAGGCGGTCTTTAAGTCTGATGTGAAATCTTGCCGGCTCAACCGCA
AGCGGTCAATTGAAACTGGGGACTTGAGTGCAGAAGAGGAAAGCGGAATTCACGTGTAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTTCCCTTCGGGACAGGATGACAGGTGATGTTGTCGTCAGCTGGG
AGCAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACAGATGAGTGCTAAGTGTGGGAGGGTTCCGCCCTT
CAGTGCTGCAGCTAACGCATTAAGCACTCCGCCTGGGGGAGTACGGTGCGAAGACTGAAACTCAAAGGAATTGAC
GGGGCCCCGCACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCT
CCTGAGCCCTGGAGACAGGGTCTTCCCTTCGGGACAGGATGACAGGTGATGACAGGTGGTGCATGTTGTCGTCAGT
GTCGTGAGATGTTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTTGTTCTAGTTGCCAGCATTCAATTGGGCACTC
TAGAGCGACTGCCGGCGACAATTCGAAGAAAGGGTGGGGGATGAAGTTCAAATCATATGCCCCCTTATGACCT
GGGGCTACACCCTGGCTACAATGGGAGGGTACAAAGGGCAGCGAAACGGCCACCCCTTATCGAATCCCAGAAAA
CTTCTTTAATTCTGAATGGCAGGGCAGCTCCTCCGGCAACCTGAAACCGAAATCCCTTGAATCCGGAAATAACAGCC
CGGTGAATACTTTCCCGGGCTTGGTAACCCCCCGTACCACGGAAATTTGTAACCCCAACCCGGAGGAAACCTAA
GAGCCCCAATGGGCGATTGGGAGTGTGTAGAGAGCTGAGGCCCCCCCCCCCTGGGGCAAGGGGAAGAGA
AGAGAAAAAGAGAAGTGTAGAAGATGATGGGGAAGGGGGGGGAACCAA

>E06_27F

CTGACTGCGCATGCCTATACATGCAAGTCGAGCGGACCGAAGGGAGCTTGCTCCTTTAGGTTAGCGGCGGACGGG
TGAGTAACACGTGGGCAACCTGCCTGCAGACTGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACACCG
AAAACCGCATGGTTTTCGGTTGAAAGCGGCTTTTAGCTGCTACTGCAGGATGGGCCCGCGGCATTACCTAGTT
GGTGAGGTAACGGCTACCAAGGCAACAATCCGACCCCAACCTGAGAGGGTACCGGCCACACTGGGACTGAAA
CACGGCCAAACTCCTACGGGAGGCAGCATTAGGGAATCTTCCGCAATGGACAAAAGTCTGACGGACCAACGCC
CGTGAGCGAAAAAGGTCTTCGGATCGTAAAGCTCTTTGTCAGGGAAAAACAAGTACCGTTGCAACAGGGCGGTAC
CTTGACGGTACCTGACAAGGAAGCCACGGCTAACTACTTGCCACCACCCGCGTAATACTTAGGTGGCAAGCGTT
GTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAATTCTGATGTGAAATCCGCGGCTCAACCGCG
AGCGGCCATTGAAAAGTGGGGAACCTTGATTGCAGGAAAGGGAGCGGAATTCACGTGTACCGTTGAAATGCGTA
AAGATGTGAAGGAACACCATTGGCAAAGGCGGCTCTCTGGCCTGTAAGTACGCTAAGGCGCGAAAGCGTGGGGA
GCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACAATAAGTGCTAAGTGTAAAGGGTATCCACCCTTAT
TGCTGCACAAACGCATTAAGCACTCCCCGGGGAGTACGGCCGCAAGGCTGAACTCAAGGAATTGACGGGA
CCCCAAAGCGGTGAAGCAGGTGGTTAATTGCAACAACCGGAAAAACCTTACCAGGTCTTGACTTCCCTGACACCCT
AGAGATAGCGTTCCCTTTCCGGACAGGTGACAGTGGTGCAGGTTGTCGTCACCTCGGTCGTGAATTTGGGTTAGT
CCCCAACGAAACAACCTTGACTTATTGCCACCTTCATTGGGCTCTAAGTGTATGCCGCTAAAAATCGAAGAAGTGGG
GATAACTCCAATCCAGCCCCCTTAACTGGGCTACCCCTGCTCCATGGGGGTCCAAGGCACCAACCCCGGTG
GAGCAATCCAAAAACCTTTATTTGGATTGAGGTGCATTCGCCCTAAACCGAATCCTTTATTCGGAAAAATGCCCGG
TAAAAATCCGGTTTTTACCCCGTCCCCAAATTTAACCCAATCTGGGACCTTAGACCCCCAGGCATTGAGGTGT
TATAAACCTAACTGAGCCCTCCCTGCTTGGATCCGATGAAGAAAGGAAAAAATGAGGTACGCC

>F06_27F

GTGCTGCGCTGCCTATAATGCAGTCGTGCGGACCTTTTAAAGCTTGCTTTTAAAGGTTAGCGGCGGAAGGGTGA
GTAACACGTGGGCAACCTGCCTGTAAGACTGGGATAACGCCCGGAAACCGGGGCTAATACCGGAAAGTTTTTCC
TCCGCATGGAGGAAAAAGGAAAGGCGGCTTCGGCTGCCACTTACAGATGGGCCCGCGGCATAGCTAGTTGG
CGGGTAACGGCCACCAAGGCAACGATGCGTAGCCCACTGAGAGGGTATCGGCCACATTGGGACTGAGACA
CGGCCCAAACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCC
GTGAGTGAAGAAAGCCTTCGGGTCGTAAAACTCTTTGCCGGGGAAGAACAAGTCCGCTTGAACAGGGCGGCGC
CCTGACGGTACCCGGCCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACTAAGTGGCAAGCGTT
GTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGCTTCTTAAGTCTGATGTGAAATCTTGGCGGCTCAACCGCA
AGCGGTCAATTGAAAAGTGGGAGGCTTGAGTGCAGAAAAAGAGAGTGAATTCACGTGTAGCGGTGAAATGCGTA
AAGATGTGGAGGAACACCAGTGGCGAAAGCGGCTCTCTGGTCTGTAAGTACGCTGAAGCGCCAAAGCGTGGGG
AGCAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACAGATGAGTGCTAAGTGTAGAGGGTTTTCCGCCCTTT
AGTCTGCAGCTAACGCATTAAGCACTCCGCTGGGAGTACGCGCCAGGCTGAAACTCAAGGAATTGACGGG
GGCCCCAACCGTGGAGCATGTGTTAATTGCAACAACCGGAAAGAACCTTACCAGGTCTTGACTTCTGACTC
CCTGGAGACAGGCTTCCCTTCGGGACAAATGACAGTGGTGCAGGGTGTGTCACCTCTGTCGTGAAATGTTGG
GTTAATCCCCACGAGGGCAACCTTGACCTATTGCCACCATTAATTGGGCTCTAAGTGAACCGCTGACAACCGAG
AAAGGGGGGAGAAAGTCAAAATCAAGCCCCCTTAACTGGGGTACCCTGCTCCAAGGAGGGGCAAGGGTGGCAA
CCCCGGTTAACCTCCCAAAACATCCCTTTGATTAGGGGGGACCCCCCTAAACCGAATCTTTAATCGAAAA
AACCCCGGAAAAATTCGGCTTTTACCCCCCCCAAAATTTAACCCAATTTGAACTAAGCCCAAGGAATG
GGGGGTTTTAAATGGGTTTAGAGCCGGTTCGGCCGACCGGAGGTACGAATTAATGAAAG

>G06_27F

GTGCATGCGCATGCCTATACATGCAAGTCGAGCGGACCGAAGGGAGCTTGCTCCTTTAGGTTAGCGGGCGGACGG
GTGAGTAACACGTGGGCAACCTGCCCTGCAGACGGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACACC
GAAAACCGCAGGGTTTTCGGTTGAAAGGCGGGCTTTAGCTGTCACTGCAGGATGGGCCCGCGGCGCATTAGCTAG
TTGGTGAGGTAACGGCTCACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGA
GACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAAGTCTGACGGAGCAACG
CCCCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTCAGGGAAGAACAAGTGCCGTTTCAACAGGGCG
GTACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAG
CGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAATCTCGCGGCTCAAC
CGCGAGCGGCCATTGGAAACTGGGGAACCTTGAGTGCAGGAAAGGGGAGCGGAATCCACGTGTAGCGGTGAAAT
GCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTACGCTGAGGCGCGAAAGCG
TGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTATCCA
CCCTTTAGTGTGCTGCAGCAAACGCATTAAGCACTCCGCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAAT
TGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCCTACAGGTCTTGAC
ATCCCTGACAACCCTAGAGATAGGGCGTTCCCTTTTCGGGGACAGGGTGACAGGTGGTGCATGGTTGTCGTCAC
CTCGTGTGCTGAAATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGACCTTAGTTGCCAGCATTCAATTGGGC
ACTCTAAGGTGACTGCCGGCTAAAATTCGGAAGAAAGGTTGGGGATGACGTCAAATCATATGCCCTTAGGACCT
GGCTACCCCTTGTACAATGGGTGGTACAAAGGGCAGCCAAAACCCCGAGGTGGAACCAATCCCAAAAAACC
CCCTCTATTTGCAATTGAAGGGGGAATTCCTCGGCTTGAACCCGAAATCCAATAAATCCGGAATAACCAGGCCCC
CGGGAAAAAATTTCCCGGTTTTTTTAAACCCCGCTGCAACCCAGGAGTTGTTAACCAAAATCCTTGGGGTAACT
TCCAGGGACACCCCCACTGGCAAATTGGGTAAGGA

>H06_27F

CTGACGTGCGCATGCCTATACATGCAAGTCGAGCGGACCGAAGGGAGCTTGCTCCTTTAGGTTAGCGGGCGGACGG
GTGAGTAACACGTGGGCAACCTGCCCTGCAGACTGGGATAACTTCGGGAAACCGGAGCTAATACCGGATAACACC
GAAAACCGCATGGTTTTCGGTTGAAAGGCGGGCTTTAGCTGTCACTGCAGGATGGGCCCGCGGCGCATTAGCTAG
TTGGTGAGGTAACGGCTCACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGA
GACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAAGTCTGACGGAGCAACG
CCGCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTCAGGGAAGAACAAGTGCCGTTTCAACAGGGCG
GTACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAG
CGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTTAAGTCTGATGTGAAATCTCGCGGCTCAAC
CGCGAGCGGCCATTGGAAACTGGGGAACCTTGAGTGCAGGAGAGGGGAGCGGAATCCACGTGTAGCGGTGAAAT
GCGTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTACGCTGAGGCGCGAAAGCG
TGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTATCCA
CCCTTTAGTGTGCTGCAGCAAACGCATTAAGCACTCCGCTGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAAT
TGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCCTACAGGTCTTGAC
ATCCCTGACACCTCTAGAAGATAGGGCGTTCCCTTTTCGGGACAGGGTGACAGGTGTTGCATGGTTGTCGTC
GCTCGTGTGCTGAAAATTTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTGACCTTAGTTGCCAGCATTCAATTTG
GCACTCTAAGGTGACGGCCGGCTAAAATCCGAAGAAAGGTGGGGGAATAACTTCAAATCTTCAGGCCCTTATAAC
CGGGGCTACACACTGCTTCAAAGGGGTGGTACAAAGGGCGCCCAACCCCGGGTGGAAGGAAACCCAAAAAAC
ACTCTTATTTTTGAATTGAGGGCGGCACTCCCCCGCTGAAAACCGAAATCCTTTAATTCCGAAAAAAGGCCCGC
GGAAAAAATTTCCCGGCTTGGTAACCCCCCTCCCCCAGATTTTTAAACCAAAATCTGGGGGAACCTCCCGG

Anexo 2

Secuencias ADNr 16S Reverse enviadas por MacroGen Inc

>A01_1492R

TCTAAGGTTCTGTCACCTTCGGCGGCTGGCTCCTTACGGTTACCTCCACCGACTTACGGGTGTTACAAACTCTCGTG
GTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCC
GGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTCTGGGATTGCTTATGCGTCGCCGCT
TCGCTGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCCAGGTCATAAGGGGCATGATGATTTGACGTCATCC
CCACCTTCTCCGACTTGTGCGCGGCAGTCGCTCTAGAGTGCCCAACTGAATGCTGGCAACTAGAACCAAGGGTT
GCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTATCCTGTCT
CCCGAAGGGAAGACCCTGTCTCCAGGGCGGTCAAGGAGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAA
TTAAACCACATGCTCCACCGCTTGTGCGGGCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCAGCCGACTCCCA
GGCGGAGTGCTTAATGCGTTAGCTGCAGCACTGAAGGGCGGAAACCCCTCCAACACTTAGCACTCATCGTTTACGG
CGTGGACTACCAGGGTATCTAATCCTGTTTGTCTCCACGCTTTCGCGCCTCAGCGTCAGTTACAGACCAGAAAGC
CGCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTTTCTCTTCTGCA
CTCAAGTCCCCAGTTCCAATGACCGCTTGCCTGAGCCGCAAGATTTACATCAGACTTAAAGGACCGCCTCC
GCGCGCTTACGCCCAATAAATCCGGACAACGCTGGCCACCTACGTATTACCGCGCTGCTGGCACGTAGTTAGC
CGTGGCTTCTCGTCAGGTACCGTCAAGGTACCGGCATTCCCTCCGGTACTTGTCTTCCCTAACAACAGAGCTTT
ACGATCCGAAGACCTTCTTCGCTCAAGCGGCGTGTCTCCGTCAGACTTTTCGTCATTGCGGAAGATTCTACTGCT
GCCTCCCCTAGGAATCTGGGGCCTGTCTCAATCCCCATGTGGGCCAACACCCTTTCAGTCGGCTACCCACCCTC
CCCCTGGTGAGCCTTTTCTCCCAACAAATTAAGGGCCCGCGGGCCACCTTTAATTGATGACAAAAACATCTTTCT
TATCTCTCTCGCGGAGAGAAAATATCCGGGTTAAACACGGTTTCCGGAATATTCCAAATTTAAGCGGGGTACACA
CGTTTTTACCCCTTCGCGCTTCCCAAAAAAGG

>B01_1492R

CAGGCTTCTGTCACCTTCGGCGGCTGGCTCCAAAGGTTACCTCCACCGACTTCCGGGTGTTACAAACTCTCGTGGT
GTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCAG
CTTCACGCAGTCGAGTTGCAGACTGCGATCCGAAGTGAAGTGGTTTTCTGGGATTGGCTTAGCCTCGCGGCTTC
GCTGCCCTTTGTTCGCCATTGTAGCACGTGTGTAGCCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCCA
ACCTTCTCCGCTTGTGTCACCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGTCAAGGGTTGCG
CTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCACTCTGCCCC
GAAGGGGAAGCCCTATCTCTAGGGTTGTCAAGAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAATTA
ACCACATGCTCCACCGCTTGTGCGGGCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCAGCCGACTCCCGAGCG
GAGTGCTTAATGCGTTTGTGCTGCAGCACTAAAGGGCGGAAACCCCTTAACACTTAGCACTCATCGTTTACGGCGTGG
ACTACCAGGGTATCTAATCCTGTTTCGCTCCCAACGCTTTCGCGCCTCAGCGTCAGTTACAGACCAGAGAGTGCCT
TCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCACCTCTCCTCTTCTGCACTCAA
GTTCCCCAGTTTCCAATGACCCCTCCCGGTTGAGCCGGGGCTTTTACATCAGACTTAAAGAACCCGCTGCGCGC
GCTTTACGCCCAATAAATCCGGACAACGCTGGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCGTG
GCTTTCTGTTAGGTACCGTCAAGGTACCGCCCTATTGCAACGGTACTTGTCTTCCCTAACAACAGAGTTTTACGA
TCCGAAAACCTTCATCACTCAAGCGGCGTGTCTCCGTCAGACTTTTCGTCATTGCGGAAGATTCTACTGCTGCCT
CCCGTAGGAATCTGGGCCGTGTCTCAATCCCCGTGTGGCCAATACCTTCTCAGGTGGCTACCCACCCTCCCCCTG
GTGAAGCCTTTTCTCCCAACAACTTAAAGGGCCCGGCTCTCTCTTAATGGGAGCTAAAACACCCTTTTATTGA
ACATGGGGTTTCATCAACCACCCGTTTAAACCCGTTTCCCGAATATTCATCTTAAAGCGGGGTACCACGGTTAT
TACCCTCCCCCTCACAAAGAGAAAAACCCATGCCTCCATCGTTAAGCCCTCCCCCCCCCGGGGTTCCATCTA
AAA

>C01_1492R

TCAGACTCTGTCCACCTTCGGCGGCTGGCTCCTTACGGTTACCTCCACCGACTTACGGGTGTTACAAACTCTCGTGG
TGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCG
GCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTCTGGGATTGCTTATGCGTCGCCGCT
CGTGGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCC
CACCTTCTCCGACTTGTGCGCGGCAGTCGCTCTAGAGTGCCCAACTGAATGCTGGCAACTAGAACCAAGGGTTG
CGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTATCCTGTCCC
CGAAGGGAAGACCCTGTCTCCAGGGCGGTCAAGGAGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAAT
AAACCACATGCTCCACCGCTTGTGCGGGCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCAGCCGACTCCCAAGG
CGGAGTGCTTAATGCGTTAGCTGCAGCACTGAAGGGCGGAAACCCCTCCAACACTTAGCACTCATCGTTTACGGCGT
GGACTACCAGGGTATCTAATCCTGTTTGTCTCCCAACGCTTTCGCGCCTCAGCGTCAGTTACAGACCAGAAAGCCGC
CTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTTTCTCTTCTGCACTC
AAGTCCCCAGTTTCCAATGACCCGCTTGCCTTGCCTGAGCCGCAAGATTTACATCAGACTTAAAGGACCGCCTGCGC
CGCTTACGCCCAATAAATCCGGACAACGCTTGCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTG
GCTTCTCGTCAGGTACCGTCAAGGTACCGGCATTCCCTCCGGTACTTGTCTTCCCTAACAACAGAGCTTTACGAT
CCGAAGACTTCTTCGCTCAAGCGGCGTGTCTCCGTCAGACTTTCGTCATTGCGGAAGATTCCCTACTGCTGCCT
CCCCTAAGAACCTTTCGCGGTGTCTCAATCCCCATGTGGCCAATACCCTTTCAGGTGGCTACACACCGCCCC
CCTGGTTAGACCTTTACTTCCCAACAACTTAAAGGGCCCGCGGGCCCCCTCTTAATTAATTGACAAAAACCACTTT
CCTATTTTTCCGGGCGGAGAGAAAATATCCGGTATAAACCAGGTTTCCGGAATATTCCAAATTTAAGCGGGGTAC

CCACGTTTATCACCTTCCCCCTTACACAAAAAATCTAATGGTCGGATCGGTAAGCCCTTCGCCCCCCCCGGA
AAAAATA

>D01_1492R

TCACCAATGCCACTTCGGCGGCTGGCTCCCGTGACGGGTTACCGTCACCGACTTTCGGGTGTTGCAAGCTTCTC
GTGGTGTGACGGGCGGTGTGTACAAGGCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGAT
TCCGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGCGGGCTTTTTGGGATTCGCTCCCCCTCGCG
GGTTCGCAGCCCTTTGTACCGCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCA
TCCCCACTTCCCGACTTGTGCGCCGGCAGTCCCTCTAGAGTGCCACCTTCGTGCTGGCACTAGAGGGCAGG
GTTGCGCTCGTTGCGGGACTTAACCCAACATCTCAGCACGAGCTGACGACAACCATGCACCACCTGTACCCT
GTCCCCCGAAGGGGGAACGCCAATCTCTGGGTTGTCAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTG
CTTCGAATTAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCCTTTCGCGCCGTA
CTCCCCAGGCGGAGTGCTTATCGGTTAGCTGCAGCACTAAAGGGTGTGACCCCTCTAACACTTAGCACTCATCGT
TTACGGCGTGGACTACCAGGGTATCTAATCCTGTTTGTCTCCCGCCCTTCGGAACGCGCTTCCTCGTCCCTCAAA
GAGAGCCGCTTCGCCACTGGTGTCTCCACATCTCTACGATTTACCGCTACACGTGGAATTCGCTCTCCTC
TCCTGCACTCAAGTCCCCCAGTTTCCAATGACCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACTTAAGGAAC
CGCTGCGCGCGCTTACGCCAATAATTCCGGACAACGCTCGCCCCCTACGTATTACCGCGGCTGCTGGCACGT
AGTTGCGCTTCTCCTCGTGAGGTACCGTACCGCGCCCGCTTCGGAACGCGCTTCCTCGAACGCGCTTCCTCGTCCCTCAAA
CAGAGCTTACGACCCGAAGGCTTCTTCGCTCAAGCGGCGTTCGCTTCGTCAGGCTTTTCCCCATTGCGGAAGAAT
TCCTACTGCTGCCCTCCCTAAGAAATCTGGGGCCGGTGTCTAAATCCCCCTTGTGGGCGGGGAACCCCTTCAGGCG
GGGCAACCAACGTTCCCCCTGTGGAAGCCTTCTCACCACAAATTAAGGGCCCCGGGGCCCTCTCCCATGT
AAAACCCAAAGGCGCCTTTTATCAAAAAAAGAGGGTTTTCTGTGTTTTTCGGGTTAAATCCCCGTTCCCGCA
AATTATCCAGGTTCTGGGGGGGGTTGGCCGCGGTTTTATTACCCTTTCGCCAACCACAAAGAGTCCAATT
GGCTCGATCCGTAAGGCCCGTTCGGTTTCGTTGGGGCGGAAACAATAAAACAAGGGAAGGGGGGTTTCGTTAA
ACATGAAAACGGG

>E01_1492R

TCACCTTTGTACCTTCGGCGGCTGGCTCCCGTAAGGGTTACCCACACCGACTTCGGGTGTTACAACTCTCGTG
GTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGATTCC
GGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTTTGGGATTCGCTCCACCTCGCGGTTT
CGCTGCCCTTTGTACCCACCTATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGCTATCCC
CACCTTCTCCGACTTTTAGCCGGCAGTACCTTAGAGTGCCCACTGAATGCTGGCACTAAGGTCAAGGGTTGC
GCTCGTTGCGGGACTTAACCCAACATCTCAGCACGAGCTGACGACAACCATGCACCACCTGTCACCCTGTCCC
CGAAAGGGGAACGCCCTATCTTAGGGTTGTAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAAT
TAAACCATGCTCCACCGCTTGTGCGGGTCCCGTCAATTCCTTTGAGTTTCAGCCTTCGCGCCGCTACCCGAG
GCGGAGTGCTTAATGCGTTTGTGTCAGCACTAAAGGGTGGATACCCTCTAACACTTAGCACTCATCGTTTACGGCG
TGGACTACCAGGGTATCTAATCCTGTTTCGCTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGGCCAGAGAGCC
GCCTTCGCCACTGGTGTCTCCACATCTCTACGATTTACCGCTACACGTGGAATTCGCTCCCTCTCCTGCA
CTCAAGTTCCTCAGTTTCCAATGGCCGCTCGCGGTTGAGCCGCGAGATTTACATCAGACTTAAGGAACCCGCTGC
GCGCGCTTACGCCAATAATTCCGGACAACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGC
CGTGGCTTCTCGTCAGGTACCGTCAAGGTACCGCCCTGTTGAACGCTACTTGTCTTCCTGACAACAGAGCTT
TACGATCCGAAGACCTTCTTCGCTCAGCGGGCGTTGCTCCGTGAGACTTTCGTCATTGCGGAAGATCCCTACTGC
TGCTCCCGTAGGAATCTGGCCGTGTCTCATTCCCATGTGGCCGTTACCCCTTCAGGTTGGGCTACCCAAGC
TCCCCCTTGTGGAACCTTACTCCACCAACAACTTAAGGGCCGCGGGGCCCTTCGCGAGTGAAGGTAACCAACC
GCCTTACCAGAAAAACAAGCGGTTTTCTGGTTATCCGGGTTTAAACCCGTTTCCAAATATCCACCTCGCCGGG
AGGGGTGCCACGTTTTCTCCCTCCCTTACCAAGAAAAGTCTTTTGGCTCATCCGTAAGCCCTCCGTTCCGC
CGCAAAAAAATCTCAGAAAAAATAATCAACCAAGGAAATGAAATTAATAAATAAATAAATAAATAAATAAATAA
ATAGATTTACGATACTAATAAACAAGTGAAGTACAGCAGAG

>F01_1492R

TCACCTACTGTCCACCTTCGGCGGCTGGCTCCCTTACCGGGTTACCTACACCGACTTTCGGGTGTTGCAAGCTCT
CGTGGTGTGACGGGCGGTGTGTACAAGGCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCG
ATTCCGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGCGGGCTTTTTGGGATTCGCTCCCCCTCG
CGGGTTCGCAGCCCTTTGTACCGCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACG
TCATCCCCACTTCTCCGACTTGTGCGCCGGCAGTCCCTCTAGAGTGCCACCTTCGTGCTGGCACTAGAGGGC
AGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCAGCACGAGCTGACGACAACCATGCACCACCTGTCAC
CCTGTCCCCCGAAGGGGGAACGCCAATCTCTGGGTTGTAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCG
TTGCTTCGAATTAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCCTTCGCGCC
GTAATCCCCAGGCGGAGTGTCTATCGGTTAGCTGCAGCACTAAAGGGTGTGACCCCTCTAACACTTAGCACTCAT
CGTTTACGGCGGGACTACCAGGGTATCTAATCCTGTTTGTCTCCCGCTTCGCGCCTCAGCGCTCAGTTGTCAGG
CCAGAGACCGCTTCGCGACTGGTTCCTCCACATCTCTACGATTTACCGCTACACGTGGAATTCGCTCTC
CTCTCTGCACTCAAGTCCCCAGTTTCCAATGACCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACTTAAGG
AACCGCCTGCGCGCGCTTACGCCAATAATTCCGGACAACGCTCGCCCCCTACGTATTACCGCGGCTGCTGGCA
CGTAGTTAGCCGGGGCTTCTCGTGAGGTACCGTACCGCGCCGCTTTCGGAACGCGCTCCTTCGTCCTCA
CAACAGACTTTACGACCCGAAGGCTTCTTCGCTCAAGGGCCGCTTCGTTCCGTCAGGCTTCGTTCCGGA
GATCTCTACTGCTGCCTCCGTAAGAATCTGGCCGTGTCTCATTCCCATTGGGGCGGTTACCTTTCAGGCGG
GTAAACACCTCCCCCTTGTGAAGCGTTTTCTCACAACAACTAAAGGGCCGCGGGGCCCTTCCCAATGAAGC
CAAAGGCGCTTTTACCAAAACATAGGGGTCTCCGGTGTTCGGGTATAATCCCGGTTCCCGAGATATCCCGCTT

GGGGCGAAGTGCCCCGGTTATCACCCCTCCCCCTCAAACCAAAAAAATCCAATTGGCTCCATCGGTAGACCC
CTCGCTTCGTCCCGCCCGGTCTATTAGGAAAAATGAATTTGGGGTGAAGAAGCGAAAACCATTTTAAAAAGTCT

>G01_1492R

TCCACGTCGTGACACCTTCGCGGGCTGGCTCCCTTGACGGGTTACCTACACCGACTTCGGGTGTTGCAAGCTCTC
GTGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTACCCGCGGCATGCTGATCCGCGATTACTAGCGAT
TCCGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGAGCGGCTTTTTGGGATTCGCTCCCCCTCGCG
GGTTCGCAGCCCTTTGTACCGCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCA
TCCCCACCTTCCCGACTTGTGCGCCGGCAGTCCCTTAGAGTGCCACCTTCGTGCTGGCAACTAGAGGCGAGG
GTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACAGAGCTGACGACAACCATGCACCACCTGTACCCT
GTCCCCCGAAGGGGGAACGCCAATCTCTGGGTTGTCAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTG
CTTCGAATTAACACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCCTTCGCGCCGTA
CTCCCCAGGCGGAGTGCTTATCGGTTAGCTGCAGCACTAAAGGGTGTGACCCCTCTAACACTTAGCACTCATCGT
TTACGGCGGGGCTTACCAGGGTATCTAATCCTGTTTGTCTCCCGCCCTTCGAGCCTTCGCGCCTCAGCGTCAGTTGCAGGCCA
GAGAGCCGCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCTCCTC
TCCTGCACTCAAGTCCCCAGTTTCCAATGACCCTCCACGGTTGAGCCGTGGGCTTTACATCAGACTTAAGGAAC
CGCTGCGCGCGCTTACGCCAATAATTCCGGACAACGCTCGCCCCCTACGTATTACCGCGGCTGCTGGCACGT
AGTTAGCGGGGCTTCTCGTGAGGTACCCTCACCGCGCCCTTAGAGTGCCACCTTCGGAACGCTGGCAACTAGAGGCGAGGG
CAGAGCTTACGACTCGAAGGCCCTTCTTCGCTCACGCGCGCTCGCTCCGTCAGGCTTTCGCCATTGCGGAAGAT
TCCTACTGCTGCCTCCGTAGGAATCTGGGCGGTGTCTCAATCCCAATGTTGGCCGGTCACCCCTTCAGGCCGG
TAACCATCGTCGCTTGGTGAGCGTTTCTCACCCTAAGTTAAGGGCCCCCGGGCCCTTCCCAATTAACACCC
AAAGGCGGGGTTTTAACAAAGAACATGGGGTCTCTGCTTATTCCGATATAAACCCCGTTTCCCGAGATATCCCG
CTTGTGGGGGGGTTGCCCTTATATAACCCCTCCCCCTCACAAAACAAAAAATCCATATTGTGCTGAATTCG
GTTAAGGCCGCTGCTTGGT

