The application of DNA taxonomy to identify fungal endophytes

by

Ryleigh A. Stack

BSc (Hons), University of New Brunswick, Saint John, 2019

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Science

in the Graduate Academic Unit of Biology, Saint John

Supervisor: Christopher A. Gray, PhD, Department of Biological Sciences and

Department of Chemistry

Examining Board: John A. Johnson, PhD, Department of Biological Sciences
Aaron Granger, PhD, Department of Biological Sciences
Alfredo Justo, PhD, New Brunswick Museum

This thesis is accepted by the
Dean of Graduate Studies

THE UNIVERSITY OF NEW BRUNSWICK

October 2020

ABSTRACT

Fungal identification has been a long-standing challenge in mycology. DNA barcoding and DNA taxonomy have been found to mitigate challenges associated with fungal identification, in this project these techniques were compared. The Natural Products Research Group (NPRG) has a collection of over 400 endophytes, 242 of which had been identified using DNA barcoding, resulting in ambiguous or missing identifications for approximately 25% of the isolates. In this study, DNA taxonomy was used to resolve and verify identifications in the NPRG's collection. To determine if DNA taxonomy is useful in identifying new DNA sequences, a test set of endophytes was isolated and a phylogenetic tree of 441 sequences was built. Analysis of this tree resulted in 55 sequences from the NPRG collection being identified (45) or re-identified (10), improving the collection's identifications. Finally, 47 of the new sequences (76%) were identified to the species (4), genus (42), or order (1) level.

DEDICATION

To my grandfather, I did not get the chance to tell you I had completed this before you had to go, but you have continued to be one of my biggest motivations to complete it. Thank you for everything.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere thanks to Dr. Christopher Gray and Dr. John Johnson. The continued support and mentorship throughout this project have encouraged me to keep going, and to dive into the project headfirst. Working with you has taught me so much, and I consider myself fortunate to have gotten to know you both.

I want to thank the past and present members of the Natural Product Research Group for the encouragement and support, and for the friendships that I will take with me when I leave. I am immensely grateful to have been welcomed into the lab group, and can say with certainty that I have enjoyed working with, and learning from all of you, you have made my experience as a graduate student one that I will remember fondly.

My appreciation goes out to Dr. Aaron Granger for taking the time to serve on my supervisory committee, examining my thesis, and for being a mentor and a friend throughout both my undergraduate and graduate degree.

Lastly, to my friends and family for believing in me, thank you. Especially to my mom, who has encouraged me to be the best version of myself, and to always give everything I do one hundred percent. This thesis, and the completion of the research within it, would not have been possible without you by my side as my greatest supporter. You have always believed in me even when I did not believe in myself, and for that I thank you.

Table of Contents

List of Tables

List of Figures

List of Symbols, Nomenclature or Abbreviations

BLAST: Basic Local Alignment Search Tool

CMA: Cornmeal agar

CZA: Czapek-dox agar

DDBJ: DNA Data Bank of Japan

DNA: Deoxyribonucleic acid

ENA: European Nucleotide Archive

FASTA: Fast-all

INSDC: International Nucleotide Sequence Database Collaboration

ITS: Internal transcribed spacer

MEA: Malt extract agar

MEGA X: Molecular Evolutionary Genetic Analysis Version X

MSA: Multiple sequence alignment

MUSCLE: Multiple Sequence Comparison by Log-Expectation

NCBI: National Center for Biotechnology Information

NPRG: Natural Products Research Group

PCR: Polymerase chain reaction

PDA: Potato dextrose agar

TBE: Tris Borate EDTA

UNBSJ: University of New Brunswick Saint John

UNITE: User-friendly Nordic ITS Ectomycorrhiza Database

UPGMA: Unweighted Pair Group Method with Arithmetic mean

Introduction

Research in a wide array of scientific fields has focused on fungi, including; ecology, plant pathology, and natural products chemistry.[1–4] Identifying fungi aids in accessing existing information for specific species and closely related species, providing valuable information.[5–8]Despite this, the ability to identify the majority of the fungal Kingdom has eluded mycologists and others interested in the fungal Kingdom. Indeed, described species represent less than ten percent of the lower range of the estimated 1.5 to 5.1 million possible species,[11–14] leaving the vast majority of the fungal kingdom to be discovered and identified. As of 2018, approximately 1.5×10^5 fungal species have been described,[9] with the majority belonging to the phyla Ascomycota (9.0×10^4), and Basidiomycota (5.0×10^4) distributed across six sub-phyla and 36 classes of fungi with comparatively few species distributed amongst the remaining phyla, Zygomycota (1.7×10^3), Chytridiomycota (980), Glomeromycota (230), and Blastocladiomycota (220) (Figure 1).[9,10]

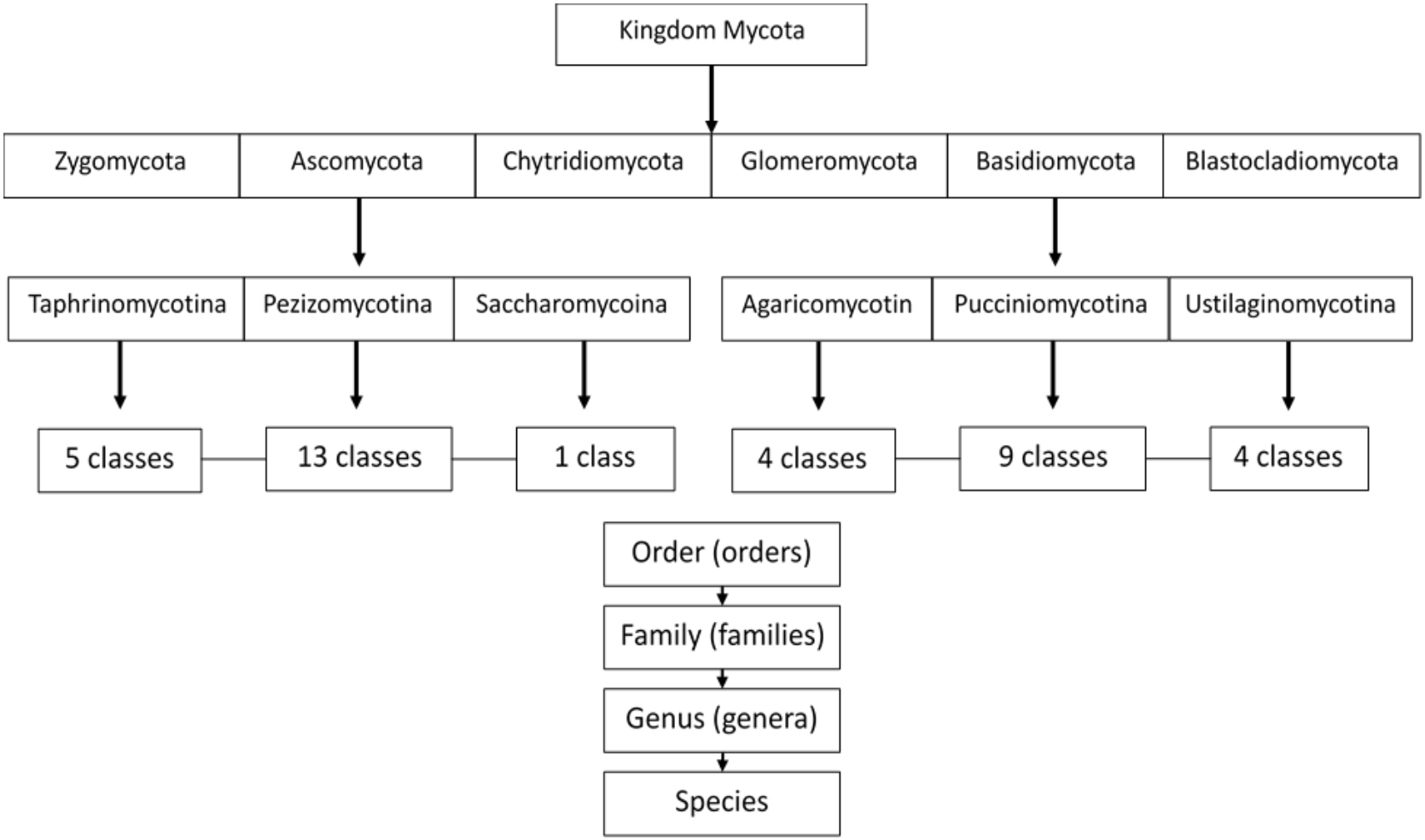

Figure 1. Taxonomic division of the fungal Kingdom from Kingdom to Species level.

Over time, several advancements have been made that changed the way fungi are identified and described. In mycology, relying only on morphological and physiological characteristics for identification was once common practice.[15–17] Unfortunately, morphological techniques are not always able to provide an accurate identification as they largely target fungi that produce fruiting bodies or spore forming structures.[9,15,18–20] While traditional taxonomy works well for some species,[3,21,22] many fungi do not produce identifiable structures when grown on artificial media (e.g. malt extract agar, potato dextrose agar, Czapek-Dox agar),[23,24] including microscopic fungi.[25–27] Fungal endophytes, for example, spend a part or all of their lifecycle living within a plant's tissue(s) and often do not produce identifiable structures.[23,28] If an identification could not be obtained, fungi were often described based on observable characteristics such as colour, hyphal growth, and colony shape when grown on artificial media. This approach is not only limiting, but can be problematic as fungi show high phenotypic plasticity[29] that has led to the presence of synonyms in fungal taxonomy, separate naming for the anamorph and teleomorph (asexual and sexual stage respectively) of the same fungus,[12,30] and misidentifications due to highly similar characteristics between unrelated species.[24] Therefore, rather than relying solely on the morphological features of a fungus, DNA sequence-based methods are proving be more amenable to the accurate identification of fungi that previous techniques were unable to identify.

DNA sequence-based identification uses short, characteristic, gene regions called barcode markers for identifying organisms.[31] These sequences are similar to the Universal Product Codes (UPC) found on consumer products and represent unique tags

that identify organisms. Because of this, and as most species have been found to have a different sequence of nucleotides that reflects genetic differences at the molecular level, these sequences have become known as DNA barcodes and identification of organisms through comparison of these representative sequence data is referred to as DNA barcoding.[31,32] In 2003, when DNA barcoding was first proposed, the mitochondrial cytochrome oxidase I (COI) gene was adopted as the barcode for all animal life,[33,34] and subsequently used as the universal barcode for all organisms.[35] However, the need for more specific genes became apparent as the COI gene was difficult to amplify in the plant, fungus, protist, and bacterial kingdoms, and did not effectively discriminate between the species in these other kingdoms.[36,37] Through the application of DNA barcoding in major taxa, gene regions that are more effective for use in other kingdoms, including *rbcL* and *matK* regions for plants,[38,39] 16S rRNA for bacteria,[37] and the 18S rRNA gene for protozoa[40] have been identified. In fungi, the internal transcribed spacer (ITS) region was designated as the universal barcode marker in 2011 by a consortium of mycologists that evaluated six possible DNA regions.[35] The ITS region was found to be effective in differentiating between fungal species,[35,41,42] with higher resolution when compared to other regions of the DNA sequence,[35,42] and has led to the most success compared to the five other investigated gene regions in obtaining identifications across a broad range of fungi.[24,41,43–45]

To identify organisms rapidly, DNA barcoding makes use of the species' specific barcode markers.[33,34] This DNA sequence is uploaded to an online database such as the National Centre for Biotechnology Information (NCBI) GenBank database that uses a basic local alignment search tool (BLAST)[46] to compare the sequence to all other

sequences within the database. This search provides users with a list of top matches based on the similarity amongst the sequences and assigns values for consideration including the bit score, E-value, and percent identity. Bit scores are calculated to reflect the degree of similarity of the hits, or matches, to the query sequence independently of the size of the search space or database. The E-value, or expected value, is calculated from bit scores and represents the number of sequences that can be retrieved with an equal or greater bit score by chance alone.[47] Therefore, the smaller the E-value, the greater the confidence in the match.[47] Finally, the percent identity describes how similar the query sequence is to the matched sequence(s) with a higher percent identity indicating a more significant match. Using these values as an indicator of homology to the query sequence, it is possible to assign a taxonomic identification using the existing sequences within a database such as INSDC[48] (International Nucleotide Sequence Database: GenBank,[49] DDBJ,[50] ENA[51]) or UNITE (User-friendly Nordic ITS Ectomycorrhiza Database).[47,52,53]

While DNA barcoding has led to the rapid accumulation of annotated DNA sequences and facilitated the identification of species in all kingdoms[38–40,54–59] there are still limitations to its use. Primarily, the databases have been found to have insufficient sequence data for a conclusive identification even in the case of well-studied groups.[60,61] Additionally, the presence of misidentified or poorly annotated sequences prohibits the identification of new sequences.[61,62] Furthermore, subjectivity surrounding an acceptable percent identity threshold for identification to the species (97-100%), or even genus level (92-96%) exists within the literature.[63–66] As such, DNA barcoding has the potential to exacerbate the challenges associated with the identification of fungi. Given that most fungi have yet to be discovered, let alone identified, the databases used to identify a

fungal ITS sequence will be incomplete by nature. In turn, an identification based on comparison with existing sequences will not be possible if it simply does not exist within the database. Thus, when similarity thresholds are not met and DNA barcoding is not able to assign an accurate identification to a sequence, other methods need to be applied in order to classify and identify the ambiguous sequence even in the absence of a formal identification.

DNA taxonomy utilises the same gene sequence(s) used for DNA barcoding, but rather than use a match generated by a database to assign a taxonomic name, the DNA sequence is incorporated into a phylogenetic tree to reveal important taxonomic relationships including fungal lineage.[67,68] The sample sequence is aligned into a multiple sequence alignment (MSA) with other sequences, either from an in-house collection, sequences previously deposited into an online database, or accepted reference or representative sequences from either NCBI's reference sequence (RefSeq) database[69–72]or UNITE.[53] Once aligned, a phylogenetic tree is built using the sequences using various techniques including Bayesian inference and maximum likelihood comparison.[73–75] By creating a phylogenetic tree, similar sequences are grouped into clades that identify groups of similar sequences and the most recent common ancestor shared as represented by a node in the phylogenetic tree (Figure 2).

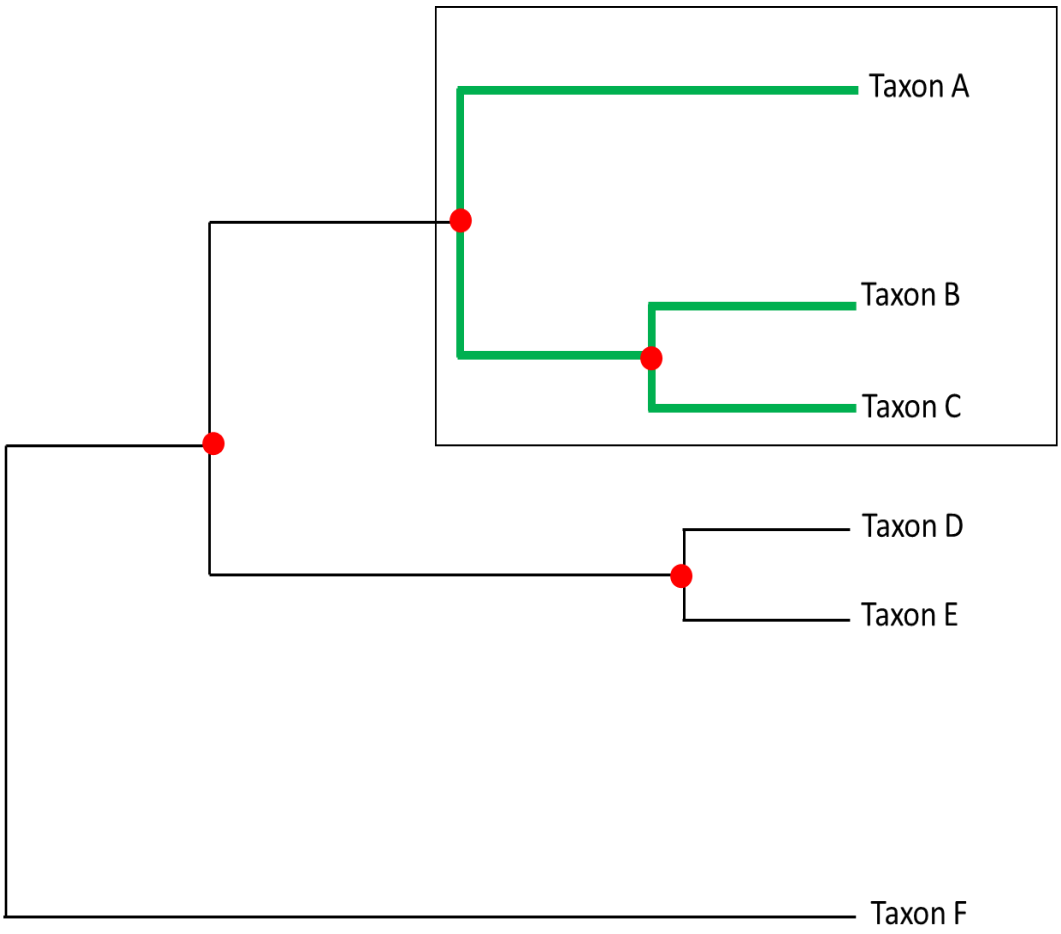

Figure 2. Depiction of a monophyletic clade forming with taxa A-C in a phylogenetic tree. A monophyletic clade includes the most recent common ancestor as depicted by a node (red), and all its descendants.

These clades, along with the distance from the most recent common ancestor, and the confidence values associated with the placement in the phylogenetic tree can be used to determine the similarity and relationships of the sequences used to build the phylogenetic tree. Placement of an unidentified sequence within this type of framework can provide valuable information as it places the previously unidentified sequence in relation to identified sequences based on the formation of various phylogenetic clades.[67,76] Using this data, it is possible to infer relationships between sequences and to gain insight into the lineage of a fungus, even when a conclusive morphological or barcoding identification is not possible. In fungi thus far, the application of DNA taxonomy has been used to help resolve several taxonomic uncertainties within the fungal kingdom while providing new insights into fungal taxonomy such as the discovery of novel genera, and the consolidation of sexual and asexual stages of the same fungus as one organism.[10,12,41,43,77]

The Natural Products Research Group (NPRG) at the University of New Brunswick Saint John (UNBSJ) has generated a fungal collection of over 400 endophytes isolated from various plant and marine algal host species. These endophytes represent a diverse group of fungi that have been subject to the many challenges associated with fungal identification. DNA had been extracted from 242 of these endophytes which were identified using DNA barcoding in conjunction with morphology prior to this project. Despite this, several sequences have returned ambiguous identifications with insufficient similarity to any sequences (≤90%) within the INSD or UNITE databases, returned no comparable identification, or have been misidentified due to the presence of poor annotations in the databases leaving the identity of fungi within the NPRG collection

unknown. As such, identification of the entirety of the NPRG collection could benefit from an alternate analysis to resolve the challenges encountered using DNA barcoding.

Objective

The objective of this project is to determine if DNA taxonomy can be used to improve identifications within an existing collection of endophytes. There are three components to the objective of this study: to use the DNA sequences from the Natural Product Research Group's (NPRG) fungal collection to compare the identification results obtained using DNA barcoding and DNA taxonomy to improve the identifications within the collection; to use a newly isolated test set of endophytes in an effort to determine if DNA taxonomy is an improvement over DNA barcoding to identify new sequences being added to the NPRG fungal collection; and finally, to determine if the addition of reference sequences to a phylogenetic tree improves the application of DNA taxonomy.

Materials and methods

General laboratory procedures and media preparation

Bacto™ malt extract media (20 g/L in distilled water), Czapek-Dox (35 g/L in distilled water), potato dextrose (24 g/L in distilled water), and cornmeal (50 g/L in distilled water) media (Difco™, Fisher Scientific, Ottawa, Ontario) were prepared in combination with dehydrated agarose crystals (14 g/L) to make solid agar media. All culture media were dissolved in distilled water and autoclaved prior to use to ensure sterility. Media were transferred to sterile Petri dishes (Fisher Scientific, Ottawa, Ontario) and stored for up to one month.

MexA agar [malt extract (20 g), yeast extract (0.1 g), glycerol (50 mL), and dehydrated agar (13 g) in distilled water (1 L)] was prepared for the cryopreservation of endophytes. Media (2.5mL) was transferred to 5mL cryopreservation tubes (VWR, Mississauga, Ontario) and the media was stored for up to one month.

Isolation of a test set of endophytes

Plant selection and collection

Four plants were selected for investigation from a plant list kept by the Natural Product Research Group for the isolation of endophytes: *Abies balsamea* (balsam fir), *Cornus canadensis* (creeping dogwood), *Juniperus communis* (common juniper), and *Vaccinium myrtilloides* (velvet-leaf blueberry). Plant leaves were collected from the University of New Brunswick, Saint John (UNBSJ) woodlot during the months of May

and July 2019. Specimens of each plant were collected, pressed, and deposited in the New Brunswick Museum Herbarium (Table 1).

Table 1. Collection data and locations for four study species

Study species	Voucher #[a]	GPS coordinates
Abies balsamea	NBM-VP-41590	N 45°30.9354 W 66°08.8801
Cornus canadensis	NBM-VP-41592	N 45°30.7417 W 66°08.8350
Juniperus communis	NBM-VP-41589	N 45°30.7550 W 66°08.1917
Vaccinium myrtilloides	NBM-VP-41591	N 45°31.0050 W 66°08.5067

[a-] NBM voucher #

Isolation of endophytic fungi

Surface sterilization of plant material

Twenty leaves from multiple plants of each species showing no apparent sign or symptoms of disease were surface sterilized. Leaf surfaces were sterilized using a combination of 70% ethanol and sodium hypochlorite at various durations of exposure. The surface sterilization conditions were adjusted until they were optimal for each plant by maximizing isolable endophytes while inhibiting the growth of epiphytes (Table 2). This was achieved by placing the leaf into a sodium hypochlorite solution (2 or 6%) (Ultra Javex® Bleach, Chlorox®) for different periods of time ranging from one to ten

seconds followed by rinsing with autoclaved distilled water. The plant tissue was then dipped into a 70% ethanol solution for durations ranging from five to 30 seconds and then rinsed with autoclaved distilled water before blotting the leaf dry on autoclaved paper towels. Once sterilized, a verification plate was made for each leaf by rubbing ("imprinting") the leaf on the surface of 2% malt extract agar (MEA) to test for the presence of epiphytes. After imprinting, the leaf was then cut with a sterile scalpel into segments (approximately 0.5 cm x 0.5 cm) and placed onto fresh 2% MEA. Successful isolation of endophytes was indicated by hyphal growth from the cut edges of surface sterilized plant tissue in conjunction with a verification plate where no epiphytic growth was observed. Isolation frequencies were determined for each plant by dividing the number of endophytes isolated from the species by the total number of leaf segments that underwent surface sterilization procedures and multiplying by 100.[28]

Table 2. Surface sterilization procedure of four plant species

Plant host	Optimal Surface sterilization procedure
Abies balsamea	Bleach (6%) with Tween 80 (1 second), ethanol (10 seconds)
Cornus canadensis	Ethanol with Tween 80 (30 seconds)
Juniperus communis	Bleach (2%) (5 seconds), ethanol (5 seconds)
Vaccinium myrtilloides	Bleach (2%) with Tween 80 (10 seconds), ethanol (10 seconds)

Endophyte subculturing

For each plant species, five surface sterilized plant pieces from the same leaf were placed in the same Petri plate, which was repeated for all of 80 leaves (20 leaves/plant).

Petri plates were incubated at room temperature in ambient light and monitored daily for the presence of fungal growth from the cut edges of surface sterilized plant segments. Once endophytic hyphae were observed, a small portion of hyphal growth was cut from the growing edge of the colony and transferred onto fresh 2% MEA. Further subcultures were performed as required until a pure culture was obtained for each isolate. Pure cultures of isolates were then transferred onto four different media types [potato dextrose agar (PDA), Czapek-Dox agar (CZA), cornmeal agar (CMA), and MEA] to aid in plate dereplication. Distinct fungal isolates from each host plant were identified through dereplication of fungal isolates showing similar morphology on all four media. Taxonomic identification was performed on all isolates that produced conidia, conidiophores, or other identifiable structures.

Cryopreservation of endophytes

Long-term storage of endophytes was achieved through cryopreservation. Cryopreservation slants were prepared by transferring fungal material onto slants made of MexA agar. Slants were incubated at room temperature until endophytic growth was observed (approximately two weeks), and then frozen in a stepwise procedure in which the isolate was chilled in a refrigerator (4℃) for two hours, transferred to a freezer (-20℃) for two hours, and finally held in permanent storage (-80℃).

Molecular identification of fungi

The internal transcribed spacer (ITS) region was used for the molecular identification of each isolate. DNA retrieval comprised four steps: DNA extraction, DNA amplification, gel electrophoresis, and DNA sequencing.

DNA extraction

Following dereplication, the genomic DNA of all fungal isolates was extracted using a DNEasy® Plant Mini Kit (QIAGEN, Toronto, ON, Canada) following the manufacturer's instructions. Extraction was attempted a maximum of two times using the Plant Mini Kit, followed by one attempt using a ThermoFisher Scientific Yeast Kit following manufacturer's directions if necessary.

Polymerase chain reaction (PCR)

PCR amplification of the extracted fungal DNA was performed using a Hotstar taq plus master mix (QIAGEN, Toronto, Ontario) and ITS1F and ITS4 universal fungal primers (Invitrogen, Burlington, Ontario; Table 3). For each DNA sample, the PCR reagent mixture was prepared as follows: 32.5 µL Hotstar taq plus master mix (DNA taq polymerase, PCR buffer, and dinucleotide base pairs), 16.5 µL RNA-free water (QIAGEN), 6.5 µL RediLoad, 4 µL of ITS1F (10 pmol/µL), 4 µL ITS4 (10 pmol/µL), and 3 µL of fungal DNA template. The mixture was added to individual PCR tubes (0.2 mL 12-strip, VWR) and amplified in a MyCycler Thermal Cycler (Bio-Rad Laboratories, Ltd., Mississauga, Ontario). The protocol for amplification was as follows: one activation step (95℃, five minutes), followed by 35 cycles of the three step process of denaturation of DNA (95℃, one minute), annealing of primers (52℃, 30 seconds), and the extension of new DNA strands (72℃, one minute), finishing with a final extension of DNA (72℃, ten minutes). Amplified DNA was stored at -20℃ until needed.

Table 3. Forward (ITS1F) and reverse (ITS4) primers

Primer	DNA sequence	Amplified rDNA region
ITS1F	5'- CTTGGTCATTTAGAGGAAGTAA - 3'	5.8S
ITS4	5'- TCCTCCGCTTATTGATATGC - 3'	5.8S

Gel electrophoresis

Once fungal DNA was obtained and amplified, each PCR product was verified through gel electrophoresis to confirm the presence of a DNA band correlating to the ITS region (approximately 550-650 base pairs). Agarose gels were prepared [(0.5g agarose in 50 mL 1X Tris Borate EDTA (TBE) buffer, 5 µL SYBR®Safe (Invitrogen, Burlington, Ontario)], and electrophoresis (120V, 25 minutes) was performed by adding PCR product (15 µL) and DNA ladder (6 µL; low mass ladder, 100-2000 bp; Invitrogen, Burlington, Ontario) to the agarose gel wells. Detection of DNA bands was performed using a gel imaging system (Gel Doc™ XR+, Bio Rad; Mississauga, Ontario). Verified PCR products (20 µL) were submitted to Génome Québec (Montréal, Québec, Canada) with copies of the ITS1F and ITS4 primers (10 µL; 5 µM) for sequencing of the nucleotide base pairs.

Fungal sequence retrieval

DNA from 242 of the fungal isolates in the NPRG collection had been previously extracted by members of the NPRG. Their ITS regions were amplified through polymerase chain reaction (PCR) using both a forward (ITS1 or ITS1F) and reverse (ITS4) sequence primer and the resultant amplicons were sent to Génome Québec for

sequencing. All forward and reverse DNA sequences were retrieved from Génome Québec and downloaded in Fast-all (FASTA) format.

Consensus sequence building

Consensus sequences were built using the BioEdit Sequence Alignment Editor©.[78]The forward (ITS1 or ITS1F) and reverse (ITS4) DNA sequences of an isolate were opened using the BioEdit software. A pairwise alignment was created to overlap the sequences and the resultant alignment was used to create a consensus sequence.

DNA barcoding of fungal sequences

DNA barcoding results were obtained for all isolates in the NPRG fungal collection by conducting a massBLASTer analysis using the PlutoF platform to search sequences within the UNITE database and all collaborators of the International Nucleotide Sequence Database (INSD) collaboration; GenBank, the European Nucleotide Archive (ENA), and the DNA DataBank of Japan (DDBJ). Program parameters were set to show one alignment and the algorithm parameter was set to similar sequences. Isolates identified to the species level had ≥99% sequence similarity and an E-value of zero, while isolates identified to the genus level had ≥96% sequence similarity and an E-value of zero. Isolates that returned sequence similarity ≥92% were identified to the class level, while isolates with <92% similarity were identified to the phyla level. Identification thresholds were selected based on the existing ranges present in the literature.[65,66]

Reference sequence retrieval

One hundred and forty-one reference and representative sequences were downloaded from the UNITE database general FASTA release[79] based on the DNA barcoding identifications. Where possible, reference sequences were selected preferentially to representative sequences.

Multiple sequence alignment

A multiple sequence alignment (MSA) was built using the consensus sequences. This was done using the Multiple Sequence Comparison by Log-Expectation (MUSCLE) alignment tool[80,81] in the Molecular Evolutionary Genetics Analysis version X (MEGA X) software.[82] The alignment was run with a maximum of 16 iterations and UPGMA (Unweighted Pair Group Method with Arithmetic mean) clustering method.

Phylogenetic tree building and annotation

The MSA was analyzed in Molecular Evolutionary Genetics analysis version X software to conduct phylogenetic analyses. A model of best fit analysis was conducted for the dataset and a phylogenetic tree was built using the maximum likelihood method and the Tamurai-Nei model as indicated by the best fit analysis with 500 bootstrap replicates.[83] Upon completion, the tree file was exported for annotation. The file was visualized using the Interactive Tree of Life (iTOL) software,[84] and sequences were annotated on the tree using identifications from the DNA barcoding results.

Phylogenetic analysis of fungal ITS sequences

Sequences for the test set were retrieved from Genome Quebec, and consensus sequences were created following the same procedure as outlined above under Consensus Sequence Building for the existing fungal collection. The consensus sequences were aligned with the identified sequences from the NPRG fungal collection, and reference and representative sequences downloaded from UNITE in an MSA, and a phylogenetic tree was built using MEGA X software.

Once the phylogenetic tree was constructed, a Newick export was uploaded into the iTOL software for analysis and annotation. The test-set of endophytes were identified using the previously annotated collection in conjunction with supportive bootstrap values (≥0.50),[85–87] and, the identifications were used to verify the DNA barcoding identifications.

DNA barcoding results were obtained for the test set by conducting an online massBLASTer search through the PlutoF platform to search the INSD and UNITE databases using the fungal sequences. Isolates identified to the species level had ≥99% sequence similarity with an e-value of zero, while isolates identified to the genus level shared ≥96% sequence similarity with an e-value of zero. Isolates that returned sequence similarity of ≥92% were identified to the class level, and finally, isolates with <92% similarity were identified to the phyla level. Identification results were confirmed using morphology and compared to those obtained using the DNA taxonomy approach, and the phylogenetic tree annotations were updated to reflect the most accurate identification.

Sequences were deposited to the NCBI GenBank database with the identifications obtained, and each sequence was assigned an accession number by NCBI.

Results

Isolation of endophytes

One hundred plant segments were used for each of *Abies balsamea*, *Cornus canadensis*, *Juniperus communis*, and *Vaccinium myrtilloides*. One hundred and nine endophytes grew from the 400 segments which resulted in a total of 62 morphologically distinct fungal isolates. The endophytes from *Vaccinium myrtilloides* had the highest isolation frequency (65%), followed by *Cornus canadensis* (22%) and *Juniperus communis* (16%), while *Abies balsamea* displayed the lowest isolation frequency (6%) (Figure 3). The number of distinct fungal isolates varied between hosts with six isolates from *Abies balsamea*, 13 isolates from *Cornus canadensis*, 11 isolates from *Juniperus communis*, and 32 isolates from *Vaccinium myrtilloides* (Figure 3).

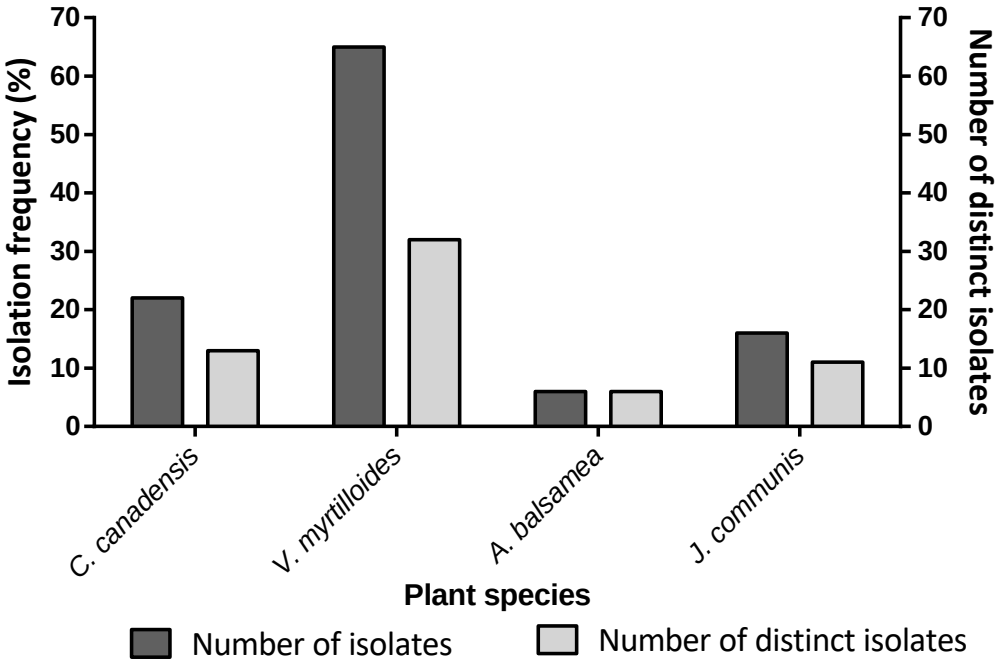

Figure 3. Isolation frequencies and number of distinct endophytes from each of the four study species.

Identification of endophytes

DNA was successfully extracted and amplified, as indicated by a visible band in gel electrophoresis (Figure 4), for 58 of the 62 isolates and sequences of the ITS region were obtained from Génome Québec for all 58 isolates (Appendix 1).

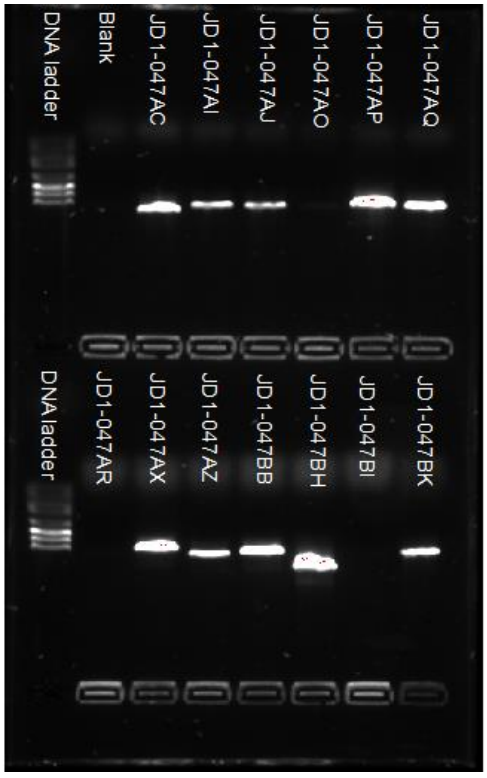

Figure 4: Example of an electrophoresed gel displaying DNA ladder (left in both upper and lower lanes), followed by a blank PCR control, DNA samples showing a visible ITS band as well as DNA samples showing no band (JD1-047AO, JD1-047AR, JD1-047BI) indicating the absence of nucleotide base pairs.

The resultant phylogenetic tree from the taxonomic analysis included 441 fungal ITS sequences; two hundred and forty-two from the existing NPRG fungal collection, 58 new sequences being added to the collection, and 141 reference or representative sequences. The final phylogenetic tree resulted in the formation of twelve major clades based on the similarity between sequences (Figure 5).

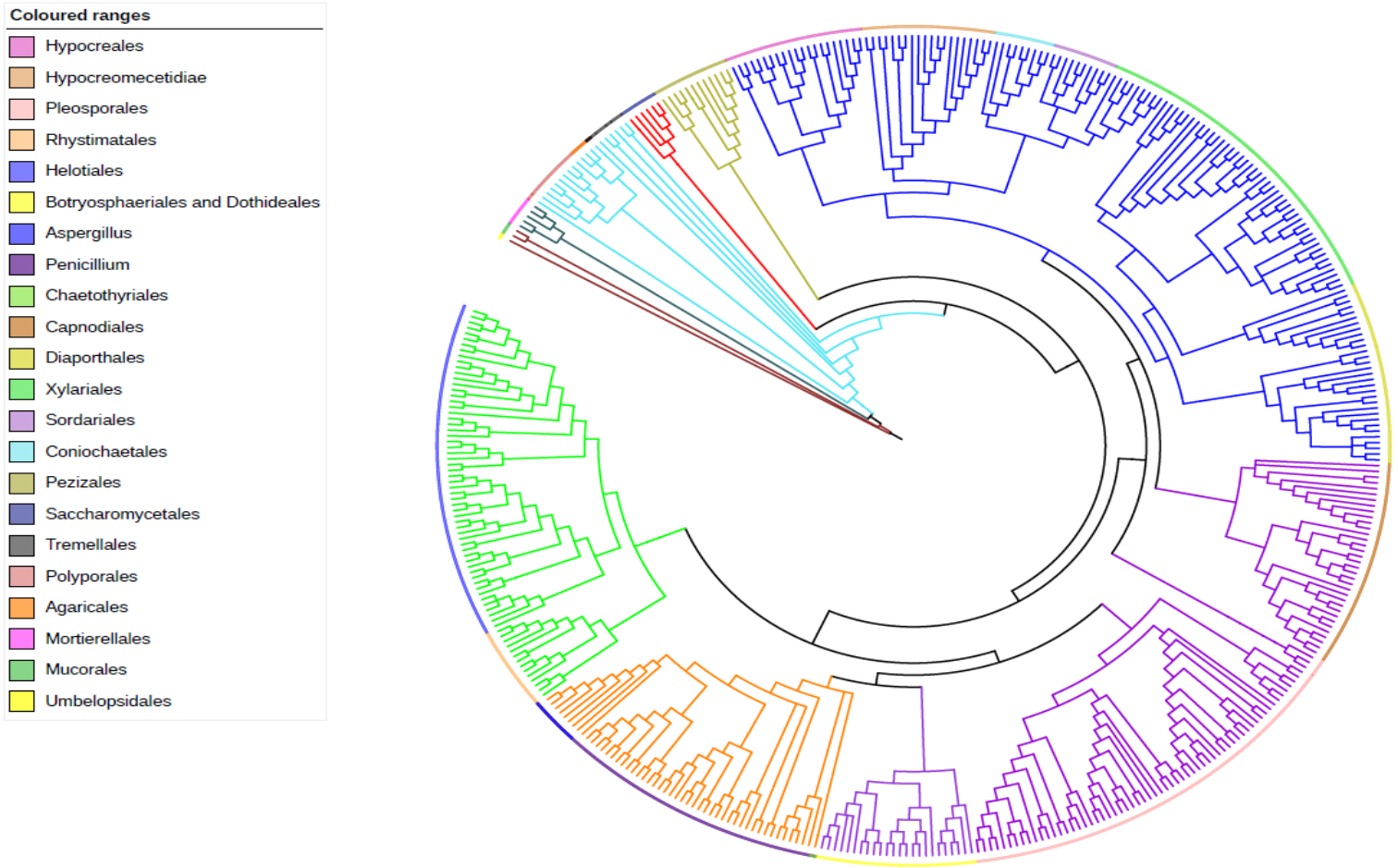

Figure 5. Phylogenetic tree of 441 fungal ITS sequences built using a maximum-likelihood approach with 500 bootstrap replicates with exterior colour ranges depicting orders and genera of fungi within the phylogenetic tree.

The identifications of 23% of the sequences within the NPRG fungal collection were resolved using DNA taxonomy. Forty-five sequences (19%) within the NPRG fungal collection were further identified to the species, genus, or order level resulting in an identification to a lower classification level than what was achieved using DNA barcoding (Table 4). A further ten sequences (4%) from the NPRG fungal collection were found to have been misidentified using DNA barcoding as indicated by their placement within the phylogenetic tree (Table 5). Additionally, 47 isolates from the test set were successfully identified based on groupings with the NPRG fungal sequences and reference or representative sequences, while eleven sequences could not be clearly identified using DNA taxonomy alone (Table 6). Of the 47 identified isolates, four were identified to the species level, 42 were identified to the genus level, and one was identified to the order level in this work (Table 6).

Table 4. Identification of 45 sequences from the NPRG fungal collection using DNA taxonomy.

Isolate code	DNA barcoding identification	DNA taxonomy identification	Final assigned identification
BRD2-125G	Helotiales	*Oidodendron* sp.	*Oidodendron* sp.
BRD2-149Y	*Phialocephala* sp.	*Phialocephala fortinii*	*Phialocephala fortinii*
BRD3-082C	*Hypoxylon* sp.	*Hypoxylon perforatum*	*Hypoxylon perforatum*
BRD3-083W	Ascomycete	Helotiales	Helotiales
BRD3-083Y	Pleosporales	*Leptospora rubella*	*Leptospora rubella*
BRD3-113P	Ascomycete	*Lophodermium* sp.	*Lophodermium* sp.
BRD3-113Y	*Chaetomium* sp.	*Chaetomium globosum*	*Chaetomium globosum*
BRD3-121F	Ascomycete	Helotiales	Helotiales
JAJ1-033	*Paraphaeosphaeria* sp.	*Paraphaeosphaeria neglecta*	*Paraphaeosphaeria neglecta*
JAJ1-057	Pleosporales	*Paraphaeosphaeria neglecta*	*Paraphaeosphaeria neglecta*
JAJ1-065	Dothideomycetes	*Mycosphaerella handelii*	*Mycosphaerella handelii*
JAJ1-117	Dothideomycetes	*Mycosphaerella handelii*	*Mycosphaerella handelii*
JAJ1-141	*Paraphaeosphaeria* sp.	*Paraphaeosphaeria sporulosa*	*Paraphaeosphaeria sporulosa*
JAJ1-149	Dothideales	*Endoconidioma populi*	*Endoconidioma populi*
KP1-017A	Helotiales	*Botryotinia* sp.	*Botryotinia* sp.
KP1-045K	*Dendryphiella* sp.	*Dendryphiella salina*	*Dendryphiella salina*
KP1-089A	Helotiales	*Botryotinia* sp.	*Botryotinia* sp.
KP1-119C	*Paraconiothyrium* sp.*	*Paraphaeosphaeria** *sporulosa*	*Paraphaeosphaeria sporulosa*
KP1-131B	Helotiales	*Mollisia* sp.	*Mollisia* sp.
KP1-131BB	Helotiales	Hyaloscyphaceae	Hyaloscyphaceae
KP1-131DA	*Thysanophora* sp.	*Thysanophora penicilloides*	*Thysanophora penicilloides*
KP1-131F	*Mollisia* sp.	*Mollisia fusca*	*Mollisia fusca*

*-Fungal synonyms

Table 4. Continued

Isolate Code	DNA barcoding identification	DNA taxonomy identification	Final assigned identification
KP1-131N	Dothideomycetes	*Paraconiothyrium brasiliense*	*Paraconiothyrium brasiliense*
KP1-131R	*Mollisia* sp.	*Mollisia fusca*	*Mollisia fusca*
KP1-131Z	Peniohoraceae	*Cryptococcus* sp.	*Cryptococcus* sp.
KP1-135B	*Trichoderma* sp.	*Trichoderma ciride*	*Trichoderma ciride*
KP1-135D	*Dendryphiella* sp.	*Dendryphiella salina*	*Dendryphiella salina*
KP2-009B	*Trametes* sp.	*Trametes versicolor*	*Trametes versicolor*
KP2-033D	Dothideomycetes	*Lophiostoma* sp.	*Lophiostoma* sp.
KP2-033E	Dothideomycetes	*Lophiostoma sp.*	*Lophiostoma sp.*
TC2-026	Ascomycete	*Drechslera* sp.	*Drechslera* sp.
TC2-035	Ascomycete	*Mollisia* sp.	*Mollisia* sp.
TC2-048	*Paraconiothyrium* sp.*	*Paraphaeosphaeria* neglecta*	*Paraphaeosphaeria neglecta*
TC2-049	Pleosporales	*Leptospora* sp.	*Leptospora* sp.
TC2-051	Ascomycete	*Drechslera* sp.	*Drechslera* sp.
TC2-052	*Phoma* sp.	*Phoma glomerata*	*Phoma glomerata*
TC2-053	Ascomycete	*Leptospora* sp.	*Leptospora* sp.
TC2-054	*Nigrospora* sp.	*Nigrospora oryzae*	*Nigrospora oryzae*
TC2-058	*Pestalotiopsis* sp.	*Pestalotiopsis microspora*	*Pestalotiopsis microspora*
TC2-082	*Ascomycete*	*Botryotinia* sp.	*Botryotinia* sp.
TC2-086	*Glomerella* sp.*	*Collototrichum* fioriniae*	*Collototrichum fioriniae*
TC2-089	Leotiomycetes	*Spaerographium* sp.	*Spaerographium* sp.
TC2-091	Fungi	*Phoma glomerata*	*Phoma glomerata*
TC3-002	*Epicoccum* sp.	*Epicoccum nigrum*	*Epicoccum nigrum*
TC3-005	Ascomycete	*Mycosphaerella* sp.	*Mycosphaerella* sp.

***-**Fungal synonyms

Table 5. Corrected identification of ten sequences from the NPRG fungal collection as obtained from DNA taxonomy.

Isolate code	DNA barcoding identification	DNA taxonomy identification	Thesis identification
JAJ1-105	*Paraphaeosphaeria neglecta*	*Mollisia* sp.	*Mollisia* sp.
KP1-119E	*Aureobasidium* sp.	*Coniochaeta* sp.	*Coniochaeta* sp.
KP1-123C	*Chrysosporium* sp.	*Paraphaeosphaeria* sp.	*Paraphaeosphaeria* sp.
KP1-131DB	*Ophiognomonia intermedia*	*Plagiostoma* sp.	*Plagiostoma* sp.
KP1-131H	*Microdochium bolleyi*	*Hypoxylon* sp.	*Hypoxylon* sp.
KP2-005A	*Phanerochaete* sp.	*Piptoporus betulinus*	*Piptoporus betulinus*
TC2-012	*Lophodermium seditiosum*	*Colletotrichum acutata*	*Colletotrichum acutata*
TC2-013	*Phacidium sp.*	*Lophodermium* sp.	*Lophodermium* sp.
TC2-029	*Seimatosporium* sp.	*Pestalotiopsis* sp.	*Pestalotiopsis* sp.
TC2-047	*Sporidiobolus* sp.	*Cryptococcus* sp.	*Cryptococcus* sp.

Table 6. Identification of newly added endophytic fungi using molecular identification

Isolate Code	DNA taxonomy identification	DNA barcoding identification	Final assigned identification
JD1-017A	*Phialocephala* sp.	*Phialocephala piceae*	*Phialocephala piceae*
JD1-017B	*Diaporthe* sp.	*Diaporthe* sp.	*Diaporthe* sp.
JD1-017C	*Hypoxylon* sp.	Ascomycete	*Hypoxylon* sp.
JD1-017E	*Podospora* sp.	Ascomycete	*Podospora* sp.
JD1-017F	■	*Pseudoplectania nigrella*	*Pseudoplectania nigrella*
JD1-017G	*Trametes* sp.	Basidiomycete	*Trametes* sp.
JD1-047AA	*Phyllosticta spinarum*	*Phyllosticta* sp.	*Phyllosticta spinarum*
JD1-047AC	■	*Pseudoplectania nigrella*	*Pseudoplectania nigrella*
JD1-047AI	*Diaporthe* sp.	*Diaporthe* sp.	*Diaporthe* sp.
JD1-047AJ	*Plagiostoma* sp.	*Plagiostoma* sp.	*Plagiostoma* sp.
JD1-047AN	*Mollisia* sp.	Ascomycete	*Mollisia* sp.
JD1-047AO	■	*Preussia* sp.	*Preussia* sp.
JD1-047AP	*Lecanicillium attenuatum*	*Lecanicillium* sp.	*Lecanicillium attenuatum*
JD1-047AQ	*Plagiostoma* sp.	*Plagiostoma* sp.	*Plagiostoma* sp.
JD1-047AR	*Plagiostoma* sp.	*Plagiostoma* sp.	*Plagiostoma* sp.
JD1-047AX	*Preussia* sp.	Ascomycete	*Preussia* sp.
JD1-047AZ	*Plagiostoma* sp.	*Plagiostoma* sp.	*Plagiostoma* sp.
JD1-047BB	*Physalospora* sp.	*Coniochaeta* sp.	*Physalospora* sp.
JD1-047BH	■	*Phialocephala* sp.	*Phialocephala* sp.
JD1-047BJ	*Diaporthe* sp.	*Diaporthe* sp.	*Diaporthe* sp.
JD1-047BK	*Plagiostoma* sp.	*Plagiostoma* sp.	*Plagiostoma* sp.

■ - identification not possible using DNA taxonomy alone

*- Fungal synonyms

Table 6. Continued

Isolate Code	DNA taxonomy identification	DNA barcoding identification	Final assigned identification
JD1-047BM	Order Helotiales	Ascomycete	Order Helotiales
JD1-047BN	*Mollisia* sp.*	*Neomollisia* sp.*	*Mollisia* sp.
JD1-047D	*Phomopsis* sp.*	*Diaporthe* sp.*	*Diaporthe* sp.
JD1-047E	*Coniochaeta* sp.	*Coniochaeta* sp.	*Coniochaeta* sp.
JD1-047F	■	*Pseudoplectania nigrella*	*Pseudoplectania nigrella*
JD1-047G	*Hypoxylon* sp.	*Hypoxylon truncatum*	*Hypoxylon truncatum*
JD1-047K	*Colletotrichum* sp.	*Colletotrichum* sp.	*Colletotrichum* sp.
JD1-047L	*Diaporthe* sp.	*Diaporthe* sp.	*Diaporthe* sp.
JD1-047O	*Diaporthe* sp.	*Diaporthe* sp.	*Diaporthe* sp.
JD1-047P	*Hypoxylon* sp.	*Hypoxylon* sp.	*Hypoxylon* sp.
JD1-047Q	*Hypoxylon fuscum*	*Hypoxylon* sp.	*Hypoxylon fuscum*
JD1-047S	*Diaporthe* sp.	*Diaporthe* sp.	*Diaporthe* sp.
JD1-047W	*Paraphaeosphaeria* sp.	*Stagonospora* sp.	*Paraphaeosphaeria* sp.
JD1-047X	*Plagiostoma* sp.	Ascomycete	*Plagiostoma* sp.
JD1-047Y	■	*Preussia* sp.	*Preussia* sp.
JD1-047Z	*Diaporthe* sp.	*Diaporthe* sp.	*Diaporthe* sp.
RAS1-015A	*Pestalotiopsis* sp.	*Pestalotiopsis* sp.	*Pestalotiopsis* sp.
RAS1-015B	*Phomopsis* sp.*	*Diaporthe* sp.*	*Diaporthe* sp.
RAS1-015H	*Pezicula* sp.	*Pezicula* sp.	*Pezicula* sp.
RAS1-015K	*Lophodermium* sp.	Ascomycete	*Lophodermium* sp.
RAS1-015L	*Lophodermium* sp.	Ascomycete	*Lophodermium* sp.

■ - identification not possible using DNA taxonomy alone

*- Fungal synonyms

Table 6. Continued

Isolate Code	DNA taxonomy identification	DNA barcoding identification	Final assigned identification
RAS1-015P	*Lophodermium* sp.	Ascomycete	*Lophodermium* sp.
RAS1-015Q	■	*Phialocephala* sp.	*Phialocephala* sp.
RAS1-015R	*Alternaria* sp.	*Phialocephala* sp.	*Phialocephala* sp.
RAS1-015U	■	*Phyllosticta* sp.	*Phyllosticta* sp.
SS1-033A	*Coniochaeta* sp.	*Coniochaeta* sp.	*Coniochaeta* sp.
SS1-033B	*Coniochaeta* sp.	*Coniochaeta* sp.	*Coniochaeta* sp.
SS1-033C	*Hypoxylon* sp.	*Hypoxylon* sp.	*Hypoxylon* sp.
SS1-033E	*Pestalotiopsis* sp.	*Pestalotiopsis* sp.	*Pestalotiopsis* sp.
SS1-033H	*Neofabraea* sp.	*Neofabraea actinidiae*	*Neofabraea actinidiae*
SS1-033I	*Pestalotiopsis microspora*	*Pestalotiopsis* sp.	*Pestalotiopsis microspore*
SS1-033L	*Mycosphaerella* sp.	Fungi	*Mycosphaerella* sp.
SS1-033M	■	*Capronia* sp.	*Capronia* sp.
SS1-033N	■	Fungi	*Pseudoplectania nigrella*
SS1-033P	*Phyllosticta* sp.	*Phyllosticta* sp.	*Phyllosticta* sp.
SS1-033R	■	Fungi	*Pseudoplectania nigrella*
SS1-033S	*Plagiostoma* sp.	*Plagiostoma* sp.	*Plagiostoma* sp.

■ - identification not possible using DNA taxonomy alone

*-Fungal synonyms

The results obtained from DNA taxonomy and DNA barcoding differed for four of the new isolate sequences from the test set; JD1-047BB, SS1-033L, JD1-047W, and RAS1-015R. Three of these sequences; JD1-047BB, SS1-033L, and JD1-047W were identified as *Physalospora* sp., *Mycosphaerella* sp., and *Paraphaeosphaeria* sp., respectively based on their location in the phylogenetic tree (Figures 6-8). The fourth sequence, RAS1-015R was identified as *Phialocephala* sp. using the DNA barcoding results.

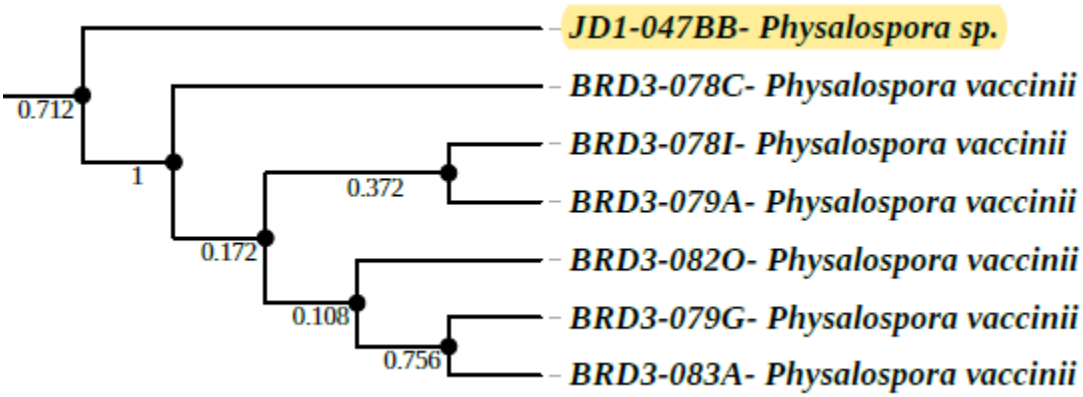

Figure 6. Phylogenetic placement of JD1-047BB leading to the identification of *Physalospora* sp.

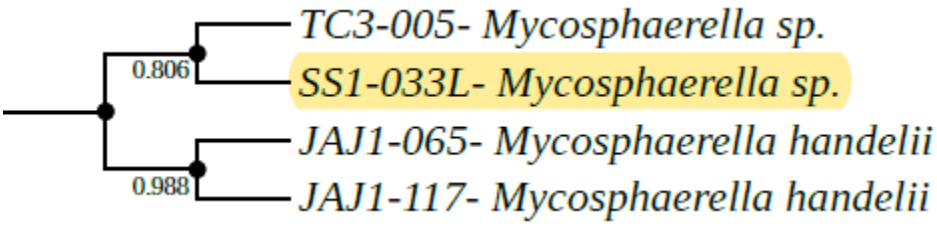

Figure 7. Phylogenetic placement of SS1-033L leading to the identification of *Mycosphaerella* sp.

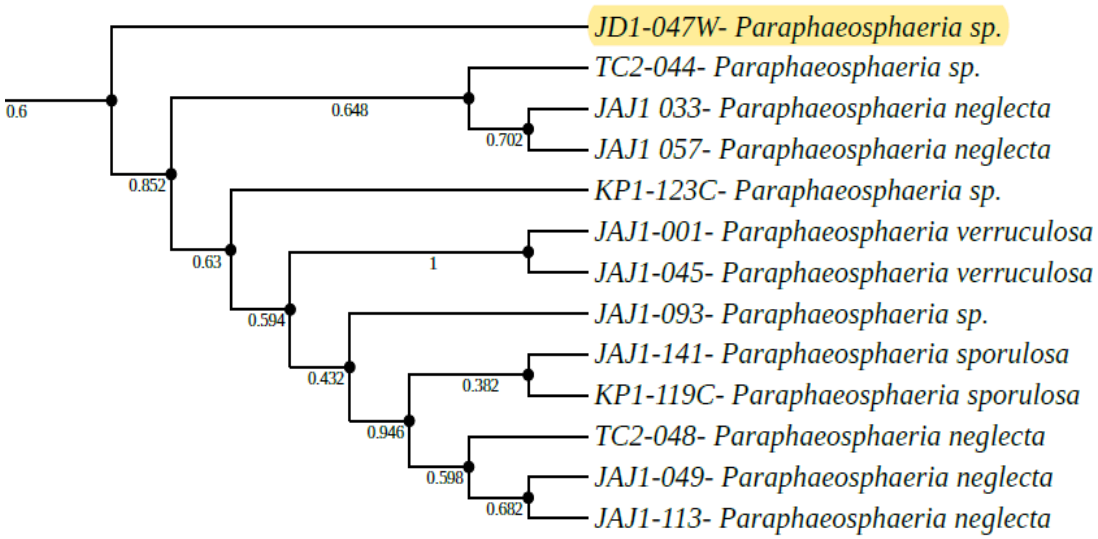

Figure 8. Phylogenetic placement of JD1-047W leading to the identification of *Paraphaeosphaeria* sp.

Finally, the four sequences that were identified to the species level using DNA taxonomy (JD1-047AA, JD1-047AP, JD1-047Q, and SS1-033I) could only be identified to the genus level using DNA barcoding as no results were obtained that met the species threshold of ≥99% (Table 7). Comparatively, DNA barcoding was able to identify three of the sequences to species level (SS1-033H, JD1-047G, and JD1-017A) that could only be identified to the genus level using DNA taxonomy (Table 8).

Table 7. Identification of sequences to the species level using DNA taxonomy.

Isolate Code	DNA taxonomy identification	DNA barcoding identification	Final assigned identification
JD1-047AA	*Phyllosticta spinarum*	*Phyllosticta* sp.	*Phyllosticta spinarum*
JD1-047AP	*Lecancillium attenuatum*	*Lecancillium* sp.	*Lecancillium attenuatum*
JF1-047Q	*Hypoxylon fuscum*	*Hypoxylon* sp.	*Hypoxylon fuscum*
SS1-033I	*Pestalotiopsis microspora*	*Pestalotiopsis* sp.	*Pestalotiopsis microspora*

Table 8. Identification of sequences to the species level using DNA barcoding.

Isolate Code	DNA taxonomy identification	DNA barcoding identification	Final assigned identification
JD1-017A	*Phialocephala* sp.	*Phialocephala piceae*	*Phialocephala piceae*
JD1-047G	*Hypoxylon* sp.	*Hypoxylon truncatum*	*Hypoxylon truncatum*
SS1-033H	*Neofabraea* sp.	*Neofabraea actinidiae*	*Neofabraea actinidiae*

Discussion

While the terms DNA barcoding and DNA taxonomy first appeared in the literature in 2003,[33,34,88] the use of DNA-based identification systems, and particularly DNA-based phylogenies is not a new concept. Work in various fields have been making use of the available technology for several decades,[89–93] and in the 1990's, DNA-based phylogenies were integral in showing the true extent of prokaryotic diversity and illustrating the division of cellular life as being divided into three domains; Archaea, Bacteria, and Eukarya.[68,94] The molecular investigation of most organisms continues to show the existence of previously unknown families, genera, and species, contributing to the classification system on a large scale.[95–100] Microbial investigation especially has benefitted from the application of DNA-based identification, given the limited and often indistinguishable morphological traits rendering an identification based on morphological data alone nearly impossible.[29,101] This includes the group of organisms known as fungal endophytes; a diverse group of species whose richness was estimated as 0.5-1 million species[102,103] representing a significant portion of the 1.5 to 5.1 million possible species that comprise the fungal kingdom,[11,13,14] and the organism of interest in this thesis.