>H01_1492R

TCCAAGTCTGACACCTTCGCGGGCTGGCTCCCGTGACGGGTTACCTGCACCGACTTCGGGTGTTGCAAGCTCTCG
TGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTACCCGCGGCATGCTGATCCGCGATTACTAGCGATT
CCGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGAGCGGCTTTTTGGGATTCGCTCCCCCTCGCGG
GTTGCGAGCCCTTTGTACCGCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCA
CCCCACCTTCCCGACTTGTGCGCCGGCAGTCCCTTAGAGTGCCACCTTCGTGCTGGCAACTAGAGGCGAGGG
TTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACAGAGCTGACGACAACCATGCACCACCTGTACCCTGT
CCCCCGAAGGGGGAACGCCAATCTCTGGGTTGTCAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCT
TCGAATTAACACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCCTTCGCGCCGTA
CCCGAGCGGGAGTGCTTATCGCGTTAGCTGCAGCACTAAAGGGTGTGACCCCTCTAACACTTAGCACTCATCGTT
ACGGCGTGGACTACCAGGGTATCTAATCCTGTTTGTCTCCCGCCCTTCGCGCCTCAGCGTCAGTTGCAGGCCAG
AGAGCCGCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCTCCTCT
CCTGCACTCAAGTCCCCAGTTTCCAATGACCCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACTTAAGGAAC
GCCACCTTCCCGGTTTGTACCGCCAAATAATTCCGGACAACGTTGAGTGCCCAACTGAATGCTGGCAACTAAGATCAAGGGT
GTTAGCCGGGGCTTCTCGTGAGGTACCCTCACCGCGCCGCTTTCGAAACGGCGCTTCTTCGCTCCCTCACAAC
AGAGCTTACGACCCGAAGGCCCTTCTTCGCTTAAAGCGGCTCGCTTCGTCAGGCTTTCGCCATTGCGGAAGAAATC
CCTACGGCTGCCTCCCCTAGGAATCTGGGGCCGGGTCTATTCCCCATGTGGGCGGGTCAACCCCTTCAGGCCGG
GAACACCGTCCCCCTGTGGAGACCTTTCTCCACACAACTTAAGGGCCCCGGGCCATCCCCAATTGAAC
CCAAAGGGCGCTTTTATCAAAGACAAAAGGGGTTTCCGGGTATTCCGGGTATAAATCCCGATATAC
CCGGCTTTGGGGGGGTTGGCCCCGTATATTCCCTCTCCCTTCCACCCAAAAAGGTCT

>A02_1492R

TCCGACGCATCTGTCGACTTCGCGGGCTGGCTCCAAAAGGTTACCTCCAACGACTTCGGGTGTTACAAACTCTCGT
GGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTACCCGCGGCATGCTGATCCGCGATTACTAGCGATT
CAGCTTCACGCAGTCGAGTTGCAGACTGCGATCCGAAGTGAAGACAGATTTGTGGGATTGGCTTAGCCTCGCGGC
TTCGCTGCCCTTTGTTCTGCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCA
CCCCACCTTCCCGGTTTGTACCGCCAGTACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGATCAAGGGT
GCGCTCGTTGCGGGACTTAACCCAACATCTCACGACAGAGCTGACGACAACCATGCACCACCTGTCACTCTGCC
CCCGAAGGGGAAGCCCTATCTTAGGTTGTGAGAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAAT
TAAACACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGTCTTCGACCGTACTCCCCAG
GCGGAGTCTAATGCGTTTGTGCGAGCACTAAAGGGCGGAAACCCCTCTAACACTTAGCACTCATCGTTTACGGCG
TGGACTACCAGGGTATCTAATCCTGTTGCTCCCGCCCTTCGCGCCTCAGCGTCAGTTACAGACCAGAGAGTGC
CCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCCTACTCTCTCTGCACT
CAAGTCCCCAGTTTCCAATGACCCTCCCCGGTTGAGCCGGGGCTTTCACATCAGACTTAAGAAACCGCTGCG
CGCGTTTTACGCCCAATAATTCCGGACAACGTTGCCACTAGTATTACCGCGGCTGCTGGCACGTAGTTAGC
GTGGCTTTCTGTTAGGTACCGTCAAGGTACCGCCCTATTGAAACGGTACTTGTCTTCCCTAACAAACAGAGTTTTA
CGATCCGAAAACCTTCACTCAAGCGGCGTTGCTCCTCAGACTTTCGTCATTGCGGAAGATTCCCTACTGCT
GCCTCCCTAGGAATCTGGGGCCGTGTCTCATTCCCCATGTGGGCAATAACCCCTCAGGTGGGGAAACCAACCG
TCCCCCTGGTGAGACCGTTTCTCACCCTAAGTTAAGGGCGCCGGGGGCTCTCTTTATTGTTGGCTAAAAACA
CCTTTTTTTAAAAACAAGGGGGTTTTTAAACAAACCCCGGTTTATACCCCGGTTTCCCCGAAATATCCACTCTTC
ACGGGGGGTCCCC

>B02_1492R

TCAGATTTGTCCACCTTCGGCGGCTGGCTCCCGTAAGGGTTACCCACCGACTTTCGGGTGTTACAAACTCTCGTG
GTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCC
GGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTTTGGGATTTCGCTCCACCTCGCGTTTT
CGCTGCCCTTTGTACCACCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCC
CACCTTCTCCGACTTTTAGCCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAGGGTTGC
GCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCACCCTGTCCC
CGAAAGGGGAACGCCCTATCTCTAGGGTTGTAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAAT
TAAACACATGCTCCACCGCTTGTGCGGGTCCCCGTCAATTCCTTGTAGTTTACAGCTTTCGCGCCGTACTCCCCAG
GCGGAGTGCTTAATGCGTTTTGCTGCAGCACTAAAGGGTGGATACCCTCTAACACTTAGCACTCATCGTTTACGGCG
TGGACTACCAGGGTATCTAATCCTGTTCCGCTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGGCCAGAGAGCC
GCCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCCTACACGTGGAATTCGCTCCCCTCTCCTGCA
CTCAAGTTCCCAAGTTTCAATGGCCGCTCGCGTTGAGCCGCGAGATTTACATCAGACTTAAGGAACCGCCTGC
GCGCGCTTACGCCCAATAATTCCGGACAACGCTGGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGC
CGTGGCTTCCGACTTTAGCCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAGGGTT
TACGATCCGAAGACCTTCTTCGCTCAAGCGGCGTTGCTCCGTCAAATTTCTGTCATTGCGGAAGAATTCCTACT
GCTGCCCCCTAAGAATCTGGGGCGTGTCTCATTCCCCCTGTGGGCCGCTCACCCCTTCAGGTGGGCTACCCAA
CCATCCCCTTGTGACCTTTCTCCCAACATATTTAAGGGCGCCGGGGCCCTCCCGCATGGAAGCTAAAAAC
GCCTTACCCAAAACAACAGCGTTTTCTGTTTTCCGGGATTAATCCCCCTTCTCCAAATATCACAACCTTCCAG
GGGAGGGTGCCCCGTTTATCACCTTCCCCCTTACCCAAAGAAAAACCCTTTCGCTCCATCGGTAAGGCCCTT
GGCTCCTTGGGGCCGACCAAATAAAGGA

>C02_1492R

TCAACAATTTGTCCACCTTCGGCGGCTGGCTCCCGTAAGGGTTACCCACCGACTTTCGGGTGTTACAAACTCTCGT
GGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCC
CGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTTTGGGATTTCGCTCCACCTCGCGTTT
TTCGCTGCCCTTTGTACCACCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATC
CCCACCTTCTCCGACTTTAGCCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAGGGTT
GCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCACCCTGTCC
CCGAAAGGGGAACGCCCTATCTCTAGGGTTGTAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCG
AATTAACACATGCTCCACCGCTTGTGCGGGTCCCCGTCAATTCCTTGTAGTTTACAGCTTTCGCGCCGTACTCCC
CAGGCGGAGTGCTTAATGCGTTTTGCTGCAGCACTAAAGGGTGGATACCCTCTAACACTTAGCACTCATCGTTTACG
GCGTGGACTACCAGGGTATCTAATCCTGTTCCGCTCCCCACGCTTTCGCGCCTCACCGTCAGTTACAGGCCAAAGA
GCCGCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCCTACACGTGGAATTCGCTCCCCTCTCCT
GCACTCAAGTTCCCAAGTTTCCAATGGCCGCTCGCGTTGAGCCGCGAGATTTACATCAGACTTAAGGAACCGC
CTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTGGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGT
TAGCCGTGGCTTCTCCTCGTCAGGTACCGTCAAGGTACCGCCCTGTTTGAACGGCACTTGTCTTCCCTGACAACAGA
GCTTTACGATCCGAAGACCTTCTTCGCTCACGCGCGTTGCTCCGTCAGACTTTCGTCATTGCGGAAGATTCCCT
ACTGCTCCCTCCCGTAGGAATCTGGGCCGTGTCTCATTCCCATGTTGGCCGGTCAACCTTTCAGGTCCGGTACCC
ACCCTCCCCTGGTGAACCGTTACCTCACAACCTAACTAAAGGCCCGGGGCCATTCCGGAATTAACACTAAAAACG
CCTTTTAAACAAAAACAAGCGTTTTTCCGGTTTTTCCGGTTAATTCCCGTTTCCCAAAATATCCACCTTTCGGGGAG
GTGGCCCCGTTTTTACCCTCTCGCCCTCCCTAAAGAAAAGACCTTCTGCCTCAATTCGGAAGGACCCCGTCCC
TTCGGCGTGGACTTAAGGGGGGGCCCCA

>D02_1492R

TCACTTTTTTGTCCACTTCGGCGGCTGGCTCCCGTAAGGGTTTACCCACCGACTTTCGGGTGTTACAAACTCTC
GTGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGAT
TCCGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTTTGGGATTTCGCTCCACCTCGCG
GTTTTGCTGCCCTTTGTACCACCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTC
TCCCCACCTTCTCCGACTTTTAGCCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAGGG
TTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTACCCTGT
CCCCGAAAGGGGAACGCCCTATCTCTAGGGTTGTAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTC
GAATTAACACATGCTCCACCGCTTGTGCGGGTCCCCGTCAATTCCTTGTAGTTTACAGCTTTCGCGCCGTACTCC
CCAGGCGGAGTGCTTAATGCGTTTTGCTGCAGCACTAAAGGGTGGATACCCTCTAACACTTAGCACTCATCGTTTAC
GGCGTGGACTACCAGGGTATCTAATCCTGTTCCGCTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGGCCAGAG
AGCCGCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCCTACACGTGGAATTCGCTCCCCTCTCC
TGCACTCAAGTTCTCAGTTTCCAATGGCCGCTCGCGTTGAGCCGCGAGATTTACATCAGACTTAAGGAACCGC
CTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTGGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGT
TAGCCGTGGCTTCTCCTCGTCAGGTACCGTCAAGGTACCGCCCTGTTTGAACGGTACTTGTCTTCCCTGACAACAGA
GCTTTACGATCCGAAGACCTTCTTCGCTCACGCGCGTTGCTCCGTCAGACTTTCGTCATTGCGGAAGATTCTCTA
CTGCTGCCCTCCCATAAGAATCTGGGCCGTGTCTCATTCCCCATTTGGCCGGTTCACCTCTCAGGTGGCTACCCA
CCCTCCCCTGGTGAACCTTTCTTACCAACAACCTAAAGGGCCGCGGGCCATCTCGCATTAAGCTAAAAACGC
CTTTACCAAAACCAACGGTTTTCTGGTTTTATCGGGTTAATCCCGTTTCCCAAAATATCCACCTTTCGGGGAG
CCCCGTATCTACCCTTCCCCCTCCCAAGAAAAGACCTTTCGCCATCGTTAAGCCCCCGGCCCTGTGGGGGG
TCCTAAGGATTTGGGGGAAGTTGTGTCTCTCGGTGAGGATAGATATATATTGTTGTCGGGTTTTCCGCT

>E02_1492R

TCAGTATCTGTCACTTCGGCGGCTGGCTCCAAAGGTTACCTCCACCGACTTCCGGGTGTTACAAACTCTCGTGGTG
TGACGGGCGGTGTGTACAAGGCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCAGC

TTCACGCAGTCGAGTTGCAGACTGCGATCCGAACAGTGAACAGATTTGTGGGATTGGCTTAGCCTCGCGGCTTCG
CTGCCCTTTGTTCTGCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCCCA
CCTTCTCCGGTTTGTACCCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGATCAAGGGTTGCGC
TCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGCTACTCTGCCCCCG
AAGGGGAAGCCCTATCTCTAGGGTTGTAGAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAATTA
CCACATGCTCCACCGCTTGTGCGGGCCCCGTCATTCCTTTGAGTTTCAGTCTTGCACCGTACTCCCCAGGCG
GAGTGTAAATGCGTTTGTGTCAGCACTAAAGGGCGGAAACCCCTCTAACACTTAGCACTCATCGTTACGGCGTGG
ACTACCAGGGTATCTAATCCTGTTCCGCTCCCCACGCTTTCCGCGCCTCAGCGTCAGTTACAGACCAGAGAGTGCCT
TCGCCACTGGTTCCTCCACATCTCTACGCATTTACCAGCTACACGTGGAATTCACCTCTCTCTCTGCACCTAA
GTTCCCCAGTTTCCAATGACCCTCCCCGGTTGAGCCGGGGGCTTTCACATCAGACTTAAGAAACCGCCTGCGCGC
GCTTACGCCCAATAAATCCGGACAACGCTGGCCACCTACGTATTACCGCGCTGCTGGCACGTAGTTAGCCGTG
GCTTCTGTTAGGTACCGTCAAGGTACCGCCCTATTGCAACGGTACTTGTCTTCCCTAACCAACAGAGTTTTACGA
TCCGAAAACCTTCATCACTCAGGCGGCGTTGCTCCGTACAGCTTTGTCATTGCGGAAGATTCCCTACTGCTGCC
TCCCGTAGGAATCTGGGGCCGTGTCTCAATCCCGTGTGGGGCAATAACCCCTCAGGTGGGCTACCCACCGTCC
CCCTGGTGAACCCCTTCTCACCACTAGCTTAAGGGCCCCGGTCTCTCTCAATGGGTACTAAAAACCCCTT
TATAATTAACAAGCGGGTTTCAATTAACAACCCGGTTTAAACCCCGGTTTCTCGGAATTATCCACCTCTTAACGCGGG
GTACCCCTGTTTCAACCCCTCCCCCTTACTTAAGGAAAAGGTCCCCTTGCCTCCCTGGTAGGCCCTCGGTT
GGTTGGTGGGGCGGCTTAGGAAAAAATAAAAAATTTCTTTAAAAAT

>F02_1492R

CCACTTATGCTGACGTTGCGGTTGCTCTGCTCCCATGATGCAGTCACCTAGACCGACTTCGGGTGTTACAAACTC
TCGTGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCG
ATTCCGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACAGAGAGTGGTTTTTTGGGATTGCTCCACCTCGC
GGTTTCGCTGCCCTTTGTACCACCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTC
ATCCCCACCTTCTCCGACTTTTAGCCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAGG
GTTGCGCTCGTTGCGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGCACCCT
GTCCCGAAAAGTGGGAACGCCCTATCTAAGGTTGTAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCT
TCGAATTAACCACATGCTCCACCGCTTGTGCGGGTCCCCGTCATTCCTTTGAGTTTCAGCCTTGCGGCCGTA
CCCCAGGCGGAGTGTAAATGCGTTTGTGTCAGCACTAAAGGGTGGATACCCTCTAACACTTAGCACTCTCTCGTT
TACGGCGTGGACTACCAGGGTATCTAATCCTGTTGCTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGGCCAA
AAAGCCGCTTTCGCCACTGGTGTCTCCACATCTCTACGCATTTTACCGCTACACGTGGAATTCGCTCCCCCTT
CCTGGCATCAAGTTCCCCAGTTTCCAATGGCCGCTCCCGTTGAACCCCGAATTTCCATCCAACTTAAGGAACCG
CCGGGCGCGCTTTACCCCAAAAAATTCGGAACACCTTGCCCCCTAGTATTACCGGGGTGTGGGACATAATTAACCG
GGGCTTCCCTCCCGGACCCCAAGGAACCCCTTTTTAAAAAGGAAATTTGTTTCTCCTGAAAAAAGATTTTAAAA
AAAAAATCTTTTCTCCCGGGTTTTCTCAAAAAATTTTTCCTTTGGGAAAAATCCCAACCCCAAGAAAA
TGGGGGGGTTTTACCCCAAGGGGGGGGCGCCCTCTAGGGGGGAAAAACCCCTGTGGGGGAGGTCTTCCC
ACAAAAAAGGGGCGCGCCCTCTGAGAAAAAACCCTCTCACACACAGAGGTGTGTGTGGGGGTTA
ACCCCGCCCTTAAATTTCCCGGGGGGGGGGCTTTTTTCCCCCCCCCAAAAGAGAACCTTCTCCCTTTAA
GGCGCTTTTGCCTCAACTCTAAAACCCCAAAACCAAAAAATAAAATGGTAACCAATATGGCCCACTAAACC
CCATAG

>G02_1492R

TCAAGCTTTGTTCCACTTCGCGGCTGGCTCCCGGAGGGTTACCTCACCGACTTCCGGGTGTTGCAAACTCTCGT
GGTGTGACGGGCGGTGTGTACAAGGCCCGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGATT
CGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACAGAGCGGCTTTTTGGGATTGCTCCCCCTCGCGGG
TTCGCGACCTTTGTACCGCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATC
CCCACCTTCTCCGACTTTTAGCCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAGGGT
GCGCTGTTGCGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGCACCGTCTC
CCCCGAAGGGGGAACGCCGATCTCTCGGGTGTGAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTT
CGAATTAACCACATGCTCCACCGCTTGTGCGGGCCCCGTCATTCCTTTGAGTTTCACTCTTGCAGCGTACTC
CCCAGGCGGAGTGTAAATGCGTTAGCTACAGCACTAAAGGGTGGATACCCTCTAACACTTAGCACTCATCGTTA
CGGCGTGGACTACCAGGGTATCTAATCCTGTTTGTCTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGACCAGA
GAGCCGCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCAGCTACACGTGGAATTCGCTCTCTCTT
CTGCACTCAAGTCCCCAGTTTCCAATGACCTCCACGTTGAGCCGTGGGCTTTCACATCAGACTTAAAAGACCG
CCTGCGCGCGCTTACGCCCAATAATTCCGGACAACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAG
TTAGCCGTGGCTTCTCGTTAGGTACCGTCAAGGTACCAGGACTGACTGCGGTAAGTGTCTTCCCTAACCAACAGA
GCTTTACGACCCGAAGGCTTCTTCGCTCAAGCGGCGTTGCTCCGTCAAATTTCTTCCATTGCCGAAGATTCTCTA
CTGCTGCCCTCCCCTAAGAAATCTGGGGCCGTGTCTATTCCCCTGTGGGCGAATAACCTCTTCAAGTGGGCTAC
CCAACTCCCCCTGGTGAAGCGTCTTCCCCAACAAAAATTAAGGGCGCCGGGGCGCTCCCTGTTGGAAGCTT
GGCGCCCTTTTAAAAAACAAGCGGCTTCTTTTATTCCGGTATAACCAATATTCTGTGTTTTCTCTCTCC
CGGGGAGTTGGCCGGTTTTATTCCCTTCCCTCAATCCAAAAAATTTAATCCCTAAATTCGGTAAGTCCCC
TCGGTCCCGTGGGGGCGGCAACTAAAAAT

>H02_1492R

TCCACTAGTGTCCACTTCGCGGCTGGCTCCCTGACGGGTTACCTCACCGACTTCCGGGTGTTGCAAGCTCTCGT
GGTGTGACGGGCGGTGTGTACAAGGCCCGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGATT
CGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACAGAGCGGCTTTTTGGGATTGCTCCCCCTCGCGGG
TTCGAGCCCTTTGTACCGCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATC

CCCACCTTCCTCCGACTTGTGCGCGGCAGTCCCTCTAGAGTGCCACCTTCGTGCTGGCAACTAGAGGCGAGGGT
TGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACAGGCTGACGACAACCATGCACCACCTGTACCCTGT
CCCCCGAAGGGGGAACGCCAATCTCTGGGTTGTCAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCT
TCGAATTAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCCTTGC GGCCGTA
CCCCAGGCGGAGTGCTTATCGCGTTAGCTGCAGCACTAAAGGGTGTGACCCCTAACTTAGCACTCATCGTTT
ACGGCGTGGACTACCAGGGTATCTAATCCTGTTTGTCCCCACGCTTTCGCGCCTCAGCGTCAGTTGCAGGCCAG
AGAGCCGCTTCGCCACTGGTGTCTCCACATCTTACGCATTTACCGCTACACGTGGAATTCGGCTCTCCTCT
CCTGCACTCAAGTCCCCAGTTTCCAATGACCCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACTTAAGGAACC
GCCTGCGCGCTTACGCCAATAATTCCGGACAACGCTCGCCCCCTACGTATTACCGCGCTGCTGGCACGTA
GTTAGCCGGGCTTCTCGTGAGGTACCGTCACCGCGCCGCCCTTTCGAACGGCGCTCCTTCGTCCTCAACAAC
AGAGCTTACGACCCGAAAGCCTTCTTCGTTAAGCGGGCTCGCTTCTTCAGGCTTTCCTCCATTGCGGAAGAATA
CCTACTGCTGCTCCCCAAGAAATCTGGGCCGGGTCTCAATCCCCCTTGGGGCGGGGCCCTTCAGGCGG
GGACGCACCTCCCCCTGTGGAGACCTTCTTCCAAACAAAGTTAAAGGGCGCCGGGGCCCTCCCCAATGAAA
AGCCAAAGGCGCCTTTACCAAAAACAAAGGGGGTTCAGGTTTTTTCGTATATAAAATCCCGTTTTCCCGGAAT
ATACCGCCTTGGGGGGAGGGGGCCCCGGTGTACTCTCCCCCTCCCAATCAAAAAGTCCCT

>A03_1492R

TCACCTACTGCTCCACTTCGGCGGCTGGCTCCCTTGACGGGTTTACCTGCACCGACTTTCGGGTGTTGCAAGCTCT
CGTGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCG
ATTCCGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACGTAGAGCGGCTTTTTGGGATTGCTCCCCCTCG
CGGGTTCGACGCCCTTTGTACCGCCATTGTAGCACGTGTGTAGCCAGGTATAAGGGGCATGATGATTTGACG
TCATCCCCACCTTCCCGACTTGTGCGCGGAGTCCCTCTAGAGTGCCACCTTCGTGCTGGCAACTAGAGGCG
AGGGTTCGCTCGTTGCGGGACTTAACCCAACATCTCACGACAGGCTGACGACAACCATGCACCACCTGTACAC
CCTGTCCCCCGAAGGGGGAACGCCAATCTCTGGGTTGTCAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCG
TTGCTTCGAATTAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCCTTGC GGCC
GTACTCCCCAGGCGGAGTGCTTATCGCGTTAGCTGCAGCACTAAAGGGTGTGACCCCTAACACTTAGCACTCAT
CGTTTAGCCGGGACTACCAAGGGTATCTAATCCTGTTTGTCTCCCCACGCTTTCGCGCCTCAGCGTCAGTTGCAGG
CCAGAGAGCCGCTTCGCCACTGGTGTCTCCACATCTTACGCATTTACCGCTACACGTGGAATTCGGCTCTC
CTCTCTGCACTCAAGTCCCCAGTTTCCAATGACCCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACTTAAGG
AACCGCCTGCGCGCGCTTACGCCAATAATTCCGGACAACGCTCGCCCCCTACGTATTACCGCGCTGCTGGCA
CGTAGTTAGCCGGGCTTCTCGTGAGGTACCGTCACCGCGCCGCCCTTTCGAACGGCGCTCCTTCGTCCTCA
CAACAGAGCTTACGACCCGAAAGCCTTCTTCGTTAAGCGGGCTCGCTTCTTCAGGCTTTCGCCCATTGCGGAA
GATTCCTACTGCTGCCCCCCGAAGAAATTTGGGGCGGGGTTTCATTCCCCCTGGTGGCCGGGTCAACCCCTTTA
GGGGGGGTAACCAACGTTCCCCCTGGTGGAGCCGTTACTCCCCAACAAAGTTAAAGGGGCCCGGGGCCATCC
CCCAGTTGAAAGGCGCCTTTAAAAAAGAAAGGGGTTTTCTGTTTTCCGGAATAATTTCCCGTTTT
CCGGAATTTTTCCGGTTTTGGGGGGGGTGGCCCCGGGTTAAC

>B03_1492R

TCAGCATCTGCTCCACTTCGGCGGCTGGCTCCCGTAAGGGTTACCTACACCGACTTTCGGGTGTTGCAAACTCTCGT
GGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTC
CGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACGTAGAAATGGTTTTTGGGATTGGCTCCACCTCGCGGTT
TCGCGACCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTATAAGGGGCATGATGATTTGACGTATCC
CCACCTTCTCCGGTTTGTACCGGCGAGTACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGTCAAGGGTTG
CGCTCGTTGCGGACTTAACCCAACATCTCACGACAGGCTGACGACAACCATGCACCACCTGTCACTCCGTTCC
CCGAAGGGGAACCTCTATCTAGAGGTAGCGAAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAA
TTAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCCTTTCGCGCCGTA
GGCGGAGTGCTTAATGCGTTAGCTGCAGCACTAAAGGGTGGAAAGCCCTAACACTTAGCACTCATCGTTACAGGC
GTGGACTACCAAGGATCTAATCCTGTTCCCTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGAGAGAGC
CGCCTTCGCCACTGGTGTCTCCACATCTTACGCATTTACCGCTACACGTGGAATTCGCTCTCCTCTCTGCA
CTCAAGTCTCCAGTTTCCAATGACCCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACTTAAGAGACCGCCGGC
GCGCGCTTACGCCAATAATTCCGGACAACGCTGGCCACCTACGTATTACCGCGCTGCTGGCACGTAGTTAGC
CGTGGCTTTCGTTGAGTACCGTCAAGGTACCGCCCTGTTGAAACGCTACTTGTTCCTCCGTTGACAACAGAGCTT
TACGATCCGAAGACCTTCTTCGCTCAAGCGCGGTTGCTCCGTTCAAACTTTCTTCCATTGCGGAAGATTCCCTACGG
CTGCCCTCCCTAAGAAATCTGGGGCCGTGTCTCAATCCCCATTGGGGCCAATACCTCTTACAGGTGGGCTACCCA
CCGTCCCCCTGGTGGAGACCTTCTCACCACAACTAAGGGGCGCCGGGGCCCTTTAAAGGGGGGGTGGGG
GCCCTTTAACTTTCTCTCCCAAGAGAAATTTCCGGTTTTAACCCCGGTTTCCGGAATAATTTCCCGTTTTCGA
GGAGGTGGCCCCGGTTATCCCCCTCCCCGCCATTTAAAAAAGAAGTGGAGGCC

>C03_1492R

TCAGTATCTGCTCCACTTCGGCGGCTGGCTCCCTACGGTTACCTACACCGACTTTCGGGTGTTACAACTCTCGTGGT
TGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCGGC
TTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACGTAGAAATGGTTTTCTGGGATTGCTTATGCGTCGCCGCTTCG
CTGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTATAAGGGGCATGATGATTTGACGTATCCCCA
CCTTCTCCGACTTGTGCGCGGCGAGTCCGCTCTAGAGTGCCCAACTGAATGCTGGCAACTAGAACCAAGGGTTGCG
CTCGTTGCGGGACTTAACCCAACATCTCACGACAGGCTGACGACAACCATGCACCACCTGTCTATCCTGTC
AAGGGAAGACCTGTCTCCAGGGCGGTCAGGAGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAATAAA
CCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCAGCCGTA
GAGTGTAAATGCGTTAGCTGCAGCACTGAAGGGCGGAACCCCTCAACACTTAGCACTCATCGTTTACGGCGTG

GACTIONCAGGGTATCTAATCCTGTTTGTCTCCACGCTTTTCGCGCCTCAGCGTCAGTTACAGACCAGAAAGCCGCC
TTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCTCTCTTCTGCACTCA
AGTCCCCCAGTTTCCAATGACCGCTTGCAGTTGAGCCGCAAGATTTACATCAGACTTAAAGGACCCGCTGCGCG
CGCTTTACGCCAATAAATCCGGACAACGCTGGCCACCTACGTATTACCGCGGCTGCTGGCACGTATTTAGCCGTG
GCTTTCTCGTACGGTACCGTCAAGGTACCGGCATTCCCTCCGGTACTTGTCTTCCCTAACAAACAGAGCTTTACGAT
CCGAAGACCTTCTCGCTCAGCGGGCTTGTCTCCGTACAGACTTTCTGTCATTGCGGAAGATTCCCTACTGCTGCCT
CCCCTAGGAATCTGGGGCCGTGTCTCAATCCCCCTGTGGGCGATAACCCCTCTCAGGTGGGCAACCCACCGTCCCC
CTGGTGAAGCGTTTCTCACCAACAACTAAAGGCCGCGGGGCCCTTTTATTTATACAAAAACATCTTTTCTTAT
TTTTCCGGGGGAGAGAAAATAATTCCGGTTTTAAACCCGGTTTCCCGGATATTCCCAATTTAAAGCGGGTTGCCCC
CGTGTATTCCCTTCCCTCCACACAAAAAAATTTTTATTTCTTTATCCGTTAAACCCCC