Endophytes display a wide range of functions, from the observed association and interactions between plants and fungi,[104–108] to the production of bioactive metabolites also known as natural products.[109,110] The latter has been the focus of the Natural Products Research Group, leading to the generation of a fungal collection of over 400 endophytes enabling the group's discovery of new and novel chemical structures often with bioactive properties. Despite the success in isolating natural products, identifying and describing the endophytes themselves presented a challenge that is all too common in

fungal research, especially as it pertains to natural products. In fact, a publication in the Journal of Natural Products in 2017 outlined fungal identification techniques, strongly encouraging those working with endophytes to provide verified identifications before publishing their work.[24] All endophytes within the NPRG fungal collection have been described based on observed morphology and plate culture characteristics and 242 had been previously identified using DNA barcoding in conjunction with morphology to verify the DNA-based identifications. While these efforts improved the collection in terms of identifying the isolates, several sequences remained ambiguous or were identified to a high taxonomic level (e.g. Phylum) due to insufficient morphological data and comparable sequence data within the search databases. DNA taxonomy had not been applied to the NPRG fungal collection prior to this work, and literature describing the approach on similar collections is limited. Most work using DNA taxonomy for fungi has focused on the fungal kingdom as a whole to resolve complicated fungal lineages and to consolidate placements of several major fungal groups.[10,12,41,43,77]

By applying DNA taxonomy to the NPRG fungal collection, it was possible to create a more comprehensive classification of the collection by being able to determine relationships between endophyte sequences. Additionally, the sequences within the fungal collection that resulted in ambiguous DNA barcoding results were identified to the genus or species level based on the groupings that were observed in the phylogenetic output, resulting in a successful application of DNA taxonomy. To verify the application of DNA taxonomy to classify and potentially identify new sequences in relation to the existing fungal collection, a test set was generated by isolating more endophytes, and 62 fungal endophytes were added to the NPRG fungal collection. Of the test set of

endophytes, genomic DNA was successfully extracted from 58 endophytes resulting in ITS sequences and it was possible to use the data in the phylogenetic tree to identify 47 of the sequences to the species (4), genus (42), or order (1) level. While obtaining a species level identification is important, complex phylogenetic relationships as well as the presence of cryptic species in some genera are both possible explanations for the limited ability to identify sequences to the species level.[44,111] Furthermore, despite the ITS region being the standard barcode marker for fungi, species level identifications using the ITS barcode alone has been difficult in several genera[112–116] which is the most likely explanation for the few species level results obtained in this analysis. By applying DNA taxonomy to the NPRG fungal collection, it was possible to create a more comprehensive understanding of the collection by being able to determine relationships between endophyte sequences. Additionally, the sequences within the fungal collection that resulted in ambiguous DNA barcoding results were identified to the genus or species level based on the groupings that were observed in the phylogenetic output, resulting in a successful application of DNA taxonomy.

There were eleven sequences that could not be identified using the DNA taxonomic approach alone on this dataset as the sequences were of similar distance and relation to multiple identified sequences or did not form a clade with identified sequences from the fungal collection. However, the combined efforts of DNA barcoding and the inclusion of references sequences allowed for the identification of five sequences as *Pseudoplectania nigrella*, a fungus that was not previously identified within the NPRG fungal collection, and six sequences belonging to various genera.

We identified twelve major clades ranging in size from three to 143 sequences in the phylogenetic output, each representing distinct groups of endophytes. One clade represented the phylum Basidiomycota and comprised 20 sequences, two clades represented the subphyla Mortierellomycotina and Mucororomycotina of the Zygomycota phylum with five and three sequences respectively, and the remaining nine clades and 413 sequences represented the phylum Ascomycota. This distribution of fungal sequences is in line with current research on the fungal kingdom with most identified fungi being Ascomycetes followed by Basidiomycetes. This distribution is also a reflection of current endophyte research as endophytes isolated from above-ground plant tissues have been found to be incredibly diverse[110] but belong primarily to the Dikarya subkingdom (Ascomycota and Basidiomycota), and particularly to the Ascomycota phylum.[110,117,118] The remaining nine clades represented six classes within the phylum Ascomycota: Eurotiomycetes, Leotiomycetes, Dothideomycetes, Sordariomycetes, Pezizomycetes, and Saccharomycetes. The largest of these clades comprised 143 sequences and represented the Sordariomycetes and was followed closely by the Dothideomycetes which comprised 123 sequences. While the class Sordariomycetes is considered to be one of the largest classes in the phylum Ascomycota,[119] the class Dothideomycetes is the largest and most diverse class of ascomycete fungi[43,120] and the slight difference between the two classes within the NPRG fungal collection is in contrast to what would be expected of the class level distribution.

The inclusion of reference sequences in the phylogenetic tree made it possible to distinguish smaller clades within the twelve major clades which represented various orders or genera of fungi. Although most classes were further resolved as representing

one or two orders of fungi, the classes Sordariomycetes and Dothideomycetes comprised six and four orders of fungi, respectively. This is again somewhat surprising given what is known about fungal diversity within the two classes. In one case, the clade which comprised the class Eurotiomycetes was further resolved to the genus level as *Aspergillus* and *Penicillium* sequences. In all cases, the ability to identify major classes, orders, and genera of fungal sequences allowed for a comprehensive classification of the NPRG fungal collection that had not previously been obtained.

Although it was not possible to identify all sequences to the species level, using DNA taxonomy resulted in valuable information regarding the relationship of sequences to one another that would not have been possible without DNA taxonomy. Within the NPRG fungal collection, 33 sequences represent various *Penicillium* species, with only seven having been successfully identified to the species level. In the phylogenetic clade for the *Eurotiales* fungi from this analysis, it is possible to see, for example, that the isolates KP1-013B and BRD2-118C are highly related having formed their own monophyletic clade, whereas KP1-013B is more distantly related to other *Penicillium* isolates, such as KP1-123A (Figure 9). While all three sequences have been identified as *Penicillium* sp., it is possible to conclude that isolates BRD2-118C and KP1-013B are likely the same species, where as isolate KP1-123A is more likely to be a different *Penicillium* species. In addition to revealing more information about an organism, the ability to infer bioactivity and chemical properties based on a closely related sequence, may be useful in targeting research efforts. This may be especially useful for highly speciose genera such as *Penicillium* that have presented as a challenge to distinguish between species,[114] but has been associated with the isolation of several diverse bioactive

metabolites from alkaloids and polyketides with plant growth promoting qualities,[121] to quinazilones such as Verrucine F that hold pharmacological relevance.[122–124]

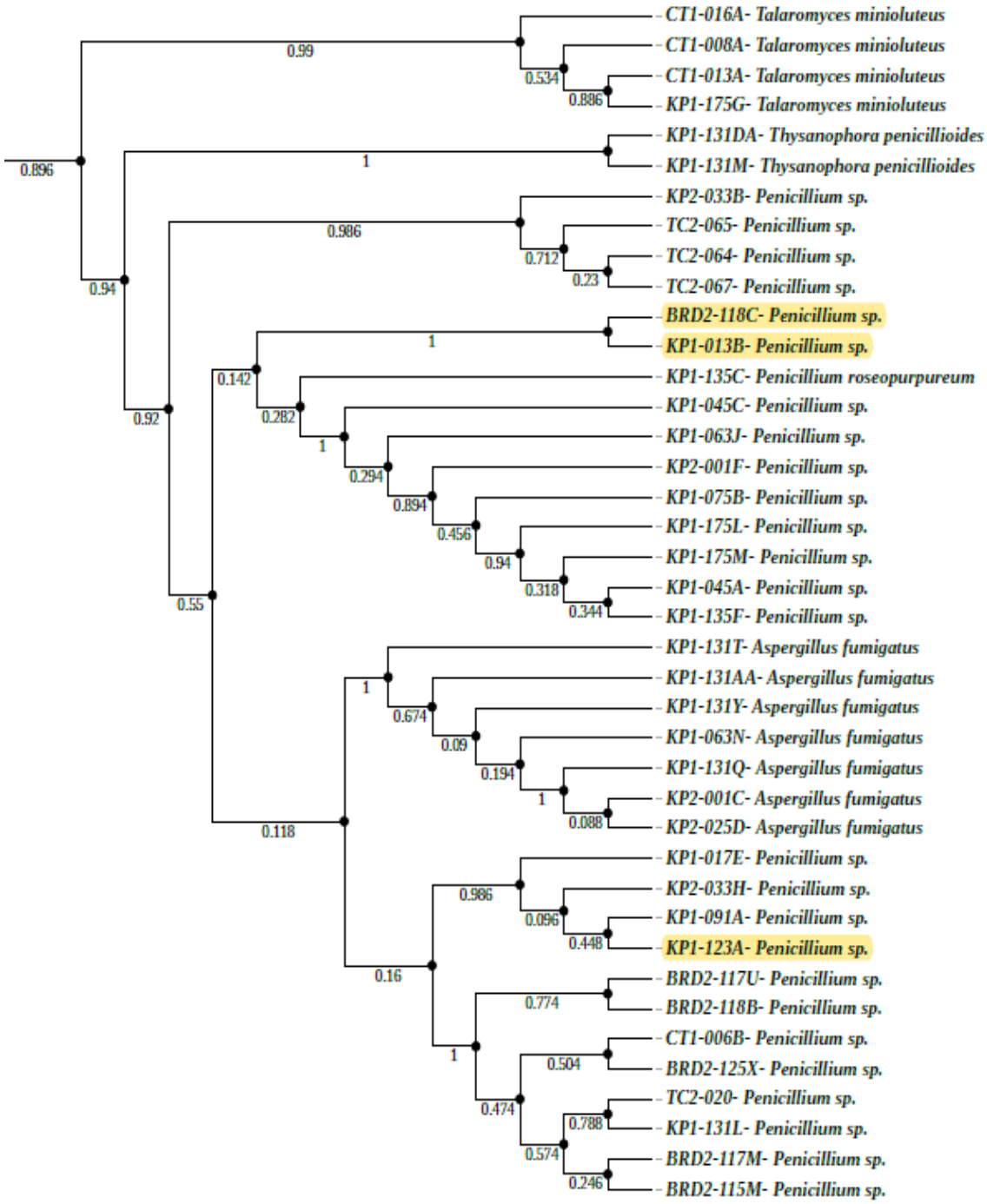

Figure 9. Phylogenetic clade depicting fungal sequences from the order *Eurotiales* within the NPRG fungal collection.

Using verified in-house data in combination with reference sequences rather than relying solely on a database such as GenBank to identify new sequences can mitigate the limitations that still exist with molecular identification, including insufficient sequence data and varying standards for identification thresholds (e.g percent similarity) that often result in ambiguous identifications.[125–127] This may be especially useful in fungal identification, as DNA-based phylogenies have previously illustrated complex relationships between different genera and species of fungi, making them difficult to identify when the sequences are not placed into a phylogenetic framework.[128–131] While it was not possible to obtain an identification for 11 of the sequences within the test set using DNA taxonomy alone, these sequences highlighted aspects of the approach that should be considered. Five of the sequences that could not be identified formed a monophyletic clade together with little relation to a previously identified sequence within the NPRG fungal collection. When identified as *Pseudoplectania nigrella* using DNA barcoding, it was found that no similar sequences existed within the collection, and thus served as a reminder that new or novel fungi to a collection may be more difficult to identify using a single approach. When a reference sequence for *Pseudoplectania nigrella* was included in the phylogenetic tree, it was possible to confirm the identification of the sequences based on the monophyletic clade that again formed and included the reference sequence. The remaining six sequences were identified to the genus level using DNA barcoding and reference sequences, as their placement within the phylogenetic tree made it difficult to obtain an identification. Within the phylogenetic tree, the sequences were of similar distance and relation to identified sequences of two or more genera and it was not

possible to determine an identification until an integrated approach was used. However, in both cases of being unable to obtain an identification, it was still possible to obtain a visual indication of where the sequence(s) were in relation to the rest of the phylogenetic tree, and thus where the endophytes were with respect to the entire fungal collection.

The overall agreement between molecular identification results between the two techniques suggests that DNA taxonomy can be used as an effective and reliable identification technique, in combination with previously identified sequences. Moreover, these results, in combination with previous research, provide a strong basis for the argument that even in the absence of a definitive identification, phylogenetic analyses are invaluable as they can reveal important information about a sequence, and thus an organism, that may otherwise not be discovered.

Significance

Endophytes play a significant role in many scientific fields, from ecology and biodiversity studies,[3,6,77] to clinical research and drug development.[4,23,132] The ability to accurately identify fungal endophytes not only impacts research directly, but also contributes to the growing body of knowledge of endophytes. In clinical research and drug development specifically, endophytes have been identified as a promising source of novel bioactive metabolites that have the capacity to serve as candidates for therapeutics.[2,121,124,132–134] As the need for new drug candidates continues to rise,[135] the need to discover novel bioactives does too. Given that the production of natural products can often be associated with specific species or genera,[123,136,137] the ability to classify sequences in relation to previously identified sequences and to accurately identify fungi

can be beneficial in the field of natural product research as obtaining an identification can enable access to existing data and knowledge in the literature, thereby allowing for more targeted research efforts.

Although this project focused on the fungal sequences from the NPRG fungal collection, the application of DNA taxonomy can, and should, be applied to other sequence types (*i.e* plant, animal, bacterium, etc.) as it offers a different perspective for interpreting not only the sequence data, but any data that may have been generated. In the case of the NPRG, other data has already been gathered for the isolates within the fungal collection including nuclear magnetic resonance (NMR) spectra, and bioactivity profiles. By being able to place the sequences associated with the existing data in relation to one another, it may be possible to find patterns in the data such as common chemical peaks in the NMR spectra, or observed biological activity against a pathogen in the bioactivity profile that correspond to phylogenetic clades that have formed. Finally, the challenges associated with identifying a diverse group of organisms is not unique to the NPRG and thus this work can be useful for other similar research groups working with a collection or organisms, especially microbial collections.

Conclusion

Identification has been a limitation that has extended far beyond endophytes, with numerous accounts of improper identifications being found in all domains of life. These misidentifications, especially at the species level, occur readily and can be attributed to several factors. One of the main causes of misidentifications is that an identification based on a taxonomic approach often does not match data when compared to a molecular

identification. As such, it has been strongly recommended to those working with fungi that whenever possible, the identification of an unknown organism be obtained by molecular methodology in association with morphology.[24]

The objective of this research was to use DNA taxonomy as an alternate technique with the sequences from the NPRG fungal collection to build a phylogenetic tree that would allow for the identification of ambiguously identified sequences, and to gain insight into the relationship between sequences. The resultant phylogenetic tree was then used as a framework to determine whether DNA taxonomy could be used as a primary method to identify new sequences using the fungal collection's existing data. The research described in this thesis has shown promising results that supports the use of DNA taxonomy as a molecular technique for the identification of fungi that are both accurate and insightful. While this work analyzed fungal sequences exclusively, the framework of the approach has the potential to be applied to all kingdoms, resulting in an opportunity to identify and classify sequences in an alternate way.

Bibliography

1. Halme, P. & Kotiaho, J. S. The importance of timing and number of surveys in fungal biodiversity research. *Biodivers. Conserv.* 21, 205–219 (2012).
2. Bamisile, B. S., Dash, C. K., Akutse, K. S., Keppanan, R. & Wang, L. Fungal Endophytes: Beyond Herbivore Management. *Front. Microbiol.* 9, (2018).
3. Mueller, G. M. *et al.* Global diversity and distribution of macrofungi. *Biodivers. Conserv.* 16, 37–48 (2007).
4. Araujo, R. Towards the Genotyping of Fungi: Methods, Benefits and Challenges. *Curr. Fungal Infect. Rep.* 8, 203–210 (2014).
5. Holzenthal, R. W., Robertson, D. R., Pauls, S. U. & Mendez, P. K. Taxonomy and systematics: contributions to benthology and J-NABS. *Freshw. Sci.* 29, 147–169 (2010).
6. Purvis, A. & Hector, A. Getting the measure of biodiversity. *Nature* 405, 212–219 (2000).
7. Michel, E. Anchoring Biodiversity Information: From Sherborn to the 21st century and beyond. *ZooKeys* 1–11 (2016).
8. Lambertz, M. Taxonomy: retain scientific autonomy. *Nature* 546, 600–600 (2017).
9. Willis, K. J. *State of the World's Fungi 2018.* (2018).
10. Spatafora, J. W. *et al.* The Fungal Tree of Life: from Molecular Systematics to Genome-Scale Phylogenies. *Microbiol. Spectr.* 5, (2017).
11. Blackwell, M. The Fungi: 1, 2, 3 … 5.1 million species? *Am. J. Bot.* 98, 426–438 (2011).

12. Hibbett, D. S. *et al.* A higher-level phylogenetic classification of the Fungi. *Mycol. Res.* 111, 509–547 (2007).

13. Hawksworth, D. L. The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycol. Res.* 105(12), (2001).

14. Hawksworth, D. L. & Lücking, R. Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiol. Spectr.* 5, (2017).

15. Efaq, A. N., Al-Gheethi, A. A., Ab. Rahman, N. N. N., Nagao, H. & Ab. Kadir, M. O. Morphological characteristics of black aspergilli isolated from clinical wastes. *Songklanakarin J. Sci. Technol.* 41, 181–191 (2019).

16. Goulding, T. C. & Dayrat, B. Integrative taxonomy: ten years of practice and looking into the future. 116-133, (2016).

17. Humber, R. A. Chapter V-1 - Fungi: Identification. in *Manual of Techniques in Insect Pathology* (ed. Lacey, L. A.) 153–185 (Academic Press, 1997).

18. Wu, B. *et al.* Current insights into fungal species diversity and perspective on naming the environmental DNA sequences of fungi. *Mycology* 10, 127–140 (2019).

19. Duong, L. M., Jeewon, R. & Lumyong, S. DGGE coupled with ribosomal DNA gene phylogenies reveal uncharacterized fungal phylotypes. *Fungal Divers.* 23, (2006).

20. Goulding, T. C. & Dayrat, B. Integrative taxonomy: Ten years of practice and looking into the future. in *Аспекты биоразнообразия* 116–133 (2016).

21. Lodge, A. & O'Dell, M. Collecting and Describing Macrofungi. *Biodiversity of Fungi.,* 128-155 (2004).

22. Tang, X. *et al.* Diversity, population genetics, and evolution of macrofungi associated with animals. *Mycology* 6, 94–109 (2015).

23. Hyde, K. D. & Soytong, K. The fungal endophyte dilemma. *Fungal Divers.* 33, 163-173 (2008).

24. Raja, H. A., Miller, A. N., Pearce, C. J. & Oberlies, N. H. Fungal Identification Using Molecular Tools: A Primer for the Natural Products Research Community. *J. Nat. Prod.* 80, 756–770 (2017).

25. Gamper, H. A., Young, J. P. W., Jones, D. L. & Hodge, A. Real-time PCR and microscopy: Are the two methods measuring the same unit of arbuscular mycorrhizal fungal abundance? *Fungal Genet. Biol.* 45, 581–596 (2008).

26. Heikkila, P., Kotimaa, M., Tuomi, T., Sami, T., & Louhelainen, K. Identification and counting of fungal spores by scanning electron microscope*. *Ann. Occup. Hyg.* (1988).

27. Naumann, A., Navarro-González, M., Peddireddi, S., Kües, U. & Polle, A. Fourier transform infrared microscopy and imaging: Detection of fungi in wood. *Fungal Genet. Biol.* 42, 829–835 (2005).

28. Petrini, O. *Fungal endophytes of tree leaves*. (Spring-Verlag, New York., 1991).

29. Slepecky, R. A. & Starmer, W. T. Phenotypic plasticity in fungi: a review with observations on Aureobasidium pullulans. *Mycologia* 101, 823–832 (2009).

30. Hawksworth, D. L. Managing and coping with names of pleomorphic fungi in a period of transition1, 2. *IMA Fungus Glob. Mycol. J.* 3, 15–24 (2012).

31. Kress, W. J. & Erickson, D. L. DNA barcodes: Genes, genomics, and bioinformatics. *Proc. Natl. Acad. Sci. U. S. A.* 105, 2761–2762 (2008).

32. Savolainen, V., Cowan, R. S., Vogler, A. P., Roderick, G. K. & Lane, R. Towards writing the encyclopaedia of life: an introduction to DNA barcoding. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 1805–1811 (2005).

33. Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B Biol. Sci.* 270, 313–321 (2003).

34. Hebert, P. D. N., Ratnasingham, S. & deWaard, J. R. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. B Biol. Sci.* 270, S96–S99 (2003).

35. Schoch, C. L. *et al.* Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proc. Natl. Acad. Sci.* 109, 6241–6246 (2012).

36. Braukmann, T. W. A., Kuzmina, M. L., Sills, J., Zakharov, E. V. & Hebert, P. D. N. Testing the Efficacy of DNA Barcodes for Identifying the Vascular Plants of Canada. *PLoS ONE* 12, (2017).

37. Lebonah, D. E. *et al.* DNA Barcoding on Bacteria: A Review. *Adv. Biol.* 2014, 1–9 (2014).

38. CBOL Plant Working Group. A DNA barcode for land plants. *Proc. Natl. Acad. Sci. U. S. A.* 106, 12794–12797 (2009).

39. de Vere, N., Rich, T. C. G., Trinder, S. A. & Long, C. DNA barcoding for plants. *Methods Mol. Biol. Clifton NJ* 1245, 101–118 (2015).

40. Chaudhary, D. K. & Dahal, R. H. DNA Bar-Code for Identification of Microbial Communities: A Mini-Review. . *EC* 6 (2017).

41. Ebersberger, I. *et al.* A Consistent Phylogenetic Backbone for the Fungi. *Mol. Biol. Evol.* 29, 1319–1334 (2012).

42. Seifert, K. A. Progress towards DNA barcoding of fungi. *Mol. Ecol. Resour.* 9, 83–89 (2009).

43. Schoch, C. L. *et al.* A class-wide phylogenetic assessment of Dothideomycetes. *Stud. Mycol.* 64, 1-15-S10 (2009).

44. Badotti, F. *et al.* Effectiveness of ITS and sub-regions as DNA barcode markers for the identification of Basidiomycota (Fungi). *BMC Microbiol.* 17, 42 (2017).

45. Nilsson, R. H., Ryberg, M., Abarenkov, K., Sjokvist, E. & Kristiansson, E. The ITS region as a target for characterization of fungal communities using emerging sequencing technologies. *FEMS Microbiol. Lett.* 296, 97–101 (2009).

46. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410 (1990).

47. Kerfeld, C. A. & Scott, K. M. Using BLAST to Teach "E-value-tionary" Concepts. *PLoS Biol.* 9, (2011).

48. Karsch-Mizrachi, I., Takagi, T. & Cochrane, G. The international nucleotide sequence database collaboration. *Nucleic Acids Res.* 46, D48–D51 (2018).

49. Benson, D. A. *et al.* GenBank. *Nucleic Acids Res.* 41, D36-42 (2013).

50. Mashima, J. *et al.* DNA Data Bank of Japan. *Nucleic Acids Res.* 45, D25–D31 (2017).

51. Leinonen, R. *et al.* The European Nucleotide Archive. *Nucleic Acids Res.* 39, D28–D31 (2011).

52. Pearson, W. R. BLAST and FASTA Similarity Searching for Multiple Sequence Alignment. in *Multiple Sequence Alignment Methods* (ed. Russell, D. J.) 75–101 (Humana Press, 2014).

53. Nilsson, R. H. *et al.* The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* 47, D259–D264 (2019).

54. Cesari, M., Guidetti, R., Rebecchi, L., Giovannini, I. & Bertolani, R. A DNA barcoding approach in the study of tardigrades. *J. Limnol.* 72, 182–198 (2013).

55. Fontaneto, D., Eckert, E. M., Anicic, N., Lara, E. & Mitchell, E. A. D. We are ready for faunistic surveys of bdelloid rotifers through DNA barcoding: The example of sphagnum bogs of the Swiss Jura Mountains. *Limnetica* 38, (2019).

56. Discovery and Identification of Meloidogyne Species Using COI DNA Barcoding. *J. Nematol.* 50, (2018).

57. Vijayan, K. & Tsou, C. H. DNA barcoding in plants: taxonomy in a new perspective. *Curr. Sci.* 99, 1530–1541 (2010).

58. Kelly, L. J. *et al.* DNA barcoding of lichenized fungi demonstrates high identification success in a floristic context. *New Phytol.* 191, 288–300 (2011).

59. Chakraborty, C., Doss, C. G. P., Patra, B. C. & Bandyopadhyay, S. DNA barcoding to map the microbial communities: current advances and future directions. *Appl. Microbiol. Biotechnol.* 98, 3425–3436 (2014).

60. Meiklejohn, K. A., Damaso, N. & Robertson, J. M. Assessment of BOLD and GenBank – Their accuracy and reliability for the identification of biological materials. *PLoS ONE* 14, 1–14 (2019).

61. Vilgalys, R. Taxonomic misidentification in public DNA databases. *New Phytol.* 160, 4–5 (2003).

62. Stavrou, A. A., Mixão, V., Boekhout, T. & Gabaldón, T. Misidentification of genome assemblies in public databases: The case of Naumovozyma dairenensis and proposal of a protocol to correct misidentifications. *Yeast Chichester Engl.* 35, 425–429 (2018).

63. Pearson, W. R. An Introduction to Sequence Similarity ("Homology") Searching. *Curr. Protoc. Bioinforma. Ed. Board Andreas Baxevanis Al* 0 3, (2013).

64. Pertsemlidis, A. & Fondon, J. W. Having a BLAST with bioinformatics (and avoiding BLASTphemy). *Genome Biol.* 2, reviews2002.1 (2001).

65. Woese, C. R. Interpreting the universal phylogenetic tree. *Proc. Natl. Acad. Sci. U. S. A.* 97, 8392–8396 (2000).

66. McInerney, J. O., Mullarkey, M., Wernecke, M. E. & Powell, R. Bacteria and Archaea: Molecular techniques reveal astonishing diversity. *Biodiversity* 3, 3–10 (2002).

67. Pruitt, K. D. RefSeq and LocusLink: NCBI gene-centered resources. *Nucleic Acids Res.* 29, 137–140 (2001).

68. Pruitt, K., Brown, G., Tatusova, T. & Maglott, D. Chapter 18: The Reference Sequence (RefSeq) Database. 24 (2008).

69. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 35, D61–D65 (2007).

70. Edgar, R. C. & Batzoglou, S. Multiple sequence alignment. *Curr. Opin. Struct. Biol.* 16, 368–373 (2006).

71. Saitou, N. & Imanshi, T. Relative Efficiencies of the Fitch-Margoliash, Maximum-Parsimony, Maximum-Likelihood, Minimum-Evolution, and Neighbor-joining

Methods of Phylogenetic Tree Construction in Obtaining the Correct Tree. *Mol. Biol. Evol.* (1989).

72. Yang, Z. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Mol. Biol. Evol.* (1993).
73. Holder, M. & Lewis, O.P. Phylogeny estimation: traditional and Bayesian approaches. *Nat. Rev. Gen.* 4, 278-284 (2003).
74. McLennan, D. A. How to Read a Phylogenetic Tree. *Evol. Educ. Outreach* 3, 506–519 (2010).
75. McLaughlin, D. J., Hibbett, D. S., Lutzoni, F., Spatafora, J. W. & Vilgalys, R. The search for the fungal tree of life. *Trends Microbiol.* 17, 488–497 (2009).
76. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. – ScienceOpen. https://www.scienceopen.com/document?vid=8b59b929-3c37-49f6-936b-f8bf6dd92ace.
77. Abarenkov, K. *et al. UNITE general FASTA release for Fungi*. (2020).
78. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113 (2004).
79. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797 (2004).
80. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* 35, 1547–1549 (2018).

81. Tamura, K., Nei, M. & Kumar, S. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci.* 101, 11030–11035 (2004).

82. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259 (2019).

83. Efron, B., Halloran, E. & Holmes, S. Bootstrap confidence levels for phylogenetic trees. *Proc. Natl. Acad. Sci.* 93, 13429–13429 (1996).

84. Felsenstein, J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution* 39, 783–791 (1985).

85. Soltis, D. E. & Soltis, P. S. Applying the Bootstrap in Phylogeny Reconstruction. *Stat. Sci.* 18, 256–267 (2003).

86. Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. A plea for DNA taxonomy. *Trends Ecol. Evol.* 18, 70–74 (2003).

87. Tenover, F. C., Arbeit, R. D., Goering, R. V. & America, M. T. W. G. of the S. for H. E. of. How to Select and Interpret Molecular Strain Typing Methods for Epidemiological Studies of Bacterial Infections A Review for Healthcare Epidemiologists. *Infect. Control Hosp. Epidemiol.* 18, 426–439 (1997).

88. Pfeifer, T. A., Humble, L. M., Ring, M. & Grigliatti, T. A. Characterization of gypsy moth populations and related species using a nuclear DNA marker. *Can. Entomol.* 127, 49–58 (1995).

89. Dibart, S. *et al.* Identification of bacterial species on or in crevicular epithelial cells from healthy and periodontally diseased patients using DNA-DNA hybridization. *Oral Microbiol. Immunol.* 13, 30–35 (1998).

90. Henry, T., Iwen, P. C. & Hinrichs, S. H. Identification of Aspergillus Species Using Internal Transcribed Spacer Regions 1 and 2. *J. Clin. Microbiol.* 38, 1510–1515 (2000).

91. Spithill, T. W. & Grumont, R. J. Identification of species, strains and clones of Leishmania by characterization of kinetoplast DNA minicircles. *Mol. Biochem. Parasitol.* 12, 217–236 (1984).

92. Cr, W., O, K. & Ml, W. Towards a Natural System of Organisms: Proposal for the Domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences of the United States of America* vol. 87 (1990).

93. Birky Jr, C. W., Ricci, C., Melone, G. & Fontaneto, D. Integrating DNA and morphological taxonomy to describe diversity in poorly studied microscopic animals: new species of the genus Abrochtha Bryce, 1910 (Rotifera: Bdelloidea: Philodinavidae): NEW CRYPTIC ROTIFER SPECIES. *Zool. J. Linn. Soc.* 161, 723–734 (2011).

94. Choi, J. & Kim, S.-H. A genome Tree of Life for the Fungi kingdom. *Proc. Natl. Acad. Sci.* 114, 9391–9396 (2017).

95. Delsuc, F., Brinkmann, H. & Philippe, H. Phylogenomics and the reconstruction of the tree of life. *Nat. Rev. Genet.* 6, 361–375 (2005).

96. Fontaneto, D. Molecular phylogenies as a tool to understand diversity in rotifers. *Int. Rev. Hydrobiol.* 99, 178–187 (2014).

97. Fontaneto, D. *et al.* Cryptic diversity in the genus Adineta Hudson & Gosse, 1886 (Rotifera: Bdelloidea: Adinetidae): A DNA taxonomy approach. *Hydrobiologia* 662, 27–33 (2011).

98. García-Morales, A. E. & Elías-Gutiérrez, M. DNA barcoding of freshwater Rotifera in Mexico: Evidence of cryptic speciation in common rotifers. *Mol. Ecol. Resour.* 13, 1097–1107 (2013).

99. Sangoi, A. R. *et al.* Challenges and pitfalls of morphologic identification of fungal infections in histologic and cytologic specimens: a ten-year retrospective review at a single institution. *Am. J. Clin. Pathol.* 131, 364–375 (2009).

100. Sieber, T. N. Endophytic fungi in forest trees: are they mutualists? *Fungal Biol. Rev.* 21, 75–89 (2007).

101. Ghimire, S. R. & Hyde, K. D. Fungal Endophytes. in *Plant Surface Microbiology* (eds. Varma, A., Abbott, L., Werner, D. & Hampp, R.) 281–292 (Springer, 2004).

102. Jia, M. *et al.* A Friendly Relationship between Endophytic Fungi and Medicinal Plants: A Systematic Review. *Front. Microbiol.* 7, (2016).

103. Devay, J. E. Mutual Relationships in Fungi. *Annu. Rev. Microbiol.* 10, 115–140 (1956).

104. Bolan, N. S. A critical review on the role of mycorrhizal fungi in the uptake of phosphorus by plants. *Plant Soil* 134, 189–207 (1991).

105. Smith, S. S. E. Mycorrhizas of Autotrophic Higher Plants. *Biol. Rev.* 55, 475–510 (1980).

106. Mohammadi, K., Khalesro, S., Sohrabi, Y. & Heidari, G. A Review: Beneficial Effects of the Mycorrhizal Fungi for Plant Growth. 10 (2011).

107. Gouda, S., Das, G., Sen, S. K., Shin, H.-S. & Patra, J. K. Endophytes: A Treasure House of Bioactive Compounds of Medicinal Importance. *Front. Microbiol.* 7, (2016).

108. Rodriguez, R. J., White, J. F., Arnold, A. E. & Redman, R. S. Fungal endophytes: diversity and functional roles. *New Phytol.* 182, 314–330 (2009).

109. Jörger, K. M. & Schrödl, M. How to describe a cryptic species? Practical challenges of molecular taxonomy. *Front. Zool.* 10, (2013).

110. Nilsson, R. H., Kristiansson, E., Ryberg, M., Hallenberg, N. & Larsson, K.-H. Intraspecific ITS Variability in the Kingdom Fungi as Expressed in the International Sequence Databases and Its Implications for Molecular Species Identification. *Evol. Bioinforma.* 4, EBO.S653 (2008).

111. Schubert, K. *et al.* Biodiversity in the Cladosporium herbarum complex (Davidiellaceae, Capnodiales), with standardisation of methods for Cladosporium taxonomy and diagnostics. *Stud. Mycol.* 58, 105–156 (2007).

112. Frisvad, J. C. & Samson, R. A. Polyphasic taxonomy of Penicillium subgenus Penicillium A guide to identification of food and air-borne terverticillate Penicillia and their mycotoxins. *Stud. Mycol.* 49, 1-174 (2004).

113. Kõljalg, U. *et al.* Towards a unified paradigm for sequence-based identification of fungi. *Mol. Ecol.* 22, 5271–5277 (2013).

114. Pryor, B. M. & Michailides, T. J. Morphological, Pathogenic, and Molecular Characterization of *Alternaria* Isolates Associated with Alternaria Late Blight of Pistachio. *Phytopathology* 92, 406–416 (2002).

115. Dickie, I. A. Host preference, niches and fungal diversity. *New Phytol.* 174, 230–233 (2007).

116. Arnold, A. E. Understanding the diversity of foliar endophytic fungi: progress, challenges, and frontiers. *Fungal Biol. Rev.* 21, 51–66 (2007).

117. Ainsworth, G. C., Bisby, G. R., Kirk, P. M., & CABI Bioscience.. *Ainsworth & bisby's dictionary of the fungi* (10th ed.). CABI. (2008).

118. Hyde, K. D. *et al.* Families of Dothideomycetes. *Fungal Divers.* 63, 1–313 (2013).

119. El-Neketi, M. *et al.* Alkaloids and polyketides from Penicillium citrinum, an endophyte isolated from the Moroccan plant Ceratonia siliqua. *J. Nat. Prod.* 76, 1099–1104 (2013).

120. Alagarsamy, V. *et al.* An overview of quinazolines: Pharmacological significance and recent developments. *Eur. J. Med. Chem.* 151, 628–685 (2018).

121. Leong, S.-L., Schnürer, J. & Broberg, A. Verrucine F, a quinazoline from Penicillium verrucosum. *J. Nat. Prod.* 71, 1455–1457 (2008).

122. Nicoletti, R. & Trincone, A. Bioactive Compounds Produced by Strains of Penicillium and Talaromyces of Marine Origin. *Mar. Drugs* 14, (2016).

123. Teletchea, F. After 7 years and 1000 citations: Comparative assessment of the DNA barcoding and the DNA taxonomy proposals for taxonomists and non-taxonomists. *Mitochondrial DNA J. DNA Mapp. Seq. Anal.* 21, 206–226 (2010).

124. Will, K. W. & Rubinoff, D. Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* 20, 47–55 (2004).

125. Townzen, J. S., Brower, A. V. Z. & Judd, D. D. Identification of mosquito bloodmeals using mitochondrial cytochrome oxidase subunit I and cytochrome b gene sequences. *Med. Vet. Entomol.* 22, 386–393 (2008).

126. Carbone, M., Agnello, C. & Alvarado, P. Phylogenetic and morphological studies in the genus Pseudoplectania (Ascomycota, Pezizales). *Ascomycetes,* 6(1), 17-33 (2014).

127. Chen, Q., Hou, L. W., Duan, W. J., Crous, P. W. & Cai, L. Didymellaceae revisited. *Stud. Mycol.* 87, 105–159 (2017).

128. Liyanage, K. K. *et al.* Morpho-Molecular Characterization of Two Ampelomyces spp. (Pleosporales) Strains Mycoparasites of Powdery Mildew of Hevea brasiliensis. *Front. Microbiol.* 9, (2018).

129. Verkley, G. J. M., Dukik, K., Renfurm, R., Göker, M. & Stielow, J. B. Novel genera and species of coniothyrium-like fungi in Montagnulaceae (Ascomycota). *Persoonia Mol. Phylogeny Evol. Fungi* 32, 25–51 (2014).

130. Gouda, S., Das, G., Sen, S. K., Shin, H.-S. & Patra, J. K. Endophytes: A Treasure House of Bioactive Compounds of Medicinal Importance. *Front. Microbiol.* 7, (2016).

131. Schulz, B., Boyle, C., Draeger, S., Römmert, A.-K. & Krohn, K. Endophytic fungi: a source of novel biologically active secondary metabolites. *Mycol. Res.* 106, 996–1004 (2002).

132. Eliwa, E. M. *et al.* New bioactive compounds from the marine-derived actinomycete *Nocardiopsis lucentensis* sp. ASMR2. *Z. Naturforschung - Sect. B J. Chem. Sci.* 72, 351–360 (2017).

133. Spížek, J., Novotná, J., Řezanka, T. & Demain, A. L. Do we need new antibiotics? The search for new targets and new compounds. *J. Ind. Microbiol. Biotechnol.* 37, 1241–1248 (2010).

134. Loreto, F., Bagnoli, F. & Fineschi, S. One species, many terpenes: matching chemical and biological diversity. *Trends Plant Sci.* 14, 416–420 (2009).

135. Vetter, W. & Janussen, D. Halogenated natural products in five species of Antarctic sponges: compounds with POP-like properties? *Environ. Sci. Technol.* 39, 3889–3895 (2005).

Appendix 1: Newly generated DNA sequences from Génome Québec

>JD1-017A_ITS1F
aaacaagtcccagcgcaagtcagcactagctggcaacacaatcgaactgcggggacgctt
taaagcccagtggtaccaacccccccggtgaccggaggggggcgggcatcgcccgaggtcag
aacccactgagatgtacaataagcaatccgcagcggtccgatcggaccgttcacagacta
agtggttgtggggagcctgctccttaagatatagtcgggcccggcgggcagtcgccgggg
gttttgcgtccgtaggtgaacctgcggaaggatcattacaagtgaggccgccgaacaccg
gtcaaacggtggggaggcggctaacacccacccgtgtctacatactcttgttgctttggc
gggccgtggcctccactgcgggctctgctctgctcgtgtgtgcccgccagaggaccaaac
tctgaatgttagtgatgtctgagtactatttaatagttaaaactttcaacaacggatctc
ttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaat
tcagtgaatcatcgaatctttgaacgcacattgcgccctgtggtattccgcagggcatgc
ctgttcgagcgtcatttaaccactcacgcctggcgtggtgttggggcacgcggtctcgcg
gccctcgaaagcagtggcggcgccggtgggctctaagcgtagtacatactcccgctatag
agttcc

>JD1-017A_ITS4
cacccgagagggggggaggcgactataaacaagtcccagcgcaagtcagcactagctggca
acacaatcgaNNNgcggggacgctttaaagcccagtggtaccaacccccggtgaccgga
gggggcgggcatcgcccgaggtcagaacccactgagatgtacaataagcaatccgcagcg
gtccgatcggaccgttcacagactaagtggttgtggggagcctgctccttaagatatagt
cgggcccggcgggcagtcgccgggggttttgcgtccgtaggtgaacctgcggaaggatca
ttacaagtgaggccgccgaacaccggtcaaacggtggggaggcggctaacacccacccgt
gtctacatactcttgttgctttggcgggccgtggcctccactgcgggctctgctctgctc
gtgtgtgcccgccagaggaccaaactctgaatgttagtgatgtctgagtactatttaata
gttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatg
cgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcg
ccctgtggtattccgcagggcatgcctgttcgagcgtcatttaaccactcacgcctggcg
tggtgttggggcacgcggtctcgcggccctcgaaagcagtggcggcgccggtgggctcta
agcgtagtacatactcc

>JD1-017B_ITS1F
gatcattgctggaacgcgccccaggcgcacccagaaaccctttgNNNNcttataccttac
tgttgcctcggcgctagctggtccttcggggcccctcaccctcgggtgttgagacggccc
gccggcggccaacctaactcttgtttttacactgaaactctgagaataaaacataaatga
atcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatg
cgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcg
ccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggct
tggtgatggggcactgcttttacccaaaagcaggccctgaaattcagtggcgagctcgcc
aggaccccgagcgcagtagttaaaccctcgctttggaaggccctggcggtgccctgccgt
taaaccccccaacttctgaaaatttgacctcggatcaggtaggaatacccgctgaactta
agcatatcaata

>JD1-017B_ITS4
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggaggga
tcattgctggaacgcgccccaggcgcacccagaaaccctttgtgaacttataccttactg
ttgcctcggcgctagctggtccttcggggcccctcaccctcgggtgttgagacggcccgc
cggcggccaacctaactcttgtttttacactgaaactctgagaataaaacataaatgaat
caaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcg
ataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcc
ctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcttg
gtgatggggcactgcttttacccaaaagcaggccctgaaattcagtggcgagctcgccag
gaccccgagcgcagtagttaaaccctcgctttNNNaggccctggcggtgccctgccgtta
aacccccccaac

>JD1-017C_ITS1F
gatcattacagagttatcaaaactcNNNacccatgtgaacatacctcgcgttgcctcggc
aggtggcgtctcgccccgtaagaacctaccctgtaggaccttacccggtagacgaccctg
ccgacggcccccgaaactctgttttatagcattaaacttctgaaaacataactaaataag
ttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgc
gataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgc
ccactagtattctggtgggcatgcctgttcgagcgtcatttcaacccttaagcccctgtt
gcttagcgttaggagcctaccggaactctctggtagctccccaaagtcagtggcggagcc
ggttcgcactccagacgtagtagcttttacacgtcgcctgtagcgcgggccggtcccctg
ccataaaacaccccaatttttataggttgacctcggatcaggtaggaatacccgctgaac
ttaagcatatca

>JD1-017C_ITS4
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattacag
agttatcaaaactcccaaacccatgtgaacatacctcgcgttgcctcggcaggtggcgtc
tcgccccgtaagaacctaccctgtaggaccttacccggtagacgaccctgccgacggccc
ccgaaactctgttttatagcattaaacttctgaaaacataactaaataagttaaaacttt
caacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaat
gtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccactagtat
tctggtgggcatgcctgttcgagcgtcatttcaacccttaagcccctgttgcttagcgtt
aggagcctaccggaactctctggtagctccccaaagtcagtggcggagccggttcgcact
ccagacgtagtagcttttacacgtcg

>JD1-017E_ITS1F
gatcattacagagttgcaaaactcNNNaccattgtgaacgtcacNNNatatcgttgcttc
ggcgggtggttcccactcgggtaccgcatctccccctcggggtaggtagcccgccggagg
cgtctcaactctcgattttaattggcatctctgagtaacttataaaataagtcaaaactt
tcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaa
tgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgctagta
ttctggcgggcatgcctgttcgagcgtcatttcaaccatcaagccccaggcttgtgttgg
ggccctgcggccgtccgcagcccccgaaaagcagtggcgggctcgctgtcacaccgagcg
cagtagtatatcttcgctgggacgtgattgcgggttccggccgtgaaacaccccaattta
tcaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatca

>JD1-017E_ITS4
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattaca
gagttgcaaaactcccaaaccattgtgaacgtcaccgtatatcgttgcttcggcgggtgg
ttcccactcgggtaccgcatctccccctcggggtaggtagcccgccggaggcgtctcaac
tctcgattttaattggcatctctgagtaacttataaaataagtcaaaactttcaacaacg
gatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattg
cagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgctagtattctggcgg
gcatgcctgttcgagcgtcatttcaaccatcaagccccaggcttgtgttggggccctgcg
gccgtccgcagcccccgaaaagcagtggcgggctcgctgtcacaccgagcgcagtagtat
atcttcgcNNNNacgtgattgcgggttccggccgtgaaacaccccaat

>JD1-017F_ITS1F
acgcgaatggccgggcctcgtcaccacccccttctgggggtggaatcctgggggggcccggt
gtccttttgattatctcccatctgtgaactcttccactgttgcttccgtgagcctgtact
ctgctctctttgggagtcggagcacctgaaaagccctggttctgccaggggcgcacggca
atggtaccacaactctggttttatgttgtatcagtctgagtggttaacacacaaacaagt
taaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatgcg
ataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcc
tcctggtattccgggaggcatgcctgttcgagcgtcatcaaacacatcctcaagtatttt
tgcttggttgtggaggaagatagtctgttgattatcccctctggaattcctcagcagaga
ccacagcagcctcaggtgtgataatactacctcgcttgagtgttattgtgggctcccgct
ggagagcagccccagcttttatcatgtttgacctcggatcaggtagggatacccgctga
acttaagcatatcaa

>JD1-017F_ITS4
aggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaattaa
acgcgaatggccgggcctcgtcaccacccccttctgggggtggaatcctgggggggcccggt
gtccttttgattatctcccatctgtgaactcttccactgttgcttccgtgagcctgtact
ctgctctctttgggagtcggagcacctgaaaagccctggttctgccaggggcgcacggca
atggtaccacaactctggttttatgttgtatcagtctgagtggttaacacacaaacaagt
taaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatgcg
ataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcc
tcctggtattccgggaggcatgcctgttcgagcgtcatcaaacacatcctcaagtatttt
tgcttggttgtggaggaagatagtctgttgattatcccctctggaattcctcagcagaga
ccacagcagcctcaggtgtgataatactacctcgcttgagtgttattgtgggctcccg

>JD1-017G_ITS1F
tgtgctcagcctcgctctaatccactcaacccctgtgcactattcggagtgttgcaagct
gagacttgaataggtcttggtcgcgtttctgatgcggctgggttctaagacagtcaaagc
ttgtcctccggtctatatacaaactcttataatgtcttgtagaatgtgttagcctctcgt
taggcgaaacttaaatacaactttcaacaacggatctcttggctctcgcatcgatgaaga
acgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttg
aacgcaccttgcgctccttggtattccgaggagcatgcctgtttgagtgtcatgttaata
tcaaccccgtcggtttgtaaatcagaccgttgggtgttggacttggaggttcttgctggc
tgcaaagtcggctcctcttgaatgtattagcttgggcctgttgtgcgtgctcgcggtgta

ataagttttattcatcgtttgggtcgtacacttattgggcttgcttctaactgtcttcgg
acaaacataacctttgactctgacctcaaatcaggtaggactacccgctgaacttaagca
tatcaat

>JD1-017G_ITS4
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaac
gagttttaagtgggcttgatgctggctttaccaaaagcatgtgctcagcctcgctctaat
ccactcaacccctgtgcactattcggagtgttgcaagctgagacttgaataggtcttggt
cgcgtttctgatgcggctgggttctaagacagtcaaagcttgtcctccggtctatataca
aactcttataatgtcttgtagaatgtgttagcctctcgttaggcgaaacttaaatacaac
tttcaacaacggatctcttggctctcgcatcgatgaagaacgcagcgaaatgcgataagt
aatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaccttgcgctccttgg
tattccgaggagcatgcctgtttgagtgtcatgttaatatcaaccccgtcggtttgtaaa
tcagaccgttgggtgttggacttggaggttcttgctggctgcaaagtcggctcctcttga
atgtattagcttgggcctgttgtgcgtgctcgcggtgtaataagttttattcatcgtttg
ggtcgtac

>SS1-033A_ITS1F
atcattacaagaagccgaaaggctacttaaaaccatcgcgaacttatccaagttgcttcg
gcggcgcggttcccctcacagggcgccgcagccccgcctctcaggaggtatggggcgccc
gccggaggtacgaaactctgtattatagtggtatctctgagtaaaaaacaaataagttaa
aactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgata
agtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgc
tagtactctagcgggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtg
ttggggccctacggctgccgtaggccctgaaaggaagtggcgggctcgctacaactccga
gcgtagtaattcattatctcgctagggaggtgcagcgtgctcctgccgttaaagacccca
tctttaaccaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatca

>SS1-033A_ITS4
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattaca
agaagccgaaaggctacttaaaaccatcgcgaacttatccaagttgcttcggcggcgcgg
ttcccctcacagggcgccgcagccccgcctctcaggaggtatggggcgcccgccggaggt
acgaaactctgtattatagtggtatctctgagtaaaaaacaaataagttaaaactttcaa
caacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtg
aattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgctagtactct
agcgggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtgttggggccc
tacggctgccgtaggccctgaaaggaagtggcgggctcgctacaactccgagcgtagtaa
ttcattatctcgctagggNNNtgcagcgtgctcctgccgttaaagaccccat

>SS1-033B_ITS1F
gatcattacaagaagccgaaaggctacttaaaaccatcgtgaacttatccaagttgcttc
ggcggcgcggctcccctcacagggcgccgcagcccccgcctctcaggaggtatggggcac
ccgccggaggtactaaactctgtattatagtggcatctctgagtaaaaaacaaataagtt
aaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcga
taagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccc

gctagtactctagcgggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttgg
tgttggggccctacggctgccgtaggccctgaaaagaagtggcgggctcgctacaactcc
gagcgtagtaattcattatctcgctagggaggtgcggcgtgctcctgccgttaaagaccc
atctttaaccaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatc

>SS1-033B_ITS4
aggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattacaaga
agccgaaaggctacttaaaaccatcgtgaacttatccaagttgcttcggcggcgcggctc
ccctcacagggcgccgcagcccccgcctctcaggaggtatggggcacccgccggaggtac
taaactctgtattatagtggcatctctgagtaaaaaacaaataagttaaaactttcaaca
acggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaa
ttgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgctagtactctag
cgggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtgttggggcccta
cggctgccgtaggccctgaaaagaagtggcgggctcgctacaactccgagcgtagtaatt
cattatctcgctagggaggtgcggcgtgctcctgccgttaaagacccatctta

>SS1-033C_ITS1F
gatcattactgagttcttacaaactccacccctttgtgaaccaNNNNaactgttgcctcgg
cgcgagctgcggctgcttggtagctacccggtagtcacctacccgggagctacccggtag
cgtctgcgtacaggcccgccgaaggaccaccaaactctgtttgacagtgtattctgaatg
cttcaactaaatagttaaaactttcaacaacggatctcttggttctggcatcgatgaaga
acgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttg
aacgcacattgcgcccattagtattctagtgggcatgcctattcgagcgtcatttcgacc
ctgaagccctagttgcttcgcgtcgggactctactggctaccctgtagttccctaatgac
agtggcggagttcaggtgtactctcagcgtagtaatttcatctcgcttttgcagtagcct
ggtcgccggccgtaaaacccccctattttctagtggttgacctcggattaggtaggaatac
ccgctgaactNaagcatatcaa

>SS1-033C_ITS4
ttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggaggg
atcattactgagttcttacaaactccaacccctttgtgaaccataccaactgttgcctcgg
cgcgagctgcggctgcttggtagctacccggtagtcacctacccgggagctacccggtag
cgtctgcgtacaggcccgccgaaggaccaccaaactctgtttgacagtgtattctgaatg
cttcaactaaatagttaaaactttcaacaacggatctcttggttctggcatcgatgaaga
acgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttg
aacgcacattgcgcccattagtattctagtgggcatgcctattcgagcgtcatttcgacc
ctgaagccctagttgcttcgcgtcgggactctactggctaccctgtagttccctaatgac
agtggcggagttcaggtgtactctcagcgtagtaatttcatctcgc

>SS1-033E_ITS1F
tgttgcctcggcagaagctacctggttaccttaccttggaacggcctaccctgtagcgcc
ctaccctggaacggcctaccctgtaacggctgccggtggactaccaaactcttgttattt
tattgtaatctgagcgtcttattttaataagtcaaaactttcaacaacggatctcttggt
tctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagt
gaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctgtt

cgagcgtcatttcaacccttaagcctagcttagtgttgggagcctactgcttttgctagc
ggtagctcctgaaatacaacggcggatctgcgatatcctctgagcgtagtaatttttatc
tcgcttttgactggagttgcagcgtctttagccgctaaacccccccaatttttaatggttg
acctcggatcaggtaggaatacccgctgaacttaagcatatca

>SS1-033E_ITS4
tgatatgcttaagttcagcgggtattcctacctgatccgaggtcaaccattaaaaattgg
ggggtttagcggctaaagacgctgcaactccagtcaaaagcgagataaaaattactacgc
tcagaggatatcgcagatccgccgttgtatttcaggagctaccgctagcaaaagcagtag
gctcccaacactaagctaggcttaagggttgaaatgacgctcgaacaggcatgcccacta
gaatactaatgggcgcaatgtgcgttcaaagattcgatgattcactgaattctgcaattc
acattacttatcgcatttcgctgcgttcttcatcgatgccagaaccaagagatccgttgt
tgaaagttttgacttattaaaataagacgctcagattacaataaaataacaagagtttgg
tagtccaccggcagccgttacagggtaggccgttccagggtagggcgctacagggtaggc
cgttccaaggtaaggtaaccaggtagcttctgccgaggcaaca

>SS1-033H_ITS1F
tactattgttgctttggcgggcccgcctcggccaccggctccggctggtgagcgcccgcc
agaggaccccaaactctgaaatttagtgtcgtctgagtactattttaatagttaaaactt
tcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaa
tgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggta
ttccgggggggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggtattggg
cgtccccggcagcggggtgccctaaaatcagtggcggtgccgtctggctctaagcgtagt
aaatctctcgctctggagacccggcggatgcttgccagcaacccccaatttttaaggtt
gacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>SS1-033H_ITS4
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggaggga
tcattacagagttcatgcccttcggggtagatctcccacccgtgttattatactattgtt
gctttggcgggcccgcctcggccaccggctccggctggtgagcgcccgccagaggacccc
aaactctgaaatttagtgtcgtctgagtactattttaatagttaaaactttcaacaacgg
atctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgc
agaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggg
catgcctgttcgagcgtcattacaaccctcaagctctgcttggtattgggcgtccccggc
agcggggtgccctaaaatcagtggcggtgccgtctggctctaagcgtagtaaatctctc

>SS1-033I_ITS1F
tgttgcctcggcagaagctgctcggtgcaccttaccctggaacggcctaccctgtagcgc
cttaccctggaacggcctaccctgtagcggctgccggtggactaccaaactcttgttatt
ttattgtaatctgagcgtcttattttaataagtcaaaactttcaacaacggatctcttgg
ttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcag
tgaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctgt
tcgagcgtcatttcaacccttaagcctagcttagtgttgggagcctactgcttttactag
ctgtagctcctgaaatacaacggcggatctgcgatatcctctgagcgtagtaaatttta
tctcgcttttgactggagttgcagcgtctttagccgctaaatcccccaatttttaatggt

tgacctcggatcaggtaggaatacccgctgaacttaagcatatcaat

>SS1-033I_ITS4
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattata
gagttttctaaactcccaacccatgtgaacttaccattgttgcctcggcagaagctgctc
ggtgcaccttaccctggaacggcctaccctgtagcgccttaccctggaacggcctaccct
gtagcggctgccggtggactaccaaactcttgttattttattgtaatctgagcgtcttat
tttaataagtcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgca
gcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgc
acattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccttaa
gcctagcttagtgttgggagcctactgcttttactagctgtagctcctgaaatacaacgg
cggatctgcgatatcctctgagcgtagtaaatttttatctcgctttt

>SS1-033L_ITS1F
ttactgagtgagggcccccgggctcgacctcccacccttttgtgaacacaacttgttgctt
cgggggcgaccctgccgtttacgcggcgagcgcccccggaggccttcaaacactgcatca
ttgcgtcggagtttaagcaaattaaacaaaactttcaacaacggatctcttggttctggc
atcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatca
tcgaatctttgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcg
tcatttcaccactcaagcctagcttggtattgggcgccgcggtgttccgcgcgcctcaaa
gtctccggctgagctgtccgtctccaagcgttgtgatttcattaatcgctctggagcgcg
ggcggccgcggccgttaaatctttcacaaggttgacctcggatcaggtagggatacccgc
tgaacttaagcatatca

>SS1-033L_ITS4
gaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattactga
gtgagggcccccgggctcgacctcccacccttttgtgaacacaacttgttgcttcgggggc
gaccctgccgtttacgcggcgagcgcccccggaggccttcaaacactgcatcattgcgtc
ggagtttaagcaaattaaacaaaactttcaacaacggatctcttggttctggcatcgatg
aagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatc
tttgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttc
accactcaagcctagcttggtattgggcgccgcggtgttccgcgcgcctcaaagtctccg
gctgagctgtccgtctccaagcgttgtgatttcattaaNcgctctggagcgcgggcggcc
gcggccgttaaat

>SS1-033M_ITS1F
gtttaccgaaccactgttgcttcggcgggcccgtcttcacgaccgccggaggaccgtcaa
accgtcctctggtccgcgcccgccgatggccccaacttaacaaattcttaaatgaatcgt
gtcattgtctgagtactatttgaaaaaataaattaaaactttcaacaacggatctcttgg
ttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgcgaattgcagaattcca
gtgagtcatcgaatctttgaacgcacattgcgccctttggtattccgaagggcatgcctg
ttcgagcgtcattatcacccctcaagccctctcttgggtttggtgttggacggcctggtc
aggcaactgacccctcctaaagacaatgacggcgtcctcgtgggatccccggtacactga
gcttcttaactgagcacgtatcggatcaagggtctacgggacacggtctgaaaacctata
tattttctcaaggttgacctcggatcaggtagggatacccgctgaacttaagcatat

>SS1-033M_ITS4
gaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaaga
gttagggtctctcaggcccgatctcccaacccattgtttaccgaaccactgttgcttcgg
cgggcccgtcttcacgaccgccggaggaccgtcaaaccgtcctctggtccgcgcccgccg
atggccccaacttaacaaattcttaaatgaatcgtgtcattgtctgagtactatttgaaa
aaataaattaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagc
gaaatgcgataagtaatgcgaattgcagaattccagtgagtcatcgaatctttgaacgca
cattgcgccctttggtattccgaagggcatgcctgttcgagcgtcattatcacccctcaa
gccctctcttgggtttggtgttggacggcctggtcaggcaactgacccctcctaaagaca
atgacggcgtcctcgtgggatccccggtacactgagcttcttaactgagcacgtatcgga
tcaagggtctacgggacacggtctgaaaaccta

>SS1-033N_ITS1F
atcgcgaaaggggccgggttgccctacaccNNNtctgggggtgatatggtgggcaactgg
ccccgcttttgattatatcccatctgtgaacttttccactgttgcttccgtgagcctgta
ctttgctctctttgagagttggagcacctgaaaagccctggttctaccaggggcgcacgg
caatggtaccacaactctggttttatgttgtataagtctgagtggttaacacacaaacaa
gttaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatg
cgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcg
cctcctggtattccgggaggcatgcctgttcgagcgtcatcaaacacatcctcaagcatt
tttgcttggtcctggaggaagatggtctcttgactatcccctctggaattcctcagcaga
gaccacagcagcctcaggtgtgataattctacctcgcttgagtgttactgtgggctcccg
ctggagaacagcccccccagcttttaccacgtttgacctcggatcaggtagggatacccgc
tgaacttaagcatatcaa

>SS1-033N_ITS4
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaat
taatcgcgaaaggggccgggttgccctacacccctttctgggggtgatatggtgggcaact
ggccccgcttttgattatatcccatctgtgaacttttccactgttgcttccgtgagcctg
tactttgctctctttgagagttggagcacctgaaaagccctggttctaccaggggcgcac
ggcaatggtaccacaactctggttttatgttgtataagtctgagtggttaacacacaaac
aagttaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaa
tgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattg
cgcctcctggtattccgggaggcatgcctgttcgagcgtcatcaaacacatcctcaagca
tttttgcttggtcctggaggaagatggtctcttgactatcccctctggaattcctcagca
gagaccacagcagcctcaggtgtgataattctacctcgcttgag

>SS1-033P_ITS1F
aattctctgaaaggtcgccggtacccNNNccccctaaacgaggggccggggaaggtcct
ctcacaccttgtgtaccttaccaccgttgctttggcgggccgacccggtttcgacccgg
gcggccggcgcccccagccctcaccggccaggacgtcaggctaagcgcccgccagtatac
aaaactcaagcgattatttcatgtagtcctgatgaattattcaattaattaaaactttca
acaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgt
gaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggtattc
cggagggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtattgggcga

cgtccgctgccggacgcgcctcgaagacctcggcgacggcgtcctagcctcgagcgtagt
agtaatatctcgctttggagtgctgggcgttggccgccggacaatcgacctttggtctat
tacttccaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaata
agcggag

>SS1-033P_ITS4
gaactagtaattctctgaaaggtcgccggtacccggtccccttaaacgaggggccgggg
aaggtcctctcacacccttgtgtaccttaccaccgttgctttggcgggccgacccggttt
cgacccgggcggccggcgccccccagccctcaccggccaggacgtcaggctaagcgcccgc
cagtatacaaaactcaagcgattatttcatgtagtcctgatgaattattcaattaattaa
aactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgata
agtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctc
tggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggta
ttgggcgacgtccgctgccggacgcgcctcgaagacctcggcgacggcgtcctagcctcg
agcgtagtagtaatatctcgctttggag

>SS1-033R_ITS1F
taattaatcgcgaaaggggccgggttgccctacaccNNNtctgggggtgatatggtgggc
aactggccccgcttttgattatatcccatctgtgaacttttccactgttgcttccgtgag
cctgtactttgctctctttgagagttggagtacctgaaaagccctggttctaccaggggc
gcacggcaatggtaccacaactctggttttatgttgtataagtctgagtggtcaacacac
aaacaagttaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagc
gaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcac
attgcgcctcctggtattccgggaggcatgcctgttcgagcgtcatcaaacccatcctca
agcatttttgcttggtcctggaggaagatggtcttctgactatcccctctggaattcctc
agcagagaccacagcagcctcaggtgtgataagtctacctcgcttgagtgttactgtggg
ctcccgctggagaacagccccccagcttttaccatgtttgacctcggatcaggtagggat
acccgctgaacttaagcatatcaata