>D03_1492R

ATTCTTTAGTTTCGTTGACCTCGCTCTGCTTCGGAGGGTTACCTACACCGACTTCGGGTGTTGCAAACCTCTCGTGGT
GTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCG
GCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTAAAGCGGCTTTTTGGGATTGGCTCCCCCTCGCGGGT
CGCGACCTTTGTACCGCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCC
CCCCCTCTCCACTTTTAGCCGGGAGTACCTTAAAGTGCCAACTGAATGCTGGCACTAAAGTTCGAGGGTTGC
GCTGAGTGCAGGACTTAACCCAAACATCTCTCGACACGAGCTGACGACAACCATGCACCACCTGTCACCCTGTCCC
CCCGAAGGGGAACGCCGATCTCTCGGGTGTGTCAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCA
AATTAACCCACATGCTCCACCGCTTGTGCGGGCCCGTCAATTCCTTTGAGTTTCACTCTTGCAGCGTACTCCC
CAGGCGGAGTGCCTAATGCGTTAGTACAGCACTAAAGGGTGGATACCCTTAACTTACTCATCGTTTACG
GCGTGGACTACCAGGGTATCTAATCCTGTTTGTCTCCACGCTTTTCGCGCCTCAGCGTCAGTTACAGCCAGAGAG
CCGCTTTCGCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCTCTCTTCTG
CACTCAAGTCCCCAGTTTCCCATGACCCTCCACGTTGAGCCGTGGGCTTTCACATCAGACTTAAAGGACCCGCT
GCGCGCGCTTACGCCAATAATTCCGGACAACCTTGGCAACTACTTATTACCGAGGTTGCTGGCACGTAGCTAG
CCGTGACTTCTCGTTGGTACCGTCAAGGTACCCCGTGGATGCGGTCTTGTCTTTCCGTACTAGTAATCTTTAA
AACCAAAAGGTTTTCTTTCCAAGGCGTTCTTCTCGAAAGGATTTTTTCTTGGAAAAATTTCCCTTCTGCCGGCAT
ACCCACTAAGGGGTCCCCTACAAAAAGGGGGGGGG

>E03_1492R

TCAAAGACTTATCAACCTACCGGCGGCTGGCTCCCCTAAGGGTTTACCCACCGACTTACGGGTGTTACAAACTCCT
CGTGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGA
TTCCGGCTTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGAGTGGTTTTTTGGGATTTCGCTCCACCTCGCG
GTTTTCGTGCCTTTGTACCACCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTC
TCCACCTTCTCCGACTTTTAGCCGGCAGTCACTTAGAGTGCCCAACTGAATGCTGGCACTAAGGTCAAGGG
TTGCGCTCGTTGCGGGACTTAACCCAAACATCTCAGACACGAGCTGACGACAACCATGCACCACCTGTACCCTGT
CCCCGAAAGGGGAACGCCCTATCTTAGGGTGTGTCAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTC
GAATTAACCCACATGCTCCACCGCTTGTGCGGGTCCCCGTCAATTCCTTTGAGTTTCACTTTCGCGCCGTACTCC
CCAGGCGGAGTGTCTAATGCGTTTGTGTCAGCACTAAAGGGTGGATACCCTTAACTTAGCACTCATCGTTTAC
GGCGTGGACTACCAGGGTATCTAATCCTGTTGCTCCCCACGCTTTTCGCGCCTCAGCGTCAGTTACAGGCCAGAG
AGCCGCTTTCGCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCCCCTCTCC
TGCACTCAAGTTCGCCAGTTTCCAATGGCCGCTCGCGGTTGAGCCGCGAGATTTACATCAGACTTAAAGGAACCCG
CTGCGCGCTTTACGCCAATAATTCCGGACAACGCTTGCACCTACGTATTACCGCGCTGCTGGCACGTAGTT
AGCCGTGCGTTTCTCGTACGGTACCGTCAAGGTACCCGCTGTTGCAACGCTACTTGTCTTCCCTGACAACAGAG
CTTTACGATCCGAAGACCTTCTTCGCTCAGGCGGCTTGTCTCCGTACAGCTTTCTGTCATTGCGGAAGATTCCCTA
CTGCTGCTCCCCGTAAGAAATCTGGGGCCGTGTCTCAATTCCTCATGTGGGCCGGCTACCCTCTCAGGTGCGGT
ACACCACCTCTCCCTTGTGAAGCCGTTCTTCACTTCACTAATTTAAAGGGCCGCGGGGCCACCCCGCGGTAA
CGCCTAAAAACGCTTTTACTAAAAACAACCCGTTTTTTGTTTTTTTCCCGGATAAATTTCCCTTCTGCCGCAAT
ATTCACACTCTCGGGGGGAGGTTGCCCGGTTAATCCACCCCCCCCCACACAAAGAAAAAGCCATTTATTGC
CTGGCATCCGTAAAAGGCCCCCGTTACAGGGCCAGGAATAAAA

> F03_1492R

TCAGTATTTGTCCACTACCGGCGGCTGGCTCCCCTAAGGGTTTACCCACACCGACTTTCCGGGTGTTACAAACTCT
CGTGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGA
TTCCGGCTTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGAGTGGTTTTTTGGGATTTCGCTCCACCTCGCG
GTTTCGCTGCCCTTTGTACCACCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTC
TCCCCACCTTCTCCGACTTTTAGCCGGCAGTCACTTAGAGTGCCCAACTGAATGCTGGCACTAAGGTCAAGGG
TTGCGCTCGTTGCGGGACTTAACCCAAACATCTCAGACACGAGCTGACGACAACCATGCACCACCTGTACCCTGT
CCCCGAAAGGGGAACGCCCTATCTTAGGGTGTGTCAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTC
GAATTAACCCACATGCTCCACCGCTTGTGCGGGTCCCCGTCAATTCCTTTGAGTTTCACTTTCGCGCCGTACTCC
CCAGGCGGAGTGTCTAATGCGTTTGTGTCAGCACTAAAGGGTGGATACCCTTAACTTAGCACTCATCGTTTAC
GGCGTGGACTACCAGGGTATCTAATCCTGTTGCTCCCCACGCTTTTCGCGCCTCAGCGTCAGTTACAGGCCAGAG
AGCCGCTTTCGCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCCCCTCTCC
TGCACTCAAGTTCGCCAGTTTCCAATGGCCGCTCGCGGTTGAGCCGCGAGATTTACATCAGACTTAAAGGAACCCG
CTGCGCGCTTTACGCCAATAATTCCGGACAACGCTTGCACCTACGTATTACCGCGCTGCTGGCACGTAGTT
TAGCCGTGGCTTCTCGTACGGTACCGTCAAGGTACCCGCTGTTGCAACGCTACTTGTCTTCCCTGACAACAGAG
GCTTTACGATCCGAAGACCTTCTTCGCTCAAGCGGCGTTGCTCCGTCAAACTTTCTGTCATTGCGGAAAAATCCCT
ACTGCTGCCTCCCCTAAGAAATCTGGGGCCGTGTCTCAATCCCCCTGTGGGCCGGTACCCTCTCAGGTGGGCTA

CCCACCCCTCCCCTTGTGGAGACCTTTCTCTACCAACTAAGTTTAAGGGCCGCGGGGCGCTTCCCTGGCATGGA
GAGCTAAAAAACCGCTCTACCCAAAAACCAAGCGGTTTTTTGTGTTTTTCCGGGATATAAACCCGGTTTTCCCAA
TATATCCATTTCCAGGGGAGGTTGCCACCGTTTATCCCTCTTCCCCCTACCAAAGAAAAAGACCTCTTGCGCC
GATCGGTAAAGCCCCTCGGGTTCGGGGCCCCCGCGGGCTTAAGAGAAGAAAA

>G03_1492R

TCACAACCTGTCCACCTTCGGCGGCTGGCTCCCCTAAGGGTTACCCACCGACTCGGGTGTACAAACTCTCGTG
GTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCC
GGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTTTGGGATTCCGCTCCACCTCGCGTTT
CGCTGCCCTTTGTACCACCCATTGTAGCAGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTATCCC
CACCTTCTCCGACTTTTAGCCGGCAGTACCTTAAAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAGGGTTGC
GCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACAACAACCATGCACCACCTGTACCCCTGTCCCC
GAAAGGGGAACGCCCTATCTCTAGGGTTGTCAAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCAATT
AAACCACATGCTCCACCGCTGGTGCAGGCTCGCGTTGAGCCCGGAGATTTACATCAGACTTAAGGAACCGCCGAG
GCGGAGTGTAAATGCGTTTGTGCAGCACTAAAGGGTGGATACCCTTAACACTTAGCACTCATCGTTTACGGCG
TGGACTACCAGGGTATCTAATCCTGTTCCGCTCCCCACGCTTTCCGCGCCTCATCGTCAGTTACAGGCCAGAGAGCCG
CCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCCCCTCTCCTGCAC
TCAAGTTCCCGAGTTTCCAATGGCCGCTCGCGTTGAGCCCGGAGATTTACATCAGACTTAAGGAACCGCCGAG
GCGCGCTTTACGCCCAATAAATTCGGACAACGCTCGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGC
CGTGGCTTCTCGTCAGGTACCGTCAAGGTACCGCCCTGTTGCAACGGTACTTGTCTTCCCTGACAACAGAGCTT
TACGATCCGAAAAATTTCTCGCTCAGGGCGGTTGCTCCATCAGACTTTCTCCATTGGGGAAAAATTCCTTACG
GCTGCCCCCGGAAGAAATTTGGGGCCCTGGTTTCAACTCCCGTGGGGCCGCGCCACCCCTTAGGGGGGGCA
ACCCACCCCCCTTGGGTGAACCTTTCTCCCCAAAAAATTAAGGGCCCCCGGGCCCTTCTTCAATGAA
ATACTAAAAACCCCTTTTACCAAAAAACAAGGGGTTTTTTGTTTTATTCCGGGAAAAACCCCGTTTCCCCA
TTTTACCCATTTCCCGGGGAGGGTGGCCCCCTTTTTCTCCCCCCCCCTTAACAAAAAGAAAAATCATTT
GGTCTTTTTCGTTAAGACCGGTCGTCAA

>H03_1492R

TCAGAATCTGTCCACTTCGGCGGCTGGCTCCTTACGGTTACCTCACCGACTTCCGGGTGTTACAAACTCTCGTGGT
GTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCG
GCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTCTGGGATTCCGTTATGCGTCCGCGCTT
CGCTGCCCTTTGTACCATCCATTGTAGCAGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTATCCC
CACCTTCTCCGACTTTGTCGCCGGCAGTGCCTTAGAGTGCCCAACTGAATGCTGGCAACTAGAACCAAGGGTTG
CGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTATCCTGTCCC
CGAAGGGGAAGACCTGTCTCCAGGGCGGTCAGGAGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCAATT
AAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCAGCCGACTCCCCAGG
CGGAGTGTAAATGCGTTAGCTGCAGCACTGAAGGGCGGAAACCTCCAACACTTAGCACTCATCGTTTACGGCGT
GGACTACCAGGGTATCTAATCCTGTTTGTCTCCCCACGCTTTCCGCGCCTCAGCGTCAGTTACAGACCAGAAAGCCG
CCTTCCGACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTTCCCTTCTGCACTC
AAGTCCCCAGTTTCCAATGACCGCTTGGCGTTGAGCCGCAAGATTTACATCAGACTTAAAGGACCGCCTGCGCG
CGCTTACGCCCAATAAATTCGGACAACGCTGGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCGT
GGCTTTCTCGTCAGGTACCGTCAAGGTACCGGCACTTCCCTCCGGTACTTGTCTTCCCTAACAACAGAGCTTTACG
ATCCGAAGACCTTTCGCTCACGCGGCTTGTCTTCAAATTTCTCCATTGCGGAAGATTTCTACTGCTCCCT
CCCGTAGGAATCTGGGGCGTGTCTATCCATGTGGGCCAATACCCTCTCAGGTGGCTACCCACCGTCCCCCTG
GTGAACCGTTATCTCCCAACAAATTAAGGGCCCCGGGGCCCTCTTAAATGTATGACAAAAACACTTTTTCTTTT
TCTTCCGCGGGAGAGAAAAATATTCCGGGTATAAACACGGTTCCCGCAATTTTCCATTTTATGGG

>A04_1492R

TCAACCTTCTGTCCACTTCGGCGGCTGGCTCCAAAGGTTACCTCCACCGACTTTCGGGTGTTACAAACTCTCGTGG
TGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCA
GCTTCACGCAGTCGAGTTGCAGACTGCGATCCGAAGTGAAGACAGATTTGTGGGATTGGCTTAGCCTCGCGGCTT
CGCTGCCCTTTGTACCACCCATTGTAGCAGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTATCCC
CACCTTCTCCGTTTGTACCCGGCAGTACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGATCAAGGGTTGC
GCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTACTCTGCCCC
CGAAGGGGAAGCCCTATCTCTAGGGTTGTCAAGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCAATTA
AACACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCAGCCGACTCCCCAGGC
GGAGTGTAAATGCGTTTGTGCAGCACTAAAGGGCGGAAACCTTAACACTTAGCACTCATCGTTTACGGCGTG
GACTACCAGGGTATCTAATCCTGTTCCGCTCCCCACGCTTTCCGCGCCTCAGCGTCAGTTACAGACCAGAGATCGCC
TTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCCACTCTCTCTTCTGCACTCA
AGTTCGCCAGTTTCCAATGACCCCTCCCCGTTGAGCCGGGGCTTTACATCAGACTTAAGAAACCGCCTGCGCG
CGTTTTACGCCCAATAAATTCGGACAACGCTGGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCGT
GGCTTTCTGGTTAGGTACCGTCAAGGTGCCGCCCTATTGCAACGGTACTTGTCTTCCCTAACAACAGAGTTTACG
ATCCGAAAACCTTCACTCAAGCCGGCGTGTCTCCGTTCCAGACTTTTCGTCATTTGCGGAAAAATTCCTACTGC
TCCCTTCCCTTGGTTGAGCCGTTTTCTCAACCAAACTTAAGAGCGCCCGGGCCCTCTTCAATGGT
AGCTAAAAGCACCCCTTATATTAGGAAACATGGCGGTTTATTAACACCCCGCTA

>B04_1492R

TCCCAGTAACTGTTAACTTTTCGGCGGCTGGCTCCTAAAGGTTACCTCACCGACTTCGGGTGTTACAAACTCTCGTG
GTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCC
AGCTTCACGCAGTCGAGTTGCAGACTGCGATCCGAAGTGAAGACAGATTTGTTGGGATTGGCTTAACCTCGCGGTTT
CGCTGCCCTTTGTTCTGTCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCC
CACCTTCTCCGGTTTGTACCGGCAGTCACCTTAGAGTGCCAACTGAATGCTGGCAACTAAGATCAAGGGTTGC
GCTCGTTGCGGGACTTAACCCAAACATCTCACGACAGAGCTGACGACAACCATGCACCACCTGTCACTCTGCCCC
CGAAGGGGACGTCCTATCTCTAGGATTGTAGAGGATGTAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAATTA
AACACATGCTCCACCGCTTGTGCGGGCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCACCGTACTCCCCAGGC
GGAGTGCTTAATGCGTTAGCTGCAGCACTAAGGGGCGGAAACCCCTAACACTTAGCACTCATCGTTTACGGCGT
GGACTACCAGGGTATCTAATCCTGTTCCGCTCCCCAGCTTTCCGCTCCTCAGCGTCAGTTACAGACCAGAGAGTCCG
CTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCACCTCTCCTTCTGCACTC
AAGTTCCCCAGTTTCAATGACCTCCCCGTTGAGCCGGGGCTTTACATCAGACTTAAGAAACCGCTGCGA
GCCCTTACGCCCAATAATTCCGACAACGCTTTCACGACAGCTATTACCGCGGCTGCTGGCACGTAGTTAGCCGT
GGCTTTCTGGTTAGGTACCGTCAAGGTACCGCCCTATTGAAACGGTACTTGTCTTCCCTAACAAACAGAGCTTTACG
ATCCGAAAACCTTCACTCACTCAAGCGGCGTTGCTCCGTCAAACCTTTCGTCATTGGCGGAAAATTCCTACTGCTG
CCTCCCCCTAGGAATCTGGGGCCGTGTCTCAATCCCCATGTGGGCCAATAACCTCCTCAGGTGGGCTACCCACCG
TTTCTTGTGAGCAATTACCTACCAACGTAATGAAAGGCCCGCCCTCTTAAATTGGTGGCCAAAGCAC
CCCTCTTATTTTTAACAAAGCGGGTTAAAAAAAACACGGGTATAAACCCGGT

>C04_1492R

TCAGCTATTGTACCCGAGGCAGCGCCCTCCTTGCGGTTAAGCTACCTGCTTCTGGTGCAACCAACTCCCATGGT
GTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCAGCAATGCTGATCTGCGATTACTAGCGATTCCG
ACTTCATGGAGTCGAGTTGCAGACTCCAATCCGACTGAGATGGGGTTTCTGGGATTGGCTCCCCCTCGCGGTT
CGCAGCCCTCTGTCCCCACCATTGTAGTACGTGTGTAGCCCTGGTCTGAAGGGCCATGATGACTTGACGTCATCCC
CACCTTCTCCGGTTTGTACCGGGCGTCTCCTTAGAGTTCCACCATTACGTGCTGGCAACTAAGGACAAGGGTT
GCCCTCGTTGCGGGACTTAACCCAAACATCTCACGACAGAGCTGACGACAGCCATGCAGCACTGTCTCGCGGCT
CCCGAAGGCACCCGCCATCTCTGGGCAGTTCGCGGGATGTCAAGACCAGGTAAGGTTCTTCGCGTTGCATCGAA
TTAAACCACATACTCCACCGCTTGTGCGGGCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCACCGTACTCCCCA
GGCGGCGAATTAACGCGTTAGCTTCGATACTGAGTCCCAAATTAACCCAAACATCCAGTTCGCATCGTTTAGGGC
GTGGACTACCAGGGTATCTAATCCTGTTTCTCCACGCTTTCGTCGCTCAGTGTCAAGTTTGGCCAGGTGGCC
GCCTTCGCCACGGGTGTTCTCCGATCTCTACGCATTTCACTGCTACACCGGGAATTCGCCACCCTCTACCACA
CTCTAGTACCCAGTATCCACTGCAATTCAGGTTGAGCCAGGGCTTTCACAACGGGCTTAAGCCACCACCTAC
GCACGCTTACGCCCAAGTAATCCGAGTACCGTTCACCCCTCGTATTACCGCGGCTGCTGGCACGAAAGTTAGCC
GGTGTATTCTTTGGGTACCGTATCCCGCGGGTATTAGCCAGCAGGATTTCTCCCAACAAAAGGGCTTTA
CAACCCGAAGGCTTCTTACCCACGCGGCGAGGGCTGGATCAGGCTTTTCGCCATTGTCCAATATTCCCCACTGCT
GCCTCCCCGTAAGAATCTGGGGCCGTGTCTCAATCCCCAGTGTGGCTGGACCATCCTCTTCAAACAAGCTACCGAT
CGTCGCCCTTGTGGGCTTTACCCCCCACAATAATTCGAAATCGGGCCATTCTTCCCGGGGAGGGCCCT
GGGGTCCCCCTTTCCCTCAGGGGTAGCGGGGTAAAGGGAAAGTTTCCCTTTTCTCCCCCAGCG
GGAAGAAATCAAAGCGCTCTACCCCTCCCCCTTCCCGGCTGGAAGAGGGTTTTCC

>D04_1492R

TCCACAGATACTGTCCACTTATCGGCGGCTGGCTCCAAAGGTTACCTCACCGACTTCGGGTGTTACAAACTCTCGT
GGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGATTCC
CAGCTTCACGCAGTCGAGTTGCAGACTGCGATCCGAAGTGAAGACAGATTTGTTGGGATTGGCTTAGCCTCGCGC
TTCGCTGCCCTTTGTTCTGCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATC
CCCACCTTCTCCGGTTTGTACCGGCAGTCACCTTAGAGTGCCAACTGAATGCTGGCAACTAAGATCAAGGGTT
GCGCTCGTTGCGGGACTTAACCCAAACATCTCACGACAGAGCTGACGACAACCATGCACCACCTGTCACTCTGCC
CCCGAAGGGGAAAGCCCTATCTCTAGGGTTGTAGAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAAT
TAAACCACATGCTCCACCGCTTGTGCGGGCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCACCGTACTCCCCAG
GCGGAGTGCTTAATGCGTTTGTGTCAGCACTAAGGGGCGGAAACCCCTAACACTTAGCACTCATCGTTTACGGCG
TGGACTACCAGGGTATCTAATCCTGTTCCGCTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGACCAGAGAGTCCG
CCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCACCTCTCCTTCTGCACT
CAAGTTCCCCAGTTTCCAATGACCTCCCCGTTGAGCCGGGGGCTTTCACATCAGACTTAAGAAACCGCCTGCG
CGCGCTTACGCCCAATAATTCCGACAACGCTTGCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCC
GTGGCTTCTGGTTAGGTACCGTCAAGGTACCGCCCTATTGAAACGGTACTTGTCTTCCCTAACAAACAGAGTTTTA
CGATCCGAAAACCTTCACTCACTCAAGCGGCGTTGCTCCGTGAGACTTTCGTCATTGCGGAAAGAAATCCCTACTGCT
GCCTCCGTAGGAATCTGGGGCCGGTCTCAATCCCCATGTGGGCCAATACCCCTCTCAGGTGCGCTACACAACGT
TCCCCCTGTTGGAGCCGTTTCTCCCCAACTAGCTTAAGGGCCGCGGGGCTTTTCTAATGTGGGCTAAAAACAC
CTTTTTAATAACAAGCGGGTGTATCAACAACACCGCGTTTAAACCCCGGTTCCCGGAATATTCCACTCTCTAA
GGCGGGTTACCCCGTTTTTTTCCCTCCCCCTCTCTAAGGGGAAAC

>E04_1492R

TCAAGTATCCTGTCCACTTTTCGGCGGCTGGCTCCTAAAGGTTTACCTTCCACCGACTTTCCGGGTTGTTTACAAACT
CCTCGTGGTTGTTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTAC
TAGCGATTCCAGCTTACGCAGTCGAGTTGCAGACTGCGATCCGAAGTGAAGACAGATTTGTTGGGATTGGCTTAAC
CTCGCGTTTTCGCTGCCCTTTGTTCTGTCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTG
ACGTCATCCCCACCTTCTCCGGTTTGTACCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGAT

CAAGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTC
ACTCTGCCCCGAAGGGGACGTCTATCTCTAGGATTGTCAGAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTG
CTTCGAATTAACACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCAGCCGTA
CTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTAAGGGGCGGAAACCCCTAACACTTAGCACTCATCG
TTTACGGCGTGACTACCAGGGTATCTAATCCTGTTGCTCCCCACGCTTTCGCTCCTCAGCGTCAGTTACAGACC
AGAGAGTCGCCTTCGCCACTGGTGTTCCTCCACATCTCTACGCATTTACCCGCTACACGTGGAATTCACCTCTCCT
CTTCTGCACTCAAGTTCGCCAGTTTCCAATGACCCTCCCCGGTTGAGCCGGGGGCTTTCACATCAGACTTAAGAAA
CCGCCCTGCGAGCCCTTACGCCCAATAATTCGGACAACGCTGGCCACCTACGTATTACCGCGGCTGCTGGCAGC
TAGTTAGCCGTGGCTTCTGTTAGGTACCGTCAAGGTACCGCCCTATTGGAACGGTACTTGTCTTCCCTAACAA
AGAGCTTACGATCCGAAAACCTTCACTCACTCAGGCGGCGTTGCTCCGTCGACTTTCTTCCCTGCGGAAGAATC
CCTACTGCTGCCTCCCCTAAGAAATCTGGGGCCGTGTCTATTCCCCTGTGGCCATAACCCTCTCAGGTCCGC
TACCAACCGTTTCCCTGGTGAGACCGTTACCTCCACAACCTAAGGGCCCCGCGGGTCTTTTCTTAATTGTT
AGCCAAACACACCTTTTTATTTTTAATCATGCCGGTTTTAAACAACATCCCGGCTATTAATCACGGGTTTCCGGAG
TATATCCATCTCTAAGGGGGGGTTACCACCGTGTATTCCCCCTTCCCCCTTCAACCCAGGGAAGAGGTTCTT
TCT

>F04_1492R

TCCAGTATCTGTTACCTTCGGCGGCTGGCTCCAAGGTTACCTCACCGACTTCGGGTGTTACAAACTCTCGTGGT
GTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCAG
CTTCACGCAGTCGAGTTGCAGACTGCGATCCGAAGTGAAGACAGATTTGTTGGGATTGGCTTAGCCTCGCGGCTTC
GCTGCCCTTTGTTCTGCCATTGTAGCACGTGTGTAGCCCAAGGTCATAAGGGGCATGATGATTTGACGTCATCCCC
ACCTTCCCTCGGTTTGTACCCGGCAGTCACTTAGAGTGCCTAAGTGAATGCTGGCAACTAAGATCAAGGGTTGCG
CTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCACTCTGCCCC
GAAGGGGAAGCCCTATCTCTAGGGTTGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTGAATTA
ACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCAGCCGACTCCCCAGGCG
GAGTGCTTAATGCGTTTGTGCGAGCACTAAGGGCGGAAACCCCTCACACTTAGCACTTACGACTCATGTTACGGCGTGG
ACTACCAAGGTATCTAATCCTGTTGCTCCCCACGCTTTCGCGCCTCATCGTCAGTTACAGACCAGAGAGTCCGCT
TCGCCACTGGTGTTCCTCCACATCTCTACGCAATTCACCGCTGCAGTGGAATCCACTCTCCTTCTGCACTCAA
GTTCCCCACTTTCCAATGACCCTCCCCGGTTGAGCCGGGGGCTTTCACATCAAACCTAAGAAACCGCCGGCGCG
GCTTACGCCCAATAATTCGGGACAACACTAGCCACTACGTATTACAGCGGCTGCTGGCACGTAATAGCCATGC
CTTTCTGGTTAGGTACCGTACCGGTACGGGCTATTGGAACGGGACGTTGTTCTTCCCTTACCACCAAGTTGAGG
ATTGAAAAACCTTCACTCCCTCCAAACGGCGGTGCTTCTTTTACTTTTATTCATTTTGGGAGAATGGCAAAACAT
GCCGCCCGCTCCCAATTAAGAAAAGGGGGGGGGGGGTTATAACCCAAGAACGACCCCAAGTATTGGTAGGGG
GGGGGCCCTTGGCCGGTAGGGTGGAAAAATAAGCAAAGCCAGCCAGGTTGGGGGTACCAGGTTAATTTTT
TATTGATTGTTAAGAAAAGGAACCCAACTAGAAAAGGCCGGAAGGAAGGATACTGCCAATTTGGGGTTAAACG
TGTCAATCCCTTTGTTTTAATTTCTGGGGCGGGGCCGGCCACTCCTCACCCAAATCCAAGTACAAAAAACG
GTTTGGCTTTTTGG

>G04_1492R

TACCAGCTTCTGTCCCTTCGGCGGCTGGCTCCAAGGTTACCTCACCGACTTCGGGTGTTACAAACTCTCGTGG
TGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCA
GCTTCACGCAGTCGAGTTGCAGACTGCGATCCGAAGTGAAGACAGATTTGTTGGGATTGGCTTAGCCTCGCGGCTT
CGTGCCCTTTGTTCTGCCATTGTAGCACGTGTGTAGCCCAAGTCAAGGGGCATGATGATTTGACGTCATCCC
CACCTTCTCCGGTTTGTACCCGGCAGTCACTTAGAGTGCCTAAGTGAATGCTGGCAACTAAGATCAAGGGTTGCG
GCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCACTCTGCCCC
CGAAGGGGAAGCCCTATCTCTAGGGTTGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTGAATTA
AACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAAGTCTTGCAGCCGACTCCCCAGG
GGAGTGCTTAATGCGTTTGTGCGAGCACTAAGGGCGGAAACCCCTCAACACTTAGCACTTACGACTCATGTTACGGCGT
GACTACCAGGGTATCTAATCCTGTTGCTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGACCAGAGAGTCCGCC
TTCGCCACTGGTGTTCCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCACACTCTCCTTCTGCACTCA
AGTTCCCAGTTTCCAATGACCCTCCCCGGTTGAGCCGGGGGCTTTCACATCAGACTTAAGAAACCGCCGGCGCG
CGTTTACGCCCAATAATTCGGGACAACGCTTGCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTG
GCTTTCTGGTTAGGTACCGTCAAGGTGCCGCCCTATTGGAACGGTACTTGTCTTCCCTAACAACAGAGTTTTACGA
TCCGAAAACCTTCACTCACTCAGGCGGCGTTGCTCCGTCCGACTTTCGTCCATTGCGGAAAAATTCCTACTGCTGC
CTCCCCGTAAGAAATCTGGGACCGTGTCTCAGTCCCATGTGGGGCCATAACCCCTCAAGTTGGGCTACCCAGG
GTCCCCCTGTGGAGCCGTTACTACCAACAAACTAAGGGCGCCCGGTCGCTCTCTATTAGGTACTTAAAAACCA
CTTTTAGATGGAACAACACGGTGGACGAAAAAACGGCGATTATGCCACGTTCTCGCGGATATTTCTCACTTTC
GGCGGGTATCCCGTGTGTGGACGAACCCAC

>H04_1492R

TCCAGACTGTCCACTTCGGCGGCTGGCTCCCCTAAGGTTTACCCACACCGACTTTCGGGTGTTACAAACTCTCGT
GGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGATTCC
CGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTTTGGGATTGCTCCACCTCGCGGT
TTCGCTGCCCTTTGTACCACCAATTGTAGCACGTGTGTAGCCCAAGTCAAGGGGCATGATGATTTGACGTCATC
CCCACCTTCCCGGACTTTTAGCCGGCAGTCACTTAGAGTGCCTAAGTGAATGCTGGCAACTAAGGTCAAGGGTT
GCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCAACCTGTC
CCCGAAAGGGGAACGCCCTATCTCTAGGGTTGTCAAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCG
AATTAACACATGCTCCACCGCTTGTGCGGGTCCCCGTCATTCCTTTGAGTTTCAAGCTTGCAGGCGGACTCCC

CAGGCGGAGTGCTTAATGCGTTTGCTGCAGCACTAAAGGGTGGATACCCTCTAACACTTAGCACTCATCGTTTACG
GCGTGGACTACCAGGGTATCTAATCCTGTTCCGCTCCCAAGCTTTCCGCGCCTCAGCGTCAGTTACAGGCCAGAGA
GCCGCCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATCCGCTCCCCTCTCCT
GCACTCAAGTTCCCCAGTTTCCAATGGCCGCTCGCGTTGAGCCGCGAGATTTACATCAGACTTAAGGAACCCG
CTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTGGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGT
TAGCCGTGGCTTCTCGTCAGGTACCGTCAAGGTACCGCCCTGTTTGAACGGTACTTGTCTTCCCTGACAACAGA
GCTTTACGATCCGAAGACCTTCTTCGCTCAAGCGCGTTGCTCCGTCAAACTTTTCGTCCATTGTGCAAAAAATCCTA
CTGCTGCCTCCCCTAAGAAATCTGGGGCCGTGTCTCAATCCCCCTGTGGGGCCGGGACCCCTTCAGGTGGGTAAC
CAAACGTCCCCTGGTTGAGCCGTTTCTCCCCACACACTTAAGGGCGCCGGGGCGCTCCCGCATTGAGAAT
AAAAAACGCCTCTCTCAAACACACAAGGGTTTTTGTGTTTTTCCGGGAAAAATCCCCGGTTCCCCCGAAATTAC
C

>A05_1492R

TCAGCATTTGCCACCTTCGGCGGCTGGCTCCCGTAAGGGTTTACCTTCCACCGACTTTCCGGGTGTTGCAAACCTCC
TCGTGGTGTGACGGGCGGTGTTGTACAAGGCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGC
GATTCGGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACCTGAGAGCGGCTTTTTGGGATTGGCTCCCCCTC
GCGGGTTCGCGACCCCTTTGTACCGCCATTGTAGCACGTGTGTAGCCCAGGTGATAAGGGGCATGATGATTTGAC
GTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTACAGCACTAAAGGGTGGATACCCTTAACACTTAGCACTATG
AGGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTAC
CCTGTCCCCCGAAGGGGGGAACGCCGATCTCTCGGGTGTGACGGGGATGTCAAGACCTGGTAAGGTTCTTCGCG
TTGCTTCGAATTAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCACTCTTGGCAGC
GTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTACAGCACTAAAGGGTGGATACCCTTAACACTTAGCACTATG
CGTTTACGGCGTGGACTACCAGGGTATCTAATCCTGTTTGTCTCCCAAGCTTTTCGCGCCTCAGCGTCAGTTACAGA
CCAGAGAGCCGCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCGCTCTC
CTCTTCTGCACTCAAGTCCCCAGTTTCCAATGACCTCCACGGTTGAGCCGTGGGCTTTACATCAGACTTAAAA
GACCGCTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCACCTACGTATTACCGCGGCTGCTGGCA
CGTATTTAGCCGTGGCTTCTCGTTAGGTACCGTCAAGGTACCGCCCGTACTGCGCGTACTTGTCTTCCCTAAC
AACAGAGCTTTACGACTCGAAAGCCTTCTCGTTAGGCGGCGTGTCTCTCAAACCTTTCTTCCATTTCCGAAGAA
TTCCCTACTGCTGCCTCCCCTAAGAAATCTGGGGCCGTGTCTCATTCCCCCTGTGGGGCGAATAACCCCTTCAGGT
GGGCCACCCACACTCCCCCTGGTTAGACGTTTCTCTCCCAACAAATTAAGGGCGCGC

>B05_1492R

TCAATCATTTCCACCTTTCCGGCGGCTGGCTCCCGTAAGGGTTTACCTACACCGACTTTCCGGGTGTTGCAAACCTCTC
GTGGTGTGACGGGCGGTGTGTACAAGGCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGCGAT
CGGTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACCTGAGAGCGGCTTTTTGGGATTGGCTCCCCCTC
GGTTTCGCGACCCCTTTGTACCGCCATTGTAGCACGTGTGTAGCCCAGGTGATAAGGGGCATGATGATTTGACGTCA
TCCCCACCTTCTCCGACTTTTAGCCGGCAGTCACTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTGAGGG
TTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTACCCCTGT
CCCCCGAAGGGGGGAACGCCGATCTCTCGGGTGTGACGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCT
TCGAATTAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCACTCTTGGCAGCGTACT
CCCCAGGCGGAGTGCTTAATGCGTTAGCTACAGCACTAAAGGGTGGATACCCTTAACACTTAACACTCATCGTTT
ACGGCGTGGACTACCAGGGTATCTAATCCTGTTTGTCTCCCAAGCTTTTCGCGCCTCAGCGTCAGTTACAGACAGA
GAGCCGCTTCGCCACTGGTGTCTCCACATCTACGCATTTACCGCTACACGTGGAATTCGCGCTCTCTCTT
CTGCACTCAAGTCCCCAGTTTCCAATGACCTCCACGGTGTGACCGGTTTTCACATCAGACTTAAAAAGCCG
CCTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCACCTACGTATTACCGCGGCTGCTGGCACGTAAT
TAGCCGTGGCTTCTCGTTAGGTACCGTCAAGGTACCGCCAGTACTGCGCGTACTTGTCTTCCCTAACAAACAG
AAGCTTTACTACCGGAAGGCCTTTCTTTTTCAGGCGGCGGTGCTTGGTTCAAACCTTTCAATCCATTGCCGAAG
GAGTCACTAAGTGGCCCTCCCCCTACCAAACTTGGGGCGGGGGGGGGCGGGGGGGGGGGGGGGGGGGGGGG
ACGGCCCTAAGGTGGGGGGGGGGGGTGGTCCCCCTTGGTTGAGTGAGGTATGATTACCTAGTTACCTCAAATAA
GTGCGGGGGGGGG

>C05_1492R

TCCCAGTAATTTGTCAACCTACGGCGGCTGGCTCCCGTAAGGGTTTACCCACACCGACTTTCCGGGTGTTACAAACT
CTCGTGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGGCATGCTGATCCGCGATTACTAGC
GATTCGGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACCTGAGAGTGGTTTTTTGGGATTCCGCTCCACCTCG
CGTTTTCGCTGCCCTTTGACCACCCATTGTAGCACGTGTGTAGCCAGGTGATAAGGGGCATGATGATTTGACGT
CATCCCCACCTTCTCCGACTTTTAGCCGGCAGTCACTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAG
GGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTACCC
TGTCCCCGAAAGGGGAACGCCATCTCTAGGGTGTGACGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTG
TTCGAATTAACCACATGCTCCACCGCTTGTGCGGGTCCCCGTAATTCCTTTGAGTTTCACTCTTGGCAGCGTAC
TCCCCAGGCGGAGTGCTTAATGCGTTTGTGCTGCAGCACTAAAGGGTGGATACCCTTAACACTTAGCACTACGT
TACGGCGTGGACTACCAGGGTATCTAATCCTGTTTCGCTCCCAAGCTTTTCGCGCCTCAGCGTCAGTTACAGGCCA
GAGAGCCGCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCCCCTC
TCCTGCACTCAAGTCTCAGTTTCCAATGGCCGCTCGCGTTGAGCCGCGAGATTTACATCAGACTTAAGGAAC
CGCCTGCGCGCGCTTTACGCCCAATAATTCCGGACAACGCTTGCACCTACGTATTACCGCGGCTGCTGGCACGT
AGTTAGCCGTGGCTTCTCGTCAGGTACCGTCAAGGTACCGCCCTGTTTGAACGGTACTTGTCTTCCCTGACAAC
AGAGCTTTACGATCCGAAGACCTTCTTCGCTCACGCGCGTGTCTCCTTCAGACTTTTCGTCCATTGCGGAAAAAT
TCCCTACGGCTGCCCGCCGTAAGAAATCTGGGGCCGTGTCTCAATCCCCATGTGGGGCCGGTGACCCCTTCA

GGGTGGGCAACCCACCACTCCCCCTTGTGAGACCCGTTACCTCACCCAATAATGCTATAGGGCCCGCCGGG
CGCATCTCCCGCGGTGAAAAGCCTAAAAAGCGCCGTTT

>D05_1492R

TCAGTTTCTGTCCACTTCGGCGGCTGGCTCCTAAAGTTACCTCACCGACTTTCGGGTGTTACAAACTCTCGTGT
GTGACGGGCGGTGTGTACAAGGCCCGGAACGTATTCACCGCGCATGCTGATCCGCGATTACTAGCGATTCCAG
CTTCACGCAGTCGAGTTGCAGACTGCGATCCGAAGTGAAGACAGATTTGTGGGATTGGCTTAACCTCGCGTTTTG
CTGCCCTTTGTTCTGTCCATTGTAGCACGTGTGTAGCCAGGTGATAAGGGGCATGATGATTTGACGTCATCCCCA
CCTTCTCCGGTTTTGTACCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGTCAAGGGTTGCGC
TCGTTGCGGGACTTAACCAACATCTCACGACAGAGCTGACGACAACCATGCACCACCTGTCACCTGCCCCG
AAGGGGACGTCCTATCTCTAGGATTGTCAGAGGATGTAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAATAAA
CCACATGCTCCACCGCTTGTGCGGGCCCCGTCATTCCTTTGAGTTTCAGTCTTGCAGCCGACTCCCCAGGCG
GAGTGCTTAATGCGTTAGCTGCAGCACTAAGGGGCGGAACCCCTAACACTTAGCACTCATCGTTTACGGCGTG
GACTTCTGGTTAGTCAATCCTGTTCCGCTCCACGCTTTCGCTCAGTCCAGTCAAGTCAAGTCAAGTCAAGTCAAG
TTCCGCACTGGTGTCTCCACATCTCTACGATTTACCGCTACACGTGGAATTCACCTCTCCTCTTCTGCACTCA
AGTTCACCGAGTTTCCAATGACCCTCCCCGTTGAGCCGGGGGCTTTACATCAGACTTAAGAAAACCGCCTGCGAG
CCCTTACGCCAATAATTCCGGACAACGTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTG
GACTTCTGGTTAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT
TCCGAAAACCTTCATCACTCAAGCGGCGTTGCTCCTCAGACTTTGCTCCATTGCGGAAAAAATCCCTACGGCTGC
CTCCCCGTAGAAATCTGGGGCGGTGTCTCATTCCCCCTGTGGGCCAATAACCCCTTCAGGTGGGCTACCCACC
CTTTCCCTTGTGTGAGCCCTTCTCCCCACCAACCTAAAGCGCCCGGGTTCCTCCTTTAATTGGTATCCAA
AACACCCCTTTTATGTTTTAACCATGCGGGTTTTAAA

>E05_1492R

TCAGTATCTGTCCACCTTCGGCGGCTGGCTCCTTACGGTTACCTCCACCGACTTTCGGGTGTTACAAACTCTCGT
GTGTGACGGGCGGTGTGTACAAGGCCCGGAACGTATTCACCGCGCATGCTGATCCGCGATTACTAGCGATTCC
GGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTCTGGGATTTCGTTATGCGTCGCGGCT
TCGCTGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTGATAAGGGGCATGATGATTTGACGTCATCC
CCACCTTCTCCGACTTGTGCGCCGGCAGTCGCTCTAGAGTGCCCAACTGAATGCTGGCAACTAGAACCAAGGGT
GCGCTCGTTGCGGGACTTAACCAACATCTCACGACAGAGCTGACGACAACCATGCACCACCTGTATCCTGTG
CCCGAAGGGAAAGACCTGTCTCCAGGGCGGTACGAGATGCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAA
TTAAACCACATGCTCCACCGCTTGTGCGGGCCCCGTCATTCCTTTGAGTTTCAGTCTTGCAGCCGACTCCCA
GGCGGAGTGCTTAATGCGTTAGCTGCAGCACTGAAGGGCGGAACCCCTCCAACACTTAGCACTCATCGTTTACGG
CGTGGACTACCAGGGTACTAATCCTGTTTGTCTCCACGCTTTCGCGCCTCAGCGTCAGTTACAGACCAGAAAGC
CGCCTTCGCACTGGTGTCTCCACATCTCTACGATTTACCGCTACACGTGGAATTCGGCTTTCCTCTTCTGCA
CTCAAGTCCCCAGTTTCCAATGACCCTTTCGCGTTGAGCCGCAAGATTTACATCAGACTTAAGGACCGCCTGC
GCGCGCTTACGCCAATAAATCCGGACAACGTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGC
CGTGGCTTTCTCGTCAGGTACCGTCAAGGTACCGGCATTCCCTCCGGTACTTGTCTTCCCTAACAAACAGAGCTTT
ACGATCCGAAGACCTTCTCGCTCAAGCGGCGTTGCTCCTCAAACCTTTCTCCATTGCGGAAGACTTACTGTG
CCTCCCCAAGAATCTGGGGCGGTGTCTCATTCCCCCTGTGGGGCAATACCCCTTCAGGTGGGCTACCCACCGTCCC
CTGGTGAACCGTTTCTCCCAACAAACTAAGGGCGCGGGCCCTCCTTTATTTAACAACCAACATCTTTCTTATT
TTCTCCGCGGGAGAAAAAATTCGGGTTAA

>F05_1492R

TCCAGTATCTGTCCACCTTTCGGCGGCTGGCTCCGTAAGGTTTACCTTCCACCGACTTTCGGGTGTTACAAAC
TTCTCGTGGTTGTTGACGGGCGGTGTTGTACAAGGCCCGGAACGTATTCACCGCGCATGCTGATCCGCGATT
CTAGCGATTCCGGCTTCATGCAGGCGGTTGCAGCCTGCAATCCGAAGTGGGATGTTTTCTGGGATTGGCTTAA
CCTCGCGCTCTCGCAGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTGATAAGGGGCATGATGATTT
GACGTCATCCCCACCTTCTCCGGTTTGTACCGGCAGTCACCTTAGAGTGCCCAACTCAATGCTGGCAACTAAGG
TCAAGGGTTGCGCTCGTTGCGGGACTTAACCAACATCTCACGACAGAGCTGACGACAACCATGCACCACCTGT
CACTCTGTCCCCGAAGGGGAAGGCCCTGTCTCCAGGGAGGTGAGAGGATGTAAGACCTGGTAAGGTTCTTCGC
GTTGCTTCAAGTAAACACATGCTCCACCGCTTGTGCGGGCCCCGTCATTCCTTTGAGTTTACGCTTTCGCGC
CGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTAAAGGGCGGAACCCCTTAACACTTAGCACTCA
TCGTTTACGGCGTGGACTACCAGGGTATCTAATCCTGTTTGTCTCCACGCTTTCGCGCCTCAGCGTCAGTTACAG
ACCAGAGAGCCGCTTCGCCACTGGTGTCTCCACATCTCTACGATTTACCGCTACACGTGGAATTCACCTCT
CCTCTTCTGCACTAAGCCTCCAGTTTCCAATGACCCTTTCGCGTTGAGCCGCAAGATTTACATCAGACTTAAG
AAGCCGCTGCGCGGCTTACGCCAATAAATCCGGACAACGTTGCCACCTACGTATTACCGCGGCTGCTGGC
ACGTATTTAGCCGTGGCTTTCGCGCGGTTACCGTCAAGGGCGCCGCTTTCGAACGGCACTTGTCTTCCCG
GCAACAAAATTTACGACCCGAAGGCCCTTTCACCTAGCGCGGCTTGTCTTCAAACCTTTCGTTCAATTCGCGAA
AAATTCCTACTGCTGCCCTCCCCAAGAAATTTGGCCCTGGTTTCAATCCCCAAGTGGGCCAAAAACCCCT
CAGGTGGGCAACCCACCTTTCCCTTTGTTGGGGCGTTTCCCCCACCACCTAAAGGCCCCCGGGGCC
CATTTTTAATTTGAGGGCCAAAAACCCCTTTTTTTTTTTTTTCTTAAAG

>G05_1492R

TCAAGAATCTGTCCACTTCGGCGGCTGGCTCCTTTGCGGTTACCTCCACCGACTTTCGGGTGTTACAAACTCTCGT
GGTGTGACGGGCGGTGTGTACAAGGCCCGGAACGTATTCACCGCGCATGCTGATCCGCGATTACTAGCGATTCC
CGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGTGGTTTTCTGGGATTTCGTTATGCGTCGCGC
CTTCGCTGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTGATAAGGGGCATGATGATTTGACGTCAT

CCCCACCTTCTCCGACTTGTGCGCCGCGAGTCGCTCTAGAGTGCCCAACTGAATGCTGGCAACTAGAACCAAGG
TTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCATCCTGT
CCCCGAAGGGAAGACCCTGTCTCCAGGCGGTGTCAGGAGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGA
ATTAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGTCTTGGCACCCTACTCCCC
AGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTGAAGGGCGGAAACCCTCCAACACTTAGCACTCATCGTTTACG
GCGTGGACTACCAGGGTATCTAATCCTGTTTGTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGACCAGAAAG
CCGCTTCGCCACTGGTGTTCCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGGCTTTCCTTCTGTC
ACTCAAGTCCCCAGTTTCCAATGACCGCTTGCAGTTGAGCCGCAAGATTTACATCAGACTTAAAGGACCGCCTG
CGCGCGTTTACGCCAATAAATCCGGACAACGCTTGCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAG
CCGTGGCTTTCGTCAGGTACCGTCAAGGTACCGGCATTCCTCCGGTACTTGTTCCTCCAAACAACAGAGCTT
TAGGATCCGAAGACCTTCTCGCTCAAGCGCGTGTCTCCTTCAAACCTTTCGTCATTGCGGAAGAATTCCTACTGC
TGCCCTCCCTAAGAAATCTGGGGCCGTGTCTCAATCCCCCTGTGGGCCAATAACCCTCTCAGGTGGGGAACCAAC
GTCCCCCTGGTGAAGCGTTTCTCCCCACAAAACCTAAAGGCCCGCGGCCCCCTCTTAATTATTAGAGAAAACAC
ATTTTCTTATTTTCTTCAA

>H05_1492R

TCCACCTATCTGACCTTCGGCGGCTGGCTCCTTACGGTTACCTCACCGACTTCGGGTGTTACAAACTCTCGTGG
TGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGATTCCG
GCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACCTGAGAATGGTTTTCTGGGATTCGCTTATGCGTCGCGCTT
CGCTGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCC
CACCTTCTCCGACTTGTGCGCGGCGAGTCGCTCTAGAGTGCCCAACTGAATGCTGGCAACTAGAACCAAGGGTTG
CGCTGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCTCCTGTCCC
CGAAGGGAAGACCCTGTCTCCAGGGCGGTGACGAGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCAATT
AAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGTCTTGGCACCCTACTCCCCAGG
CGGAGTGCTTAATGCGTTAGCTGCAGCACTGAAGGGCGGAAACCCTCCAACACTTAGCACTCATCGTTTACGGCGT
GGACTACCAGGGTATCTAATCCTGTTTGTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGACCAGAAAGCCGC
CTTCGCCACTGGTGTTCCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGGCTTTCCTTCTGCACTC
AAGTCCCCAGTTTCCAATGACCGCTTGCAGTTGAGCCGCAAGATTTACATCAGACTTAAAGGACCGCCTGCGCG
CGCTTACGCCAATAAATCCGGACAACGCTTGCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTG
GCTTTCGCTCAGGTACCGTCAAGGTACCGGCATTCCTCCGGTACTTGTTCCTCCTAACAACAGAGCTTACGAT
CCGAAGAACTTTCGCTTAAGCGGCGTGTCTCCGTCAAAACCTTTCGTCATTGCGGAAGATTCCCTACCGCTGCC
CCCCAAGAAATCTGGGGCGGTGTCTCATTCCCCCTGTGGGCGAATAACCCTCTAGGTGGGCTACCCAACGTCC
CCCTGTGAGACCGTTCCTCCACAACAAAACCTAAGGGGCGCCGGGGCGCCTCTTTTTTTGTTTGACAAAAAACACTTT

>A06_1492R

TCATAGACTGAACCTATTACAGGCGGCTGGCTCCCGTAAGGGTTTACCCACCGACTTTCGGGGTGTACAAACTC
TCGTGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCG
ATTCCGGCTTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACCTGAGAGTGGTTTTTTGGGATTCGCTCCACCTCGC
GGTTTCGCTGCCCTTTGTACCACCCATTGTAGCAGTGTGTAGCCAGGTGATAAGGGGCATGATGATTTGACGTC
ATCCCCACCTTCTCCGACTTTTGTAGCCGGCAGTACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAGG
GTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTACCCCT
GTCCCCGAAAGGGGAACGCCCTATCTTAGGGTTGTACAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTGTCT
TCGAATTAACCACATGCTCCACCGCTTGTGCGGGTCCCCGTTCAATTCCTTTGAGTTTCAGCTTCCGCGGTACT
CCCCAAGCCGGAGTGCTTAATGCGTTTGTGCGGACTCAAGGGTGGATACCCTCTAACACTTAGCACTCATCGTTT
ACGGCGTGGACTACCAGGGTATCTAATCCTGTTTCGCTCCACGCTTTCGCGCCTCAGCGTCAGTTACAGGCCAG
AGAGCCGCTTCGCCACTGGTGTTCCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGGCTCCCCCTCT
CCTGCACTCAAGTTCGCCAGTTTCCAATGGCCGGCTCGCGTTGAGCCGCGAGATTTACATCAGACTTAAAGGAAC
CGCTGCGCGCGTTCACGCCAATAAATCCGGACAACGCTTGCACCTACGTATTACCGCGGCTGCTGGCACGT
AGTTAGCCGTGGCTTTCCTCGTCAGGTACCGTCAAGGTACCGCCCTGTTCCGAACGGTACTTGTTCCTCCCTGACAAC
AGAGCTTACGATCCGAAAAACCTTCTTCGCTCACGCCGCGTTTTTCTTCCTTCGGACTTTTCTTCCCTTTCGCGAA
AAATTCCTACAGGTGCTCCCCCTAAGAAATCTGGGGCCCCGGTCTCCATTCCCCCTGTGGGGCC

>B06_1492R

TTCTAGGTATCTGTACCTTCGGCGGCTGGCTCCGTAAGGTTACCTCCACCGACTTTCGGGTGTTACAAACTCTCG
TGGTGTGACGGGCGGTGTGTACAAGGCCCGGGAACGTATTACCGCGGCATGCTGATCCGCGATTACTAGCGATT
CCGGCTTCATGCAGGCGGTTGCAGCCTGCAATCCGAACCTGGGAATGGTTTTCTGGGATTGGCTAACCTCGCGG
TCTCGCAGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTGATAAGGGGCATGATGATTTGACGTCAT
CCCCACCTTCTCCGGTTTGTACCGGCGAGTCACCTTAGAGTGCCCAACTCAATGCTGGCAACTAAGGTCAAGGGT
TGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCATCTGTG
CCCCGAAGGGGAAGCCCTGTCTCCAGGGAGTCCAGAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCCG
AATTAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCTTTCGCGCGTACTCCC
CAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTAAAGGGCGGAAACCCTCTAACACTTAGCACTCATCGTTTACG
GCGTGGACTACCAGGGTATCTAATCCTGTTTGTCCCCACGCTTTCGCGCCTCAGCGTCAGTTACAGACCAGAGAG
CCGCTTCGCCACTGGTGTTCCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCACCTCTCCTCTTCTGC
ACTCAAGCTCCAGTTTCCAATGACCGCTTGCAGTTGAGCCGCAAGATTTACATCAGACTTAAAGAACCGCCCTG
CGCGCGCTTACGCCAATAAATCCGGACAACGCTGGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAG
CCGTGGCTTTCGCGCGGTTACCGTCAAGGCGCCGCCCTGTTCCGAACGGCACTTGTTCCTCCCGGCAACAGAGT
TTTACGACCCGAAGGCTTCTTCACTCAGGCGGCGTTCCTTCTTCGAACTTTCGTCATTGCCGAAAAATTCCTA

CTGCTGCCCTCCCCCTAAGAAATTTTGGGCCCGGGTCTCAATTCCCCAAGTGGGCCGAATAACCCCCCTTCAGGTTGG
GCAACCCA

>C06_1492R

TCAACTTTCCCTGTCCACCTTCGGCGGGTGGCTCCTTTACGGTTTACCTTCCACCGACTTTCGGGTGTTACAAACTCT
CGTGGTGTGACGGGCGGTGTTGTACAAGGCCGGGAACGTATTACCAGCGGCATGCTGATCCGCGATTACTAGC
GATTCGGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGATGGTTTTCTGGGATTTCGCTTATGCGTC
GCCGCTTCGCTGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACG
TCATCCCCACCTTCTCCGACTTGTCCGGCAGTCGCTCTAGAGTGCCCAACTGAATGCTGGCAACTAGAACCAA
GGGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCAGCACGAGCTGACGACAACCATGCACCACCTGTCATC
CTGTCCCCGAAGGGAAGACCCTGTCTCCAGGGCGGTGAGGAGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCT
TCGAATTAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCAGCCGACT
CCCCAGGCGGAGTGCTTAATGCGTTAGCTGCAGCACTGAAGGGCGGAAACCCTCCAACACTTAGCACTCATCGTT
TAGCCGCTGGACTACCAGGGTATCTAATCCTGTTTGTCCCCACGCTTTCGCGCCTCAGCGTCAAGTTACAGACCAG
AAAGCCGCTTCGCCACTGGTGTTCCTCCACATCTCTACGCATTTACCAGCTACACGTGGAATTCGCGTTTCTCTT
CTGCACTCAAGTCCCCCAGTTTCCAATGACCGCTTGCAGTTGAGCCGCAAGATTTACATCAGACTTAAAGGACCG
CCTGCGCGCTTTACGCCAATAAATCCGGACAACGCTGGCCACCTACGTATTACCAGCGGCTGCTGGCACGTAG
TTAGCCGCTGGACTACCAGGGTATCTAATCCTGTTTGTCCCCACGCTTTCGCGCCTCAGCGTCAAGTTACAGACCAG
GCTTTACGATCCGAAGAACTTTTTCGCTTAAGCGGGCGTTTCTTCTTCAAATTTCTTCCATTGCGGAAAAATTC
TACGGCTGCCCTCCCCCTAAGAAATCTGGGGCCGTGTCTCAATCCCCCTTGTGGGCCGAATAACCCCTCCTCAGGGTGG
GCTACCCACCCCTCCCCCTGTTTGAACCC

>D06_1492R

TCAATCATCTGTCCACCTTTGGCGGGTGGCTCCTACGGTTACCTACCGACTTAGGGTGTACAAACTACTCGTGG
TGTGACGGGCGGTGTGTACAAGGCCGGGAACGTATTACCAGCGGCATGCTGATCCGCGATTACTAGCGATTCCG
GCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGATGGTTTTCTGGGATTTCGCTAATGCGTCGCGCT
TCGCTGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATCC
CCACCTTCTCCGACTTGTCCGGCAGTCGCTCTAGAGTGCCCAACTGAATGCTGGCAACTAGAACCAAGGGT
GCGCTCGTTGCGGGACTTAACCCAACATCTCAGCACGAGCTGACGACAACCATGCACCACCTGTCATCCTGTG
CCCCAAGGGAAGACCCTGTCTCCAGGGCGGTGAGGAGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAA
TTAAACCACATGCTTCCAGCGCTTGTCCGGGCCCCCGTCAATTCCTTTGAGTTTCAGTCTTGCAGCCGACTCCCCA
GGCGGAGTGCTTAATGCGTTAGCTGCAGCACTGAAGGGCGGAAACCCTCCAACACTTAGCACTCATCGTTTACGG
CGTGGACTACCAGGGTATCTAATCCTGTTTGTCCCCACGCTTTCGCGCCTCAGCGTCAAGTTACAGACCAGAAAGC
CGCCTTCGCCACTGGTGTTCCTCCACATCTCTACGCATTTACCAGCTACACGTGGAATTCGCTTTCCTCTTCTGCA
CTCAAGTCCCCCAGTTTCCAATGACCGCTTGCAGTTGAGCCGCAAGATTTACATCAGACTTAAAGGACCCGCTGC
GCGCGCTTTACGCCAATAAATCCGGACAACGCTTGCCACCTACGTATTACCAGCGGCTGCTGGCACGTAATTAGCC
GTGGCTTTTCTCGTCAAGTACCGTCAAGGTACCGGCAATTCCTTCCGCTACTTGTCTTTCCTTAAACAACAGAAGCT
TTTCCAATCCGAAGAACTTCTTCGCTTCAAGCCGCGTGTGCTTCTTCAAATTTTCATCCCTTTCGCGGAAAAAT
TCCCTACTGCTGCCCTCCCCGTAGAGAATTTGGGCACGTGTTCTCAGATTCCCCCTTGTGGCCCAATACACCC
TCCTTAAGGTTGGGCTAACGCCACCCCTTACCTCTGTTGTGAGAGCCC