>SS1-033R_ITS4
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaat
taatcgcgaaaggggccgggttgccctacaccccttctgggggtgatatggtgggcaact
ggccccgcttttgattatatcccatctgtgaacttttccactgttgcttccgtgagcctg
tactttgctctctttgagagttggagtacctgaaaagccctggttctaccaggggcgcac
ggcaatggtaccacaactctggttttatgttgtataagtctgagtggtcaacacacaaac
aagttaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaa
tgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattg
cgcctcctggtattccgggaggcatgcctgttcgagcgtcatcaaacccatcctcaagca
tttttgcttggtcctggaggaagatggtcttctgactatcccctctggaattcctcagca
gagaccacagcagcctcaggtgtgataagtctacctcgcttga

>SS1-033S_ITS1F
gtgaatactacctaaaaaatgttgcctcggcgcaggtNNNNccttttacagggtcccttc
tcccctaggggtttaggagcagaccggccggcggccctataaactcttgtttttgtaaaa
tcatctgagtaaaaaaaataaatgaatcaaaactttcaacaacggatctcttggttctgg

catcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatc
atcgaatctttgaacgcacattgcgcccggtggtattccaccgggcatgcctgttcgagc
gtcatttcaaccctcaaagcctagctttggtgttggaggaagacctgtaaaaaggtcccc
tcccaaatttagtggcgggctcgctagaattttgagcgtagtaattttacctcgttttta
aagactagtgggacttcttgccgtaaaacccccccaactttctgaaaattgacctcgaatc
aggtaggaatacccgctgaacttaagcatatcaataagcggag

>SS1-033S_ITS4
gtcatttagaggaagtaaaagtcgtaacaaggNNNccgttggtgaaccagcggagggatc
attgctggaacaaacgccctcacgggtgctacccagaaaccctttgtgaatactacctaa
aaaatgttgcctcggcgcaggttggccctttacagggtcccttctcccctaggggttta
ggagcagaccggccggcggccctataaNNNNNtgtttttgtaaaatcatctgagtaaaaa
aaataaatgaatcaaaactttcaacaacggatNNNNtggttctggcatcgatgaagaacg
cagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaac
gcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccctc
aaagcctaNNNNtggtgttggaggaagacctgtaaaaaggtcccctcccaaattt

>RAS1-015A_ITS1F
tgttgcctcggcagaagctacctggttaccctaccttggaacggcctaccctgtagcgcc
ttaccctggaacggcctaccctgtaacggctgccggtggactaccaaactcttgttattt
tattgtaatctgagcgtcttattttaataagtcaaaactttcaacaacggatctcttggt
tctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagt
gaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctgtt
cgagcgtcatttcaacccttaagcctagcttagtgttgggagcctaccgcttttgctagc
ggtagctcctgaaatacaacggcggatctgcgatatcctctgagcgtagtaatttttatc
tcgcttttgactggagttgcagcgtctttagccgctaaaccccccaaatttttaatggt
tgacctcggatcaggtaggaatacccgctgaacttaagcatatca

>RAS1-015A_ITS4
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattata
gagttttctaaactcccaacccatgtgaacttaccattgttgcctcggcagaagctacct
ggttaccctaccttggaacggcctaccctgtagcgccttaccctggaacggcctaccctg
taacggctgccggtggactaccaaactcttgttattttattgtaatctgagcgtcttatt
ttaataagtcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcag
cgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgca
cattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccttaag
cctagcttagtgttgggagcctaccgcttttgctagcggtagctcctgaaatacaacggc
ggatctgcgatatcctctgagcgtagtaatttttatctcgctttt

>RAS1-015B_ITS1F
ttgagacagcccgtcggcggccaacctaactcttgtttttacactgaaactctgNNNaca
aaacataaatgaatcaaaactttcaacaacggatctcttggttctggcatNNNtgaaaaa
cgcagcgaaatgcgataagtaatgtgaattgcNNNattcagNgaatcatcgaatctttga
acgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccc
tcaagcctggcttggtgatggggcactgcttcttac

>RAS1-015B_ITS4
gttggtgaaccagcggagggatcattNctggaacgcgccccaggcgcacccagaaaccct
ttgtaaacttataccttactgttgNNNcggcgctagctggtcccNcNNNgcccctcaccN
NcgggtgttgagacagcccgtcggcggccaacctaacNNNtgtttttacNcNgaaactNt
gagcacaaaacataaatgaatcaaaactttcaacaNcggatctcttggttctggcatcNa
tgaagaNNNcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcaNNNNa
tctttgaacgcacattgcgccctNNggtattcNNgagggcatgcctgttcgagcgtcatt
tcaaccctcaagcctggcttggtgatggggcNNNgcttcttacccaagaagcaggccctg
aaattcagtgg

>RAS1-015H_ITS1F
gatcattacagagactctgccctttgggtagacctcccaccctgtgtcgttatacctttg
ttgctttggcgggccgcggggctccggccctgcccctggctccggctagggcgcgcccgc
cagaggactccaaaacctgaatgttagtgtcgtctgagtactatataatagttaaaactt
tcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaa
tgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggta
ttccgggggggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggccttggg
cgtcaccggtccccggtgtgccttaaaatcagtggcggcgccgtctggctctaagcgtag
taaacactctcgctacagacgcccggcggatgctggccagcaacccccaatctatcaagg
ttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>RAS1-015H_ITS4
gaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacaga
gactctgccctttgggtagacctcccaccctgtgtcgttatacctttgttgctttggcgg
gccgcggggctccggccctgcccctggctccggctagggcgcgcccgccagaggactcca
aaacctgaatgttagtgtcgtctgagtactatataatagttaaaactttcaacaacggat
ctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcag
aattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggca
tgcctgttcgagcgtcattacaaccctcaagctctgcttggccttgggcgtcaccggtcc
ccggtgtgccttaaaatcagtggcggcgccgtctggctctaagcgtagtaaacactctcg
ctNcNgacgcccggcggatgctggccagcaacccccaat

>RAS1-015K_ITS1F
accgaactttgttgctttggcaggcccctggtcactagctgactcgcgcagctaccagcc
atgcctgccaaaggacccaaactctgaaatttttgctgtctgagtactatgtaatagtta
aaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgat
aagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccct
ctggtattccgggggggcatgcctgttcgagcgtcattacaaccctcacgcctgacgtggt
attgggcagacccttttggggcctgcctcaaagtcagtggcggtaccgtgtgaccccgaag
cgtagtactactcgtcgcttctgggcgtcaggcggcagcctgccagcaaacccccacttt
acaaggttgacctcgaatcaggtagggatacccgctgaacttaagcatatcaataagcgg
ag

>RAS1-015K_ITS4
tcttggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaa
ggatcattaaagaatatcaggcctcacggccctattctcacccttttgtctaccgaacttt
gttgctttggcaggcccctggtcactagctgactcgcgcagctaccagccatgcctgcca
aaggacccaaactctgaaatttttgctgtctgagtactatgtaatagttaaaactttcaa
caacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtg
aattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggtattcc
ggggggcatgcctgttcgagcgtcattacaaccctcacgcctgacgtggtattgggcaga
ccctttggggcctgcctcaaagtcagtggcggtaccgtgtgaccccgaagcgtagtacta
ctcgtc

>RAS1-015L_ITS1F
ccctctggtattccgggggggcatgcctgttcgagcgtcattacaaccctcacgcctggNN
NNNtattgggcagacccttggggcctgcctcaaagtcagtggcggcaccgtgtgac

>RAS1-015L_ITS4
cttggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaag
gatcattaaagaatatcaggcctcacggccctattctcacccttttgtataccgaactttg
ttgctttggcaggcccctggtcactagctgactcgcgcagctaccagctatgcctgccaa
aagacccaaactctgaaatttttgctgtctgagtactatgtaatagttaaaactttcaac
aacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtga
attgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggtattccg
gggggcatgcctgttcgagcgtcattacaaccctcacgcctggcgtggtattgggcagac
cctttggggcctgcctcaaagtcagtggcggcaccgtgtgaccccgaagcgtagtactac
tcgtc

>RAS1-015P_ITS1F
tactttgttgccttggcgccgtgcgccagtgNNNNNaaacccgtgaatcattgctgtctg
agtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaag
aacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatcttt
gaacgcacattgcgccccctggcattccgggggggcatgcttgttcgagcgtcattacaac
cctcaagctccgcttggtgttgggctcgcctttctcggaaggcctgcctcaaagtcagtg
gcggccgccgtccgaccttcagcgcagtaatgctcgtcgctgcaagggaagggtagcaag
cgccgtcatacaacccccacacacaaggttgacctcggatcaagtagggatacccgctga
acttaagcatatcaataagcggag

>RAS1-015P_ITS4
cttggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaag
gatcattaaagaataacacgcctccgggcgctattctcacccttgcgtaccatactttg
ttgccttggcgccgtgcgccagtggctcgaaacccgtgaatcattgctgtctgagtacta
tataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcag
cgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgca
cattgcgccccctggcattccgggggggcatgcttgttcgagcgtcattacaaccctcaag
ctccgcttggtgttgggctcgcctttctcggaaggcctgcctcaaagtca

>RAS1-015Q_ITS1F
aaacaagcttccagcgcaagtcagcactactggcaacacaatcgaattgcggggacgctt
taaagcccaccggtaccaacagcccttctaagggctgggcgagcatagctcgaggttaca
acccagtgggatggtacaataagcaatccgcagcggccgagccgtccacagactaagtgg
ttgtgggtaggagaaccctatctaagatatagtcgggcctgaggagagatccccggggct
tctgcgtccgtaggtgaacctgcggaaggatcattaataaggataccgggaaaccggtag
acccacccgtgtctatctactcttgttgctttggcaggccgtggcctccaccgtgggctc
tgcctgcgcgtgcctgccagaggaccaaactctgaattttagtgatgtctgagtactata
taatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcg
aaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaca
ttgcgcccggtggtattccgccgggcatgcctgttcgagcgtcattataaccactcaagc
ctagcttggtattggggttcgcggttccgcggctcctaaaatcggtggcggtgccggtgg
gctctaagcgtagtaaatctcctcgctatagggtccctccggtcacttgccagaaccccc
aatttttacaggttgacctcggat

>RAS1-015Q_ITS4
gaggaagtaaaagtcgtaacaaggttttgctaagcaacagtcNNNctgtagcagagctta
acggagaccttgcacccgagaggggggaggcgactataaacaagcttccagcgcaagtca
gcactactggcaacacaatcgaattgcggggacgctttaaagcccaccggtaccaacagc
ccttctaagggctgggcgagcatagctcgaggttacaacccagtgggatggtacaataag
caatccgcagcggccgagccgtccacagactaagtggttgtgggtaggagaaccctatct
aagatatagtcgggcctgaggagagatccccggggcttctgcgtccgtaggtgaacctgc
ggaaggatcattaataaggataccgggaaaccggtagacccacccgtgtctatctactct
tgttgctttggcaggccgtggcctccaccgtgggctctgcctgcgcgtgcctgccagagg
accaaactctgaattttagtgatgtctgagtactatataatagttaaaactttcaacaac
ggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaatt
gcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccggtggtattccgccg
ggcatgcctgttcgagcgtcattataaccactcaagcctaNNNtggtattggggttcgcg
gttccgcggctcctaaaatcggtggcggtgccggtgggctctaagcgtagtaaatctc

>RAS1-015R_ITS1F
cattgcgccctgtggtattccgcagggcatgcctgttcgagcgtcattatagccaattaa
gcctagcttagtattggggctcgcggtctcgcggcctctaaaaccagtggcggtgccggt
aggctctgagcgtagtaatactcctcgctatagagtcctgtcggttccctgccagaaccc
ccatattttatcaggttgacct

>RAS1-015R_ITS4
tgtagcagagcttaacggagaccttgcacccgagaggggggaggcgactataaaagaact
cccggcgcaagtcagcaccgctggcaacacaatcgaattgcggggacgctttaaagctca
ccagtaccaactgctcttccagggcagggctagcttagctagaggtcacaacctggtagg
atgctacaataagc

>RAS1-015U_ITS1F
atcattaaggagtataagacgaagcttcggcttctgactccccaNNNNttgcttaccatc
ccccatacttcggcgggcgatatagagggaatgaacttctagcgatagaggatcctccct

ctgtcNNNgctcgccggcattacaaaacctcaagctataacagtgctgtctgagcatttg
caagcaaaagaataaaaactttcaacaacggatctcttggttctggcatcgatgaagaac
gcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaa
cgcacattgcgccctctggtattccggagggcacacctgttcgagcgctagtacaacaat
caagctcagcttggtgttggatctggccctttcctcggaagggacctttccaaaactcat
gggcgccgccgtccggctccaagcgtagcaagcgaaagccctcgctttggatagtcgggt
cgatggccgccggtaaacacttggggtaacccaaattactacagtttagcctcggatcag
gtggggatacccgctgaacttaagcatatcaataagcggag

>RAS1-015U_ITS4
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaagga
tcattaaggagtataagacgaagcttcggcttctgactccccatcctttgcttaccatcc
cccatacttcggcgggcgatatagagggaatgaacttctagcgatagaggatcctccctc
tgtcccagctcgccggcattacaaaacctcaagctataacagtgctgtctgagcatttgc
aagcaaaagaataaaaactttcaacaacggatctcttggttctggcatcgatgaagaacg
cagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaac
gcacattgcgccctctggtattccggagggcacacctgttcgagcgctagtacaacaatc
aagctcagcttggtgttggatctggccctttcctcNgaagggacctttccaaaactcatg
ggcgccgccgtccggctccaagcgtagcaagcgaaagccctcgctttggat

>JD1-047D_ITS1F
cttataccttactgttgcctcggcgctagctggtccctcggggcccctcaccctcgggtg
ttgagacagcccgtcggcggccaacctaactcttgtttttacactgaaactctgagcaca
aaacataaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaa
cgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttga
acgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccc
tcaagcctggcttggtgatggggcactgcttcttacccaagaagcaggccctgaaattca
gtggcgagctcgccaggaccccgagcgcagtagttaaaccctcgctctggaaggccctgg
cggtgccctgccgttaaaccccccaacttctgaaaatttgacctcggatcaggtaggaata
cccgctgaacttaagcatatcaataagcggagga

>JD1-047D_ITS4
cttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagg
gatcattgctggaacgcgccccaggcgcacccagaaaccctttgtgaacttataccttac
tgttgcctcggcgctagctggtccctcggggcccctcaccctcgggtgttgagacagccc
gtcggcggccaacctaactcttgtttttacactgaaactctgagcacaaaacataaatga
atcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatg
cgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcg
ccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggct
tggtgatggggcactgcttcttacccaagaagcaggccctgaaattcagtggcgagctcg
ccaggaccccgagcgcagtag

>JD1-047E_ITS1F
gatcattattagaagccgaaaggctacttaaaaccatcgcgaactNNNccaagttgcttc
ggcggcgcggcctccctcacgggggcgccgcagccccgcctctccggaggtgtggggcgc

ccgccggaggtacgaaactctgtattatagtggcatctctgagtaaaaaacaaataagtt
aaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcga
taagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccc
gctagtactctagcgggcatgcctgttcgagcgtcatttcaaccctcaagccctgcttgg
tgttggggccctacggctgccgtaggccctgaaaggaagtggcgggctcgctacaactcc
gagcgtagtaattcattatctcgctagggacgttgcggcgcgctcctgccgttaaagacc
atctttaactcaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatc
aa

>JD1-047E_ITS4
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattatta
gaagccgaaaggctacttaaaaccatcgcgaactcgtccaagttgcttcggcggcgcggc
ctccctcacgggggcgccgcagccccgcctctccggaggtgtggggcgcccgccggaggt
acgaaactctgtattatagtggcatctctgagtaaaaaacaaataagttaaaactttcaa
caacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtg
aattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgctagtactct
agcgggcatgcctgttcgagcgtcatttcaaccctcaagccctgcttggtgttggggccc
tacggctgccgtaggccctgaaaggaagtggcgggctcgctacaactccgagcgtagtaa
ttcattatctcgctagggacgttgcggcgcgctcctgccgttaaagaccatctt

>JD1-047F_ITS1F
tcctgggccgtggggacgggtcccccgtgtggggataccgccgcatgggccttcgcctcc
cttcgattatatcccatctgtgaacacttccactgttgcttccgtgagcctgctgctgac
tccttctgggggtgagcgcctgaaaagccctggctctgctaggggcgcacggcaatgcta
caatcaccctggttttatgtatctccagtctgagtggtgaaacacatcaaacaagttaaa
actttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatgcgataa
gtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcctcct
ggtattccgggaggcatgcctgttcgagcgtcatgaacaacatcctcaagcattcttgct
tggttgtggaggaagagggtccttggaccgtcccctctggaatttctcagcagagaccgc
aacagcctcaggtgtgataacattacctcgcttgagggctgcagtgggctcccgctgcgt
tgatccccccagcaatactatgtttgacctcggatcaggtagggatacccgctgaactta
agcatatcaataagcggag

>JD1-047F_ITS4
aggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaattaa
tcgcgaaatgcgtcgtgccatgtaagactcacggcctcctgggccgtggggacgggtccc
ccgtgtggggataccgccgcatgggccttcgcctcccttcgattatatcccatctgtgaa
cacttccactgttgcttccgtgagcctgctgctgactccttctgggggtgagcgcctgaa
aagccctggctctgctaggggcgcacggcaatgctacaatcaccctggttttatgtatct
ccagtctgagtggtgaaacacatcaaacaagttaaaactttcaacaacggatctcttggt
tctcgcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagt
gaatcatcgaatctttgaacgcacattgcgcctcctggtattccgggaggcatgcctgtt
cgagcgtcatgaacaacatcctcaagcattcttgcttggttgtggaggaagagggtcctt
ggaccgtcccctctggaatttctcagcagagaccgcaacagcctcaggtgtgataacatt
acctcgctt

>JD1-047G_ITS1F
gatcattaccgagttataacaactccaaccctttgtgaaccttaccgtcgtttcctcggc
gcactgctgcgaccgccccgcaggcggccctcgcggccccgcgggcagctacttaccctg
tacggatctaccctgcagcagctgaaggcgcagcggcgccgaaaggaccgctacaaactc
ttttatccaagctacgtctgaacaatttactatgcaaaatagtttaaaactttcaacaac
ggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaatt
gcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccattagtattctagtg
ggcatgcctattcgagcgtcatttcgacccttaagccctgttgcttagcgttgggagtct
gcggcttcggccgcagttccttaaagtcagtggcggagctgtggcacactctaggcgtag
tagttttaccgcctcgcctccagagtggccccggctgcctgccgtaaaaccccctaatttt
ctagtggttgacctcggattaggtaggaatacccgctgaacttaagcatatcaa

>JD1-047G_ITS4
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattaccg
agttataacaactccaaccctttgtgaaccttaccgtcgtttcctcggcgcactgctgcg
accgccccgcaggcggccctcgcggccccgcgggcagctacttaccctgtacggatctac
cctgcagcagctgaaggcgcagcggcgccgaaaggaccgctacaaactcttttatccaag
ctacgtctgaacaatttactatgcaaaatagtttaaaactttcaacaacggatctcttgg
ttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcag
tgaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctat
tcgagcgtcatttcgacccttaagccctgttgcttagcgttgggagtctgcggcttcggc
cgcagttccttaaagtcagtggcggagctgtggcacactctaggcgtagtagttttaccg
cctcgcctccagagtggccccggctgcctgccgtaaaaccccctaatt

>JD1-047K_ITS1F
ttgcttcggcgggcaggggaagcctctcgcgggcctcccctcccggcgccggcccccacc
agggggacggggcgcccgccggaggaaaccaaactctatttacacgacgtctcttctgag
tggcacaagcaaataattaaaacttttaacaacggatctcttggttctggcatcgatgaa
gaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctt
tgaacgcacattgcgctcgccagcattctggcgagcatgcctgttcgagcgtcatttcaa
ccctcaagcaccgcttggttttggggccccacggcacacgtgggcccttaaaggtagtgg
cggaccctcccggagcctcctttgcgtagtaactaacgtctcgcactgggattcggaggg
actcttgccgtaaaaccccccaatttttttacaggttgacctcggatcaggtaggaatac
ccgctgaacttaagcatatcaa

>JD1-047K_ITS4
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggaggga
tcattactgagttaccgctctataacccttttgtgaacatacctaaccgttgcttcggcgg
gcaggggaagcctctcgcgggcctcccctcccggcgccggcccccaccaggggggacgggg
cgcccgccggaggaaaccaaactctatttacacgacgtctcttctgagtggcacaagcaa
ataattaaaacttttaacaacggatctcttggttctggcatcgatgaagaacgcagcgaa
atgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacatt
gcgctcgccagcattctggcgagcatgcctgttcgagcgtcatttcaaccctcaagcacc
gcttggttttggggccccacggcacacgtgggcccttaaaggtagtggcggaccctcccg
gagcctcctttgcgtagtaactaacgtctcg

>JD1-047L_ITS1F
gtttttacaccgaaactctgagaataaaacataaatgaatcaNNactttcaac

>JD1-047L_ITS4
cttggtcatttagaggaagtaaaagtcgtaacaaggtctNNNttggtgaaccagcggagg
gatcattgctggaacgcgcNccaggcgcacccagaaaccctttgtgaacttataccttac
tgttgNNNNNgcgctggctggtcNNNNgggggcccctcaccNNcgggtgttgagacggccc
gNNNgNggccaacctgacNattgtttttacaccNNNacNNNgagaataaaacataaatga
atcaaaactttcaacaacggatNNNNtggttctggcatcgatgaagaacgcagNNaaatg
cgataagtaatgtgaattgcagaattcagtgaatcaNNNaatctttgaacgcacattgaN
cNcNctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcN
tggtgatggggcactgcttttacccaaaagcaggccctgaaatt

>JD1-047O_ITS1F
cccctcaccctcgggtgttgagacggcccgccggcggccaacctaactcttgtttttaca
ctgaaactctgagaataaaacataaatgaatcaaaactttcaacaacggatctcttggtt
ctggcatcgatgaNNaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtg
aatcatcgaatctttgaaNNcacattgcgccctctggtattccgNNgggcatgcctgttN
NNNNgtcatttcaaccctcaagcctggcttNNNgatggggcactgcttttaccca

>JD1-047O_ITS4
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggaggga
tcattgctggaacgcgccccaggcgcacccagaaacNNNttgtgaacttataNcttactg
ttgcctcggcgctagctggtccttcggggcccctcacNcNcgggtgttgagacggcccgc
cggcggccaacctaacNNNtgtttttacactgaaactctgagaataaaacataaatgaat
caaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgNN
ataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcc
cNNNggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcttg
gtgatggggcactgcttttacccaaaagcaNNccctgaaattcag

>JD1-047P_ITS1F
gttgcctcggcgtgcgctgcggctacccggtttccggccccgcagacgggtggttaccct
gtagccggtcaacagcccgccgaaggacccctaaactctgtttaaaatggtgtattctga
attacttaaactaaaataagttaaaactttcaacaacggatctcttggttctggcatcga
tgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaa
tctttgaacgcacattgcgcccattagtactctagtgggcatgcctattcgagcgtcatt
tcgaccattaagccctgtagcttagcgttgggagcctacgcccttgcggcgtagctcctt
aaagtcagtggcggagttggagcacactctaagcgtagtaaacttatctcgcttctgtag
ttgccccggctgcctgccgtaaaaccccctatatttctagtggttgacctcggattaggt
aggaatacccgctgaacttaagcatatcaata

>JD1-047P_ITS4
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattacag
agttaaacaaactccaaacccttgtgaaccttaccataagttgcctcggcgtgcgctgc
ggctacccggtttccggccccgcagacgggtggttaccctgtagccggtcaacagcccgc

cgaaggacccctaaactctgtttaaaatggtgtattctgaattacttaaactaaaataag
ttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgc
gataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgc
ccattagtactctagtgggcatgcctattcgagcgtcatttcgaccattaagccctgtag
cttagcgttgggagcctacgcccttgcggcgtagctccttaaagtcagtggcggagttgg
agcacactctaagcgtagtaaacttatctcgc

>JD1-047Q_ITS1F
actccacccttttgtgaaccataccaactgttgcctcggcgcgagctgcggctgcttggta
gctacccggtagtcacctacccgggagctacccggtagcgtctgcgtacaggcccgccga
aggaccaccaaactctgtttgacagtgtattctgaatgcttcaactaaatagttaaaact
ttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagta
atgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccattagt
attctagtgggcatgcctattcgagcgtcatttcgaccctgaagccctagttgcttcgcg
ttgggactctactggctaccctgtagttccctaatgacagtggcggagttcaggtgtact
ctcagcgtagtaatttcatctcgcttttgcagtagcctggtcgccggccgtaaaaccccc
tattttctagtggttgacctcggattaggtaggaatacccgctgaacttaagcatatca

>JD1-047Q_ITS4
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattact
gagttcttacaaactccaaccctttgtgaaccataccaactgttgcctcggcgcgagctg
cggctgcttggtagctacccggtagtcacctacccgggagctacccggtagcgtctgcgt
acaggcccgccgaaggaccaccaaactctgtttgacagtgtattctgaatgcttcaacta
aatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcga
aatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacat
tgcgcccattagtattctagtgggcatgcctattcgagcgtcatttcgaccctgaagccc
tagttgcttcgcgttgggactctactggctaccctgtagttccctaatgacagtggcgga
gttcaggtgtactctcagcgtagtaatttcatctcgNNNNtgcagtagcctggtcgccgg
ccgtaaaacccccc

>JD1-047S_ITS1F
gcagaggccccctggagacagggagcagcccgccggtggccaatcaaactctgtttctat
agtgaatctctgagtaaaaaaaacataaatgaatcaaaactttcaacaacggatctcttg
gttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattca
gtgaatcatcgaatctttgaacgcacattgcgccctctggtattccggagggcatgcctg
ttcgagcgtcatttcaaccctcaagcctggcttggtgatggggcactgcctgtgaaaggg
caggccctgaaatctagtggcgagctcgccaggaccccgagcgtagtagttatatctcgc
tctggaaggccctggcggtgccctgccgttaaacccccaacttctgaaaatttgacctcg
gatcaggtaggaatacccgctgaacttaagcatatcaataagcggag

>JD1-047S_P1917325_ITS4
tggtcatttagaggaagtaaaagttgtaacaaggtctccgttggtgaaccagcggaggga
tcattgctggaacgcgccccggcgcacccagaaacccttgtgaacttatacccatactg
ttgcctcggcgcaggccggcctctgctgcagaggccccctggagacagggagcagcccgc
cggtggccaatcaaactctgtttctatagtgaatctctgagtaaaaaaaacataaatgaa

tcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgc
gataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgc
cctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggctt
ggtgatggggcactgcctgtgaaagggcaggccctgaaatctagtggcgNNNNcgccagg
accccgagcgtagtagttatatctcgc

>JD1-047W_ITS1F
taccaaaactttaggtttattccccccttcttaggggggtNNNgccatacccttttgtctacg
tgtacctatcgttgtttcctcgggggggctatttgccccccgataggaatcccttaaaccc
tttgtattatagtattaaacctctgataactacctaaataatcacaactttcaacaatgg
atctcttggttctggcatcgatgaagaacgcagcgaaatgcgaaaagtagtgtgaattgc
agaattccgtgaatcatcgaatctttgaacgcacattgcgcccctcggtattccgtgggg
catgcctgttcgagcgtcatttacaccctcaagctctgcttggtgttgggcgtctgtcct
gcttcatgcatggactcgccccaaagtcattggcagcggtcgtgccagcttctcgcgcag
cacatttgcgtttctcgaagctctggtggatcagcgtccagtaagctctctttatgactt
gacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>JD1-047W_ITS4
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaagga
tcattaccaaaactttaggtttattccccccttcttagggggttgagccatacccttttgtc
tacgtgtacctatcgttgtttcctcgggggggctatttgccccccgataggaatcccttaa
acccttttgtattatagtattaaacctctgataactacctaaataatcacaactttcaaca
atggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgaaaagtagtgtgaa
ttgcagaattccgtgaatcatcgaatctttgaacgcacattgcgcccctcggtattccgt
ggggcatgcctgttcgagcgtcatttacaccctcaagctctgcttggtgttgggcgtctg
tcctgcttcatgcatggactcgccccaaagtcattggcagcggtcgtgccagcttctcgc
gcagcacatttg

>JD1-047X_ITS1F
acgccctcgcgggcgctacccagaaacccttttgtgaacttatacctatcactgttgcctc
ggcatggttggcctcttatgaggtcccttctgtacaggaggagcaggccggccggtggcc
ctataaactcttgttttttttgtaatatcatctgagcaaacaactataaatgaatcaaaac
tttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagt
aatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccggtgg
tattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagcctggctttggtgt
tggaggaccacttttcgttaagtggcctctgaaatttagtggcgggctcgctaggatttt
gagcgtagtaattatacctcgtttttaaagactagtggttctcttgccgtaaaaccccccc
aacttttctgaaaattgacctcggattaggtaggaatacccgctgaacttaagcatatca
ataagcggag

>JD1-047X_ITS4
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgct
ggaacaaacgccctcgcgggcgctacccagaaacccttttgtgaacttatacctatcactg
ttgcctcggcatggttggcctcttatgaggtcccttctgtacaggaggagcaggccggcc
ggtggccctataaactcttgttttttttgtaatatcatctgagcaaacaactataaatgaa

tcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgc
gataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgc
ccggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagcctggct
ttggtgttggaggaccacttttcgttaagtggcctctgaaatttagtggcgggctcgcta
ggattttgagcgtagtaattatacctcgtt

>JD1-047Y_ITS1F
ttatcgtggggcttcggccccgtcgagaaaataccccttgcctttatgagaaccatggtgt
aggtcttccctgggattgtccgggatcggagcgtccatgtacctttttcctcggccggct
tgccggccaatggggaacacaaaaacactgtcattgctgtaactgtctgataaaacaaac
aaaaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcga
aatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacat
tgcgccctttggtattccttagggcatgcctgttcgagcgtcatttaaaccttcaagcac
tgcttggtgttgggtgcctgtcccgcctccgcgcgtggactcacctcaaattcattggcg
gccctcacgtcggctacgagcgcagcagaaacgcgaactcgtgggcccgacggggaggct
cccagaagctacaattcaccattttgacctcggatcaggtagggataccgctgaactta
agcatatcaa

>JD1-047Y_ITS4
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattatc
gtggggcttcggccccgtcgagaaaataccccttgcctttatgagaaccatggtgtaggtc
ttccctgggattgtccgggatcggagcgtccatgtacctttttcctcggccggcttgccg
gccaatggggaacacaaaaacactgtcattgctgtaactgtctgataaaacaaacaaaaa
tcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgc
gataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgc
cctttggtattccttagggcatgcctgttcgagcgtcatttaaaccttcaagcactgctt
ggtgttgggtgcctgtcccgcctccgcgcgtggactcacctcaaattcattggcggccct
cacgtcggctacgagcgcagcagaaacgcgaactcgtgggcccgacggggaggctcccag
aagctacaattcaccat