>E06_1492R

TCAATTTTGTCCACTTCGGCGGGTGGCTCCTTAAGGGTTTACCCACACCGACTTTCGGGTGTTACAAACTCTCG
TGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTACCAGCGGCATGCTGATCCGCGATTACTAGCGATT
CCGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAAGTGAAGATGGTTTTTTGGGATTTCGCTCCACCTCGCGG
TTTTCGCTGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCAT
CCCCACCTTCTCCGACTTTTAGCCGGCAGTACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAGGGT
TGCCTCGTTGCGGGACTTAACCCAACATCTCAGCACGAGCTGACGACAACCATGCACCACCTGTCACCCTGT
CCCCGAAAGGGGAACGCCCTATCTAGGGTTGAGCCGATGTCAAGACCTGGTAAGGTTCTTCGCTTGTCTT
GAATTAACCACATGCTCCACCGCTTGTGCGGGTCCCCGTCAATTCCTTTGAGTTTCAGCCTTGCAGCCGACTCC
CCAGGCGGAGTGCTTAATGCGTTTGTGTCAGCACTAAGGGTGGATAACCTTAAACACTTAGCACTCATCGTTTAC
GGCGTGGACTACCAGGGTATCTAATCCTGTTGCTCCCCACGCTTTCGCGCCTCAGCGTCAAGTTACAGGCCAGAG
AGCCGCTTTCGCCACTGGTGTTCCTCCACATCTCTACGCATTTACCAGCTACACGTGGAATTCGCTCCCTCTCC
TGCACTCAAGTTCCCCAGTTTCCAATGGCCGCTCGCGTTGAGCCGCGAGATTTACATCAGACTTAAAGGAACCGC
CTGCGCGCGCTTTACGCCAATAAATCCGGACAACGCTGGCCACCTACGTATTACCAGCGGCTGCTGGCACGTAGT
TAGCCGTTGGCTTCTCGTCAAGTACCGTCAAGGTACCCGCTGTTTCAAGCGGTACTTGTCTTCCCTGACAACAGA
GCTTTACGATCCGAAACCTTCTTCGCTCAAGCGCGGTTGTTTCTTCCAGACTTTCCTTCCCTTGGCGGAAAAATTC
CTACGGCTGCCCTCCCCCTAAGAAATCTGGGGCCGTGTCTCAATCCCCCTTGTGGGGCGGGGGACCCCTTTCAGG
GTGGGCAACCCACCCCTCCCCCTGTTTGAACCC

>F06_1492R

TCGGCATTCTGTCACTTCGGCGGGTGGCTCCTTAAGGGTTACCTACCGACTTTCGGGTGTTACAAACTCTCGTGG
AGAGACGGGCGGTGTGTACAAGGCCGGGAACGTATTACCAGCGGCATGCTGATCCGCGATTACTAGCGATTCCG
GCTTCATGCAGGCGGGTGTGAGCCTGCAATCCGAAGTGGAAATGGTTTTCTGGGATTTCGCTTAACTCGCGGTCT
CGCAGCCCTTTGTACCATCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCC
CACCTTCTCCGGTTTGTACCCGCGAGTACCTTAGAGTGCCCAACTCAATGCTGGCAACTAAGGTCAAGGGTTGC
GCTCGTTGCGGGACTTAACCCAACATCTCAGCACGAGCTGACGACAACCATGCACCACCTGTCACCTCTGCCCC
CGAAGGGGAAGGCCCTGTCTCCAGGGAGGTCAAGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCAAT
TAAACCACATGCTCCACCGCTTGTGCGGGCCCCCGTCAATTCCTTTGAGTTTCAGCCTTGCAGCCGACTCCCCAG

GCGGAGTGCTTAATGCGTTAGCTGCAGCACTAAAGGGCGGAAACCCCTAACAACCTTAGCACTCATCGTTTACGGCG
TGGACTACCAGGGTATCTAATCCTGTTTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATCCACTCTCCTCTTCTGCACT
CCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATCCACTCTCCTCTTCTGCACT
CAAGCCTCCCAGTTTCCAATGACCGCTTGGGTTGAGCCGCAAGATTTACATCAGACTTAAGAAGCCGCCTGCGC
GCGCTTTACGCCAATAATTCCGGACAACGCTTGCCACCTACGTATTACCGCGGCTGCTGGGCACGTAGTTTGCC
GTGGCTTTCTGGCCCGGTACCGTCAAGGCGCCGCCCTGTTTCAACGGCACTTGTCTTCCCCGGCAACAGAGTT
TTTACGAACCGAAAGGCTTTCTTCACTACGCGGGCTTTGCTTCCGTGAGAATTTTCGTTTATTGGCGAAAAATTC
CCCTACTGGCTGCCCTCCCCGTAAGGAATTTGGGGCCGTGTTCTCAATTTCCAATGGGGGCCGAATCACCCCTTT
TCAGGGTCGGGCTAACCCATCCGTTTGCCTTGGGTGGGGCGGTTAAC

>G06_1492R

TCCCAATGTCACCTTTTCGGCGGCTGGCTCCCGTAAGGGGTACCCACCGACTTCGGGTGTTACAAACTCTCGTG
GTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGCATGCTGATCCGCGATTACTAGCGATTCC
GGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACCTGAGAGTGGTTTTTTGGGATTTCGCTCCACCTCGCGGTTT
CGCTGCCCTTTGTACCACCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGTATCCC
CACCTTCTCCGACTTTTAGCCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAGGGTTGC
GCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCACCCTGTCCC
CGAAAGGGGAACGCCCTATCTCTAGGGTTGTAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGCTTCGAAT
TAAACCACATGCTCCACCGCTTGTGCGGGTCCCCGTC AATTCCTTTGAGTTTACGCTTTCGCGCCGTACTCCCCAG
GCGGAGTGCTTAATGCGTTTGTGTCAGCACTAAAGGGTGGATACCCTCTAACACTTAGCACTCATCGTTTACGGCG
TGGACTACCAGGGTATCTAATCCTGTTTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCCCTCTC
GCACTCAAGTTCCCCAGTTTCCAATGGCCGGCTCGCGGTTGAGCCGCGAGATTTACATCAGACTTAAGGAACCGCT
GCGCGCGCTTACGCCAATAATTCCGGACAACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTA
GCCGTGGCTTTCTCGTCAAGTACCGTCAAGGTACCGCCCTGTTTCAACGGTACTTGGTTCTTCCCTGACCACAG
AAGTTTTACGATCCGAAAAACCTTCTTCTGCTTACGCGCGCTTGTCTCCGTTACAGACTTTCGTTCCATTTGCGGA
AGAATTCCTAATGGCTGCCCTCCCCGTAAGAAATCTGGGGCGGTTGCTCAATTTCCCAATTTGGCCCGGTTCA
CCCTCTTACAGTCCGGCTTACCCACTCTTGCCTTGGTGGAGGCCGTTTACGTCGCGCAACTAACTAAC

>H06_1492R

TCAGTATTTGTCACCTACCGCGGCTGGCTCCCGTAAGGGTTTACCCACACCGACTTTCCGGGTGTTACAAACTT
CTCGTGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGCGCATGCTGATCCGCGATTACTAGC
GATTCGGGCTTCATGCAGGCGAGTTGCAGCCTGCAATCCGAACCTGAGAGTGGTTTTTTGGGATTTCGCTCCACCTCG
CGTTTTCGCTGCCCTTTGTACCACCCATTGTAGCACGTGTGTAGCCAGGTCATAAGGGGCATGATGATTTGACGT
CATCCCCACCTTCTCCGACTTTTAGCCGGCAGTCACCTTAGAGTGCCCAACTGAATGCTGGCAACTAAGGTCAAG
GGTTGCGCTCGTTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTCACCC
TGTCCCCGAAAGGGGAACGCCCTATCTCTAGGGTTGTAGGGGATGTCAAGACCTGGTAAGGTTCTTCGCGTTGC
TTCGAATTAACACATGCTCCACCGCTGGTGCAGGGTCCCCGTC AATTCCTTTGAGTTTACGCTTTCGCGCCGTAC
TCCCAGGCGGAGTGCTTAATGCGTTTGTGTCAGCACTAAAGGGTGGATACCCTCTAACACTTAGCACTCATCGTT
TACGGCGTGGACTACCAGGGTATCTAATCCTGTTTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCCCTCT
AGAGCCGCTTCGCCACTGGTGTCTCCACATCTCTACGCATTTACCGCTACACGTGGAATTCGCTCCCTCT
CCTGCACTCAAGTTCCCCAGTTTCCAATGGCCGGCTCGCGGTTGAGCCGCGAGATTTACATCAGACTTAAGGAAC
CGCTGCGCGCGCTTACGCCAATAATTCCGGACAACGCTTGCCACCTACGTATTACCGCGGCTGCTGGCACGT
AGTTAGCCGTGGCTTCTCGTCAAGTACCGTCAAGGTACCGCCCTGTTTCAACGGCACTTGTCTTCCCTGACAAC
AGAGCTTACGATCCGAAAAACCTTCTTCTGCTTACGCGCGCTTTTTTCTTCAAACCTTTCGTTCCCTTTGGGGAAAA
ATTCCCTACTGGTGGCTCCCCGCTAGGAAATCTGGGGCCCGGTTTTTCATTTCCCCCGGGGGGGCGGGGACCC
CCCTCAGGTGGGGCAACCCACCCTCCCCCTGGGGGAGGCCGTTTCCCTCCACACCAAGTTA

Anexo 3

Secuencias consenso de bacterias termófilas cultivables

>Bacillus_hisashii_A01

GGTGCTAATACCGGATAGATTATCTTTCCGCCTGGAGAGATAAGGAAAGATGGCTTTTGTACACTTACAGATGGGCCCGCGGCATTAGCTA
GTTGGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCAGACTCCT
ACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGCGAAGAAGGTTCTCGGATCGTAAAGCT
CTGTTGTTAGGGAAGAACAAGTACCGGAGGGGAATGCCGGTACCTTACGGTACCTGACGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGC
GGTAATACGTAGGTGGCMAGCGTTGTCCGGATTTATTGGGCGTAAAGCGCGCGCAGGCGGTCTTTAAGTCTGATGTGAAATCTTGGGCTCA
ACCGCAAGCGGTCATTGAAACTGGGGGACTTGAAGTGCAGAAGAGGAAAGCGGAATCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAG
GAACACCAAGTGGCGAAGGCGGCTTCTGGTCTGTAACCTGACGCTGAGCGCGAAAGCGTGGGAGCAAAACAGGATTAGATACCTCGTAGTCC
ACGCCGTAACGATGAGTGTAAAGTGTGGAGGGTTCCGCCCTTCAAGTGTGACGCTAACGCATTAAGCACTCCGCTGGGGAGTACGGTCCG
AAGACTGAAACTCAAAGGAATTGACGGGGGCCGACAAGCGGTGGAGCATGTGGTTAATTGAAAGCAACCGGAAGAACCTTACCAGGTCTT
GACATCTCTGACCGCCTGGAGACAGGGTCTTCCCTTCCGGGACAGGATGACAGGTGGTGCATGGTTGTCGTACGCTCGTGTGATGATGTT
GGGTTAAGTCCCGCAACGAGCGCAACCTTGGGTTCTAGTGGCAGCATTAGTTGGGCACTCTAGAGCGACTGCCGGCGACAAGTCGGAGGA
AGGTGGGGATGACGTCAAATCATATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGATGGTACAAAGGGCAGCGAAGCGCGCACGC
ATAAGCGAATCCAGAAAACCTTCTAGTTCGGATTGACAGGCTGCAACTCGCTGCATGAAGCCGGAATCGTAGTAATCGCGGATCAGCATG
CCGCGGTGAATACGTTCCCGGCTTGTACACACCGCCCTCACACCAGAGAG

>Bacillus_licheniformis_A02

GGAGCGGAGCGACGGGAGCTTGTCCCTTAGGTGACGGCGGACGGGTGAGTAACCGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCG
GGAAACCGGGGCTAATACCGGATGCTTGAATTGAACCGCATGTTCAATCATAAAAGGTGGCTTTTAGCTACCACTTACAGATGGACCCGCGCG
CATTAGCTAGTTGGTGAAGTAAACGGCTCACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCC
CAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGTGTGAAGGTTTTCGGATC
GTAAACTCTGTTGTTAGGGAAGAACAAGTACCGTTCGAATAGGGCGGTACCTTACGGTACCTAACCCAGAAAGCCACGGCTAACTACGTGCCA
GCAGCCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCAGGCGGTTTTAAGTCTGATGTGAAAGCC
CCGGCTCAACCGGGGAGGGTCAATTGAAACTGGGGAACCTTGAAGTGCAGAAGAGGAGAGTGGAAATCCACGTGTAGCGGTGAAATGCGTAGAG
ATGTGGAGGAACACCAAGTGGCGAAGGCGACTCTTGGTCTGTAACCTGACGCTGAGGCGCGAAAGCGTGGGAGCGAACAGGATTAGATACCC
TGGTAGTCCACGCGTAAACGATGAGTGTAAAGTGTAGAGGGTTTTCCGCCCTTAGTGTGACGAAACGCATTAAGCACTCCGCTGGGGGA
GTACGGTCCGCAAGACTGAACTCAAAGGAATTGACGGGGGCCGACAAGCGGTGGAGCATGTGGTTAATTGAAAGCAACCGGAAGAACCTT
ACCAGGTCTTACATCTCTGACAACCTAGAGATAGGGCTTCCCTTCCGGGGCAGAGTACAGGTGGTGCATGGTTGTCGTACGCTCGTGTG
GTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTATGCTTAGTTGACGCAATTCAGTTGGGCACTTAAGGTGACTGCCGGTGACAAA
CCGGAGGAAGGTGGGGATGACGTCAAATCATATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGCAGAACAAAGGGCAGCGAAGC
CGGAGGCTAAGCCAATCCACAAATCTGTTCTAGTTCGGATCGAGTCTGCAACTCGACTGCGTGAAGCTGGAATCGTAGTAATCGCGGATC
AGCATGCCGCGTGAATACGTTCCCGGCTTGTACACACCGCCCTCACACCAGAGAGTTTGAACACCCGAAAGTGC

>Bacillus_haynesii_A04

TTACAACTCTCGTGGTGTGACGGCGGTGTGTACAAGGCCGGGAACGATTACCGCGCATGCTGATCCGCGATTACTAGCGATTCCAGCT
TCACGCGATCGAGTTGACAGACTGCGATCCGAACTGAGAACAGATTTGTTGGGATTGGCTTAGCCTCGCGGCTTCCGCTGCCCTTTGTTCCGCAAT
GTAGCACGTGTAGCCAGGTCATAAGGGGCATGATGTTGACGTCATCCACCTTCTCCGGTTGTACCAGGCGATCACCTTAGAGTGGC
CAACTGAATGCTGGCAACTAAGTCAAGGGTTGCGCTCGTTGCGGGACTTAACCAACATCTCACGACAGGCTGACGACAACCATGACCCAC
CTGTACTCTGCCCGAAGGGGAAGCCCTATCTTAGGGTTGTCAGAGATGTCAAGACCTGGTAAGGTTCTTCCGCTTGTCTGCAATTAACCC
ACATCTCCCTTGTGCGGGCCCGCTCAATCTTGTAGTTGATTTGACGACCGTACTCCCGGCGGCTTAACTGCTGCA
GCACTAAAGGGCGGAAACCTTAACACTTAGCACTCATGTTTACGGCGTGGACTACCAGGGTATCTAATCTGTTGCTCCCGCAGCTTCCGCG
CCTCAGCGTCAGTTACAGACCAGAGAGTGCCTTCCGCACTGGTGTCTCCACATCTTACGCAATTCACCGCTACAGTGGAAATCCACTCTCT
CTTCTGCACTCAAGTTCCCGGTTTCCAATGACCTCCCGGTTGAGCGGGGGCTTACATCAGACTTAAGAAACCGCTGCGCGCGCTTACG
CCCAATAATCCGGACAACGCTGGCCACCTACGATTACCGCGGCTGCTGGCACGT

>Aeribacillus_compostii_A06

TGCTCCTTAGGTTAGCGGCGGACGGGTGAGTAACCGTGGGCAACCTGCCCTGCAGACGGGATAACTTCGGGAAACCGGAGCTAATACCGG
ATAACACCGAAAACCGCAGGGTTTTCGGTTGAAAGGCGGCTTTAGCTGTCACTGACGGATGGGCCCGCGCGCATTAGCTAGTTGGTGAGGTA
ACGGCTCACCAAGGCGACAATGCGTAGCCAACTGAGAGGGTGACCGGCCACACTGGGACTGAGACACGGCCAAACTCCTACGGGAGGCAGC
AGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGCGAAGAAGTCTTCGGATCGTAAAGCTCTGTTGTCAGGGA
AGAACAAGTCCGTTGAAACAGGGCGGTACCTTACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGGTAATACGTAG
GTGGCAAGCGTTGTCGGGAATATTGGGCGTAAAGCGCGCAGGCGGTTCTTAAGTCTGATGTGAAATCTCGCGCTCAACCGGAGCCGG
CCATTGAAACTGGGGAACCTTGAAGTGCAGGARAGGGGAGCGGAATTCACGTGTAGCGGTGAAATGCGTARAGATGTGGAGGAACACCAAGT
GCGAAGGCGGCTCTGCGCTGTAACGACGCTGAGCGCGAAAGCGTGGGAGCGAACAGGATTAGATACCTGGTAGTCCACGCGCTAAAC
GATGAGTGTAAAGTGTAGAGGGTATCCACCTTATGCTGACGCAACGCATTAAGCACTCCGCTGGGGAGTACGGCCGCAAGGCTGAA
ACTCAAAGGAATTGACGGGGACCCGACAAGCGGTGGARCCATGTGTTTAACTCGAAGCAACCGGAAGAACCTTACCAGGCTTGTGACATCCC
TGACAACCTAGAAGATAGGGGCTTCCCTTTCCGGGACAGGGTACAGGTGGGTGCATGGTTGTCGTACSTCGTGTGATGATGTTGGGTT
AAGTCCCGCAACGAGCGCAACCTTACCTTATGTTCCAGCATTAGTTGGGCACTTAAGGTGACTGCCGGTAAAGTCGGAGGAAGGTTGG
GGATGACGTCAAATCATATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGTGGTACAAAGGGCAGCGAAACCGCGAGGTGGAGCG

AATCCAAAAAACCACTCTCAGTTCGGATTGCAGGCTGCAACTCGCTGCATGAAGCCGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGT
GAATACGTTCCCGGGTCTGTACACACCCGCCGTACACCACGAGAGTTTGTAAACCC

>Bacillus_licheniformis_B01

CTTGCTCCCTTAGGTACGCGGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACC
GGATGCTTGAATGAACCGCATGGTTCAATCATAAAAGGTGGCTTTCAGCTACCACTTACAGATGGACCCGCGGCATTAGCTAGTTGGTGAGG
TAACGGCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTATCGGCCACTGGACTGAGACACGCCCCAGACTCCTACGGGAGGC
AGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGTGATGAAGGTTTTCCGATCGTAAAACCTCTGTTGTTAGG
GAAGAACAAGTACCGTTCGAATAGGGCGGTACCTTGACGGTACCTAACCGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAAATACGT
AGGTGGCMAGCGTTGTCGGGAATTATTGGGCGTAAAGCGCGCCAGGCGGTTTCTTAAGTCTGATGTAAAGCCCCCGCTCAACCGGGGAG
GGTCATTGAAACTGGGGAACCTGAGTGCAGAAGAGGAGAGTGAATTCACGTTGAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCAG
TGGCGAAGGCGACTCTGCTGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCTGGTAGTCCACGCCGTAA
ACGATGAGTGCTAAGTGTAGAGGGTTCCGCCCTTAGTGTGACGCAAAACGCATTAAGCACTCCGCTGGGGGAGTACGGTGCAGAACTGA
AACTCAAAGGAATTGACGGGGGCCGCAACGCGGTGGARCCATGTGGTTAATTGAAAGCAACGCAAGAACCTTACCAGGTCTTGACATCCT
CTGACAACCTAGAGATAGGGCTTCCCTTCCGGGGCAGAGTGACAGGTGGTTGCATGGTTGCTGAGCTGCTGCTGAGATGTTGGGTTA
AGTCCCGCAACGAGCGCAACCTTATGACCTGGGCTACACACGTGCTACAATGGGCGAGAACAAGGGCAGCGAAGCCGCGAGGCTAAGCCA
ATCCACAATCTGTTCTCAGTTCGGATCGCAGTGTCAAACCTGACGCTGCGTGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGA
ATACGTTCCCGGGCTTGTACACACCCGCCGTACACCACGAGAGTTTGTAAACCCCGGAAGTCCGGTGGAGT

>Bacillus_smithii_B03

AAGAAGCTTGCTTTTTGAAAGTTAGCGGGACGGGTGAGTAACACGTGGGCAACCTGCCTGCAAGACGGGGATAACTCCGGGAAACCGGGG
CTAATACCGGATAATATCTTCTCGCATGAAGGAAGTTGAAAGGCGCGCAAGCTGCCGCTGCAAGTGGGCCCGCGGCATTAGCTAGTT
GGTGAGGTAACGGCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTATCGGCCACTGGACTGAGACACGGCCAGACTCCTAC
GGGAGGCGAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGCGAAGAAGGTTCTCGGATCGTAAAGCTCT
GTTGTGAGGGAAGAACAAGTACCGTTCGAACAGGGCGGTACCTTGACGGTACCTGACCAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCG
GTAATACGTAGGTGGCMAGCGTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCMGGCGGCTCTTAAGTCTGATGTAAAGCCACCGGCTCAA
CCGTGGAGGGTCAATGGAAACTGGGAGACTTGAGTGCAGAAGAGGAGAGCGGAATTCACGTTGAGCGGTGAAATGCGTAGAGATGTGGAGG
AACACCAGTGGCGAAGGCGGCTCTGCTGTAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCTGGTAGTCCA
CGCCGTAACAGATGAGTCTAAGTGTAGAGGGTTCACCCCTTAGTGTGACGTAACGCATTAAGCACTCCGCTGGGGGAGTACGGCCGC
AAGGCTGAAACTCAAAGGAATTGACGGGGGCCGCAACGCGGTGGARCCAKGTGGTTAATTGAAAGCAACCGAAGAACCTTACCAGGTCT
TGACATCTTCTGACTCTAGAGGATAGAGGGTCCCCTTCCGGGGACGGAKTGACAGGTGGTTGCATGGTTGCTGCTGAGTCTGCTGAGT
ATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTACCTTAGTTGCCAGATTGAGTGGGCACTTAAGGTGACTGCCGGTACAAAACCGGA
GGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGATGGTACAAAAGGTCGCGAAACCGCGA
GGTGGAGCAATCCCAAAAACCATCTCAGTTCGGATTGACAGGCTGCAACTGCCTGCATGAAGCCGGAATCGCTAGTAATCGCGGATCAGCA
TGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCCGCCGTACACCACGAGAGTTTGCAACACCCGAAGTCCGGT

>Bacillus_subtilis_B04

GCAAGTTCGAGCGGACAGATGGGAGCTGTCCCTGATGTTAGCGGCGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAAC
TCCGGGAAACCGGGGCTAATACCGGATGCTTGTGTTGAAACCGCATGGTTCAAACATAAAAGGTGGCTTCGGCTACCACTTACAGATGGACCCGCG
GCGCATTAGCTAGTTGGTGTAGGTAATGGCTACCAAGGCAACGATGCGTAGCCGACCTGAGAGGGTATCGGCCACTGGGACTGAGACACG
GCCAGACTCCTACGGGAGGCGAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGTGATGAAGGTTTTCGG
ATCGTAAAGCTCTGTTGTAGGGAAGAACAAGTACCGTTCGAATAGGGCGGTACCTTGACGGTACCTAACCGAAGCCACGGCTAACTACGTG
CCAGCAGCCGCGTAAATACGTAGGTGCGAAGCGTTGTCGGGAATTATTGGGCGTAAAGGGCTCGCAGGGTTCCTTAAGTCTGATGTGAAAG
CCCCGGCTCAACCGGGGAGGGTCAATGGAAACTGGGGAACCTGAGTGCAGAAGAGGAGAGTGAATTCACGTTGAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACAGTGGCGAAGGCGACTCTGCTGTTGTAAGTGTGAGGCGTGAAGGAGCGAAGCGTGGGGAGCGAACAGGATTAGATA
CCCTGGTAGTCCACGCCGTAACGATGAGTGTAAAGTGTAAAGGGGTTTTCCGCCCTTAGTGTGCTGAGCTAACGCATTAAGCACTCCGCTGG
GGGAGTACGGTCCGAAGACTGAAACTCAAAGGAATTGACGGGGGCCGCAACGCGGTGGARCCATGTGGTTAATTCGAAGCAACCGGAAG
AACCTTACCAGGTCCTGACATCCTGACAATCCTAGAAGATAGGACGCTCCCTTCCGGGGCAGAGTGAACAGGTGGTTGATGGTTGCTGCTCAG
CTGCTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTATGCTTAAAGTGGCAGCATTAGTTGGGCACTTAAGGTGACTGCCG
GTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGACAGAACAAGGGC
AGCGAAACCGCGAGGTTAAGCAATCCCAAAATCTGTTCTCAGTTCGGATCGCAGTGTCAAACCTGACTGCGTGAAGCTGGAATCGCTAGTAAT
CGCGGATCAGCATGCCGCGTGAATACGTTCCCGGGCCTTGTACACACCCGCCGTACACCACGAGAGTTTGTAAACCCCGAAGTCCGGTGAAGT
AACA

>Anoxybacillus_kamchatkensis_B05

TCGTTAGCGGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCCTGTAGACGGGGATAACACCGAGAAATCGGTGCTAATACCGGATAATACGA
AAGGCCCGCATGGTCTTTCGTTGAAAGGCGCGCAAGCTGTCGCTACAGGATGGGCCCGCGGCATTAGCTAGTTGGTGAGGTAACGGCTCAC
CAAGGCGACGATGCTAGCCGACCTGAGAGGGTATCGGCCACTGGGACTGAGACACGGCCAGACTCCTACGGGAGGCGAGCAGTAGGGA
ATCTTCCGAATGGACGAAAGTCTGACGGAGCAACCGCGGAGGAGGCGCTTCGRGTGTAAGCTGTTGTTAGGGAAGAACAAG
TACCGGACGTACKGGCGGTACCTGACGGTACCTAACGAGGAAGCCACGGCTAAMTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAG
CGTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGTCTTTAAGTCTGATGTAAAGCCACGGCTCAACCGTGGAGGGTCAATGGAAA
CTGGGGGACTTGAAGTGCAGAAGAGGAGAGCGGAATTCACGTTGAGCGGTGAAATGCGTAGAGATGTGGAGGAACCAAGTGGCGAAGGC
GGCTCTGCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGGAGCAACAGGATTAGATACCTGGTAGTCCACGCCGTAACAGTGAAGT

CTAAGTGTAGAGGGTATCCACCCTTTAGTGCTGTAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGCTCGCAAGAGTAACTCAAAGGA
ATTGACGGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACCGGAAGAACCTTACCAGGTCTTGACATCCCCTGACAAACCGCA
GAGATCGGGCTTCCCTTTCGGGGGACAGGGTGACAGGTGGTGCATGGTGTGCTCAGCTCGTGTGAGATGTTGGTTAAGTCCCAGCA
ACGAGCGCAACCTCGACCTTAGTTGCCAGCATTAGTTGGGCACTTAAGGTGACTGCCGGCTAAAAGTCGGAGGAAGGTGGGGATGACGTC
AAATCATCATGCCCTTATGACCTGGGCTACACAGTGTACAATGGCGGTACAAAGGTGCGGAACCCGCGAGGGGGAGCCAATCCAAAAA
AGCCGCTCAGTTCGATTGACGGCTGCAACTCGCTGCATGAAGCCGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTT
CCGGCCCTGTACAACACCCGCCCTCACACCAG

>Bacillus_hisashii_C01

TTGGTTAGCGGCAACCGGTGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGATTA
TCTTTCGCTGGAGAGATAAGGAAAAGATGGCTTTTGTATCATTACAGATGGGCCCGCGGCATTAGCTAGTTGGTGGAGTAACGGCTCAC
CAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGCCACACTGGGACTGAGACACGGCCAGACTCTACGGGAGGACGAGTAGGGA
ATCTTCCGCAATGACGAAAGTCTGACGGAGCAACCGCGTGGAGCGAAGAAGTCTTGGATCGTAAAGCTGTTGTTAGGGAAGAACAAG
TACCGGAGGGGATGCCGGTACCTTACGGTACCTGACGAGAAAGCCAGGGTAACTACGTGCCAGCAGCCGGTAAATACGTAGGTGGCAAGC
GTTGTCCGATTTATTGGGCGTAAAGCGCGCGCAGGCGGTCTTTAAGTCTGATGTGAAATCTTCCGGCTCAACCGCAAGCGGTATTGGAAAC
TGGGGGACTTGAGTGCAGAAAGAGGAAAGCGGAATCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCAAGTGGGGAAGCGG
CTTTCTGGTCTGTAACCTGACGCTGAGGGCGGAAAGCGTGGGGAGCAACAGGATTAGATACCCTGGTAGTCCACGCGTAAACGATGAGTGT
AGTGTGGAGGGTTCCGCCCTTCAAGTGTGCTGACGCTAACGCATTAAGCACTCCGCTGGGGAGTACGGTGCAGAACTGAAACTCAAAGGAAT
TGACGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACCGGAAGAACCTTACCAGGTCTTGACATCTCTGACCGCCCTGGA
GACAGGGTCTTCCCTTCCGGGACAGGATGACAGGTGGTGCATGTTGTGCTCAGCTCGTGTGAGATGTTGGTTAAGTCCCAGCAACGAGC
GCAACCTTGGTTCTAGTTGCCAGCATTAGTTGGGCACTTAGAGCGACTGCCGGCGACAAGTCGGAGGAAGGTGGGGATGACGTCAAATCAT
CATGCCCTTATGACCTGGGCTACACAGTGTACAATGGATGGTACAAGGGCAGCGAAGCGCGCACGCATAAGCGAATCCAGAAAACCTT
CTCAGTTCGATTGACGGCTGCAACTCGCTGCATGAAGCCGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCC
TTGTACACACCCGCTCACACCAGAGATTTGTAACCCCTAAGTCGGTGGAGTA

>Aeribacillus_compostii_C02

CAGTCGAGCGGACCGAAGGAGCTTGTCTTTAGTGTAGCGCGGACGGGTGAGTAACACGTGGGCAACCTGCCTGCAGACTGGGATAACT
TCGGGAAACCGGAGCTAATACCGGATAAACACCGAAAACCGCATGGTTTTCGGTTGAAAGCGGCTTTTAGCTGTCACTGCAGGATGGGCCCGC
GCGCATTAGCTAGTTGGTGGAGTAACGGCTCACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACTGGGACTGAGACAG
GCCAGACTCCTACGGGAGGCGAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCCGTGGAGCAAGAAGGTCTCGG
ATCGTAAAGCTCTGTTGTCAGGGAAGAACAAGTCCGCTTGAACAGGGCGGTACCTTACGGTACCTGACGAGGAAGCCACGGTAACTACGT
GCCAGCAGCCGGTAAATACGTAGGTGGCMAGCGTTGTCCGGAATTTGGGCGTAAAGCGCGCGCAGGCGGTTCCCTAAGTCTGATGTGAAA
TCTCGCGGCTCAACCGCGAGCGGCCATTGGAAACTGGGGAATTTAGTGTGACGAGARAGGGGAGCGGAATTCACCGTGTAGCGGTGAAATGCGT
ARAGATGTGGAGGAACACAGTGGCGAAGGCGGCTCTYTGCCCTGTAAGTACGCGSTRAGGCGGAAAGCGTGGGGAGCGAAGCAGGATTAGAT
ACCCTGGTAGTCCAGCCGTAACAGATGAGTGTAAAGTGTAGAGGGTATCCACCTTTAGTGTGCTGACGAAACGCATTAAGCACTCCGCKGG
GGAGTACGGCGCAAGGCTGAAACTCAAAGGAATTGACGGGACCCGCAACAGCGGTGGAGCATGTGGTTAATTCGAAGCAACCGCAAGAA
CCTTACCAGGTCTTGACATCCCCTGACAACCTAGAGTAGGGCGTTCCTTTCGGGARAMACSSYSWCAKKTGRATKYAKGKGGWMRCT
CMSCYSTARWCSYGARAYSYTGTTAAKTCCSGAAMAACRAGCGCAACCCCTGACMWWWARTTKCCMGSATCAKTKKRCAMYCYAMS
GACTSCRCWAAAARTYKGAKKAASGTGGGGATGACGTCAAATCATCATGCCCCCTTATGACCTGGGCTACACACGTGCTACAATGGGTGGTACA
AAGGGCAGCGAAACCGGAGGTGGAGCGAATCCAAAAACCACTCTCAGTTCGATTGACGGCTGCAACTCGCTGCATGAAGCCGGAATCG
CTAGTAATCGCGATCAGCATGCCGCGGTGAATACGTTCCCGGGTCTGTACACACCCCGCTCACACCAGAGAGTTGTAACACCCGAAAGTGC
GTGGGGTAACC

>Bacillus_hisashii_C03

TGCAGTCGAGCGAACCAATAAGAAGCTTGTCTTTTGGTGTAGCGCGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATA
ACTCCGGGAAACCGGTGCTAATACCGGATAGATTATCTTCCGCTGGAGAGATAAGGAAAGATGGCTTTTGTATCATTACAGATGGGCCCG
CGGCGCATTAGCTAGTTGGTGGAGTAACGGCTCACCAAGGCGACGATGCGTAGCCAACCTGAGAGGGTGATCGCCACACTGGGACTGAGACA
CGGCCAGACTCCTACGGGAGGCGAGCAGTAGGGAATCTCCGCAATGGACGAAAAGTCTGACGGAGCAACGCCCGTGGAGCAAGAAGGTCTTC
GGATCGTAAAGCTGTTGTTAGGGAAGAACAAGTACCGGAGGGAATGCCGTTACCTTACCGGTACCTGACGAGAAAGCCACGGTAAATAC
GTGCCAGCAGCCGCTAATACGTAGGTGGCMAGCGTTGTCCGATTTATTGGGCGTAAAGCGCGCGCAGGCGGCTCTTAAAGTCTGATGTGA
AATCTTCCGCTCAACCGCAAGCGGTCAATTGGAAACTGGGGACTTGAAGTGCAGAAAGGAGARAGCGGAATTCACGTGTAGCGGTGAAATGCG
TAGAGATGTGGAGGAACACAGTGGCGAAGCGGCTTTCTGGTCTGTAAGTACGCTGAGGCGGAAAGCGTGGGGAGCAACAGGATTAGA
TACCTGGTAGTCCAGCCGTAACAGATGAGTGTAAAGTGTGGGAGGGTTCCGCCCTTCAAGTGTGACGTAACGCATTAAGCACTCCGCCTG
GGGAGTACGGTGCAGACTGAAACTCAAAGGAATTGACGGGGCCCGCAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACCGGAAG
AACCTTACCAGGTCTTGACATCTCTGACCGCCCTGGAGACAGGGTCTTCCCTTCCGGGACAGGATGACAGGTGGTGCATGTTGTGCTCAGCTC
GTGCTGAGATGTTGGTTAAGTCCCAGCAGCGCAACCTTGGTTCTAGTGTCCAGCATTAGTTGGGCACTTAGAGCGACTGCCGGC
GACAAGTTCGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGATGGTACAAGGGCAG
CGAAGCGGCGACGATAAGCGAATCCAGAAAACCTTCTCAGTTCGATTGACGGCTGCAACTCGCTGCATGAAGCCGGAATCGCTAGTAAT
CGCGGATCAGCATGCCGCGTGAATACGTTCCCGGGCTGTACACACCCCGCTCACACCAGAGAGTTGTAACACCCGAAAGTGGTGGGT
AACCGTAAGGAGCCAGCCGCC

>Pseudoxanthomonas_taiwanensis_C04

CGGACGCGGGTAGGAAAGCTTGTCTTCTATGCCGGCAGTGGCGGACGGGTGAGGAGCGCATCGGAATCTACTCCGTCGTGGGGGATAACGTA
GGGAACTTACGCTAATACCGCATAACGCTAAGGGTGAAGTGGGGGACCGCAAGCCCTCACCGATGGAATGAGCCGATGTCCGATTAGCT

AGTTGGCGGGTAAAGGCCACCAAGGCGACGATCGGTAGTGGTCTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCAGACT
CTACGGGAGGCGAGCAGTGGGAATATTGGACAATGGCGAAAGCCTGATCCAGCMTGCCGCTGGGTGAAGAAGCCCTCGGGTTGTAAG
CCCTTTGTTGGGGAAGAAATCTGCTGGCTAATACCCGGCGGGGATGACGGTACCCAAAGAATAAGCACCAGGCTAATTCGTGCCAGCAGCCG
CGGTAATACGAAGGGTGAAGCGTTACTCGGAATTAAGTGGCGTAAAGCGTGCCTAGGTGGTAAAGTCCGTTGTGAAAGCCCTGGGCTC
AACCTGGGAATTGCACTGGTACTGGGTACTAGAGTGTGGTAGAGGGTGGCGGAATCCCGGTGTAGCAGTGAATGCGTAGAGATCGGGA
GGAACACCCGTGGCGAAGGCGGCCACCTGGCCCAACTGACACTGAGGCACGAAAGCGTGGGAGCAACAGGATTAGATACCCCTGGTGT
CCACGCCCTAAACGATGCGAACTGGATGTTGGGTTCAATTTGGGACTCAGTATCGAAGCTAACGCGTTAAGTTCGCCGCTGGGGGAGTACGGT
CGAAGACTGAACTCAAAGGAATTGACGGGGGCCGCACAAGCGGTGGAAKTAKGTGGGTTAATTCGATGCAACCGGAAGAACCTTACCTG
GTCTTGACATCCGCGAACTGCCAGAARATGGCGGGTGCCTTCGGGAGCCGCGAGACAGGTGCTGCATGGCTGCTGACGCTCGTGTGCTG
AGATGTTGGGTTAAGTCCGCAACGAGCGCAACCCCTTGCTTAGTGGCAGCAGTAATGGTGGAACTCTAAGGAGACCAGCGGTGACAAA
CCGGAGGAAGGTGGGATGACGTCAAGTCAATGACGCTTACGACAGGGCTACACAGTACTACAATGGTGGGACAGAGGGTGCAGAAC
CCGCGAGGGGAGCCAAATCCAGAAACCCCTCTCAGTCCGGATTGGAGTCTGCAACTCGACTCCATGAAGTCCGAATCGCTAGTAATCGCAGA
TCAGCATTGCTGCGGTGAATACGTTCCCGGCCCTGTACACACCCCGCTCACACCATGGGAGTTGGTTGCACC

>Aeribacillus_composti_C05

GAGCTTGCTCTTTAGGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCCTGCAGACTGGGATAACTTCGGGAAACCGGAGCTAAT
ACCGGATAACACCGAAACCGCAGGGTTTTCGGTTGAAAGCGGCTTTAGCTGTACTGCAGGATGGGCCCGCGCGCATTAGCTAGTTGGTG
AGGTAACGGCTCACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACTGGGACTGAGACACGGCCAGACTCCTACGGGA
GGCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCCGTGAAGCAAGAAGGTCTTCGGATCGTAAAGCTCTGTTG
TCAGGGAAGAACAAGTRCCGTTGCAACAGGGCGGTACCTTGACGGTACTGACGAGGAAGCCACGGCTAATCAGTGCAGCAGCCGCGGTAA
TACGTAGGTGGCMAGCGTTGTCCGGAATTATTGGCGTAAAGCGCGCGCAGGGCTTCTTAAGTCTGATGTGAAATCTCGCGCTCAACCG
GAGCGCCATTGGAACTGRGGAACCTTGGTGCAGGARAGGGGAGGGAATCCACGTGTAGCGGTGAAATGCGTARAGATGTGGAGGAACA
CCAGTGGCGAAGGGCTCTGCGCTGTAACGCTGACGCTRAGGCGCGAAAGCGTGGGAGCGAAGCAGGATTAGTACCTCGTGGTAGCCAGCC
GTAACGATGAGTCTAAGTGTAGAGGGTATCCACCCTTAGTGTGACGAAACGCATTAAGCACTCCGCTGGGAGTACGGCCGCAAGGC
TGAAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCCTTACCAGGTCTGACAT
CCCCTGACAACCTAGAGATAGGGCGTTCCTTTCGGGGACAGGGTGCAGGTGGTGCATGGTTGTCGTCAGCTCGTGTGAGATGTTGGG
TTAAGTCCCGCAACGAGCGCAACCCCTTACCTTAGTGGCAGCATTAGTGGGCACTAAGGTGACTGCCGCTAAAAGTCCGAGGAAGGTG
GGGATGACGTCAAATCATCATGCCCTTATGACCTGGGTACACACGTGCTACAATGGGTGGTACAAGGGCAGCGAAACCGCGAGGTGGAGC
GAATCCCAAAAACCACTCTCAGTTCGGATTGACGGTGCACACTGCCTGCATGAAGCCGGAATCGCTAGTAATCGCGGATCAGCATGCCGCG
TGAATACGTTCCCGGGTCTGTACACACCCCGCTCACACCAGAGAGTTGTAACACCCGAAAGTCGGTGT

>Bacillus_hisashii_C06

GATAAGAAGCTTGCTTTTTGTTGTTAGCGGCGAACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGGAAACCGGT
GCTAATACCGGATAGATTATCTTCCGCTGGAGAGATAAGGAAAGATGGCTTTGTATCACTTACAGATGGGCCCGCGCGCATTAGCTAGTT
GGTGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGACTGAGAGGGTGTGCGCCACTGGGACTGAGACACGGCCAGACTCCTAC
GGGAGGAGCAGTAGGGAATCTCCGCAATGACGAAAGTCTGACGAGCAACGCCGCTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCT
GTTGTTAGGGAAGAACAAGTACCGGAGGGAATGCCGTTACCTGACGGTACTGACGAGAAAGCCACGGCTAATCAGTGCAGCAGCCGCG
TAATACGTAGGTGGCMAGCCTTGTCCGATTTATTGGGCTAAAGCGCGCAGGGCTCTTTAAGTCTGATGTGAAATCTTCGGCTCAACC
GCAAGCGGTCAATGGAACTGGGGACTTGAAGTGCAGAAGAGGAAAGCGGAATCCAGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAA
CACCAGTGGCGAAGGCGCTTCTGCTGTGTAACGACGCTGAGGCGGAAAGCGTGGGAGCAACAGGATTAGATACCCTGGTAGTCCAGC
CCGTAAACGATGAGTGTAAAGTGTGGGAGGTTTTCCGCCCTTCAAGTGTGACGCTAACGCATTAAGCACTCCGCTGGGGAGTACGGTGC
AGACTGAACTCAAAGGAATTGACGGGGGCCGCACAAGCGGTGGARCCAKGTGGTTAATTCGAAGCAACCGGAAGAACCCTTACCAGGTCT
GACATCTCTGACCCCTGGAAGACAGGGTCTTCCCTTCGGGGACAGGATGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTGAGATGT
TGGTTAAGTCCCGCAACGAGCGCAACCCCTTGGTTCTAGTTGCCAGCATTAGTTGGGCACTTACAGCGACTCCGCGCAACAAGTCCGAGGAA
GGTGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGTACACACGTGCTACAATGGATGGTACAAGGGCAGCGAAGCGCGACGCA
TAAGCGAATCCAGAAAACCTTCTCAGTTCGGATTGACGGTGCACACTGCCTGCATGAAGCCGGAATCGCTAGTAATCGCGGATCAGCATGC
CGCGTGAATACGTTCCCGGGCTGTACAACACCCCGCTCACACCAGAGAGTTGTAACACCCGAAAGTCGGTGT

>Geobacillus_kaustophilus_D01

GGAGCTTGCTGTTGGTTCAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCGCAAGACCGGGATAACTCCGGAAACCGGAGCTA
ATACCGGATAACACCGAAGACCGCATGGTCTTTGGTTGAAAGCGGCTTTGGCTGTACTTCCGATGGGCCCGCGCGCAATAGCTAGTTGG
TGAGGTAACGGCTCACCAAGGCGACGATGCGTAGCCGCTGAGAGGGTACCGGCCACTGGGACTGAGACACGGCCAGACTCCTACGG
GAGGACGAGTAGGGAATCTCCGCAATGGGCGAAAGCCTGACGGAGCGACGCCGCTGAGCGAAGAAGGCCTTCGGGTGTAAGTCTGT
TGTGAGGGAGCAAGGAGCGCCGTTGAAAGGGCGGCGGTTGACGGTACTCAGGAGGAAGCCCGGCTAATCAGTGCAGCAGCCGCG
TAATACGTAGGGGCGAGCGTTGTCGGAATTTATTGGGCTAAAGCGCGCGCAGGGCTTCTTAAGTCTGATGTGAAAGCCACGGCTCAAC
CGTGGAGGGTATTGAAACTGGGGACTTGAAGTGCAGGAGAGGAGCGGAATCCAGTGTAGCGGTGAAATGCGTAGAGATGTGGAGG
AACACAGTGGCGAAGGCGGCTCTGCGCTGCAACTGACGCTGAGGCGCGAAGCGTGGGAGCAACAGGATTAGATACCCTGGTAGTCCA
CGCCGTAACGATGAGTGTAAAGTGTAGAGGGTGCACCCCTTAGTGTGACGCTAACCGGATAAGCACTCCGCTGGGGAGTACGGCCG
CAAGGCTGAACTCAAAGGAATTGACGGGGGCCGCACAAGCGGTGGARCCAKGTGGTTAATTCGAAGCAACCGGAAGAACCCTTACCAGGT
TTGACATCCCCTGACAACCAAGAAGATTGGGCTTCCCTTCGGGGGACAGGGTACAGGTGGTTGATGGTTGTCGTCAGCTCGTGTCT
GAGAWGTTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTGCCTCTGTTGCCAGCACGAAGGTGGGCACTTACAGGGACTGCCGGGACAA
GTCGGAGGAAGTGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGTACACACGTGCTACAATGGGCGGTACAAGGGCTGCGAAC
CCGCGAGGGGAGCGAATCCAAAAAGCCGCTCTCAGTTCGGATTGACGGTGCACACTCGCTGCATGAAGCCGGAATCGCTAGTAATCGCGG
ATCAGCATGCCGCGTGAATACGTTCCCGGGCTGTACACACCCCGCTCACACCAGAGAGTTGTAACACCCGAAAGTCGGTGT

>Aeribacillus_composti_D02

GCTTGCTCCTTTAGGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCCTGCAGACTGGGATAACTTCGGGAAACCGGAGCTAATACC
GGATAACACCCGAAAACCGCAGGGTTTTTCGGTTGAAAGGCGGGCTTTTAGCTGTCACTGCAGGATGGGCCCGCGGCATTAGCTAGTTGGTGAG
GTAACGGCTCACCAGGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACTGGGACTGAGACACGGCCAGACTCTACGGGAGG
CAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCSCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTGTCA
GGGAAGAACAAGTRCCGTTTCGAACAGGGCGGTACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATAC
GTAGGTGGCMAGCGTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAGTCTGATGTGAAATCTCGGGCTCAACCGCGAG
CGGCCATTGAAACTGRGGAAGTGTAGTGCAGGARAGGGGAGCGGAATCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCA
GTGGCGAAGGCGGCTCTCTGGCCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCTGGTAGTCCACGCCGTA
AACGATGAGTGTAAAGTGTAGAGGGTATCCACCCTTTAGTGTGACGCAAAACGCATTAAGCACTCCGCTGGGGGAGTACGGCCGCAAGGCTG
AAACTCAAAGGAATTGACGGGGACCCGCACAAGCGGTGARCCATGTGGTTTTAATTGAAAGCAACCGGAAGAACCTTACCAGGTCTTGACATC
CCCTGACAACCTAGAAGATAGGGCGTTCCTCTTCGGGGACAGGGTACAGGTGGGTGCATGGTTGTCGTASCTCGTGTCTGARAATKTTG
GGTTAAGTCCCGCAACGAGCGCAACCTTACCTTAGTTGCCAGCATTAGTTGGGCACTTAAGGTGACTGCCGGTAAAAGTGGAGGAAAGG
TGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGTACAATGGGTGGTACAAGGGCAGCGAAACCGCGAGGTGGA
GCGAATCCAAAAAACACTCTCAGTTCGGATTGACAGGCTGCAACTCGCTGCATGAAGCCGGAATCGTAGTAATCGCGGATCAGCATGCCGC
GGTGAATACGTTCCCGGTTCTGTACACACCGCCCGTACACCACGAGAGTTTGTAAACCCC

>Anoxybacillus_kamchatkensis_D03

GCTAATACCGGATAATACGAAAGGCCGCATGGTCTTTCGTTGAAAGGCGCGCAAGCTGTGCTACAGGATGGGCCCGCGCGCATTAGCTAG
TTGGTGAGGTAACGGCTCACCAGGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACTGGGACTGAGACACGGCCAGACTCTCA
CGGGAGGCGAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCCGTGTAGCGAAGAAGCCCTTCGGGTCTGTAAGCTC
TGTGTTAGGGAAAGCAAGTACCGCAGTCACTGGCGGTACCTTGACGGTACCTAACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGG
GTAATACGTAGGTGGCAAGCGTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAGTCTGATGTGAAAGCCACGGCTCAAC
CGTGGAGGGTCATTGAAACTGGGGGACTTGAAGTGCAGAAGAGGAGCGGAATCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGG
AACACAGTGGCGAAGGCGGCTCTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCTGGTAGTCCA
CGCCGTAACGATGAGTGMATAAGTGTAGAGGGTATCCACCCTTTAGTGTGAGTAAACGCATTAAGCACTCCGCTGGGGAGTACGCTCGCA
AGAGTGAACCTCAAAGGAATTGACGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTTAATTGAAAGCAACCGGAAGAACCTTACCAGGTCTTG
ACATCCCCTGACAACCGAGAGATCGGGGCTTCCCTTCGGGGGACAGGGTACAGGTGGTGCATGGTTGTCGTAGCTCGTGTGAGAGAGA
TGTTGGGTTAAGTCCCGCAACGAGCGCAACCTCGACTTTAGTTCGACGATTCAGTTGGGCACTTTAAGGTGACTCCCGGCTAAAAGTGGGAG
GAAGGGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGTACAATGGGCGGTACAAGGGTCCGGAACCCGCGAG
GGGGA

>Bacillus_licheniformis_D04

CAGTCGAGCGGACCGACGGGAGCTTGCTCCCTTAGTGCAGCGGCGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACT
CCGGGAAACCGGGGTAATACCGGATGCTTGATTGAACCGCATGGTTCAATCATAAAAGGTGGCTTTAGCTACCACCTACAGATGGACCCGCG
GCGCATTAGCTAGTTGGTGAGGTAACGGCTCACCAGGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACTGGGACTGAGACACG
GCCAGACTCTACGGGAGGCGAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCCGCTGAGTGTAGAGTTTCGG
ATCGTAAAACCTGTTGTTAGGGAAAGCAAGTACCGTTCGAATAGGGCGGTACCTTGACGGTACCTAACAGAAAGCCACGGCTAACTACGTG
CCAGCAGCCGCGTAATACGTAGGTGGCAAGCGTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTTCTTAAGTCTGATGTGAAAG
CCCCCGGCTCAACCGGGGAGGGTCATTGAAACTGGGGAACCTTGAAGTGCAGAAGAGGAGAGTGGAAATCCACGTGTAGCGGTGAAATGCGTA
GAGATGTGGAGGAACACCAAGTGGCGAAGGCGACTCTCTGGTCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGGAGCGAACAGGATTAGATA
CCCTGGTAGTCCACCGGTAACGATGAGTGTAAAGTGTAGAGGGTTTCCGCCCTTATGCTGCAGCAAACGCATTAAGCACTCCGCTGGG
GGAGTACGGTCGCAAGACTGAAACTCAAAGGAATTGACGGGGCCCGCACAAGCGGTGARCCATGTGGTTTTAATTGAAAGCAACCGGAAGA
ACCTTACCAGGTCTTGACATCCTGACAACCTAGAGATAGGGCTTCCCTTCGGGGGACAGGTGACAGGTGGTGCATGGTTGTCGTAGCTC
GTGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTATGCTTATGTTGCCAGCATTAGTTGGGCACTTAAGGTGACTCCCGGTG
ACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGTACAATGGGCGAACAAGGGGACG
GAAGCCGCGAGGCTAAGCCAATCCCAAACTCTGTTCTCAGTTCGGATCGCAGTCTGCAACTCGACTGCGTGAAGCTGGAATCGTAGTAATCG
CGGATCAGCATGCCGCGGTGAATACGTTCCCGGGCTGTACACACCGCCCGTACACCACGAGAGTTTGTAAACCCC

>Bacillus_subtilis_D05

GATGGGAGCTTGCTCCCTGATGTTAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGG
CTAATACCGGATGGTTGTTGAACCGCATGGTTCAAACATAAAAAGGTGGCTTCGGCTACCCTTACAGATGGACCCGCGCGCATTAGCTAGTTG
GTGAGGTAACGGCTCACCAGGCAACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACTGGGACTGAGACACGGCCAGACTCTACGG
GAGGACAGTAGGGAAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCKTGAGTGTGAAAGGTTTTTCGGATCGTAAAGCTCTGTT
GTTAGGGAAGAACAAGTACCGTTCGAATAGGGCGGTACCTTGACGGTACCTAACAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTA
ATACGTAGGTGGCAAGCGTTGTCGGGAATTATTGGGCGTAAAGGGCTCGCAGGCGGTTTCTTAAGTCTGATGTGAAAGCCCGGCTCAACCGG
GGAGGGTCATTGAAACTGGGGAACCTTGAAGTGCAGAAGAGGAGAGTGGAAATCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAAC
ACCAGTGGCGAAGGCGACTCTCTGGTCTGTAAGTACGCTGAGGAGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCTGGTAGTCCACGC
CGTAAACGATGAGTGTAAAGGTTAAGGGGTTTTCCGCCCTTATGCTGCAGCTAACGCATTAAGCACTCCGCTGGGGGAGTACGGTCCGCA
GACTGAAACTCAAAGGAATTGACGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTTAATTGAAAGCAACCGGAAGAACCTTACCAGGTCTTGA
CATCCTCTGACAACTAGAGATAGGACGTCCCTTCGGGGGACAGTGCAGAGTGGTGCATGGTTGTCGTAGCTCGTGTGAGATGTTG
GGTTAAGTCCCGCAACGAGCGCAACCTTATGCTTATGTTGCCAGCATTAGTTGGGCACTTAAGGTGACTGCCGGTGAACAACCGGAGGAAGG
TGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGTACAATGGGCGAACAAGGGCAGCGAAACCGCGAGGTTAA

GCCAATCCCAAAATCTGTTCTCAGTTCGGATCGCAGTCTGCAACTCGACTGCGTGAAGCTGGAATCGTAGTAATCGCGGATCAGCATGCCGG
GTGAATACGTTCCCGGGCCTTGTACACACCGCCGTCACACCACGAGAGTTTGTAAACCCGAAAGTCGGTGAGGTAAC

>Bacillus_hisashii_D06

TAAAAAGCTTGCTTTTTGTTGGTTAGCGGGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACCGGGATAACTCCGGGAAACCGGTGC
TAATACCGGATAGATTATCTTCCGCTGGAGAGATAAGGAAAGATGGCTTTTCCATCACTTACAGATGGGCCCGGGCGCATTAGCTAGTTGG
TGAGGTAACGGCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGG
GAGGCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTT
GTTAGGGAAAGAACAGTACCGGARGGAAATGCCGGTACCTGACGGTACCTGACGAGAAAAGCCACGGGTAAYTACGTGCCAGCAGCCGCGG
TAATACGTAGGTGCAAGCGTTGTCGGATTTATTGGGCGTAAAGCGCGCGCAGGCGGTCTTAAAGTCTGATGTGAAATCTTCGGCTCAACC
GCAAGCGTCAATTGAAACTGGGGGACTTGAGTGCAGAAGAGGAAAGCGGAATCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAA
CACCAGTGGCGAAGGCGGCTTCTGGTCTGTAACCTGACGCTGAGGCGGAAAGCGTGGGGAGCAAACAGGATTAGATACCCCTGGTAGTCCACG
CCGTAACAGTATGAGTCTAAGTGTGGGAGGTTTCCGCCCTTCACTGCTGACGTAACGCATTAAGCACTCCGCTGGGGAGTACGGTCGCA
AGACTGAACTCAAAGGAATTGACGGGGGCCGACAAGCGGTGGAGCATGTGGTTAATTGAAAGCAACGCGAAGAACCTTACCAGGTCTTG
ACATCTCTGACCCGCTGGAGACAGGGTCTCCCTTCGGGGACAGGATGACAGGTGGTGCATGTTGTGCTGACGCTGCTGCTGAGATGTTG
GGTAAAGTCCCGCAACGAGCGCAACCTTGGTCTAGTTCAGCTGACGATTCAGTTGGGCACTTAGAGCGACTGCCGGCGAAGTCCGAGGAA
GTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGATGGTACAAAGGGCAGCGAAGCGGCGACGCAT
TAGCGAATCCCAAGAAAACCTTCTCAGTTCGGATTGACGGCTGCAACTCGCTGCATGAAGCCGGAATCGCTAGTAATCGCGGATCAGCATGC
CGGGTGAATACGTTCCCGGGCCTTGTACACACCGCCGTC