>JD1-047Z_ITS1F
gatcattgctggaacgcgccccggcgcacccagaaacccttNNNNNactcataccttac
tgttgcctcggcgcaggccggccccccctgggggcccctcaccccggtgaggagcaggccc
gccggcggccaagccaactcttgtttttacactgaagctctgagaaacaaacacaaatga
atcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatg
cgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcg
ccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggct
tggtgatggggcgctgcttcctgcccagggagcaggccctgaaattcagtggcgggctcg
ccaggaccccgagcgcagtagttaaaccctcgctctggaaggccctggcggcgccctgcc
gttaaaccccccaactcctgaaaatttgacctcggatcaggtaggaatacccgctgaactt
aagcatatcaataagcggag

>JD1-047Z_ITS4
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggaggga
tcattgctggaacgcgccccggcgcacccagaaacccttgtgaactcataccttactg

ttgcctcggcgcaggccggccccccctgggggcccctcaccccggtgaggagcaggcccgc
cggcggccaagccaactcttgtttttacactgaagctctgagaaacaaacacaaatgaat
caaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcg
ataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcc
ctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcttg
gtgatggggcgctgcttcctgcccagggagcaggccctgaaattcagtggcgggctcgcc
aggaccccgagcgcagtagttaaaccctcgctc

>JD1-047AA_ITS1F
ccgtcaaaagcttccctggtacatgcctcacccctttgtatatctaccatgttgctttggc
gggccgacccggtttcgacccgggcggccggcgccccccagcctgcttgccaggccaggac
gcccggccaagtgcccgccagtatacaaaactccagcgattattttgtgtagtcctgaga
atttattcaataaattaaaactttcaacaacggatctcttggttctggcatcgatgaaga
acgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttg
aacgcacattgcgccctctggcattccggagggcatgcctgttcgagcgtcatttcaacc
ctcaagctctgcttggtattgggcgacgtctgctgtcagacgcgcctggaagacctcggc
gacggcattccagcctcgagcgtagtagtaaaatatctcgctttggaggatggggtgacg
gcttgccggacaaccgacctctggtcatttttttccaaggttgacctcggatcaggtaggg
atacccgctgaacttaagcatatcaataagcggag

>JD1-047AA_ITS4
aggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattactgaa
aatgtaataaaccccttcaggttttggaaggggggagccgtcaaaagcttccctggtacatg
cctcacccccttgtatatctaccatgttgctttggcgggccgacccggtttcgacccgggc
ggccggcgcccccagcctgcttgccaggccaggacgcccggccaagtgcccgccagtata
caaaactccagcgattattttgtgtagtcctgagaatttattcaataaattaaaactttc
aacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatg
tgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggcatt
ccggagggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtattgggcg
acgtctgctgtcagacgcgcctggaagacctcggcgacggcattccagcctcgagcgtag
tagtaaaatatctcgct

>JD1-047AC_ITS1F
taatcgcgaaaggggccgggttgccctacacccctttctgggggtgatatggtgggcaact
ggccccgcttttgattatatcccatctgtgaacttttccactgttgcttccgtgagcctg
tactttgctctctttgagagttggagtacctgaaaagccctggttctaccaggggcgcac
ggcaatggtaccacaactctggttttatgttgtataagtctgagtggttaacacacaaac
aagttaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaa
tgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattg
cgcctcctggtattccgggaggcatgcctgttcgagcgtcatcaaacacatcctcaagca
tttttgcttggtcctggaggaagatggtctcttgactatcccctctggaattcctcagNN
NNNaccacagcagcctNNNNtgtgataattctacctcgcttgagtgttactgtgggctcc
cgctggagaacagcccccagcttttaccatgtttgacc

>JD1-047AC_ITS4
aaggatcattaattaatcgcgaaaggggccgggttgccctaNacccctttctgggggtgat
atggtgggcaactggccccgcttttgattatatcccatctgtgaacttttccactgtNNN
Ntccgtgagcctgtactttgctctctttgagagttggagtacctgaaaagccctggttct
accaggggcgcacggcaatggtaccacaactctggttttatgttgtataagtctgagtgg
ttaacacacaaacaagttaaaactttcaacaacggatctcttggttctcgcatcgatgaa
gaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctt
tgaacgcacattgcgcctcctggtattccgggaggcatgcctgttcgagcgtcatcaaac
acatcctcaagcatttttgcttggtcctggaggaagatggtctcttgactatcccctctg
gaattcctcagcagagaccacagcagcctcaggtgtgataattctacctcgcttg

>JD1-047AI_ITS1F
ttataccttactgttgcctcggcgctagctggtccttcggggcccctcaccctcgggtgt
tgagacggcccgccggcggccaacctaactcttgtttttacactgaaactctgagaataa
aacataaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaac
gcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaa
cgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccct
caagcctggcttggtgatggggcactgcttttacccaaaagcaggccctgaaattcagtg
gcgagctcgccaggaccccgagcgcNNNNNNtaaaccctcgctttggaaggccctggcgg
tgccctgccgttaaaccccccaacttctgaaaatttgacctcggatcaggtaggaatacc
cgctgaacttaagcatatcaata

>JD1-047AI_ITS4
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggaggga
tcattgctggaacgcgccccaggcgcacccagaaaccctttgtgaacttataccttactg
ttgcctcggcgctagctggtccttcggggcccctcaccctcgggtgttgagacggcccgc
cggcggccaacctaactcttgtttttacactgaaactctgagaataaaacataaatgaat
caaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcg
ataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcc
ctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcttg
gtgatggggcactgcttttacccaaaagcaggccctgaaattcagtggcgagctcgccag
gaccccgagcgcagtagttaaaccctcgct

>JD1-047AJ_ITS1F
tcattgctggaacaaacgccctaacgggtgctacccagaaaccctttgtgaatactacct
aaaaaatgttgcctcggcgcaggttggccctttacagggtcccttctcccctaggggtt
taggagcagaccggccggcggccctataaactcttgtttttgtaaaatcatctgagtaaa
aaaaataaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaa
cgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttga
acgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccc
tcaaagcctagctttggtgttggaggaagacctgtaaaaaggtcccctcccaaatttagt
ggcgggctcgctagaattttgagcgtagtaattttacctcgtttttaaagactagtggga
cttcttgccgtaaaaccccccaactttctgaaaattgacctcgaatcaggtaggaatacc
cgctgaacttaagcatatcaataagcggag

>JD1-047AJ_ITS4
ggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggat
cattgctggaacaaacgccctaacgggtgctacccagaaaccctttgtgaatactaccta
aaaaatgttgcctcggcgcaggttggccctttacagggtcccttctcccctaggggttt
aggagcagaccggccggcggccctataaactNNtgtttttgtaaaatcatctgagtaaaa
aaaataaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaac
gcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaa
cgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccct
caaagcctagctttggtgttggaggaagaccNgtaaaaaggtcccctcccaaatttagtg
gcgggctcgctagaattttgagNNtagtaattttacctcgttNNNaagactagtgggact
tcttg

>JD1-047AN_ITS1F
ataaaagaactcccagcgcaagtcagcccagctggcaacacaatcgaattgcggggacgc
tttaaagcccaccagtaccaactgcccccaagggcagggccagctatgctggaggtcaca
acctggtaggatattacaataagcaatccgcagcggctcagccgtccacagactaagtgg
ttgtgggtagggatcccctacttaagatatagtcgggccctgtgacaagtcatagggct
tctgcgtccgtaggtgaacctgcggaaggatcattaaccattggataccttcgggtatat
cccatccgtgtctacatactcttgttgctttggcaggccgtggtcacccactgtgggcta
tgcctgcatgtgcctgccagaggaccaaactctgaattttagtgatgtctgagtactata
taatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcg
aaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaca
ttgcacccggtggtattccgccgggtatgcctgttcgagcgtcattataaccactcaagc
ctgtcttggtgttggggttgcgaatattttgcagccctcgagtctcgtagcgccacctgt
gggttctaagcgtagtaatttctcctcgctataggacctgctcggggaaaagtataattc
gtagcctggttctatgtgtccgct

>JD1-047AN_ITS4
aagcaacagtccttctgtagcagagcttaacggagaccttgcacccgagagNgggggagg
cgactataaaagaactcccagcgcaagtcagcccagctggcaacacaatcgaattgcggg
gacgctttaaagcccaccagtaccaactgcccccaagggcagggccagctatgctggagg
tcacaacctggtaggatattacaataagcaatccgcagcggctcagccgtccacagacta
agtggttgtgggtagggatcccctacttaagatatagtcgggccctgtgacaagtcatag
gggcttctgcgtccgtaggtgaacctgcggaaggatcattaaccattggataccttcggg
tatatcccatccgtgtctacatactcttgttgctttggcaggccgtggtcacccactgtg
ggctatgcctgcatgtgcctgccagaggaccaaactctgaattttagtgatgtctgagta
ctatataatagttaaaactttcaacaacggatctcNNNNttctggcatcgatgaagaacg
cagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaac
gcacattgcacccggtggtattccgccgggtatgcctgttcgagcgtcattataaccact
caagcctgtcttggtgttggggttgcgaatattttgcagccctcgagtctcgtagcgcca
cctgtgggttctaagcgtagtaatttctcctcgctataggacctgctcggggaaaagtat
aat

>JD1-047AO_ITS1F
aacccttgcctttatgagtaccgtccgtttcctcggcaggtccgcctgccgatggacaaa
aaaaataaaccctttatcacctgtatcagtctgaaaaaacaaataaaaattaaaactttc
aacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtg
tgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtatt
ccttagggcatgcctgttcgagcgtcatttaaaccttcaagctctgcttggtgttgggtg
tctgtcccgcctcggcgcggactcgccttaaattcattggcggccggtacattggcttcg
agcgcagcagaaacgcgaactcggagactttggatcagctcccagaagctattttctatt
tttgacctcggatcaggtagggatacccgctgaacttaagcatatca

>JD1-047AO_ITS4
gaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattatcgt
agggcttcggccctgtcgagatagaacccttgcctttatgagtaccgtccgtttcctcgg
caggtccgcctgccgatggacaaaaaaaataaacccttttatcacctgtatcagtctgaaa
aaacaaataaaaattaaaactttcaacaacggatctcttggttctggcatcgatgaagaa
cgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttga
acgcacattgcgccctttggtattccttagggcatgcctgttcgagcgtcatttaaacct
tcaagctctgcttggtgttgggtgtctgtcccgcctcggcgcggactcgccttaaattca
ttggcggccggtacattggcttcgagcgcagcagaaacgcgaactcggagactttgga

>JD1-047AP_ITS1F
acccttatgtgaacatacctactgttgcttcggcggactcgccccggcgtccggacggcc
tcgcgccgcccgcggcccggacccaggcggccgccggagacccccaaactctgtattatc
agcattttctgaatccgccgcaaggcaaaacaaatgaatcaaaactttcaacaacggatc
tcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcaga
attcagtgaatcatcgaatctttgaacgcacattgcgcccgccagcattctggcgggcat
gcctgttcgagcgtcatttcaaccctcgacttccctttggggaaatcggcgttggggaac
ggcagcataccgccggccccgaaatgaagtggcggcccgtccgcggcgacctctgcgtag
taatccaactcgcaccggaaccccgacgtggccacgccgtaaaacacccaacttctgaac
gttgacctcggatcaggtaggaatacccgctgaacttaagcatatca

>JD1-047AP_ITS4
aggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattatcgag
tttacaactcccaaaccttatgtgaacatacctactgttgcttcggcggactcgccccg
gcgtccggacggcctcgcgccgcccgcggcccggacccaggcggccgccggagaccccca
aactctgtattatcagcattttctgaatccgccgcaaggcaaaacaaatgaatcaaaact
ttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagta
atgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccagc
attctggcgggcatgcctgttcgagcgtcatttcaaccctcgacttccctttggggaaat
cggcgttggggaacggcagcataccgccggccccgaaatgaagtggcggcccgtccgcgg
cgacctctgcgtagtaatccaactcgcaccggaaccccgacgtggccacgcc

>JD1-047AQ_ITS1F
acgccctcacgggtgctacccagaaacccttgtgaatactacctcaaaaatgttgcctc
ggcgcaggttggccctttacagggtcccttctcccctaggggtttaggagcagaccggc

cggcggccctataaactcttgtttttgtaaaatcatctgagtaaaaaaaataaatgaatc
aaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcga
taagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccc
ggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagcctagcttt
ggtgttggaggaagacctgtaaaaaggtcccctcccaaatttagtggcgggctcgctaga
attttgagcgtagtaattttacctcgtttttaaagactagtgggacttcttgccgtaaaa
ccccccaactttctgaaaattgacctcgaatcaggtaggaatacccgctgaacttaagca
tatcaata

>JD1-047AQ_ITS4
ttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggaggg
atcattgctggaacaaacgccctcacgggtgctacccagaaaccctttgtgaatactacc
tcaaaaatgttgcctcggcgcaggttggccctttacagggtcccttctccctaggggt
ttaggagcagaccggccggcggccctataaactcttgtttttgtaaaatcatctgagtaa
aaaaaataaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaaga
acgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttg
aacgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaacc
ctcaaagcctagctttggtgttggaggaagacctgtaaaaaggtcccctcccaaatttag
tggcgggctcgctagaattttgagcgtagtaattttacctcgttt

>JD1-047AR_ITS1F
cccagaaacccttttgtgaatactacctaaaaatgttgcctcggcgcaggttggccctttt
acagggtcccttctcccctaggggtttaggagcagaccggccggcggccctataaactct
tgtttttgtaaaatcatctgagtaaaaaaaaaaaaatgaatcaaaactttcaacaacgga
tctcttggttctggc

>JD1-047AR_ITS4
taaaaaaaaaaaaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatg
aagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatc
tttgaacgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttc
aaccctcaaagcctagctttggtgttggaggaagacctgtaaaaaggtcccctcccaaat
ttagtggcgggctcgctagaattttgagcgtagtaattttacctcgttt

>JD1-047AX_ITS1F
atcattatcgtagggcttcggccctgtcgagatagaacccttgNNNttatgagtaccgtc
cgtttcctcggcaggtccgcctgccgatggacaaaaaaaataaacccttttatcacctgta
tcagtctgaaaaaacaaataaaaattaaaactttcaacaacggatctcttggttctggca
tcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcat
cgaatctttgaacgcacattgcgccctttggtattccttagggcatgcctgttcgagcgt
catttaaaccttcaagctctgcttggtgttgggtgtctgtcccgcctcggcgcggactcg
ccttaaattcattggcggccggtacattggcttcgagcgcagcagaaacgcgaactcgga
gactttggatcagctcccagaagctattttctatttttgacctcggatcaggtagggata
cccgctgaacttaagcatatcaataagcggagga

>JD1-047AX_ITS4
cttggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaag
gatcattatcgtagggcttcggccctgtcgagatagaacccttgcctttatgagtaccgt
ccgtttcctcggcaggtccgcctgccgatggacaaaaaaaataaaccctttatcacctgt
atcagtctgaaaaaacaaataaaaattaaaactttcaacaacggatctcttggttctggc
atcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatca
tcgaatctttgaacgcacattgcgccctttggtattccttagggcatgcctgttcgagcg
tcatttaaaccttcaagctctgcttggtgttgggtgtctgtcccgcctcggcgcggactc
gccttaaattcattggcggccggtacattggcttcgagcgcagcagaa

>JD1-047AZ_ITS1F
acgccctcacgggtgctacccagaaaccctttgtgaatactacctaaaaatgttgcctcg
gcgcaggttggccctttacagggtcccttctcccctaggggttaaggagcagaccggcc
ggcggcccttataaactcttgtttttgtaaaatcatctgagtaaaaaaaaataaatgaat
caaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcg
ataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcc
cggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagcctagctt
tggtgttggaggaagacccgtaaaaaggtcccctcccaaatttagtggcgggctcgctag
aattttgagcgtagtaattttacctcgtttttaaagactagtgggacttcttgccgtaaa
acccccccaactttctgaaaattgacctcgaatcaggtaggaatacccgctgaacttaagc
atatcaataagcggag

>JD1-047AZ_ITS4
ttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggaggg
atcattgctggaacaaacgccctcacgggtgctacccagaaaccctttgtgaatactacc
taaaaatgttgcctcggcgcaggttggccctttacagggtcccttctcccctaggggtt
aaggagcagaccggccggcggccccttataaactcttgtttttgtaaaatcatctgagtaa
aaaaaaataaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaag
aacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatcttt
gaacgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaac
cctcaaagcctagctttggtgttggaggaagacccgtaaaaaggtcccctcccaaattta
gtggcgggctcgctagaattttgagcgtagtaattttacctcgtttttaaagactagtgg
gacttcttgcc
>JD1-047BB_ITS1F
gatcattagaaaacgagttgcgagactcgaaaaccatcgcgaacatcacctcgtcgtcgt
tgcttcggcgggcggccccgcgaaggggcccgcaaggcccgcgagggcgcccgccggagg
agcgtgaacactctgttttcttagcgcctctctgagtggatacagcgaaatgaatcaaaa
ctttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataag
taatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgcca
gtattctggcgggcatgcctgttcgagcgtcatttcaacccctcaagccccgcttggtgc
tggggccctacgtgcgccgtaggccccgaattacagtggcgggctcgccgcgatcccgag
cgtagtaacacacctcgctctggggctccgcggcgggtgccggccgtgaaaacgaccaaa
acacaaacgcaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatca
a

>JD1-047BB_ITS4
tagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattaga
aaacgagttgcgagactcgaaaaccatcgcgaacatcacctcgtcgtcgttgcttcggcg
ggcggccccgcgaaggggcccgcaaggcccgcgagggcgcccgccggaggagcgtgaaca
ctctgttttcttagcgcctctctgagtggatacagcgaaatgaatcaaaactttcaacaa
cggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaat
tgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccagtattctggc
gggcatgcctgttcgagcgtcatttcaacccctcaagccccgcttggtgctggggcccta
cgtgcgccgtaggccccgaattacagtggcgggctcgccgcgatcccgagcgtagtaaca
cacctcgctctgggg

>JD1-047BH_ITS1F
gagcttaacggagaccttgcacccgagaggggggaggcgactataaacaagcttccagcg
caagtcagcactgctggcaacacaatcgaattgcggggacgctttaaagcccaccggtac
caacggctccctttaggggccgggcgagcatgctcgaggtcacaatctggtgggatagca
caataagcaatccgcagcggcagagccgtccacagactaagtggttgtgggtaggagaac
cctatctaagatatagtcgggcccgaggagagatccccggggcttctgcgtccgtaggtg
aacctgcggaaggatcattaaaaggatgccggaaacggtagaccccacccgtgtttacat
actcttgttgctttggcaggccgtggcctccaccgtgggttctgcctgcgtgtgcctgcc
agaggaccaaactctgaatgttagtgatgtctgagtactatataatagttaaaactttca
acaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgt
gaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccggtggtattc
cgccgggcatgcctgttcgagcgtcattataaccactcaagcctcgcttggtattggggt
acgcggttccgcggctcctaaaatcagtggcggtgc

>JD1-047BH_ITS4
aggaagtaaaagtcgtaacaaggttttgtgaagcaacagtccttctgtagcagagcttaa
cggagaccttgcacccgagaggggggaggcgactataaacaagcttccagcgcaagtcag
cactgctggcaacacaatcgaattgcggggacgctttaaagcccaccggtaccaacggct
ccctttaggggccgggcgagcatgctcgaggtcacaatctggtgggatagcacaataagc
aatccgcagcggcagagccgtccacagactaagtggttgtgggtaggagaaccctatcta
agatatagtcgggcccgaggagagatccccggggcttctgcgtccgtaggtgaacctgcg
gaaggatcattaaaaggatgccggaaacggtagaccccacccgtgtttacatactcttgt
tgctttggcaggccgtggcctccaccgtgggttctgcctgcgtgtgcctgccagaggacc
aaactctgaatgttagtgatgtctgagtactatataatagttaaaactttcaacaacgga
tctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgca
gaattcagtgaatcatcgaatctttgaacgcacattgcgcccggtggtattccgccgggc
atgcctgttcgagcgtcattataaccactcaagcctcgcttggtattggggtacgcggtt
ccgcggctcctaaaatcagtggcggtgccggtgggctctaagcgtagtacatctcctcgc
tacagggtccttccggttgtctgccagaaccccccattttacaggtg

>JD1-047BJ_ITS1F
gaaattcagtggcgagctcgccaggaccccgagcgcagtagttaaaccctcgctttggaa
ggccctggcggtgccctgccgttaaaccccccaacttctgaaaatttgacctcggatcNN
NtNggaatacccgctgaacttaagNatatcaataagcggag

>JD1-047BJ_ITS4
cttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagg
gatcattgctggaacgcgccccaggcgcacccagaaaccctttgtgaacttataccttac
tgttgcctcggcgctagctggtNNtNNNgggcccctcaccctcgggtgttgagacggccc
gccggcggccaacctaacNNNtgtttttacactgaaacNcNgagaataaaacataaatga
atcaaaactttcaacaacggatctcttggtNNNggcatcgatgaagaacgcagcgaaatg
cgataagtaatgtgaattgcagaattcagtgaatcaNNNaatctttgaacgcacattgcg
ccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggct
tggtgatggggcNNNgcttttacccaa

>JD1-047BK_ITS1F
acgccctcacgggtgctacccagaaaccctttgtgaatactacctaaaaatgttgcctcg
gcgcaggttggccctttacagggtcccttctcccctaggggtttaggagcagaccggcc
ggcggccctataaactcttgtttttgtaaaatcatctgagtaaaaaaaaaaatgaatcaa
aactttcaacaacggatctcttggttctggcatcgatgaaaaacgcagcgaaatgcgata
agtaatgtgaattgcaaaattcagtgaatcatcgaatctttgaacgcacattgcgcccgg
tggtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagcctNNNtttgg
tgttggaggaagacctgtaaaaaggtcccctcccaaatttaNNggcgggctcgctagaat
tttgagcgtagtaattttacctcgtttttaaagactaNNgggacttcttgccgtaaaacc
ccccaactttctgaaaattgacctNNNatcaggtaggaatacccgctgaacttaagcata

>JD1-047BK_ITS4
ccgttggtgaaccagNNgagggatcattgcNggaacaaacgccctcacgggtgctaccca
gaaaccctttgtgaatactacctaaaaatgttgcctcggcgcaggttggccctttacag
ggtccctNNNcccctaggggtttaggagcagaccggccggcggccctataaactcttgtt
tttgtaaaatcaNNNgagtaaaaaaaaaaatgaatcaaaactttcaacaacggatctctt
ggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattc
agtgaatcatcgaatctttgaacgcacattgcgcccggtggtattccaccgggcatgcct
gttcgagcgtcatttcaaccctcaaagcctagctttggtgttggaggaagacctgtaaaa
aggtcccctcccaaatttagtggcgggctcgctagaattttgagcgtagtaattttacct
cgtttttaaagactagtgggacttcttgccgtaaaaccccccaact

>JD1-047BM_ITS1F
acaggatgcaagtcgcgatggcggcaacataatcgaattgcggggatctcctaaagccaa
gcagtaccaacaccgtccagaaatgtagggtgggccagcgttaatcccgctggaggttag
aatctgctaggatgccacaatggatgatccgcagcgaagcccctaacggcgccagcctac
ggggaacgttcacagactaagtggttgtgggtaggctcgtcctgcttaagatatagtcgg
gccctgtaggaaactacaggggctagccgttccgtaggtgaacctgcggaaggatcatta
cagagttcatgcccttcggggtagacctcccacccttgtgtatcattatagaatgttgct
ttggcgggccgcgcctcgtgcgcctagattcgcgtctagcgtgcccgccagaggacccct
aaactctgaatgttagtgtcgtctgagtactattaaatagttaaaactttcaacaacgga
tctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgca
gaattcagtgaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggc
atgcctgttcgagcgtcatttataccaatctagcctggctaggtgttgggcgccgccgcc
tggcggaccttaaaactagtggcggtgctctcaagctctacgcgtagtaatcttctcgct

acagggtcttgggagatgctagccagcaaccccaaatttttctaggttgacctcgga

>JD1-047BM_ITS4
cggggcgcttgccgaagNNNatgcaaccctcacggggtccgcgactataaacacaggatg
caagtcgcgatggcggcaacataatcgaattgcggggatctcctaaagccaagcagtacc
aacaccgtccagaaatgtagggtgggccagcgttaatcccgctggaggttagaatctgct
aggatgccacaatggatgatccgcagcgaagcccctaacggcgccagcctacggggaacg
ttcacagactaagtggttgtgggtaggctcgtcctgcttaagatatagtcgggccctgta
ggaaactacaggggctagccgttccgtaggtgaacctgcggaaggatcattacagagttc
atgcccttcggggtagacctcccacccttgtgtatcattatagaatgttgctttggcggg
ccgcgcctcgtgcgcctagattcgcgtctagcgtgcccgccagaggacccctaaactctg
aatgttagtgtcgtctgagtactattaaatagttaaaactttcaacaacggatctcttgg
ttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcag
tgaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgt
tcgagcgtcatttataccaatctagcctggctaggtgttgggcgccgccgcctggcggac
cttaaaactagtggcggtgctctcaagctctacgcgtagtaatcttct

>JD1-047BN_ITS1F
cggagaccttgcacccgagagggggggaggcgactataagagaactcccagcgcaagtcag
cctagctggcaacacaatcgaattgcggggacgctttaaagcccaccggtaccaactgct
cctctggggcagggccagttatgctggaggtcacaacccggtaggatgctacaataagca
atccgcagcggctccgccgtccacagactaagtggttgtgggtagggatcccctacttaa
gatatagtcgggccctgtgataagtcatagggggcttctgcgtccgtaggtgaacctgcgg
aaggatcattaaccattggataccttcgggtatatcccatccgtgtctacatactcttgt
tgctttggcaggccgtggtcacccactgtgggctatgcctgcatgtgcctgccagaggac
caaactctgaatattagtgatgtctgagtactatacaatagttaaaactttcaacaacgg
atctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgc
agaattcagtgaatcatcgaatctttgaacgcacattgcacccggtggtattccgccggg
tatgcctgttcgagcgtcattataaccactcaagcctgtcttggtgttggggttgcgaat
cttttgcagccctcgagtctcgtagcgccacctgtgggttctaagcgtagttatttctcc
tcgctatagaatctgct

>JD1-047BN_ITS4
tgtagcagagcttaacggagaccttgcacccgagagggggggaggcgactataagagaact
cccagcgcaagtcagcctagctggcaacacaatcgaattgcggggacgctttaaagccca
ccggtaccaactgctcctctggggcagggccagttatgctggaggtcacaacccggtagg
atgctacaataagcaatccgcagcggctccgccgtccacagactaagtggttgtgggtag
ggatcccctacttaagatatagtcgggccctgtgataagtcatagggggcttctgcgtccg
taggtgaacctgcggaaggatcattaaccattggataccttcgggtatatcccatccgtg
tctacatactcttgttgctttggcaggccgtggtcacccactgtgggctatgcctgcatg
tgcctgccagaggaccaaactctgaatattagtgatgtctgagtactatacaatagttaa
aactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgata
agtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcacccgg
tggtattccgccgggtatgcctgttcgagcgtcattataaccactcaagcctgtcttggt
gttggggttgcgaatcttttgcagccctcgagtctcgtagcgccacctgtgggttctaag

cgtagttatttctcctcgctatagaatctgctgagggaaaagtataatNNNtagcctggt
tctatgtgtccgctatcaaacccccaatttttaagg

Appendix 2: DNA consensus sequences of endophytic fungi

>TC2-011
ccgtaggtgaacctgcggagggatcattacacaaatatgaaggcgggctggaacctctcggggttacagccttgctgaattattc
acccttgtcttttgcgtacttcttgtttccttggtgggttcgcccaccactaggacaaacataaaccttttgtaattgcaatcagcgtca
gtaacaaattaataattacaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgt
gaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtattccaaagggcatgcctgttcgagcgtcatt
tgtaccctcaagctttgcttggtgttgggcgtcttgtctctagctttgctggagactcgccttaaagtaattggcagccggcctactg
gtttcggagcgcagcacaagtcgcactctctatcagcaaaggtctagcatccattaagccttttttcaacttttgacctcggatcag
gtagggatacccgctgaacttaagcatat

>TC2-012
ctaaccgttgcttcggcgggcaggggaagcctctcgcgggcttccycwccckkygccgkmccmyacywygkkgmck
kggcgccgtgcgccrgwggamwcsaaactcyaktkacaysaykkctcktctgagtggcacwakmwaatarttaaaactt
tyaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatca
tcgaatctttgaacgcacattgcgcycscyrgcattcyggsgrgcatgcytgttcgagcgtcattwcaaccctcaagcwccgctt
ggtkttgggctcgccttycwcggmmgrcstgggccctyaaagkyagtggcggmcgcckyccggagcctcctywgcgy
agtaaykmwcgtckctgcacwgggatmsggwrgsamkcttgccgtcatacaacccccamattcwytacaggttgacct
cggatcargtaggratacccgctgaacttaagcatatcaataagcggag

>TC2-013
ctgaccagcnngagccggttgtcttgcgacggggccgccaaagcaacaaaggtattgtatacagagggtgggaggtctaccc
cggagggcatgatctcattaatgatncttccgtaggtgaacctgcggaaggatcattaatgagatcatgccctccggggtagacc
tcccaccctctgtatacaataccttttgttgctttggcggccccgtcgcaagacaaccggctccggctggtcagcggccgccagag
gaatcaaaactcttattattattgtcgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaag
aacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccctggtatt
ccgggggggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggtattgggtgtcacccccgggtgcgccttaaaatc
agtggcggtgccgtctggcttcaagcgtagtaatacttctcgctttggagtccgggcgagcgtcctgccaaaaccccccatatttttt
caggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggagga

>TC2-014
cgagggtttgactactgcgctcggggtcctggcrasctcrmcwctkrwtttyasrgcctgmwwctctgrgyacaraasmtaa
rtgmmycawmactttcaagcmacggmtctswkggttgaamtgrcgmtcgawsargcawgcscwscgraatascaga
krgmgyaatgtgmrttgcaragattcrrtgawtcayygaatctytgcaaykcacattgcgmcytmtsgyattycggaskgcr
tkcytswtcgagygycatttsaaccmwsagakccgtkgcttgaaagtkwtgrkkcatttmtgyttyktrcycagagwwkc
aggccctgtaaawwcagtggmgagytmgsywggmcsccgacgsgcwgtastcaacaccckmgckstggarggsccy
sgmggtgmccwgccgtwrmrccsagscaacagtwmkgtawaakttsacmwmggrtcwkstrggwryrccygskg
mrcgttcmagcaatgatcmmtmcgcaggttcacctacgg