>Aeribacillus_compostii_E01

AGCTTGCTCCTTAGGTTAGCGGGGACGGGTGAGTAACACGTGGGCAACCTGCCTGCAGACTGGGATAACTCCGGGAAACCGGAGCTAATA
CCGGATAACACCGAAAACCGCAGGGTTTTCGGTTGAAAGGCGGCTTTAGCTGCTACTGCAGGATGGGCCCGGGCGCATTAGCTAGTTGGTGA
GGTAACGGCTACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGAG
GCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCCGTGAGCGAAGAAGTCTTCGGATCGTAAAGCTCTGTTGTC
AGGGAAGAACAAATRCGGTTCGAACAGGGCGGTACCTGACGGTACCTGACGAGGAAGCCACGGTAACTACGTGCCAGCAGCCGCGGTAATA
CGTAGGTGGCAAGCGTTTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAAGTCTGATGTGAAATCTCCGGCTCAACCCGGA
GCGGCCATTGGAAACTGRGGAACCTGAGTGCAGGARAGGGGAGCGGAATCCACGTGTAGCGGTGAAATGCGTARAGATGTGGAGGAACACC
AGTGGCGAAGGCGGCTCTGCGCTGTAACCTGACGCTGAGGCGGAAAGCGTGGGGAGCGAACAGGATTAGATACCTGGTAGTCCACGCCGT
AAACGATGAGTGTAAGTGTAGAGGGTATCCACCTTATGCTGCTGACGAAACGCATTAAGCACTCCGCTGGGGAGTACGCGCCGCAAGGT
GAAACTCAAAGGAATTGACGGGGGACCCGCAACAGCGGTGGARCCATGTGGTTTTAATTGAAAGCAACGCGAAGAACCTTACCAGGTCTTGAC
ATCCCTGACAACCTAGAAGATAGGGCGTTCCTTTCGGGGACAGGGTGACAGGTGGTGCATGTTGTGCTCASCTCGTGCATGAGATGTT
GGGTTAAGTCCCGCAACGAGCGCAACCTTACCTTAGTTCAGGATTCAGTTGGGCACTCAAAGGTGACTGCCGGTCAAAGTCCGAGGAAAG
GTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGTGGTACAAAGGGCAGCGAAGCCGCGAGGTGG
AGCGAATCCCAAAAACCACTCTCAGTTCGGATTGACGGCTGCAACTCGCTGCATGAAGCCGGAATCGCTAGTAATCGCGGATCAGCATGCCG
CGGTGAATACGTTCCCGGGTCTTGTACACACCGCCGTCACACCACGAGAGTTTGTAAACCCG

>Bacillus_licheniformis_E02

GGAGCTTGCTCCCTTAGGTTAGCGGGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAA
TACCGGATGCTTGATTGAACCGCATGGTTCAATCATAAAAGGTGGCTTTACGCTACCCTTGCAGATGGACCCGCGGGCGCATTAGCTAGTTGGTG
AGGTAACGGCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGGA
GGCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGTGTGAAGGTTTTCCGGATCGTAAAACCTCTGTTGTT
AGGGAAGAACAAAGTACCGTTCGAATAGGGCGGTACCTTACGGTACCTAACAGAAAAGCCACGGTAACTACGTGCCAGCAGCCGCGGTAATA
CGTAGGTGGCMAGCGTTTCCGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTTTCTTAAAGTCTGATGTGAAAGCCCGGCTCAACCGGGG
AGGGTCAATTGAAACTGGGGAACCTGAGTGCAGAAGAGGAGTGGAAATCCACGTGTAGCGGTGAAATGCGTARAGATGTGGAGGAACACC
AGTGGCGAAGGCGACTCTGCTGCTGTAACCTGACGCTGAGGCGGAAAGCGTGGGGAGCGAACAGGATTAGATACCTGGTAGTCCACGCCGT
AAACGATGAGTGTAAGTGTAGAGGGTTTTCCGCCCTTATGCTGCTGACGAAACGCATTAAGCACTCCGCCYGGGGAGTACGGTCGCAAGACT
GAAACTCAAAGGAATTGACGGGGGCCGACAAGCGGTGGARCCAKGTGGTTAATTGAAAGCAACGCGAAGAACCTTACCAGGTCTTGACAT
CCTCTGACAACCTAGAGATAGGGCTTCCCTTTCGGGGCAGAGTGACAGTGGTGGTGCATGTTGTGCTGACGCTGCTGCTGAGATGTTGGGT
TAAGTCCCGCAACGAGCGCAACCTTATGCTTATGTTGCCAGCATTAGTTGGGCACTTAAAGGTGACTGCCGGTGCACAAACCGGAGGAAAGGTGG
GGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGGCGAACAAGGGCAGCGAAGCCGCGAGGCTAAGCC
AATCCCAAAATCTGTTCTCAGTTCGGATCGCAGTCTGCAACTCGACTGCGTGAAGCTGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTG
AATACGTTCCCGGGCCTTGTACACACCGCCGTCACACCACGAGAGTTTGTAAACCC

>Aeribacillus_compostii_E03

GGAGCTTGCTCCTTAGGTTAGCGGGGACGGGTGAGTAACACGTGGGCAACCTGCCTGCAGACGGGGATAACTCCGGGAAACCGGAGCTAA
TACCGGATAACACCGAAAACCGCAGGGTTTTCGGTTGAAAGGCGGCTTTAGCTGCTACTGCAGGATGGGCCCGGGCGCATTAGCTAGTTGGT
GAGGTAACGGCTACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGAGACACGGCCAGACTCCTACGGG
AGGCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCCGCTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCTGTTG
TCAGGGAAAGAACAAATRCGGTTCGAACAGGGCGGTACCTTACGGTACCTGACGAGTGGTGGTGCATGTTGTGCTGACGCTGCTGCTGAGATGTTGGGT
TACGTAGGTGGCAAGCGTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAAGTCTGATGTGAAATCTCCGGCTCAACCGG
AGCGGCCATTGAAACTGGGGAACCTGAGTGCAGGARAGGGGAGCGGAATCCACGTGTAGCGGTGAAATGCGTARAGATGTGGAGGAACAC
CAGTGGCGAAGGCGGCTCTGCGCTGTAACCTGACGCTGAGGCGGAAAGCGTGGGGAGCGAACAGGATTAGATACCCCTGGTAGTCCACGCCG
TAAACGATGAGTGTAAAGTGTAGAGGGTATCCACCTTATGCTGCTGACGAAACGCATTAAGCACTCCGCTGGGGAGTACGGCCGCAAGG

TGAAACTCAAAGGAATTGACGGGGACCCGCAACAAGCGGTGGARCCATGTGGTTAAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGAC
ATCCCTGACAACCTAGAAAGATAGGGCGTTCCCTTTTCGGGGACAGGGTGACAGGTGGTGCATGGTTGTCGTAGCTCGTGTGATGATGTT
GGGTTAAGTCCCGCAACGAGCGCAACCCCTTGACCTTAGTGGCAGCATTAGTGGGCACTCTAAGGTGACTGCCGGCTAAAAGTCGGAGGAAG
GTGGGGATGACGTCAAATCATATGCCCTTATGACCTGGCTACACACGTGCTACAATGGTGGTACAAAGGGCAGCGAAACCCGCGAGGTGG
AGCGAATCCCAAAAACCACTCTCAGTTCGGATTGACAGGCTGCAACTCGCTGCATGAAGCCGGAATCGTAGTAATCGCGGATCAGCATGCCG
CGGTGAATACGTTCCCGGGTCTGTACACACCCCGCTCACACCACGAGGAGT

>Bacillus_subtilis_E04

AGCGGACAGATGGGAGCTTGCTCCCTGATGTAGCGGGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGA
AACCGGGGCTAATACCGGATGGTTGTTTGAACCGCATGGTTCAAACATAAAAAGGTGGCTTCGGCTACCACTTACAGATGGACCCGCGGCGCATT
AGCTAGTTGGTGAAGTACCGGCTACCAAGGCGACGATGCGTAGCCAACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCGAGCTCTACGGGAGGCA
ACTCCTACGGGAGGACAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCGCGTGAGTGAAGGTTTTCCGGATCGTA
AAGCTCTGTTGTTAGGGAAGAACAAGTACCCTGCAATAGGGCGGTACCTTGACGGTACCTAACAGAAAGCCACGGCTAACTACGTGCCAGCA
GCCCGGTAATACGTAGTGGCMAGCGTTGTCGGGAATTATTGGGCGTAAAGGGCTCGCAGGCGGTTCTTAAGTCTGATGTGAAAGCCCCG
GCTCAACCGGGGAGGGTCAATGGAACTGGGGAACCTTGAAGTGCAGAAAGAGGAGTGGAAATCCACGTGTAGCGGTGAAATGCGTAGAGATG
TGGAGGAACACCACTGGCGAAGGCGACTCTCGTCTGTAAGTACGCTGAGGAGCGAAAGCGTGGGAGCGCAACAGGATTAGATACCTGG
TAGTCCACGCGTAAACGATGAGTGTAAAGTGTAAAGGGGTTTTCCGCCCTTAGTGTGCTGACGTAACGCTAAAGCACTCCGCTGGGGGAGT
ACGGTCGCAAGACTGAAACTCAAAGGAATTGACGGGGGCCGCAACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTAC
CAGGCTTGACATCTCTGACAATCTAGAGATAGGACGCTCCCTTCGGGGGACAGTGCAGGTGGTGCATGGTTGTCGTAGCTCGTGTGCTG
GAGATGTTGGTTAAGTCCCGCAACGAGCGCAACCCCTTATGCTTAGTTCAGCATTAGTGGGCACTCTAAGGTGACTGCCGGTGCAGAAACC
GGAGGAAGTGGGGATGACGTCAAATCATATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGACAGAAACAAAGGGCAGCGAAACCG
CGAGGTTAAGCAATCCCAAAATCTGTTCTCAGTTCGGATGCGAGTCTGCAACTCGACTGCGTGAAGCTGGAATCGTAGTAATCGCGGATCAG
CATGCCCGGTGAATACGTTCCCGGGCTTGT

>Bacillus_hisashii_E05

CTTGCTTTTTGTTGGTTAGCGGGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGGGAAACCGGTGCTAATACC
GGATAGATTATCTTCCGCTGGAGAGATAAGGAAAGATGGCTTTTGTATCACTTACAGATGGGCCCGCGGCATTAGCTAGTTGGTGAAGT
AACGGCTACCAAGGCGACGATGCGTAGCCAACCTGAGAGGGTGATCGGCCACACTGGGACTGAGACACGGCCAGACTCTACGGGAGGCA
GCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCGCGTGAGCGAAGAAGGTTCTCGGATCGTAAAGCTCTGTTGTTAGG
GAAGAACAAGTACCAGGGAATGCGGTACCTTGACGGTACCTGACGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGGTAAATACGTA
GGTGGCAAGCGTTGTCGGATTTATTGGGCGTAAAGCGCGCGAGGCGGCTTTAAGTCTGATGTGAAATCTTCCGGCTCAACCGCAAGCGGT
CATTGGAACTGGGGACTTGAAGTGCAGAAAGGAAAGCGGAAATCCACGTGTAGCGGTGAAATGCGTAGAGATGTTGGAGGAACACCACTGG
CGAAGCGGCTTTCTGGTCTGTAAGTACGCTGAGGCGGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACCG
ATGAGTGTAAAGTGTGGGAGGGTTTTCCGCCCTTAGTGTGCTGACGTAACGCTAAAGCACTCCGCCYGGGGGAGTACGGTCGCAAGACTGAAA
CTCAAAGGAATTGACGGGGGCCGCAACAAGCGGTGGARCCAKTGGTTAATTCGAAGCAACCGGAAGAACCTTACCAGGTCTTGACATCTCT
GACCGCCTGGAGACAGGCTTCCCTTCGGGGACAGGATGACAGGTGGTTGCATGGTTGTCGTAGCTGCTGTGAGATGTTGGGTTAAG
TCCGCAACGAGCGCAACCCCTTGGTTCTAGTTCAGCATTAGTGGGCACTCTAGAGCGACTGCCGGCGCAAGTCCGGAGGAAGTGGGGA
TGACGTCAAATCATATGCCCTTATGACCTGGGCTACACAGTGTACTAATGGATGGTACAAAGGGCAGCGAAGCGGCGACGATAAGCGAAT
CCGAGAAACCTTCTCAGTTCGGATTGACAGGCTGCAACTCGCTGCATGAAGCCGGAATCGTAGTAATCGCGGATCAGCATGCCCGGTGAA
TACGTTCCCGGCTTGTACACACCCCGCTCACACCAGAGGTTTTGAACCCGGAAGTCGGTA

>Aeribacillus_compostii_E06

AGCTTGCTCTTTAGTTAGCGGGGACGGGTGAGTAACACGTGGGCAACCTGCCTGCAGACTGGGATAACTCCGGGAAACCGGAGCTAATA
CCGGATAACACCGAAAACCGCATGGTTTTTCGGTTGAAAGCGGCTTTTAGCTGTACTGACAGGATGGGCCCGCGGCATTACCTAGTTGGTGA
GGTAACGGCTACCAAGGCAACAATCCGTACCAACCTGAGAGGGTGACCGGCCACACTGGGACTGAAACACGGCCCAAACTCTACGGGAGG
CAGCATTAGGGAATCTCCGCAATGGACAAAAGTCTGACGGACCAACGCCCGTGAAGCAAAAAGGTTCTGGATCGTAAAGCTCTKTGTGAG
GGAARAACAAGTACCGTTGCAACAGGGCGGTACCTTGACGGTACCTGACRAGGAAGCCACGGCTAACTACKTGCCASCASCSCGGTAAATACT
AGGTGGCMAGCGTTGTCGGGAATATTGGGCGTAAAGCGCGCAGGCGGTTCTTAAKTCTGATGTGAAATCTCGCGGCTCAACCGGAGCG
GCCATTGAAAACCTGGGAACTTGAKTGACAGGARAGGGGAGCGGAATCCACGTGTASCCKTGAATGCGTARAGATGTGRAGGAACACCAKTG
GCRAAGGCGGCTCTGCGCTGTAAGTACGCTRAGGCGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCGTAAAC
RATRAGTGCTAAGTGTARAGGATCCACCCTTTAKTGTGCASCAACGCATTAAGCACTCCSCKGGGGAGTACGGCCGAAGGCTGAAACT
CAAAGGAATTGACGGGGACCCGCAACGCGGTGRAGCAKGTGGTTAATTCGAAGCAACCGGAAGAACCTTACCAGGTCTTGACWTCCCTGA
CAACCTAGAGATAGGGCGTTCCCTTTTCGGGGACAGGGTGACAGGTGGTGCATGGTTGTCGTAGCTCGTGTGAGATGTTGGGTTAAGTC
CCGCAACGAGCGCAACCTTACCTTAGTGGCAGCATTCAAGTGGGCACTAAGGTGACTGCCGGCTAAAAGTCGGAGGAAGGTGGGGATG
ACGTCAAATCATATGCCCTTATGACCTGGGCTACACAGTGTACAATGGTGGTACAAAGGGCAGCGAAGCCGAGGTGGAGCGAATCCC
AAAAAACCTCTCAGTTCGGATTGACAGGCTGCAACTCGCTGCATGAAGCCGGAATCGTAGTAATCGCGGATCAGCATGCCCGGTGAATAC
GTTCCCGGCTTGTACACACCCCGCTCACACCAGAGTTTTGAACACC

>Geobacillus_kaustophilus_F01

GTCAGCGGGGACGGGTGAGTAACACGTGGGCAACCTGCCCGCAAGACCGGGATAACTCCGGGAAACCGGAGCTAATACCGGATAACACCGA
AGACCGCATGGCTTTGGTTGAAAGCGGCTTTGGCTGTCACTTGGGATGGGCCCGCGGCATTAGCTAGTTGGTGAAGTAAAGGCTCACC
AAGGCGACGATGCGTAGCCGGCTGAGAGGGTGACCGGCCACACTGGGACTGAGACACGGCCAGACTCTACGGGAGGCAAGCAGTAGGGAA
TCTTCCGCAATGGGCGAAAGCCTGACGGAGCGACGCCGCGTGAAGCAAGAAGGCTTCGGGTGTAAGCTCTGTTGTGAGGGACGAAGGAG
CGCGTTCGAAGAGGGCGGCGGGTACCGGTACCTCAGAGGAAGCCCGGCTAACTACGTGCCAGCAGCCGGTAACTAGTGGGGGCGA

GCGTTGTCCGGAATTATTGGGCGTAAAGCGCGCGAGCGGTTCTTAAGTCTGATGTGAAAGCCACGGCTCAACCGTGGAGGGTCTATTGGA
AACTGGGGGACTTGAGTGACAGGAGAGGAGAGCGGAATTCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCACTGGCGAAGG
CGGCTCTGGCTGCAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAACAGGATTAGATACCTGGTAGTCCACGCCGTAACCGATGAGT
GCTAAGTGTAGAGGGGTACACCCCTTATGCTGACGCTAACCGGATAAGCACTCCGCCYGGGGGAGTACGGCCGAAGGCTGAAACTCAAA
GGAATTGACGGGGGCCCGCAACAAGCGGTGGARCCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCCTGACAAC
CCAAGAGATTGGCGTTCCTCCCTCGGGGGACAGGGTGACAGGTGGTTGCATGGTTGTCGTGAGTGTGAGATGTTGGGTTAAGTC
CCGCAACGACGCGCAACCTCGCTCTAGTTGCCAGCCAGGAAGTGGGCACTARAAGGGACTGCCSGCSAAKTCARGAAAAGGTGGGGG
ATGACGTCAAATCATCAKCGCCCTTATGACCTGGGCTACACAGTGTACAATGGGCGGTACAAAGGGCTGCAACCCGCGAGGGGGAGCGA
ATCCCAAAAAGCCGCTCTAGTTCGATTGACAGGCTGCAACTCGCTGCATGAAGCCGGAATCGTAGTAAATCGCGGATCAGCATGCCGCGGT
AATACGTTCCCGGCTGTACACACCGCCGTCACACCAGAGACTTGCAACACCCG

>Aeribacillus_composti_F02

TGCTCCTTAGGTTAGCGGCGACGGGTGAGTAACAGTGGGCAACCTGCCCTGCAGACTGGGATAACTTCGGGAAACCGGAGCTAATACCGG
ATAACACCGAAAACCGCAGGGTTTTCGGTTGAAAGCGGCTTTAGCTGTACTGCAGGATGGCCCGCGGCGATTAGCTAGTTGGTGAGGTA
ACGGCTCACAAGGCGAGATGCGTAGCCAACTGAGAGGGTGACGGCCCACTGGGACTGAGACACGGCCAGACTCTACGGGAGGCAG
CAGTAGGGAATCTCCGAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGCGAAGAAGTCTTCGGATCGTAAAGTCTGTTGTCAGGG
AAGAACAAGTGGCGTTCCGAACAGGGCGGTACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACGTA
GGTGGCAAGCGTTGTCGGGAATTATTGGGCGTAAAGCGCGCAGGCGGTTCTTAAGTCTGATGTGAAATCTCGCGGCTCAACCCGCGAGCGG
CCATTGAAAACCTGGGAACTTGAGTGACGAAAGGGGAGCGGAATTCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCACTG
GCGAAGGCGGCTCTCGGCTGTAACGTGACGCTGAGCGCGAAAAGCGTGGGGAGCGAACAGGATTAGATACCTGGTAGTCCACGCCGTAAC
GATGAGTGCTAAGTGTAGAGGGTATCCACCTTATGCTGTCAGCAAACGCATTAAGCACTCCGCTGGGGAGTACGGCCGCAAGGCTGAAAC
TCAAAGGAATTGACGGGGACCCGCAACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCCTGA
CAACCTAGAGATAGGGCGTTCCCTTTGCGGGACAGGGTGACAGGTGGTGCATGGTTGTCGTGAGTGTGAGATGTTGGGTTAAGTC
CCGCAACGAGCGCAACCTTACCTTAGTTGCCAGCATTAGTTGGGCACTCTAAGGTGACTGCCGCTAAAAGTCCGARGAAAAGGGKGGGGA
TGACGTCAAATCATATGCCCTTATGACCTGGGCTACACAGTGTACAATGGGTGGTACAAAGGGCAGCGAAAACCCGCGAGGTGGAGCGAAT
CCCAAAAACCACTCTAGTTCGATTGACAGGCTGCAACTCGCTGCATGAAGCCGGAATCGTAGTAAATCGCGGATCAGCATGCCGCGGTGAA
TAGTTCGCGGCTGTACACACCGCCGTCACACCAGAGAGTTTGAACCCCGAAGTGGTCTAGTGACTGCATCATGGGAGCAGAGCG
AACCAGCACTCAGCATAAGTGG

>Aeribacillus_composti_F03

GTGAAGTGCATGCCTATACATGACGTGAGCGGACCGAAGGGAGCTTGCTCCTTAGGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAA
CCTGCCCTGCAGACGGGATAACTTCGGGAAACCGGAGCTAATACGGGATAACCCGAAAACCGCAGGGTTTTCGGTTGAAAGCGGCTTTTA
GCTGTCACTGCAGGATGGCCGCGCGCATTAGCTAGTTGGTAAAGGTAACAGGCTCACCAAGGCGACAATGCGTAGCCAACTGAGAGGGTGA
CCGGCCCACTGGGACTGAGACACGGCCAGACTCTACGGGAGGCAGCAGTAGGGAATCTTCGCAATGGACGAAAGTCTGACGGAGCAAC
GCCCGTGAGCGAAGAAGTCTTCGGATCGTAAAGTCTGTTGTCAGGGAAGAACAAGTCCGTTGCAACAGGGCGGTACCTGACGGTACCT
GACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGTGGCMAGCGTTGTCGGGAATTTGGGCGTAAAGCGCGCGC
AGGCGGTTCTTAAGTCTGATGTGAAATCTCGCGGCTCAACCGGAGCGGCCATTGAAAACCTGGGAACTTGAGTGCAGGARAGGGGAGCGGGA
ATCCACGTGTAGCGGTGAAATGCGTARAGATGTGGAGGAACACCACTGCGGAGGGCGCTCTGGCCTGTAACGACTRAGCGCGCGAAA
GCGTGGGGAGCGAACAGGATTAGATACCTGGTAGTCCACGCGTAAACGATGAGTGTGAGTGTAGAGGGTATCCACCTTTAGTCTGAG
CAAACGCATTAAGCACTCCGCTGGGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGACCCGCAACAAGCGGTGGARCCATGT
GGTTAATTCGAAGCAACCGCAAGAACCCTTACCAGTCTGACATCCCTGACAACCTAGAAGATAGGGCGTTCCCTTTGCGGGACAGGGTG
ACAGGTGGTGCATGGTTGTCGTGAGTCTGTGCTGARAWGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTACCTTAGTTGCCAGCATTC
AGTTGGGCACTCTAAGGTGACTGCCGCTAAAAGTCCGAGGAAGGTGGGGATGACGCTCAAATCATATGCCCTTATGACCTGGGCTACACAC
GTGCTACAATGGGTGATACAAAGGGCAGCGAAAACCCGAGGTGGAGCGAATCCCAAAAACCACTCTAGTTCGATTGACGGCTCAACTCG
CCTGCATGAAGCCGGAATCGTAGTAAATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGCTGTACACACCGCCGTCACACCAGAG
AGTTTGAACACC

>Bacillus_coagulans_F05

GGGCTGGCGCTGCCCTAACTGACGTGCGGACCTTTAAAGCTTGCTTTTAAAGGTTAACGGCGAAGGGTGAAGTAAACAGTGGGCCACTGC
CTGTAAGACTGGGATAACACCCGGAAACCGGGGCTAATACCGGAAAGTTTTCTCCCATGGAGGAAAAGGAAAGCGGCTCTGCTGCCA
CTTACAGATGGGCCCGCGCGCAATAGCTAGTTGGCGGGTAAACGGCCCAAGGCAACGATGCGTAGCCACTGAGAGGGTATCGGCCA
CATTGGGACTGAGACACGGCCAACTCTACCAGGAGCAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCCGTG
AGTGAAGAAAGCCTTCGGGTCGTAAYTYKTTGCCGGGGAAGAACAAGTCCGTTGCAACAGGGSGGCGCCYTACGGTACCCSGCCAGAA
AGCCACGGTAAMTACGTGCCAGCAGCCGGTAATACGTARGTGGCAAGCGTTGTCSSGAATTTAGGGSGTAAAGCGCGCAGGCGGCTT
CTTAARTCTGATGTAATCTGCGGCTCAACCGCAAGCGGTCTTGGAAACTGGGAGGCTTGAAGTGCAGAAARAGARTGGAATTCACGTG
TAGCGGTGAAATGCGTARARATGTGGAGGAACACCACTGGCGAARGCGGCTCTYTGTTGCTGTAACGACTGACGCTKARGCGGAAAGCGTGGGGAG
CAAACAGGATTAGATACCTGGTAGTCCACGCGTAAACGATGAGTGTAAAGTGTAKARGTTTCCGCCCTTATGCTGACGTAACGCATTA
AGCACTCCGCCYGGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGGCCGCAACRCGTTGGAGCAKGTGGTTAATTKAA
GCAACGCRAGAACCTTACCAGTCTGACATCTGACCTYCTGGAGACAGGGCTTCCCTTKGGGGACAGARTGACARGTGGTGCATSG
TTGTCASCTCGTGTGCTGARARITGGGTTAAGTCCCGMACSAGCSMACCTTACCTTARTTGCCASCATTGARTTGGGCACVCTAAGGT
GACTGSCGGTGACAAACSGARGAAGGKGGGGRWAMSTCMAAWCAYCATGCCCYWTRACCTGGGSTACACACSTGCTWCAATGGATGG
TACAAGGGCTSCGARACCGGAGGTTAAGCCAATCCASAAAACCAATCCCAAGTYKATTGACAGGSKGCAACCCGCTGCMTGAARCCGGAA
WCGCTAKTAWTCGGGATCARCAWGGCGGKGAATAMKTTCCCGGGCTTKTACAMCMCCSCSSWMAATTYAATCCAAATCTGRGAASTT
AWGTCCAATGGGAATGGATGGGTTAATGCTTGGGGTGGGGGGTTTCGCCAAAGTCAAAA

>Bacillus_coagulans_F06

GTGCAGTCGTGCGGACCTTTAAAGCTTGCTTTAAAAGTTAGCGGGCGGAAGGGTGAGTAACACGTGGGCAACCTGCCTGTAAGACTGGGATA
ACGCCCGAAACCGGGGCTAATACCGGAAAGTTTTTCTCCGCATGGAGGAAAAAGGAAAGGCGGCTTCGGCTGCCACTTACAGATGGGCC
CGCGGCAATAGCTAGTTGGCGGGTAACGGCCCAAGGCAACGATGCGTAGCCACCTGAGAGGGTGATCGGCCACATTGGGACTGAGA
CACGGCCAACTCTACGGGAGGCAGCAGTAGGGAATCTCCGCAATGGACGAAAAGTCTGACGGAGCAACGCCCTGAGTGAAGAAAGCCT
TCGGGTCGTAATACTTTTCCGGGGGAAGAACAAGTCCGTTGCAACAGGGCGGCCCTGACGGTACCCGGCCAGAAAGCCACGGCTAACT
ACGTGCCAGCAGCCGGTAATACSTARGTGGCAAGCGTTTCCGGAATATTGGGCGTAAAGCGCGCAGGCGGCTTCTTAAGTCTGATGT
GAAATCTTGC GGCTCAACCGCAAGCGGTCAATTGGAACTGGGAGGCTTGAAGTGCAGAAARARGAGAGTGGAAATCCACGTGTAGCGGTGAAATG
CGTARAGATGTGGAGAACACCAAGTGGCGAARGCGGCTCTCGTGTGTAAGTGCAGCTGARGCGSAAAGCGTGGGGAGCAAACAGGATTAG
ATACCCTGGTAGTCCACGCCGTAACGATGAGTCTAAGTGTAGAGGGTTCCGCCCTTATGCTGCAGCTAACGCATTAAGCACTCCGCCTG
GGGAGTACGGCCGAAGGCTGAACTCAAAGGAATTGACGGGGGCCGCAACGCGGTGGAGCATGTGGTTAATTGCAAGCAACCGGAAGA
ACCTTACCAGGCTTACATCCTCTGACCTCCCTGGAGACAGGGCTTCCCTTCGGGGGACAGARTGACAGGTGGTGCATGGKGTCTGTCAGCT
CGTGTCTGARATGTTGGGTTAAGTCCCGCMACGAGSGCAACCTTGACCTAGTTGCCASCATTGARTGGGCACTTAAGGTGACTCCGGT
GACAAACCGAGRAAGGKGGGGATGAMGTCAAATCATCAWGCCCTTATGAMCTGGGSTACACASTGCTMCAATGGATGKRCAAAGGGG
TGCGAGACCGSMGGTTAAGCCAAATCCAGAAAACCAATCCAGATTYGGATTGACAGSKGSAMCCSCCTRMACYGAAKCTYKKAATCGMWAR
WAAAYCSCGGGAAATTCGGCCTTTACCCCCCCCCCAATTTAAACCAACTTGAAGTAAAGCCCAAGGAATGGGGGGTTAAATGGGTT
TAGAGCCGGTTCGGCCGACCGGAGGTACGAATAAAATGAAAG

>Geobacillus_kaustophilus_G01

GCGGCGACGGTGAGTAACACGTGGGCAACCTGCCGCAAGACCGGGATAACTCCGGAAACCGGAGCTAATACCGGATAACACCGAAGAC
CGCAGTCTTTGGTTGAAAGCGGCCTTTGGCTGTACTTTCGGATGGGCCCGGGCGCATAGCTAGTTGGTGAGGTAACGGCTACCAAGGC
GACGATGCGTAGCCGGCTGAGAGGGTACCGGCCACTGGGACTGAGACACGGCCAGACTCTACGGGAGGCAGCAGTAGGGAATCTTC
CGCAATGGGCGAAACCTGACGGAGCGACGCCGTGAGCGAAGAAGGCTTCGRGTCTAAAGCTCTGTTGTGAGGGACGAAGGAGCGCCG
TTCGAAGAGGGCGCGCGGTGACGGTACCTCACGAGGAAGCCCGGCTAACTACGTGCCAGCAGCCGGTAATACGTAGGGGGCGAGCGTT
GTCCGGAATATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAGTCTGATGTAAAGCCACGGCTCAACCGTGGAGGGTCAATTGAAACTG
GGGACTTGAGTGCAGGAGAGGAGAGCGGAATCCACGTGTAGCGGTGAAATGCTAGAGATGTGGAGAACACCAAGTGGCGAAGGCGGCT
CTCTGGCTGCAACTGACGCTGAGGCGGAAAGCGTGGGGAGCAACAGGATTAGATACCCTGGTAGTCCACGCCGTAACAGATGAGTGTAA
GTGTTAGAGGGGTACACCCCTTAGTGTGCTGACGTAACGCGATAAGCACTCCGCCYGGGGAGTACGCCCGAAGGCTGAAACTCAAAGGAAT
TGACGGGGGCCGCAACGCGGTGGARCAKGTGGTTAATTGAAAGCAACCGGAAGAACCTTACCAGGTCTTGACATCCCCTGACAACCCAAAG
AGATTGGGCGTTCCTCCCTTCGGGGGACAGGGTACAGGTGGTGCATGGTTGTCGTGAGTCTGTGCTGAGATGTTGGGTTAAGTCCCGCAA
CGAGCCAAACCTTCCTCTAGTTGCCAGCACGAAGGTGGGCACTTAGAGGGACTGCCGGCGACAAGTCCGGAGGAAGTGGGGATGACGTC
AAATCATCATGCCCTTATGACCTGGGCTACACAGTGTACAATGGGCGGTACAAAGGGTGCGAACCCGCGAGGGGGAGCGAATCCAAAA
AGCCGCTCTCAGTTCGGATTGACGGTCAACTCGCTGCATGAAGCCGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGAATACGTT
CCGGGCTGTACACACCGCCGTACACACGAGAGCTTCAACACCCGAAGTCCGGT

>Anoxybacillus_kamchatkensis_G02

CTTGCTTTGAATCGTTAGCGGGGACGGGTGAGTAACACGTGGGCAACCTGCCCTGTAGACGGGATAACACCGAAGAAATCGGTGCTAATACC
GGATAATACGAAAGGCCGATGTTCTTCTGTTGAAAGCGCGCAGGCTGTGCTACAGGATGGGCCCGGGCGCATTAGCTAGTTGGTGAGG
TAACGGCTACCAAGGCACAATGCGTAGCCAACCTGAGAGGGTGTGCGCCACTGGGACTGAGACACGGCCAGACTCTACGGGAGGCA
GCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGCGAAGAAGGCTTCGGGTCGTAAGCTGTTGTTAGG
GAAGAACAAGTACCGCAGTCACTGGCGGTACCTGACGGTACCTAACGAGGAAGCCAGGCTAACTACGTGCCAGCAGCCGGTAAATACGTA
GGTGGCAAGCGTTTCCGGAATATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTTAAGTCTGATGTGAAAGCCCAAGGCTCAACCGTGGAGGG
TCATTGAAACTGGGGACTTGAGTGCAGAAGAGGAGAGCGGAATCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCAAGT
GCCAAGCGGGTCTCTGGTCTGTAAGTACGCTGAGGCGGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAC
GATGAGTGTAAAGTGTAGAGGGTATCCACCTTATGCTGTAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGCTCGCAAGAGTGAAC
TCAAAGGAATTGACGGGGGCCGCAACGCGGTGGAGCATGTGGTTAATTGAAAGCAACCGGAAGAACCTTACCAGGTCTTGACATCCCCTGA
CAACCCGAGAGATCGGGCTTCCCTTCGGGGGACAGGGTACAGGTGGTGCATGTTGTCGTGAGTCTGTGCTGAGATGTTGGGTTAA
GTCCCGCAACGAGCGCAACCTCGACCTTAGTGGCAGCATTAGTGGGCACTTAAGGTGACTGCCGGCTAAAAGTCCGGAGGAAGTGGGG
ATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACAGTGTACAATGGGCGGTACAAAGGGTGCGAACCCGCGAGGGGGAGCCAA
TCCAAAAAGCCGCTCAGTTCCGATTGACGGTCAACTCGCTGCATGAAGCCGAATCGCTAGTAATCGCGGATCAGCATGCCGCGGTGA
ATACGTTCCCGGGCTTGTACACACCGCCGTACACACGAGAGTTGCAACA

>Aeribacillus_composti_G03

TGCAGTCGAGCGGACCGAAGGGAGCTTGCTCCTTAAAGTTAGCGGGGACGGGTGAGTAACACGTGGGCAACCTGCCCTGCAACAGGGATAA
CTTCGGGAAACCGGAGCTAATACCGGATAACACCGAACCCGACCGTTCGGTTGAAAGGGCGGCTTTAGCTGCTACTGCACGATGGGCCCG
CGGGCATTAGCTAGTTGGTGAAGTAAACGGCTACCAAGGGCAGCATTCTGATGCCACCTGAAAGGGTACCGGCCACTGGGACTGAAACA
CGGCCAACTCTACGGGAGGCAGCAGTAGGGAATCTCCGCAATGGACGAAAAGTCTGACGGAACAACGCCCTGAGCGAAAAAGGCTTTC
GGATCGTAAAGTCTTTGTCAGGGAATAACAAGTCCGTTTCGACAGGGCGGTACCTTGACGGTACCTGACGAGGAAGGCTAACTAACT
GTGCCAGCAGCCGCTAATACGTAGGTGGCAAGCGTTGTCGGAAATATTGGGCGTAAAGCGCGCAGGCGGTTCTTAAGTCTGATGTGA
AATCTCGGGCTCAACCGCAGCGGCAATTGAAACTGGGAACTTGAAGTGCAGGAAGGGGAGCGGAATCCACGTGTAGCGGTGAAATGC
GTAGAGATGTGGAGGAACCAAGTGGCGAAGGGCGGCTCTGTCGCTGTAAGTGCAGCTGAGGCGGAAAGCGTGGGGAGCGAACAGGATTAG
ATACCCTGGTAGTCCACGCCGTAACGATGAGTCTAAGTGTAGAGGGTATCCACCTTATGCTGCAGCAACGCATTAAGCACTCCGCTG

GGGAGTACGGCCGAAGGCTGAAACTCAAAGGAATTGACGGGGACCCGACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACCGGAAGA
ACCTTACCAGGCTTGTACATCCCTGACAACCCTAGAGATAGGGCGTTCCCTTTCCGGGAACAAGGTGACAGGTGGTGCATGGTTGTCGTACCT
CGTGTCTGAAAATGTTGGGTTAATCCCGCAACGAGCGCACCTTGGACCTTAGTCCACCATTATTTGGCCCTTAAGGTGACTGCCGG

>Bacillus_haynesii_G04

TACCTCACCGACTTCGGGTGTTACAAACTCTCGTGGTGTGACGGGCGGTGTGTACAAGGCCGGGAACGTATTACC CGCGCATGCTGATCCGC
GATTACTAGCGATTCCAGCTTACGCACTGAGTTGACAGCTGCGATCCGAACTGAGAACAGATTTGTGGGATTGGCTAGCCTCGCGCTTCG
TGCCCTTTGTTCTGCCATTGTAGCACGTGTAGCCAGGTATAAGGGGCATGATGATTTGACGTCATCCCCACCTTCTCCGGTTTGTACCCG
GCAGTACCTTAGAGTGCCTCAACTGAATGCTGGCAACTAAGATCAAGGGTTGCGCTCGTTGCGGGACTTAACCAACATCTCACGACACGAGCT
GACGACAACCATGACCACCTGTCACTCTGCCCGAAGGGGAAGCCCTATCTAGGGTTGTCAGAGGATGTCAAGACCTGGTAAGGTCTTC
GCGTTGCTTGAATTAACACATGCTCCACCGCTTGTGCGGGCCCCGCTCAATTCCTTTGAGTTTCACTTTCGACCGTACTCCAGGGCGAG
TGCTTAATGCGTTTGTGCTGACGACTAAAGGGCGGAAACCCTCTAACACTTAGCACTCATCGTTTACGGCGTGGACTACCAGGGTATCTAATCTCG
TTCGCTCCCCACGCTTTCGCGCTCAGCGTCAGTTACAGACCAGAGAGTGCCTTCCGCACTGGTGTCTCCACATCTACGCATTTACCAGCTA
CACGTGGAATTCACCTCTCTCTTCTGCACTCAAGTCCCGAGTTTCAATGACCTCCCGGTTGAGCCGGGGGCTTTCACATCAGACTTAAGAA
ACCGCCGGCGCGCTTACGCCAATAATCCGGACAACGCTTCCACCTACGATTACC CGGGCTGCTGCGACGTAGTTAGCCGTGGCTTCT
GGTAGGTACCGTCAAGG

>Aeribacillus_compostii_G06

GAGCGGACCGAAGGGAGCTTGTCTTTAGTTAGCGGGGACGGGTGAGTAACACGTGGGCAACCTGCCTGACAGCGGGGATAACTTCGG
GAAACCGGAGCTAATACCGGATAACACCGGAAACCGCAGGGTTTTCCGGTTGAAAGCGGGCTTTTAGCTGTCACTGCAGGATGGGCCCGCGGCG
CATTAGCTAGTTGGTGAAGTAACGGCTCACCAAGGGCACAATGCGTAGCCAACTGAGAGGGTGACCGGCCACTGGGACTGAGACACCGGCC
CAGACTCTACGGGAGGACAGTGGGAATCTCCGCAATGGACAAAGTCTACCGGAGCAAAAGCCGCGTGAAGCGAAGAGGTTTTCG
GATCGTAAAGCTTCTGTGKTCAGGGGAAGAACCAAGTRCCGTTCCGAAACAGGGCGGTACCTTACCGTACCTGACGAGGAAAGCCACGGCTAA
CTACGTGCCAGCAGCGCGGTAATACGTAGGTGGCAAGCGTTGTCCGGAATATTGGGCGTAAAGCGCGCAGGGGTTCTTAAGTCTGAT
GTGAAATCTCGCGCTCAACCGCAGCCGCTTGGAAACTGGGGAACCTGAGTGCAGGARAGGGGAGCGGAAATCCACGTGTAGCGGTGA
AATGCGTAGAGATGTGGAGGAACACAGTGGCGAAGGGCGCTCTGGCTGTAAGTGCAGCTGAGGCGCGAAAGCGTGGGGAGCGAAGCG
ATTAGATACCCTGGTAGTCCACCGCTAAACGATGAGTGTAAAGTGTAGAGGGTATCCACCCTTAGTGTGCAGCAAACGCATTAAGCACTCC
GCCTGGGGAGTACGGCCCAAGGCTGAAACTCAAAGGAATTGACGGGGACCCGACAAGCGGTGGAGCATGTGGTTAATTCGAAGCAACCG
GAAGAACCTTACCAGGTCTGACATCCCTGACAACCCTAGAGATAGGGCGTTCCCTTTCCGGGACAGGGTACAGGTGGTGCATGGTTGTCG
TCASCTGTGTGATGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCTTACCTTAGTTGCCAGCATTAGTTGGGCACTTAAGGTGACTG
CCGGCTAAAAGTCCGAGGAAGGTGGGGATGACGTCAAATCATATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGGTGGTACAAAG
GGCAGCGAAACCGGAGGTGGAGCGAATCCAAAAAACCTCTCAGTTCCGATTGACGGCTGCAACTCGCTGCATGAAGCCGGAATCGTA
GTAATCGCGGATCAGCATGCCCGGTGAATACGTTCCCGGCTTGTACACACCGCCGTCACACCACGAGAGTTTGAACCCCGAAGTC

>Bacillus_licheniformis_H01

CTTGCTCCCTTAGGTACGCGGGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGACTGGGATAACTCCGGGAAACCGGGGCTAATACC
GGATGCTTATTGAACCCGATGGTTCAATCATAAAAGGTGGCTTTCAGCTACCCTTACAGATGGACCCGCGGCGCATTAGCTAGTTGGTGAGG
TAACGGCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGTAGGGCCACTGGGACTGAGACACCGGCCACTCTACGGGAGGC
AGCAGTAGGGAATCTCCGCAATGGACGAAAAGTCTGACGGAGCAACGCCGCKTGAGTGATGAAGGTTTTCCGGATCGTAAAACCTGTTGTTAGG
GAAGAACAAGTACCCTTGAATAGGGCGGTACCTTACGGTACCTAACAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAATACGT
AGGTGCMAGCGTTGTCGGGAATATTGGGCGTAAAGCGCGCGCAGGGCGTTTCTAAGTCTGATGTGAAGCCCGGCTCAACCGGGGAG
GGTATTGGAACCTGGGGAACCTGAGTGCAGAAGAGGAGAGTGGAAATCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCCAG
TGGCGAAGGCGACTCTGGTCTGTAAGTGCAGCTGAGGCGCGAAAGCGTGGGGAGCGAACAGGATTAGATACCTGGTAGTCCACGCCGTA
ACGATGAGTGCTAAGTGTAGAGGGTTTTCCGCCCTTAGTGTGCTGACGAAACGATTAAGCACTCCGCTGGGGGAGTACGGTCCGAAGACTGA
AACTCAAAGGAATTGACGGGGCCCGACAAGCGGTGGARCATGTGGTTAATTCGAAGCAACCGGAAGAACCTTACCAGGTCTTACATCCT
CTGACAACCTAGAGATAGGGCTTCCCTTCCGGGGCAGAGTACAGGTGGTTGCATGGTTGTCGTAGCTCGTGTGATGTTGGGTTA
AGTCCCGCAACGAGCGCAACCTTATGTTAGTGGCAGCATTAGTTGGGCACTTAAGTGTACTGCCGTTGACAAACCGGAGGAAGGTGGG
GATGACGTCAAATCATATGCCCTTATGACCTGGGCTACACAGTGTACAATGGGCGAAGAAAGGGCAGCGAAGCCGCGAGGCTAAGCCA
ATCCCAAAATCTGTTCTCAGTTCCGATCGAGTCTGCAACTGACTGCGTGAAGCTGGAATCGTAGTAATCCGCGGATCAGCATGCCCGGTTGA
ATACGTTCCCGGCTTGTACACACCGCCGTCACACCACGAGAGTTTGAACCCCGAAGTCCGGTGGAGT

>Geobacillus_kaustophilus_H02

GTCAGCGGGGACGGGTGAGTAACACGTGGGCAACCTGCCCGAAGACCGGGATAACTCCGGGAAACCGGAGCTAATACCGGATAACACCGA
AGACCGCAGGTCTTTGGTTGAAAGGGCGCCTTTGGTGTCACTTCCGGATGGGCCCGCGCGCATTAGCTAGTTGGTGAAGTAAACGGCTACCA
AGGCGACGATGCGTAGCCGGCCTGAGAGGGTGACCGGCCACTGGGACTGAGACACCGCCAGACTCTACGGGAGGCAGCAGTAGGKAW
TCTTCCGCAATGGSGAAAGCCTGAMGRAGCGACCGGCKTRAGCGAAGAAGGCTTCCGGTCTGTAAGTCTGTTGTGAGGGACGAAGGAG
CGCCGTTGGAAGAGGGCGCGCGGTGACGGTACTCACAGGAAGCCCCGGCTAACTACGTGCCAGCAGCCGCGTAATACGTAGGGGGCGA
GCGTTGTCGGGAATATTGGGCGTAAAGCGCGCGCAGGGGTTCTTAAGTCTGATGTGAAAGCCACGGCTCAACCGTGGAGGGTCAATTGGA
AACTGGGGGACTTGAAGTGCAGGAGAGGAGAGCGGAATCCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACACCACTGGCGAAGG
CGGCTCTGGCCTGCAACTGACGCTGAGGCGCGAAAGCGTGGGGAGCAAAACAGGATTAGATACCTGGTAGTCCACGGCTAAACGATGAGT
GCTAAGTGTAGAGGGGTACACCTTTAGTGTGACGCTAACCGGATAAGCACTCCGCCCKGGGGAGTACGGCCGCAAGGCTGAAACTCAA
GGAAATGACGGGGCCCGACAAGCGGTGGAASCATGTGGTTAATTCGAAGCAACCGGAAGAACCTTACCAGGTCTTACATCCCTGACAAC
CCAAGAAGATTGGGCGTTCCCTTCCGGGGGACAGGGTACAGGTGGTTGCAKGGTTGTCGTAGCTCGTGTGAGAAATGTTGGGTTAA
GTCCCGCAACGAGCGCAACCTCGCTCTWGTGCCASCACTRAAGTGGGGCACTCTTARAGGGACTGCCGGCGACAARTCCGRAGRAAAGG

TGGGGGATGAAAMKTCAAATMATCAAKGCCCTTATRAMCTGGGGSTACMMACSTGGYTACAATGGGGCGGGTACAAAGGTYTGGCGAA
CCCGGAGGGGGAGCGAATCCCAAAAAGCCGCTCTCAGTTCCGATTGCGAGGTGCAACTCGCCTGCATGAAGCCGGAATCGCTAGTAATCGCG
GATCAGCATGCCGCGTGAATACGTTCCCGGCCCTGTACACACCCCGCTCACACCAGAGACTTGCAACACCCGAAAGTCGGT

>Bacillus_hisashii_H03

AGCGAACCAATAAGAACCTTGCTTTTTGTTGGTTAGCGGCGGACGGGTGAGTAACACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGG
GAAACCGGTGCTAATACCGGATAGATTATCTTTCCGCTGGAGAGATAAGGAAAGATGGCTTTTGCATCACTTACAGATGGGCCCGCGCGCA
TTAGCTAGTTGGTGGGTAACGGCTACCAAGGCGACGATGCGTAGCCGACCTGAGAGGGTGATYGGCCACACTGGGAMTGAGACACGSCCC
AGAYTCTACGGGAGGACGAGTAGGGAATCTCCGCAATGGAMGAAAGTYTAMGGAGCAACGCCGCTGAGCGAAGAAGGTCTTCGGATC
GTAAAGCTCTGTTGTTAGGGAAGAACAAGTACCGGAGGGAATGCCGGTACCTTGACGGTACCTGACGAGAAAGCCACGGCTAACTACGTGCCA
GCAGCCGCGTAATACGTAGGTGGCMAGCGTTGTCGGGATTTATTGGGCGTAAAGCGCGCAGGCGGCTTTAAGTCTGATGTGAAATCTT
GCGGCTCAACCGCAAGCGGTCAATTGGAAACTGGGGACTTGAGTGCAGAAGAGGAAAGCGGAATTCACGTGTAGCGGTGAAATGCGTAGAG
ATGTGGAGGAACACCAAGTGGCGAAGGCGCTTTCTGGTCTGTAAGTACGCTGAGGCGGAAAGCGTGGGGAGCAAAACAGGATTAGATACCT
GGTAGTCCACCGCTAAACGATGAGTCTAAGTGTGGAGGGTTCCGCCCTTCACTGCTGCAGTAAACGCATTAAGCACTCCGCTGGGGGA
GTACGGTCCGCAAGACTGAACTCAAAGGAATTGACGGGGGCCGACAGCGGTGGAGCATGTGTTAATTGAAAGCAACCGGAAGAACCTT
ACCAGGTTGACATCTCTGACCGCCTGGAGACAGGGTCTTCYCTTGGGGACAGGATGACAGGTGGTGCATGGTTGCTGACGTCGTCGTG
GTGAGATGTTGGGTTAAGTCCCGCAACSASCGCAACCTTGTTGCTAGTTGCCASCATTAGTTGGCACTCTAGAGCGACTGCCGCGCAACA
CGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGATGTAACAAGGGCAGCGAAGCG
GCGACGCATAAGCGAATCCAGAAAACCTTCTCAGTTCCGATTGCGAGGTGCAACTCGCCTGCATGAAGCCGGAATCGCTAGTAATCGCGGAT
CAGCATGCCGCGTGAATACGTTCCCGGCCCTGTACACACCCCGCTCACACCAGAGACTTGTAACACCCGGAAGTCGGTGGGTAACCGT
AAGGAGCCAGCCG

>Aeribacillus_composti_H04

GAAGGGAGCTTGCTCCTTAGGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCCTGCAGACGGGGATAACTTCGGGAAACCGGA
GCTAATACCGGATAACACCGAAAACCGCAGGGTTTTCCGTTGAAAGCGGGCTTTAGCTGTCACTGCAGGATGGGCCCGCGCGCATTAGCTAG
TTGGTGGGTAACGGCTACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGAGACACGGCCAGACTCCTA
CGGGAGGCGAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCCGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTCT
GTTGTCAGGGGAAGAACAAGTRCCGTTGAAACAGGGCGGTACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCG
GTAATACGTAGGTGGCMAGCGTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAGTCTGATGTGAAATCTCGCGGCTCAA
CCGCGAGCGGCATTGAAACTGGGGAACCTTGAGTGCAGGARAGGGGAGCGGAATTCACGTGTAGCGGTGAAATCGTARAGATGTGGAGG
AACACAGTGGCGAAGGCGGCTCTCTGGCCTGTAAGTACGCTRAGCGCGGAAAGCGTGGGGAGCGAAGCAGGATTAGATACCTGGTGTAGTCCA
CGCGTAAACGATGAGTCTAAGTGTAGAGGGTATCCACCCTTATGCTGCAGCAAACGCATTAAGCACTCCGCTGGGGAGTACGGCCGCA
AGGTGAAACTCAAAGGAATTGACGGGACCCGCAACGCGGTGGAGCATGTGGTTAATTCGAAGCAACCGGAAGAACCTTACCAGGCTTG
ACATCCCCTGACAACCTAGAGATAGGGGCTTCCCTTTCCGGGACAGGGTGACAGGTGGTGCATGGTTGCTGCASCTCGTGCCTGARATGTT
GGGTTAAGTCCCGCAACGAGCGCAACCTTGACCTTAGTTCAGCATTCAKTTGGGCACTTAAGGTGACTGCCGCTAAAAGTCGGAGGAAG
GTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACACGTGCTACAATGGTGGTACAAGGGGAGCGAAGCCGCGAGGTGG
AGCGAATCCCAAAAACCACTCTCAGTTCCGATTGCGAGGTGCAACTCGCCTGCATGAAGCCGGAATCGCTAGTAATCGCGGATCAGCATGCC
CGGTGAATACGTTCCCGGGTCTGTACACACCCCGCTCACACCAGAGACTTGTAACACCCGAAAGTCG

>Bacillus_hisashii_H05

ACGTGGGTAACCTGCCTGTAAGATCGGGATAACTCCGGGAAACCGGTGCTAATACCGGATAGATTATCTTTCCGCTGGAGAGATAAGGAAAGA
TGGCTTTTGCATCACTTACAGATGGGCCCGCGCGCATTAGCTAGTTGGTGGGTAACGGCTACCAAGGCGACGATGCGTAGCCGACCTGAG
AGGGTATCGGCCACACTGGGACTGAGACACGGCCAGACTCTACGGGAGGCGAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACG
GAGCAACCGCGGTGAGCGAAGAAGGTCTTCGGATCGTAAAGCTGTTGTTAGGGAAGAACAAGTACCAGGAGGAATGCGGTACCTTGACG
GTACCTGACGAGAAAGCCACGGCTAACTACGTGCCAGCAGCCGCGTAACTAGTGGCAAGCGTTGTCGGGATTTATTGGGCGTAAAGCG
CGCGCAGGCGTCTTTAAGTCTGATGTGAAATCTGCGGCTCAACCGCAAGCGGTCAATTGAAACTGGGGACTTGAGTGCAGAAGAGGAAA
GCGGAATTCACGTGTAGCGGTGAAATGCGTAGAGATGTGGAGGAACCAAGTGGCGAAGGCGGCTTTCTGGTCTGTAAGTACGCTGAGGCG
CGAAAGCGTGGGAGCAAACAGGATTAGATACCCTGGTGTAGTCCACCGCTAAACGATGAGTGTAAAGTGTGGAGGGTTCCGCCCTTCACTG
CTGCAGCTAACGCATTAAGCACTCCGCTGGGAGTACGGTGCAGACTGAAACTCAAAGGAATTGACGGGGCCCGCAACAAGCGGTGGAGC
ATGTGGTTAATTCGAAGCAACCGGAAGAACCTTACCAGGCTTGACATCTCTGACCGCCTGGAGACAGGGTCTTCCCTTCCGGGACAGGAT
GACAGGTGGTGCATGGTTGCTGCTGAGTCTGCTGAGATGTTGGTTAAGTCCCGCAACGAGCGCAACCTTGGTTCTAGTTGCCAGCATT
AGTTGGGCACTTAGAGCGACTGCCGCGCAAGTGGAGGAGGTGGGGATGACGTCAAATCATCATGCCCTTATGACCTGGGCTACACA
CTGTACAATGGAWGTTACAAGGGCAGCGAAGCGCGCAGCATAAGCGAATCCAGAAAACCTTCTCAGTTCCGATTGCGAGGTGCAACT
CGCCTGCATGAAGCCGGAATCGCTAGTAATCGCGGATCAGCATGCCGCGTGAATACGTTCCCGGGCTTGTACACACCCCGCTCACACCAG
AGAGTTTGAACACCCGAAAGTCGGTGGGTAACCGTAAGGAGCCAGCCG

>Aeribacillus_composti_H06

AGTCGAGCGGACCGAAGGGAGCTTGCTCCTTAGGTTAGCGGCGGACGGGTGAGTAACACGTGGGCAACCTGCCCTGCAGACTGGGATAACTT
CGGAAACCGGAGCTAATACCGGATAACACCGAAAACCGCATGTTTTCCGTTGAAAGGGCGCTTTTAGCTGTCACTGCAGGATGGGCCCGCG
GCGCTTAGCTAGTTGGTGGGTAACGGCTACCAAGGCGACAATGCGTAGCCAACCTGAGAGGGTGACCGGCCACACTGGGACTGAGACACG
GCCAGACTCTACGGGAGGCGAGCAGTAGGGAATCTCCGCAATGGACGAAAGTCTGACGGAGCAACGCCGCTGAGCGAAGAAGGTCTTCG
GATCGTAAAGCTCTGTTGACAGGGAAGAACAAGTCCGTTGCAACAGGGCGGTACCTTGACGGTACCTGACGAGGAAGCCACGGCTAACTACG
TGCCAGCAGCCGCGTAATACGTAGGTGGCMAGCGTTGTCGGGAATTATTGGGCGTAAAGCGCGCGCAGGCGGTTCTTAAGTCTGATGTGAA
ATCTCGGGCTCAACCGGAGCCGCCATTGAAACTGGGGAACCTTGAGTGCAGGAGAGGGGAGCGGAATTCACGTGTAGCGGTGAAATGC

GTAGAGATGTGGAGGAACACCAGTGGCGAAGGCGGCTCTYTGGCCTGTAAGTACGCTGAGGCGCGAAAGCGTGGGGAGCGAACAGGATTAG
ATACCCTGGTAGTCCACGCCGTAACGATGAGTGCTAAGTGTAGAGGGTATCCACCCTTTAGTGTGAGCAAAACGCATTAAGCACTCCGCCTG
GGGAGTACGGCCGCAAGGCTGAAACTCAAAGGAATTGACGGGGACCCGCACMAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCCGAAGA
ACCTTACCAGGTCTTGACATCCCCTGACAMCYCTAGAAGATAGGGCGTTCCTTCGGGGACAGGGTGACAGGTGKTGCATGGTTGTCGTCAG
CTCGTGTGCGTARAATKTTGGGTTAAGTCCCAGCAACGAGCGCAACCCTTGACCTTAGTTGCCAGCATTCAKTTGGGCACTCTAAGGTGACKGCCG
GCTAAAAGTCGGAGGAAGTGGGGATGACGTCAAATCATGCCCCTTATGACCTGGGCTACACAGTGCTACAATGGGTGGTACAAAGGCC
AGCGAAACCGCGAGGTGGAGCGAATCCAAAAAACCCTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTGCATGAAGCCGGAATCGCTAGTA
ATCGCGGATCAGCATGCCGCGGTGAATACGTTCCCGGCTTGTACACACCGCCCGTCACACCACGAGAAGTTTGTAAACCCCA

Nota. Datos generados por el autor.