>TC2-017
acctgcggaaggatcattaaagaataacacgcctccgggcgctattctcacccttgcgtaccatactttgttgccttggcgccgtg
cgccagtggctcgaaacccgtgaatcattgctgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggca
tcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccc
cctggcattccgggggggcatgcttgttcgagcgtcattacaaccctcaagctccgcttggtgttgggctcgcctttctcggaaggc

ctgcctcaaagtcagtggcggccgccgtccgaccttcagcgcartaatgctcgtcgctgcaagggaagggtagcaagcgccgt
catacaacccccacacacaaggttgacctcggatcaagtagggatacccgctgaacttaagcatatcaataagcggagg

>TC2-018
ccgtaggtgaacctgcggagggatcattacaagtgaccccggtctaaccaccgggatgttcataaccctttgttgtccgactctgt
tgcctccggggcgaccctgccttcgggcgggggctccgggtggacacttcaaactcttgcgtaactttgcagtctgagtaaactt
aattaataaattaaaactttaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattg
cagaattcagtgaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgttcgagcgtcatttcacc
actcaagcctcgcttggtattgggcaacgcggtccgccgcgtgcctcaaatcgaccggctgggtcttctgtcccctaagcgttgt
ggaaactattcgctaaagggtgttcgggaggctacgccgtaaaacaaccccatttctaaggttgacctcggatcaggtagggata
cccgctgaacttaagcatatcaataagcggagga

>TC2-019
ccgtaggtgaacctgcggagggatcattactgagttaccgctctataacccttttgtgaacgtacctaaccgttgcttcggcgggca
ggggaagcctctcgcgggcctcccctcccggcgccggcccccaccacggggacggggcgcccgccggaggaaaccaaac
tctatttacacgacgtctcttctgagtggcacaagcaaataattaaaacttttaacaacggatctcttggttctggcatcgatgaagaa
cgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgctcgccagcattct
ggcgagcatgcctgttcgagcgtcatttcaaccctcaagcaccgcttggttttggggccccacggccgacgtgggcccttaaag
gtagtggcggaccctcccggagcctcctttgcgtagtaactaacgtctcgcactgggatccggagggactcttgccgttaaaccc
ccaaattctttacaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatcaataagcggagga

>TC2-020
gcggaaggatcattactgagtgagggccctctgggtccaacctcccacccgtgtttattgtaccttgttgcttcggtgcgcccgcct
cacggccgccgggggggcttctgccccccgggtccgcgcgcaccggagacaccattgaactctgtctgaagattgcagtctgag
cataaactaaataagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgataactaatg
tgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggcatgcctgtccgagcgtc
attgctgccctcaagcacggcttgtgtgttgggctccgtccccccgggggacgggtccgaaaggcagcggcggcaccgagtcc
ggtcctcgagcgtatggggctttgtcacccgctctgtaggcccggccggcgccagccgacaaccaatcatccttttttcaggttg
acctcggatcaggtagggatacccgctgaacttaagcatatcaat

>TC2-022
cgtaggtgaacctgcggagggatcattacacaaatatgaaagcagnnnctcattggggcgtaacgtcacgccgtgtccgcaat
agcgccattgctgttgtgnngacgcggcgggggggttaacgccctttgggccagtctgactccatattcacccatgnnntttgcg
tattacttgtttccttggcgggttcgcccgccaattggacccaattaaaccttttttgtaattgcaatcancgtcagcaaaacaatgta
ataaattacaactttcaacaacggatcttttggtnnnggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcag
aattcagtgaatcannnaatcnttgaacgcacattgcgccctttggtatnccaaagggcatgcctgttcgagcgtcatttgtaccct
caagctttgcttggtgttgggcgtttttgnnncgggtccgcccnnagnnncgccttaaaatcattggcagccggnnnactggttt
cggagcgcagcacatttttgcgc

>TC2-025
gggatcattatagagtttttaaactcccaacccatgtgnacttaccattgttgcctcggcagaagctacctggttaccttaccttgga
acggcctaccctgtagcgccttaccctggaacggcctaccctgtaacggctgccggtggactaccaaactcttgttatttttattgta
atctgagcgtcttattttaataagtcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgata
agtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcga
gcgtcatttcaacccttaagcctagcttagtgttgggagcctactgcttttgctagcggtagctcctgaaatacaacggcggatctg

cgatatcctctgagcgtagtaatttttatctcgcttttgactggagttgcagcgtctttagccgctaaaccccccaatttttaatggttga
cctcggatcaggtaggaatacccgctgaacttaagcatatcaat

>TC2-026
gtaggtgaacctgcggagggatcattacacaaatatgaaagcagaccctcattggnnnntaacgtcacgccgtgtccgcaata
gcgccattgctgttgtgctgacgcggcgggggttaacgccctttgggccagtctgactccatattcacccatgtcttttgcgtact
acttgtttccttggcgggttcgcccgccaattggacccaattaaacctttttgtaattgcaatcagcgtcagcaaaacaatgtaataa
attacaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattc
agtgaatcatcgaatctttgaacgcacattgcgccctttggtattccaaagggcatgcctgttcgagcgtcatttgtaccctcaagct
ttgcttggtgttgggcgtttttgtctcgggtccgccccgagactcgccttaaaatcattggcagccggcctactggtttcggagcgc
agcacatttttgcgctttgtccagtcgtggtcttgcgtccatgaagcgaatattttcaacgtttgacctcggatcaggtagggatacc
cgctgaacttaagcatatcaat

>TC2-027
acctgcggagggatcattactgagttaccgctctataacccttttgtgaacgtacctaaccgttgcttcggcgggcaggggaagcc
tctcgcgggcctcccctcccggcgccggcccccaccacggggacggggcgcccgccggaggaaaccaaactctatttacac
gacgtctcttctgagtggcacaagcaaataattaaaacttttaacaacggatctcttggttctggcatcgatgaagaacgcagcga
aatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgctcgccagcattctggcgagcat
gcctgttcgagcgtcatttcaaccctcaagcaccgcttggttttggggccccacggccgacgtgggcccttaaaggtagtggcg
gaccctcccggagcctcctttgcgtagtaactaacgtctcgcactgggatccggagggactcttgccgttaaaccccccaaattctt
tacaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatcaat

>TC2-028
ccgtaggtgaacctgcggagggatcattatcgagttaccgctctacaacccttttgtnaacatacctaactgttgcttcggcgggca
gggggtccccccggggacgccctcccggccccgcccccctaccaggggacgtggcgcccgccggaggatacctaactctat
tttaacgacgtttcttctgagtggcacaagcaaataattaaaacttttaacaacggatctcttggttctggcatcgatgaagaacgca
gcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccagcattctggcg
ggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtgttggggccctacggttgacgtakgccctcaaaggtagtg
gcggaccctcccggagcctcctttgcgtagtaactaacgtctcgcatcgggatccggagggactcttgccgtaaaacccccaa
tttttcaatggttgacctcggatcaggtaggaatacccgctgaacttaagcatatcaataagcggag

>TC2-029
ttyrytrcctckgyagcracytacccgggwacggcyaccytgtarcgwgctgmcggtgrastkstraactcttgttatttattgta
atctgagcgtcttattttaataagtcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgata
cgtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcga
gcgtcatttcaacccttaagcctagcttagtgttgagaatctactgtattgtagttctctaaatccaacggcggatctgtggtatcctct
gagcgtagtaatttttatctcgcttttgttaggtgctgcagctctcggccgctaaaccccccaattttaat

>TC2-030
aacctgcggagggatcattacaagaacgcccgggcttcggcctggttattcataacccttttgttgtccgactctgttgcctccggg
gcgaccctgccttcgggcgggggctccgggtggacacttcaaactcttgcgtaactttgcagtctgagtaaacttaattaataaatt
aaaactttttaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagt
gaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgttcgagcgtcatttcaccactcaagcct
cgcttggtattgggcaacgcggtccgccgcgtgcctcaaatcgtccggctgggtcttctgtcccctaagcgttgtggaaactattc

gctaaagggtgttcgggaggctacgccgtaaaacaaccccatttctaaggttgacctcggatcaggtagggatacccgctgaac
ttaagcatatcaata

>TC2-031
ttccgtaggtgaacctgcggaaggatcattacagagttcatgcccgaaagggtagacctcccacccttgtgtattattactttgttgc
tttggcgagctgccttcgggccttgtatgctcgccagagaataccaaaactctttttattaatgtcgtctgagtactatataatagttaa
aactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtg
aatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggcatgcctgttcgagcgtcatttcaaccctcaagcttag
cttggtattgagtctatgtcagtaatggcaggctctaaaatcagtggcggcgccgctgggtcctgaacgtagtaatatctctcgtta
caggttctcggtgtgcttctgccaaaacccaaatttttctatggttgacctcggatcaggtagggatacccgctgaacttaagcatat
caataagcggag

>TC2-032
gcggaaggatcattacctagagttgtaggctttgcctgctatctcttacccatgtcttttgagtaccttcgtttcctcggcgggtccgc
ccgccgattggacaatttaaaccatttgcagttgcaatcagcgtctgaaaaaacttaatagttacaactttcaacaacggatctcttg
gttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcac
attgcgcccttggtattccatggggcatgcctgttcgagcgtcatttgtaccttcaagctctgcttggtgttgggtgtttgtctcgcct
ctgcgcgtagactcgcctcaaaacaattggcagccggcgtattgatttcggagcgcagtacatctcgcgctttgcactcataacg
acgacgtccaaaagtacttttttacactcttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>TC2-033
agccttgctgaattattcacccttgtcttttgcgtacttcttgtttccttggtgggttcgcccaccactaggacaaacataaaccttttgt
aattgcaatcagcgtcagtaacaaattaataattacaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcga
aatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtattccaaagggcat
gcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtcttgtctctagctttgctggagactcgccttaaagtaatt
ggcagccggcctactggtttcggagcgcagcacaagtcgcactctctatcagcaaaggtctagcatccattaagccttttt

>TC2-034
ccgtaggtgaacctgcggagggatcattacagagttatccaactcccaaacccatgtgracttatctctttgttgcctcggcgcaag
ctacccgggacccagcgcccccgggcggcccgccggcggacaaaccaaactcttgttattcaatggattatctgagcgtcttattt
aataagtcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgca
gaattcagtgaatcatcgaatctttgaacgcacattgcgcccatcagtattctggtgggcatgcctgttcgagcgtcatttcaacccc
taagcacagcttactgttgggactctacggcctccgtagttccccaaagccattggcggagtggcagtagtcctctgagcgtagt
aattctttatctcgcttttgttaggtgctgccccccggccgttaaacccccaatttttctggttgacctcggatcaggtaggaatac
ccgctgaacttaagcatatcaataagcggag

>TC2-035
acctgcggaaggatcattaataaaaggannncttcgggtataccccatccgtgtctacatactcttgttgctttggcaggccgtgtt
aacccactgtgggctctgcctgcatgtgcctgccagaggaccaaactcygaattttagtgatgtctgagtactatataatagttaaa
actttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtga
atcatcgaatctttgaacgcacattgcacccggtggtattccgccgggtatgcctgttcgagcgtctgtagaacaacaccaatata
gcgagcttcggcttgattttggagttgcggcgacaagctgcagcttttaagtccacgggcggttccgataggttctacgcgtagta
atttctcctcgctactgaactttgtcggcgcatttggtaataagtccagcaatggccactccaagcccctaacacccaattctagttt
gacctcggatcaagtagggatacccgctgaacttaagcatatc

>TC2-036
cggagggatcattacaagtgaccccggtctaaccaccgggatgttcataaccctttgttgtccgactctgttgcctccggggcgac
cctgccttcgggcgggggctccgggtggacacttcaaactcttgcgtaactttgcagtctgagtaaacttaattaataaattaaaac
ttttaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatc
atcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgttcgagcgtcatttcaccactcaagcctcgcttg
gtattgggcaacgcggtccgccgcgtgcctcaaatcgaccggctgggtcttctgtcccctaagcgttgtggaaactattcgctaa
agggtgttcgggaggctacgccgtaaaacaaccccatttctaaggttgacctcggatcaggtagggatacccgctgaacttaag
cata

>TC2-039
tgtcttttgagtaccttcgtttcctcggcgggtccgcccgccgattggacaaaacttaaacccttttgtaattgaaatcagcgtctgaa
aaaacttaatagttacaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtga
attgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccatggggcatgcctgttcgagcgtcatttg
taccttcaagctttgcttggtgttgggtgtttgtctcgcctctgcgcgcagactcgcctcaaaacaattggcagccggcgtattgattt
cggagcgcagtacatctcgcgctttgcactcataacgacgacgtccaaaaagtacatttttacactcttgacctcggatcaggtag
ggatacccgctgaacttaagcatat

>TC2-041
tgcggagggatcattacagagttgcaaaactccctaaamcattgtgaacgttacctataccgttgcttcggcgggcggccccgg
ggtttaccccccgggcgcccctgggccccaccgcgggcgcccgccggaggtcaccaaactcttgataatttatggcctctctga
gtcttctgtactgaataagtcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagta
atgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccagcattctggcgggcatgcctgttcgagc
gtcatttcaaccatcaagcccccgggcttgtgttggggacctgcggctgccgcaggccctgaaaagcagtggcgggctcgctg
tcgcaccgagcgtagtagcatacatctcgctctggtcgcgccgcgggttccggccgttaaaccacctttaacccaaggttgacc
tcggatcaggtaggaagacccgctgaacttaagcatatcaat

>TC2-042
ttccgtaggtgaacctgcggaaggatcattacctagagttgcgggctttgcctgccatctcttacccatgtcttttgagtaccttcgttt
cctcggcgggtccgcccgccgattggacaaaacttaaacccttttgtaattgaaatcagcgtctgaaaaaacttaatagttacaactt
tcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatca
tcgaatctttgaacgcacattgcgccccttggtattccatggggcatg

>TC2-043
ctgcggaaggatcattaaagaataacacgcctccgggcgctattctcacccttgcgtaccatactttgttgccttggcgccgtgcg
ccagtggctcgaaacccgtgaatcattgctgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatc
gatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccc
ctggcattccgggggggcatgcttgttcgagcgtcattacaaccctcaagctccgcttggtgttgggctcgcctttctcggaaggcc
tgcctcaaagtcagtggcggccgccgtccgaccttcagcgcagtaatgctcgtcgctgcaagggaagggtagcaagcgccgt
catacaacccccacacacaaggttgacctcggatcaagtagggatacccgctgaacttaagcatatc

>TC2-044
tcattaacctttcaaaacagggtgcgtcgcggcccccgaggtgcggaaacaagcctcgagcggtttggggtcctccccttcacg
cgcgcacgactgccatccttactttacragcaccttctgttctccctcggcggggcaacctgccgttggaaccgaataaactcttttt
gcatctagcattacctgttccgaaacaaacaatcgttacaactttcaacaatggatctcttggctctggcatcgatgaagaacgcag
cgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccctcggtattccgtggg

gcatgcctgttcgagcgtcatctacaccctcaagctctgcttggtgttgggcgtctgtcccgcctccgcgcgtggactcgccccaa
attcattggcagcggtccttgcctcctctcgcgcagcacattgcgcttcttgaggggctccggttcgcgtccaacaagcacat

>TC2-045
gaacctgcggagggatcattacaagaacgcccgggcttcggcctggttattcataacccttttgttgtccgactctgttgcctccgg
ggcgaccctgccttcgggcgggggctccgggtggacacttcaaactcttgcgtaactttgcagtctgagtaaacttaattaataaa
ttaaaacttttaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattca
gtgaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgttcgagcgtcatttcaccactcaagc
ctcgcttggtattgggcaacgcggtccgccgcgtgcctcaaatcgtccggctgggtcttctgtcccctaagcgttgtggaaactat
tcgctaaagggtgttcgggaggctacgccgtaaaacaaccccatttctaaggttgacctcggatcaggtagggatacccgctga
acttaagcata

>TC2-046
accctgtagtccgccctgtaaaacgggcggctacnnngtagccggctcacggcccgccgaaggacccctaaactctgtttaaa
atagtgtatttctgaatgctccaaacaaaaataagttaaaactttcaacaanggatctcttggtnntggcatcgatgaagaacgcag
cgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatnnaatctttgaacgcacattgcgcccattagtatnnnagtgg
gcatgcctattcgagcgtcatttcgaccattaagccctgtagcttagcgttgggagnnnnnggttgtagcgtagctcctcaaagtc
agtggcggagttatggcgtactctgagcgtagtaaatcttttctcgct

>TC2-047
ttcggattgtgagccttcacttacaaaacatcagtctatgaatgtaaanttttataacaaataaaaactttcaacaacggatctcttgg
ctctcgcatcgatgaagaacgcatcgaaatgcgatnnntaatgtgaattgcagaattcagtgaatcatnnnatctttgaacgcatc
ttgcgctctctggtattccggagagcatgtcnnnttgagtnnnnnnaattcttcaacccaatcttttcttgtaatcgattggtgtttggt
ttctgagcgttgctggcttcggccaagctcgttcgtaatacattnncatccctcatacaagtttggattgacttggcgtaatnnnctat
tcgctaaggattnnatggaaacatcgagccaacttaatgaaggaagcttctaattgaaaagtctacctttgattagatctcaaatca
ggcaggattacccgctgaacttangcatatcaataagcggag

>TC2-048
ctgcggaaggatcattatccatctcaaaccaggtgcggtcgcggcccccgggggcttgctcccgggtggtaggggtaacaccc
tcacgcgccgcctgcctgtaccctcttttacgagcacctttcgttctccttcggcggggcaacctgccgctggaaccaaaataaa
acctttttgcatctagcattacctgttctgatacaaacaatcgttacaactttcaacaatggatctcttggctctggcatcgatgaaga
acgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccttggtattcc
atggggcatgcctgttcgagcgtcatctacaccctcaagctctgcttggtgttgggcgtctgtcccgcctctgcgcgcggactcgc
cccaaattcattggcagcggtctttgcctcctctcgcgcagcacaattgcgtctgcgggggggcgcggcccgcgtccacgaag
caacattaccgtctttgacctcggatcaggtagggatacccgctgaacttaagcatatcaat

>TC2-049
tgcggaaggatcattacattcagtagcctagctacttgtttacacccttgtttttgcgtacctattgtttcctcggcgggcttgctcgc
cggttggacaaaactataacctttttaaatcttcaatcagcgtctgaattatacataataattacaactttcaacaacggatctcttggt
tctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacat
tgcgccccttggtattccatggggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtcttgtcttttatta
gactcgccttaaagtcattggcagccagtgtttttggtagtaagcgcagcacattttgcgtcttagtccttgaatcgtggcatccatg
aagcctctttttcacttttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>TC2-050
cggagggatcattgctggaacgcgccccaggcgcacccagaaaccctttgtgaacttataccttactgttgcctcggcgctagct
ggtccttcggggcccctcaccctcgggtgttgagacagcccgccggcggccaacccaactcttgtttttacactgaaactctgag
aataaaacataaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaa
tgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgt
catttcaaccctcaagcctggcttggtgatggggcactgcttttacccaagagcaggccctgaaattcagtggcgagctcgccag
gaccccgagcgcagtagttaaaccctcgctctggaaggccctggcggtgccctgccgttaaaccccccaacttctgaaaatttga
cctcggatcaggtaggaatacccgctgaacttaagcatatc

>TC2-051
ggagggatcattacacaaatatgaaagcagaccctcattggggcgtaacgtcacgccgtgtccgcaatagcgccattgctgttgt
gctgacgcggcgggggtaacgccctttgggccagtctgactccatattcacccatgtcttttgcgtactacttgtttccttggcgg
gttcgcccgccaattggacccaattaaacctttttgtaattgcaatcagcgtcagcaaaacaatgtaataaattacaactttcaacaa
cggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatc
tttgaacgcacattgcgcccttggtattccaaagggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggc
gtttttgtctcgggtccgccccgagactcgccttaaaatcattggcagccggcctactggtttcggagcgcagcacatttttgcgct
ttgtccagtcgtggtcttgcgtccatgaagcgaatattttcaacgtttgacctcggatcaggtagggatacccgctgaacttaagcat
at

>TC2-052
ctgcggaaggatcattacnnngagttgtaggctttgcctgctatctcttacccatgtcttttgagtaccttcgtttcctcggcgggtcc
gcccgccgattggacaatttaaaccatttgcagttgcaatcagcgtctgaaaaaacttaatagttacaactttcaacaacggatctct
tggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgc
acattgcgcccttggtattccatggggcatgcctgttcgagcgtcatttgtaccttcaagctctgcttggtgttgggtgtttgtctcg
cctctgcgcgtagactcgcctcaaaacaattggcnnccggcgtattgatttcggagcgcagtacatctcgcgctttgcactcataa
cnnngacgtccaaaagtacattttacactctg

>TC2-053
gcggaaggatcattaaattgaaacaaacaaaaaaacaccacgtgggcttgcccatgtgctgcaagttgtaagtctgcatagttttg
aaaaaaactgatgagcagcggcttgcgtgttttgttttaacccttgttttttgcgtacctattgtttcctcggcgagtttgctcgccggct
ggacaaaacttataacctttttaatcttcaatcagcgtctgaattatacataataattacaactttcaacaacggatctcttggttctgg
catcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcg
ccccttggtattccatggggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtcttgtcttttctttttgtg
attagactcgccttaaagtcattggcagccagtgtttttggtagtaagcgcagcacattttgcgtcttggccctaaaacacttgcgtc
cattaagctactttctcacttttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>TC2-054
ccgtaggtgaacctgcggagggatcattacnnngttatccaactcccaaacccatgtgaacttatctctttgttgcctcggcgcaa
gctacccgggacccagcgccccgggcggcccgccggcggacaaaccaaactctgttatcttcgttgattatctgagcgtcttatt
taataagtcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgc
agaattcagtgaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaaccc
ctaagcacagcttactgttgggaccctacggcctccgtagttccccaaagccattggcggagtggcagtagtcctctgagcgca
gtaattctttatctcgcttttgttaggtgctgcccccccggccgttaaacccccaatttttctggttgacctcggatcaggtaggaat
acccgctgaacttaagcatatcaataagcggag

>TC2-057
tacctcaccgttgcttcggcgggtggcgcccctacccggggcgccgcgccggccccactccggggccggcaacccgtcag
aggcccgcaactcttnntcatcattggcctctctgagtnacttatacaataagtcaaaactttcaacaacggatctcttggttctggc
atcgatgaagaacnnagcgaaatgcgatacgtaatgtgaattgcagatttcagtgaatcatcgaatctttgaacgcacattgcgcc
cgctagtattctggcgggcatgcctgttcgagcgtcatttcaaccatcaagccccaggcttgcgttggagccctgcggctgccgc
nnnctcccaaatccagtggcgggctcgtcgtcgtaccgagtgcagtaaacatcctcgctcnnggcacgtgacgggttcttgcc
gtgaaaccccccacatatcaaggttgacctcg

>TC2-058
tgcggagggatcattatagagttttccaaactcccaacccatgtgaacttaccattgttgcctcggcagaggctacccggtaccta
ccctggaacggcctaccctgtagcgccttacccgggaacggattaccctgtagtggctgccggtggactaccaaactcttgttatt
ttattgttatctgagcgtcttattttaataagtcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaat
gcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccattagtactctagtgggcatgcc
tgttcgagcgtcatttcaacccttaagcctagcttagtgttgggagcctactgcttttgctagctgtagctcctaaaatacaacggcg
gatctgcgatatcctctgagcgtagtaatttttttctcgcttttgactggagttgcagcgtccttagccgctaaacccccccaatttttaat
ggttgacctcggatcaggtaggaatacccgctgaacttaagcata

>TC2-061
ccgtaggtgaacctgcggagggatcattatcgagtttacaactcccaaacccttatgtgaacatacctactgttgcttcggcggact
cgccccggcgtccggacggcctcgcgccgcccgcggcccggacccaggcggccgccggagacccccaaactctgtattat
cagcattttctgaatccgccgcaaggcaaaacaaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaa
cgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccagcattct
ggcgggcatgcctgttcgagcgtcatttcaaccctcgacttccctttggggaaatcggcgttggggaacggcagcataccgccg
gccccgaaatgaagtggcggcccgtccgcggcgacctctgcgtagtaatccaactcgcaccggaaccccgacgtggccacg
ccgtaaaacacccaacttctgaacgttgacctcggatcaggtaggaatacccgctgaacttaagcatatcaataagcg

>TC2-062
tccgtaggtgaacctgcggagggatcattaccgagtttacaactcccaaacccctgtgaacataccttaatgttgcctcggcggat
cagcccgcgccccgtaaaacgggacggcccgccagaggacccaaactctaatgtttcttattgtaacttctgagtaaaacaaac
aaataaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcaaaatgcgataagtaatgtgaattgc
agaattcagtgaatcatcgaatctttgaacgcacattgcgcccgctggtattccggcgggcatgcctgttcgagcgtcatttcaacc
ctcaagcccccgggtttggtgttggggatcggctctgcccttctgggcggtgccgcccccgaaatacattggcggtctcgctgca
gcctccattgcgtagtagctaacacctcgcaactggaacgcggcgcggccatgccgtaaaaccccaacttctgaatgttgacctc
ggatcaggtaggaatacccgctgaacttaagcatatcaataagcggag

>TC2-063
ccgtaggtgaacctgcggagggatcattacaagtgaccccggtctaaccaccgggatgttcataaccctttgttgtccgactctgt
tgcctccggggcgaccctgccttcgggcgggggctccgggtggacacttcaaactcttgcgtaactttgcagtctgagtaaactt
aattaataaattaaaactttaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattg
cagaattcagtgaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgttcgagcgtcatttcacc
actcaagcctcgcttggtattgggcaacgcggtccgccgcgtgcctcaaatcgaccggstgggtcttctgtcccctaagcgttgt
ggaaactattcgctaaagggtgttcgggaggctacgccgtaaaacaaccccatttctaaggttgacctcggatcagntngggata
cccgctgaacttaagcatatcaataagcggagga

>TC2-064
ttccgtaggtgaacctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtttattttaccttgttgct
tcggcgggcccgccttaactggccgccgggggcttacgcccccgggcccgcgcccgccgaagacaccctcgaactctgtc
tgaagattgaagtctgagtgaaaatataaattatttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagc
gaaatgcgatacgtaatgtgaattgcaaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggc
atgcctgtccgagcgtcattgctgccctcaagcccggcttgtgtgttgggccccgtccccgatctccgggggacgggcccgaa
aggcagcggcggcaccgcgtccggtcctcgagngtatggggctttgtcacccgctctgtaggcccggccggcgcttgccgat
caacccaaatttttatccaggttgacctcgnatcaggtagggatacccgctgaacttaagcatatcaataagcggagga

>TC2-065
ttccgtaggtgaacctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtttattttaccttgttgct
tcggcgggcccgccttaactggccgccgggggcttacgcccccgggcccgcgcccgccgaagacaccctcgaactctgtc
tgaagattgaagtctgagtgaaaatataaattatttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagc
gaaatgcgatacgtaatgtgaattgcaaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggc
atgcctgtccgagcgtcattnctgccctcaagcccggcttgtgtgttgggccccgtccccgatctccgggggacgggcccgaa
aggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctctgtaggcccggccggcgcttgccgat
caacccaaatttttatccaggttgacctcggatc

>TC2-067
ttccgtaggtgaacctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtttattttaccttgttgct
tcggcgggcccgccttaactggccgccgggggcttacgcccccgggcccgcgcccgccgaagacaccctcgaactctgtc
tgaagattgaagtctgagtgaaaatataaattatttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagc
gaaatgcgatacgtaatgtgaattgcaaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggc
atgcctgtccgagcgtcattgctgccctcaagcccggcttgtgtgttgggccccgtccccgatctccgggggacgggcccgaa
aggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctctgtaggcccggccggcgcttgccgat
caacccaaatttttatccaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>TC2-069
ccgtaggtgaacctgcggagggatcattgatggaataaacgccctcacaggcgctacccataaacccttttgtgaatactacctaa
aatgttgcctcggcattggttggcctcttgtgaggtcccttctttcgggaaggagcagaccggccggtggccctataaactcttgtt
tttgtaatatcatctgagtaaaacaactaaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagc
gaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccggtggtattccaccggg
catgcctgttcgagcgtcatttcaaccctcaaagccttggctttggtgttggangaataccgtgtaaaaagggtaccctctgaaattt
agtggcgggctcgctagaattttgagcgtagtaattttacctcgt

>TC2-070
gattctaccacctctgtgaaccgttgacctccgggttaataatcaaacatcagtgtaacgaacgtaagagtatcttaattaaacaaa
acttttaacaacggatctcttggctctcgcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtga
atcatcgaatctttgaacgcaccttgcgccttttggtattccgaaaggcatgcctgtttcagtgtcatgaaatctcaatctaatatgtttt
ctgaacatgttaggcttggacttgggtgtctgccagcaatggctcacctcaaatgacttagtggaacatcccacatcagtgttagac
gtaataagtttcgtctctccttgtggtgatgactgctcaaaacctgccatcgcgcacctttgactttgacctgaaatcaggtagggct
acccgctgaacttaagcatatcaa

>TC2-072
cgtaggtgaacctgcggagggatcattatagagttttctaaactcccaacccatgtgaacttaccattgttgcctcggcagaagcta
cctggttaccctaccttggaacggcctaccctgtagcgccttaccctggaacggcctaccctgtaacggctgccggtggactacc
aaactcttgttattttattgtaatctgagcgtcttattttaataagtcaaaactttcaacaacggatctcttggttctggcatcgatgaaga
acgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccattagtattct
agtgggcatgcctgttcgagcgtcatttcaacccttaagcctagcttagtgttgggagcctaccgcttttgctagcggtagctcctg
aaatacaacggcggannnncgatatcctctgagcgtagtaatt

>TC2-073
gagggatcattacaagtgaccccggtctaaccaccgggatgttcataacccttttgttgtccgactctgttgcctccggggcgaccc
tgccttcgggcgggggctccgggtggacacttcaaactcttgcgtaactttgcagtctgagtaaacttaattaataaattaaaactttt
aacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatc
gaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgttcgagcgtcatttcaccactcaagcctcgcttggt
attgggcaacgcggtccgccgcgtgcctcaaatcgaccggctgggtcttctgtcccctaagcgttgtggaaactattcgctaaag
ggtgttcgggaggctacgccgtaaaacaaccccatttctaaggttgacctcggatcaggtagggatacccgctgaacttaagcat
atcaataagcggag

>TC2-074
ccctttgttgtccgactctgttgcctccggggcgaccctgccttcgggcgggggctccgggtggacacttcaaactcttgcgtaac
tttgcagtctgagtaaacttaattaataaattaaaactttaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatg
cgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcc
tgttcgagcgtcatttcaccactcaagcctcgcttggtattgggcaacgcggtccgccgcgtgcctcaaatcgaccggctgggtc
ttctgtcccctaagcgttgtggaaactattcgctaaagggtgttcgggaggctacgccgtaaaacaaccccatttctaaggttgacc
tcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>TC2-078
accgttgcttcggcgggcasggsaagmstcyckcsggcmtcccmkyscggmgmcgkymsyyacyacgsrrasgrgg
cgyccgccggaggrwmcsmmactmymtttamrsgmcswcgtckkcygwgkggccccamaascaagcgrtrmttr
arrsttkwaayracggatctcttggwwcwggcatcgmtcgcmagaaygcagygaaatgcgatrmgyaatgtgmrttgca
ragattcrrtgawtcayygaatctytgcaattcrcaytacttakcgcatttcgccwgcrttctgkcgakcgatgccwgawycaa
gagatccgtyrttwmaasyytyaakyaccgcttkgtyttggkgccmcwcrgmmgacgwsgkcsyktaaakrkagtkks
gkwycctccggcgggmgccycstyyscgtrgtrrskrmcgkcgycgsrmkgggagkccsgmgrgaskcttgccsytar
mccsccraatkcttwacaggttgasstmsgwtcasrwagggtwatagmscgstgaactyarkmatgatcmctccgcagg
ttc

>TC2-082
cacccttgtgtattattacyttgkkgytgktggcrgwgcacgcgkggaccctgtatcggcmgmgmmttaskrckctcrrgm
ccsagygtgcgccgccagmkgaawyytarasycygamyatykytgkcgggkctsartacyatakmakagcttrarrsttk
maayracggatctcttggwwcwggcatgccccsmkgaakaccamgsrgcgcaawgtgcgwtcaarkawtcgatgaw
tyrcwgaattcwgyrawtcatcgawtmyttgawcgcacattkcgcyscttgktmttccggrksgcatgccwgawycaag
agatccgtyrttkmaasyytyaagctmtkmtatrgtaytsagccmcgmcarmratakkcrgrstytagrwwtccwstggc
ggcgcacrctsggkcyygagmgtmsyaakkckctsgctacrrrgtccmcgcgtgctwcygccaaagcaacaassymata
atwywcwakggtgkgasstcgkaycmsgwagggcatgaactcygtamtgawccttmmgcakrtcaa

>TC2-084
ccgtaggtgaacctgcggagggatcattaccgagtttacaactcccaaaccccctgtgnacataccttaatgttgcctcggcggat
cagcccgcgccccgtaaaacgggacggcccgccagaggacccaaactctaatgtttcttattgtaacttctgagtaaaacaaac
aaataaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcaaaatgcgataagtaatgtgaattgc
agaattcagtgaatcatcgaatctttgaacgcacattgcgcccgctggtattccggcgggcatgcctgttcgagcgtcatttcaacc
ctcaagcccccgggtttggtgttggggatcggctctgcccttctgggcggtgccgcccccgaaatacattggcggtctcgctgca
gcctccattgcgtagtagctaacacctcgcaactggaacgcggcgcggccatgccgtaaaaccccaacttctgaatgttgacctc
ggatcaggtaggaatacccgctgaacttaagcatatcaataagcggag

>TC2-086
tgcggagggatcattnnngagttaccgctctataacccttttgtgaacgtacctaaccgttgcttcggcgggcaggggaagcctct
cgcgggcctcccctcccggcgccggcccccaccacggggacggggcgcccgccggaggaaaccaaactctatttacacga
cgtctcttctgagtggcacaagcaaataattaaaacttttaanaacggatctcttggttctggcatcgatgaagaacgcagcgaaat
gcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgctcgccagcattctggcgagcatgc
ctgttcgagcgtcatttcaaccctcaagcaccgcttggttttggggccccacggccgacgtgggcccttaaaggtagtggngga
ccctcccggagcctcctttgcgtagtaactaacgtctcgcactgggatccggagggactcttgccgttaaaccccccaaattctttac
aggttgacctc

>TC2-089
ccgtaggtgaacctgcggagggatcattaccgagttgtcgcaagacaatctccctttcttgtggttgactatgttgccttgccttcgg
gcaggattattctgaatcagtgtcgtctgagttttatttaataaatcaaaactttcaacaacggatctcttggttctggcatcgatgaag
aacgcagcgaaatgcgataagtaatgtgaattgcagaattccgtgaatcatcgaatctttgaacgcacattgcgccctttggtattc
cgaagggcatgcctgttcgagcgtcattacaccttctatcaaggtattgggagatgcactcgcactccttaaaaacagtggtgata
actgtctactcacgcatagtaatctctgcgacggatacacagtgaccgcaccttaattctttgtttgacctcggatcaggtagggata
cccgctgaacttaagcatatca

>TC2-091
cgtaggtgaacctgcggaaggatcattacctagagttgtaggctttgcctgctatctcttacccatgtcttttgagtaccttcgtttcct
cggcgggtccgcccgccgattggacaatttaaaccatttgcagttgcaatcagcgtctgaaaaaacttaatagttacaactttcaac
aacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaa
tctttgaacgcacattgcgcccttggtattccatggggcatgcctgttcgagcgtcatttgtaccttcaagctttgcttggtgttgggt
gtttgtctcgcctctgcgcgtagactcgcctcaaaacaattggcagccggcgtattgatttcggagcgcagtacatctcgcactct
gccttcaccacgacggcgtccacaagtacattttacactcttgacctcggatcaggtagggatacccgctgaacttaagcatatca
ataagcggaggaa

>TC2-092
aacctgcggaaggatcattagagaaacaatacttgggttggtggtcgacgtcgcaagactagaccctacccttttattatccacac
acctgtgcactgtagagcctttgatctttattgatcgaggttcgccctttatacaaacattagtagtcctaaacgtaacgtcttattataa
cctaataaaactttcaacaacggatctcttggctctcgcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcaga
attcagtgaatcatcgaatctttgaacgcaccttgcgcccttggtattccgaagagcatgcctgtttgagtgtcatgaaactatcaa
ccctgacttggtttactcctcgtcttggcttggacttgggtgtttgctgctgttgaagctgctcgccttaaaaggattagctggacctg
ctttgcggataggtttgactcggcgtaataagttatttcgctgaggacatcttcggatggcctgactgctttgcgactgtctgctttcta
accctgttccctttattgggacatgactttattatttggcctcaaatcaggtaggactacccgctgaacttaagcatatcaa

>TC3-002
ccgtagngtgaacctgcggaaggatcattacctagagtttgcgggctttgcctgctatctcttacccatgtcttttgagtaccttcgtt
tcctcggcgggtccgcccgccgattggacaacatttaaacccttttgtagttgcaatcagcgtctgaaaaaacataatagttacaact
ttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatc
atcgaatctttgaacgcacattgcgccccttggtattccatggggcatgcctgttcgagcgtcatttgtaccttcaagctctgcttggt
gttgggtgtttgtctcgcctctgcgtgtagactcgccttaaaacaattggcagccggcgtattgatttcggagcgcagtacatctcg
cgctttgcactcataacgacgacgtccaaaagtacatttttacactcttgacctcggatcaggtagggatacccgctgaacttaagc
atatca

>TC3-003
taggccggccttagctgaattattttacccgtgtcttttgcgtactttctgtttccttggtgggttcgcccaccactaggaccaaccata
aaaccttttgtaatagcagtcagcgtcagtaacaaacataataattacaactttcaacaacggatctcttggttctggcatcgatgaa
gaacgcagcgaaatgcgatacgtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtatt
ccaaagggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtctttgtctctagttttgctggagactcg
ccttaaagtcattggcagccggcctactggtttcggagcgcagcacaagtcgcgctctcttccagccaaggtcagcattccagta
agc

>TC3-004
ccgtaggtgaacctgcggagggatcattacacaaatatgaaggcgggctggaacctctcggggttacagccttgctgaattattc
acccttgtcttttgcgtacttcttgtttccttggtgggttcgcccaccactaggacaaacataaaccttttgtaattgcaatcagcgtca
gtaacaaattaataattacaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgt
gaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtattccaaagggcatgcctgttcgagcgtcatt
tgtaccctcaagctttgcttggtgttgggcgtcttgtctctagctttgctggagactcgccttaaagtaattggcagccggcctactg
gtttcggagcgcagcacaagtcgcactctctatcagcaaaggtctagcatccattaagccttttttcaacttttgacctcggatcag
gtagggatacccgctgaacttaagcatatcaataagc

>TC3-005
gcgcggcgccccggaggcccatcaaacactgcatnnntgcgtcggagtttaagtnnatttaaacaaaactttcaacaacggat
ctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaa
cgcacattgcgccctgtggtattccgcnnngcatgcctgtatngagcgtaatttcaccactcaagcctggcttggtattgggcgtc
gcggctcgcnnngcgnnncaaagtttccggctgagccgtccgtctccaagcgttgtgatacatattcgcttcggggcgcgggc
ggcgtcgcggccgttaaatctttcacaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcgga
g

>TC3-006
aaaccatcgcgaactcttccaagttgcttcggcggcgcggcacccctcgcggggcgccgcagcccggcctctccggaggtgc
ggggcgcccgccggaggtactaaactctgtattatagtggcatctctgagtaaaacacaaataagttaaaactttcaacaacggat
ctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaa
cgcacattgcgcccgctagtactctagcgggcatgcctgttckagcgtcatttcaaccctcaagccctgcttggtgttggggccct
acggctgccgtaggccctgaaaggaagtggcgggctcgctacaactccgagcgtagtaattcattatctcgctagggaggttgc
ggcgtgctcctgccgttaaagaccccatctttaaccaaggttgacctcggatcag

>TC3-009
ccgtaggtgaacctgcggagggatcattacagagttcatgcccttcggggtagatctcccacccgtgtatttatacctttgttgcttt
ggcaggcctgcctcggccaccggctccggctggtgagcgcctgccagaggaccccaaactctgaatttaatgtcgtctgagta

ctatataatagttaaaactttcaacaacggatctcttggttctnnnatcnatgaannncgcagcgaaatgcgataagtaatgtgaat
tgcagaattcagtgaatcatnnaatctttgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcattaca
accctcaagctctgcttggtattgggcctcaccggcaacggtgtgccctaaaatcagtggcggtgccatctggctctaagcgtag
taaatctctcgctctnggtgcccggtggacgcttgccagcaacccccaatttttacaggttgacct

>TC3-014
gatgtgtaaattactackytatgkwrsctcgrcggatcagmccgcgccmmkyaawtwsgggactgtgyyygcacrcagg
aawccctaaacwctgtttywagctggarcttswgrgtwkaaawracarmtmrawcaarrcwtkcamcrmcrgatmtm
ytggttckggcakcratgawgaacgcwkcaaagatkcgatraktmatstgaattgcwgcaattcacrtkamtyatcgmatc
tttgcwrcgttctwcatygaygccmgaaccaarwawtcygkygtkgmawgtyytgattyrwkygtywtttmwacycw
saagttccactaaaarcwkrgtktwgggtwtcctgygtgcrrrcacagtcccswwwtwsrkkgcgcgggctsatccgycg
agsywmcatarmgtagtaatktwcacaksgktaytggtratcgtygyggmmacgycgktaawsmycmacttcygma
kgttsacctacggatc

>CT1-005A
ccgtaggtgaacctgcggagggatcattatagagttttctaaactcccaacccatgtgaacttacctttgttgcctcggcagaagtt
ataggtcttcttatagctgctgccggtggaccattaaactcttgttattttatgtaatctgagcgtcttattttaataagtcaaaactttcaa
caacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcga
atctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccttaagcctagcttagtgttgg
gaatctacttctcttaggagttgtagttcctgaaatacaacggcggatttgtagtatcctctgagcgtagtaatttttttctcgcttttgtt
aggtgctataactcccagccgctaaaccccccaatttttttgtggttgacctcggatcaggtaggaatacccgctgaacttaagcata
tcaataagcggag

>CT1-006A
acctgcggagggatcattacacaaatatgaaggcgggctggaacctctcggggttacagccttgctgaattattcacccttgtcttt
tgcgtacttcttgtttccttggtgggttcgcccaccactaggacaaacataaaccttttgtaattgcaatcagcgtcagtaacaaatta
ataattacaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcaga
attcagtgaatcatcgaatctttgaacgcacattgcgcccttggtattccaaagggcatgcctgttcgagcgtcatttgtaccctca
agctttgcttggtgttgggcgtcttgtctctagctttgctggagactcgccttaaagtaattggcagccggcctactggtttcggagc
gcagcacaagtcgcactctctatcagcaaaggtctagcatccattaagccttttttcaacttttgacctcggatcaggtagggatac
ccgctgaacttaagcatatcaataagcggag

>CT1-006B
ccgtaggtgaacctgcggaaggatcattactgagtgagggccctctgggtccaacctcccacccgtgtttattgtaccttgttgctt
cggtgcgcccgcctcacggccgccgggggggcttctgcccccgggtccgcgcgcaccggagacaccattgaactctgtctgaa
gattgcagtctgagcataaactaaataagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaa
tgcgataactaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggcatg
cctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggctccgtccccccggggacgggtccgaaaggcagcgg
cggcaccgagtccggtcctcgagcgtatggggctttgtcacccgctctgtaggcccggccggcgccagccgacaaccaatcat
ccttttcaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>CT1-007A
cttccgtaggtgaacctgcggaaggatcattaccgagttgattcgagctccggctcgactctcccaccctatgtgtacctacctctg
ttgctttggcgggccgcggtcctccgcaccggcgcccttcgggggggctggccagcgcccgccagaggaccataaaactccag
tcagtgaacttcgcagtctgaaaaacaagttaataaactaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgc

agcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgagg
ggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtattgggccccgtcctccacggacgcgccttaaagacctcg
gcggtggcgtcttgcctcaagcgtagtagaaaacacctcgctttggagcgcacggcgtcgcccgccggacgaacctttgaatta
tttctcaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggagga

>CT1-008A
atctttgaacgcacattgcgccccctggcattcnnnggggcatgcctgtccgagcgtcatttctgccctcaagcccggcttgtgtg
tngggcgtggtccccccggtgtcgggggggacctgccccaaaggcagcggcgacgttccgcctaggtcctcgagcgtatggg
gctttgtcacccgctcgggaggggccymcgggcsyksgccmsccacsarykyytttkamsgytracctmgraksackgt
mggagtctacccgmtrwamttatcarmactwtcaayaagyggagga

>CT1-011A
cttccgtaggtgaacctgcggaaggatcattaccgagttgattcgagctccggctcgactctcccaccccatgtgtacctacctct
gttgctttggcgggccgcggtcctccgcaccgacccccgttcgggggccggccagcgcccgccagaggaccacaaaactcc
agtcagtaaacgtcgcagtctgagaaacaagttaataaactaaaactttcaacaacggatctcttggttctggcatcgatgaagaa
cgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccg
aggggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtattgggctccgtcctccgcggacgcgcctcgaagac
ctcggcggtggcgtcttgcctcaagcgtagtagaaaacacctcgctttggagcgcacggcgtcgcccgccggacgaacctttg
aatttttctcaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggagga

>CT1-013A
gaaggatcattactgagtgcgggcccccnnncgggtccaacctcccacccgtgtctcttgaataccctgttgctttggcgggccca
ccgggccacccccggtcgccgggggggcactgcgcccccgggcccgcgcccgccagagcgcctctgaaccctaatgaagaa
ggactgtctgagtctacgatataattatcaaaactttcaacaatggatctcttggttccggcatcgatgaagaacgcagcgaaatgc
gataagtaatgtgaattgcagaattccgtgaatcatcgaatctttgaacgcacattgcgccccctggcattccgggggggcatgcct
gtccgagcgtcatttctgccctcaagcccggcttgtgtgttgggcgtggtccccccggtgtcgggggggacctgccccaaaggca
gcggcgacgttccgcctaggtcctcgagcgtatggggctttgtcacccgctcgggaggggcctacgggcgttggccacccac
caatttttttacggttgacctcggatcaggtaggagttacccgctgaacttaagcatatcaataagcggagga

>CT1-016A
ggtccaacctcccacccgtgtctcttgaataccctgttgctttggcgggcccaccgggccacccccggtcgccgggggggcact
gcgcccccgggcccgcgcccgccagagcgcctctgaaccctaatgaagaaggactgtctgagtctacgatataattatcaaaa
ctttcaacaatggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattccgtgaat
catcgaatctttgaacgcacattgcgccccctggcattccgggggggcatgcctgtccgagcgtcatttctgccctcaagcccggc
ttgtgtgttgggcgtggtccccccggtgtcgggggggacctgccccaaaggcagcg

>BRD2-121C
agaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacagagactctgccctttgggtagacct
cccaccctgtgtcgttatacctttgttgctttggcgggccgcggggcttaggccctgcccctggctccggctagggcgcgcccgc
cagaggacctcaaaacctgaatgttagtgtcgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatc
gatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccct
tggtattccgggggggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggtcttgggcgtcaccggtccccggtgtgc
cttaaactcagtggcggcgccatctggctctaagcgtagtacatactctcgmtacagacgtccggtggatgctggccagcaacc
cccaatttatcaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaat

>BRD2-145K
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaagtaatttgatatacttcggt
atattctcttttactgtgaactgtattattgcatgacgcttgagggatgtctatagatcatatgggtaggtctagagaatgttaaccgag
tcatgatcaagcttaggcttgttgccctattattatttaccaaaagaattcagtattaatattgtaacatagacgtaaaaaatctataaaa
caacttttaacaacggatctcttggttctcgcatcgatgaagaacgtagcaaagtgcgataactagtgtgaattgcatattcagtgaa
tcatcgagtctttgaacgcatcttgcactcattggtattccaatgagtacgcctgtttcagtatcaaaaacaccccacattcaaatcttt
ttttgaatggaacttggagttacagttttacgaccgtactctttaattatattaggcctgatcttgtttagactgcctgaacttttttttaata
taaaggaaagttttgacaaaataaactgtgctggggcctttaaacaatctttcttaaacttgatctgaaatcaggcgggattacccgc
tgaacttaagcatatc

>BRD2-118A
tgttgcctcggcggggtcacgccccgggtgcgtcgcagccccggaaccaggcgcccgccggagggaccaaccaaactctttt
ctgtagtcccctcgcggacgttatttttacagctctgagcaaaaattcaaaatgaatcaaaactttcaacaacggatctcttggttctg
gcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgc
gcccgccagtattctggcgggcatgcctgtccgagcgtcatttcaaccctcgaacccctccgggggggtcggcgttggggacttc
gggaaccccctaagacgggatcccggccccgaaatacagtggcggtctcgccgcagcctctcatgcgcagtagtttgcacaact
cgcaccgggagcgcggcgcgtccacgtccgtaaaacacccaacttctgaaatgttgacctcggatcaggtaggaatacccgct
gaacttaagcatatc

>BRD2-125F
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattacaagtgaggctaccgaacgttggaaa
cagcggttaggagcttacacccacccgtgtttacatactattgttgctttggcgggccgtggcctccactgcgggctctgctcgtgt
gtgcccgccagagaaccaaactctgaatgttagtgatgtctgagtactatataatagttaaaactttcaacaacggatctcttggttct
ggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattg
cgccctgtggtattccgcagggcatgcctgttcgagcgtcatttaaccactcacgcctagcgtggtattggggcacgcgggtctc
cgcggccctcaaaattagtggcggcgccggtgggctctaagcgtagtacatactcccgctatagagttcccccggtggctcgcc
agaacccccaatttttacaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggagga

>BRD2-133H
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattaccgagtttacaactcccaaaccccctgt
gaacataccatattgttgcctcggcggtgtctgtttcggcagcccgccagaggacccaaaccctagattacattaaagcattttctg
agtcaatgattaaatcaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagta
atgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccagtattctggcgggcatgcctgtccgagc
gtcatttcaaccctcaagcccccgggcttggtgttggagatcggcgagccccccggggcgcgccgtctcccaaatatagtggc
ggtcccgctgtagcttcctctgcgtagtagcacacctcgcactgggaaacagcgtggccacgccgtaaaaccccccacttctga
aaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatca

>BRD2-149R
agaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaatgaattatgaaaggggttgtcgctggc
tgttagcagcatgtgcacgctctgatcattatccatcttacacacctgtgcacacactgtaagtcggcttttgatgcaaagtaagggt
cttcattgactctgctttaaattgggagcctgcttatgttttatcacacactacttcagtttaaagaatgtcaaatcgcgtttaacgcattt
aaatacaactttcagcaacggatctcttggctctcgcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaat
tcagtgaatcatcgaatctttgaacgcaccttgcgctccttggtattccgaggagcatgcctgtttgagtgtcatggaatcatcaact
ctatttacttttgtgaatagggcttggacttggaggttttgccggtacttgtgatcggctcctcttgaatgcattagctcgaacctttgtg

gatcagcttatcggtgtgataattgtctacgccgttactgtgaagcatatattaaagggctcggcttctaatcgtccttcacaggaca
ataactttgacctttgacctcaaatcaggtaggattacccgctgaacttaagcatat

>BRD2-117U
agaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattactgagtgagggccctctgggtccaacc
tcccacccgtgtttattgtaccttgttgcttcggtgcgcccgcctcacggccgccgggggggcttctgcccccgggtccgcgcgca
ccggagacaccattgaactctgtctgaagattgcagtctgagcataaactaaataagttaaaactttcaacaacggatctcttggttc
cggcatcgatgaagaacgcagcgaaatgcgataactaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacatt
gcgccccctggtattccgggggggcatgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggctccgtccccc
cggggacgggtccgaaaggcagcggcggcaccgagtccggtcctcgagcgtatggggctttgtcacccgctctgtaggccc
ggccggcgccagccgacaaccaatcatccttttcaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaA

>BRD2-117M
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattactgagtgagggccctctgg
gtccaacctcccacccgtgtttattgtaccttgttgcttcggtgcgcccgcctcacggccgccgggggggcttctgcccccgggtc
cgcgcgcaccggagacaccattgaactctgtctgaagattgcagtctgagcataaactaaataagttaaaactttcaacaacggat
ctcttggttccggcatcgatgaagaacgcagcgaaatgcgataactaatgtgaattgcagaattcagtgaatcatcgagtctttgaa
cgcacattgcgccccctggtattccgggggggcatgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggctc
cgtccccccgggggacgggtccgaaaggcagcggcggcaccgagtccggtcctcgagcgtatggggctttgtcacccgctctg
taggcccggccggcgccagccgacaaccaatcatccttttttcaggttgacctcggatcaggtagggatacccgctgaacttaag
catatcaataagcggag

>BRD2-118B
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattactgagtgagggccctctgg
gtccaacctcccacccgtgtttattgtaccttgttgcttcggtgcgcccgcctcacggccgccgggggggcttctgcccccgggtc
cgcgcgcaccggagacaccattgaactctgtctgaagattgcagtctgagcataaactaaataagttaaaactttcaacaacggat
ctcttggttccggcatcgatgaagaacgcagcgaaatgcgataactaatgtgaattgcagaattcagtgaatcatcgagtctttgaa
cgcacattgcgccccctggtattccgggggggcatgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggctc
cgtccccccgggggacgggtccgaaaggcagcggcggcaccgagtccggtcctcgagcgtatggggctttgtcacccgctctg
taggcccggccggcgccagccgacaaccaatcatccttttcaggttgacctcggatcaggtagggatacccgctgaacttaagc
atatcaataagcggagga

>BRD2-118C
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaccgagtgagggccctctgg
gtccaacctcccacccgtgtttatcgtaccttgttgcttcggcgggcccgcctcacggccgccgggggggcttctgccctctggtc
cgcgcccgccgaagacaccattgaacgctgtctgaagattgcagtctgagcaattagctaaataagttaaaactttcaacaacgg
atctcttggttccggcatcgatgaagaacgcagcgaaatgcgatacgtaatgtgaattgcagaattcagtgaatcatcgagtctttg
aacgcacattgcgccccttggtattccgggggggcatgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggc
cccgtcctccttcccgggggacgggcccgaaaggcagcggcggcaccgcgtccggtcctcgagcgtatggggcttcgtcttc
cgctcttgtaggcccggccggcgcttgccgacaacaatcaatctttttcaggttgacctcggatcaggtagggatacccgctgaa
cttaagcatatcaataagcggagga

>BRD2-129F
actcttgctttggtttggcctatagataggttgggccagaggtttaacaaaacacaatttaattatttttattgatagtcaaattttgaatt
aatcttcaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatgcgataagtaatatgaattgcaga

ttttcgtgaatcatcgaatctttgaacgcacattgcgccctctggtattccagagggcatgcctgtttgagcgtcatttctctctcaaac
ccccgggtttggtattgagtgatactcttagtcgaactaggcgtttgcttgaaaagtattggcatgggtagtactggatagtgctgtc
gacctctcaatgtattaggtttatccaactcgttgaatggtgtggcgggatatttctggtattgttggcccggccttacaacaaccaa
acaagtttgacctcaaatcnnntaggaatacccgctgaacttaagcatatcaataagcggag

>BRD2-115M
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattactgagtgagggccctctgg
gtccaacctcccacccgtgtttattgtaccttgttgcttcggtgcgcccgcctcacggccgccgggggggcttctgcccccgggtc
cgcgcgcaccggagacaccattgaactctgtctgaagattgcagtctgagcataaactaaataagttaaaactttcaacaacggat
ctcttggttccggcatcgatgaagaangcagcgaaatgcnnnaactaatgtgaattgcagaattcagtgaatcatcgagtctttga
acgcacattgcgcccccctggtattccgggggggcatgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggct
ccgtccccccggggacgggtccgaaaggcagcggcggcaccgagtccggtcctcgagcgtatggggctttgtcacccgctct
gtaggcccggccggcgccagccgacaaccaatcatccttttttcaggttgacctcggatcaggtagggatacccgctgaacttaa
gcatatcaataagcgga

>BRD2-114L
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattcacaatgagtgttttatggcac
tctttaaaaatccatatccaccttgtgtgcaatgtttgttgggagagaggcaactctttcccttcacaaatatcaacctatatctttaaca
acattcgtctgataacatattatgaatatacttaattcaaaatataactttcaacaacggatctcttggctctcgcatcgatgaagaacg
cagcgaaatgcgatacgtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcatattgcgctctttggtattccgaag
agcatgcttgtttgagtatcagtaaacacctcaacctcctcttgtttttcaaaaggagggtggacttgagctatcccaacaaccttca
ccggtaggcgggcggcttgaaatgcaggtgcagctggacttttatctgagctaaaagcatatctatttagtcctcgtcaaacagga
ttattactattgctgcagctaacataaaggataattgtcctcattgctgactgatgcaggattttacgacactttatgtgttgttcaactc
gatctcaaatcaagtaagactacccgctgaacttaagcatatca

>BRD2-113O
gaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattcacaatgagtgttttatggcactctttaaaa
atccatatccaccttgtgtgcaatgtttgttgggagagaggcaactctttcccttcacaaatatcaacctatatctttaacaacattcgt
ctgataacatattatgaatatacttaattcaaaatataactttcaacaacggatctcttggctctcgcatcgatgaagaacgcagcga
aatgcgatacgtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcatattgcgctctttggtattccgaagagcatg
cttgtttgagtatcagtaaacacctcaacctcctcttgtttttcaaaaggagggtggacttgagctatcccaacaaccttcaccggta
ggcgggcggcttgaaatgcaggtgcagctggacttttatctgagctaaaagcatatctatttagtcctcgtcaaacaggattattact
attgctgcagctaacataaaggataattgtcctcattgctgactgatgcaggattttacgacactttatgtgttgttcaactcgatctca
aatcaagtaagactacccgctgaacttaagcatat

>BRD2-149Y
gaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattacaagtgaggctaccgaacgctggaaac
agcggttaggagcttacacccacccgtgtttacatactattgttgctttggcgggccgtggcctccactgcgggctctgctcgtgtg
tgcccgccagagaaccaaactctgaatgttagtgatgtctgagtactatctaatagttaaaactttcaacaacggatctcttggttctg
gcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgc
gccctgtggtattccgcagggcatgcctgttcgagcgtcatttaaccactcacgcctagcgtggtattggggcacgcggtctccg
cggccctcaaaattagtggcggcgccggtgggctctaagcgtagtacatactcccgctatagagttcccccggtggctcgccag
aacccccaatttttacaggttgacctcggatcaggtagggatacccgctgaacttaagcatatca

>BRD2-125O
agaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacagagttcatgccctccgggtagatct
cccacccactgctatcactactctcgttgctttggcgggccgctgggccctgcccggccgccggccccggctggcgcgtgccc
gccagagacctcacagactctgaatgttagtgtcgtccgagtaactatacaatcgttaaaactttcaacaacggatctcttggttctg
gcatcgatgaagaacgcagcgaaatgcgataagtaatgcgaattgcagaattcagtgagtcatcgaatctttgaacgcacattgc
gccctgtggtattccgcagggcatgcctgttcgagcgtcatttcaaccctcaagcctcgcttggtgttgggccctgcccgccgcg
gccggccctaaagatagtggcggcgccgcctggccctcagcgtagtacagctctcgctccagggtccggcggcagcctgcca
gaaccccccaactctgtggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>BRD2-121E
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacagagactctgccctttgg
gtagacctcccaccctgtgtcgttatacctttgttgctttggcgggccgcggggcttaggccctgcccctggctccggctagggc
gcgcccgccagaggacctcaaaacctgaatgttagtgtcgtctgagtactatataatagttaaaactttcaacaacggatctcttgg
ttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaca
ttgcgccccttggtattccgggggggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggtcttgggcgtcaccggtc
cccggtgtgccttaaactcagtggcggcgccatctggctctaagcgtagtacatactctcgctacagacgtccggtggatgctgg
ccagcaaccccccaatttatcaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>BRD2-125G
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattaccgagttcatgcccttacg
ggtagatctcccaccctatgttatattaccttcgttgctttggcaggccgtcaggcttcggccaggctaccggctccggctggtaag
cgcctgccagaggaccctaaactctgaatattagtgttgtctgagtactataaaatagttaaaactttcaacaacggatctcttggttc
tggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattg
cgccccttggcattccgaggggcatgcctgttcgagcgtcattacaaccctcaagcactgcttggtattgggctccgctagtttca
gcgggccttaaaatcagtggcagcgccgtctgagctctgagcgtagtaaatatcctcgctatagagacccaggtagctcttgtcat
caacccccaactttctaagtttgacctcggatcaggtagggatacccgctgaacttaagcatatca

>BRD2-145L
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattacaagtgaggctaccgaacg
ctggaaacagcggttaggagcttacacccacccgtgtttacatactattgttgctttggcgggccgtggcctccactgcgggctct
gctcgtgtgtgcccgccagagaaccaaactctgaatgttagtgatgtctgagtactatctaatagttaaaactttcaacaacggatct
cttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaac
gcacattgcgccctgtggtattccgcagggcatgcctgttcgagcgtcatttaaccactcacgcctagcgtggtattggggcacg
cggtctccgcggccctcaaaattagtggcggcgccggtgggctctaagcgtagtacatactcccgctatagagttcccccggtg
gctcgccagaaccccccaatttttacaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggagg
a

>BRD2-125X
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattactgagtgagggccctctgg
gtccaacctcccacccgtgtttattgtaccttgttgcttcggtgcgcccgcctcacggccgccgggggggcttctgccccccgggtc
cgcgcgcaccggagacaccattgaactctgtctgaagattgcagtctgagcataaactaaataagttaaaactttcaacaacggat
ctcttggttccggcatcgatgaagaacgcagcgaaatgcgataactaatgtgaattgcagaattcagtgaatcatcgagtctttgaa
cgcacattgcgccccctggtattccgggggggcatgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggctc
cgtccccccggggacgggtccgaaaggcagcggcggcaccgagtccggtcctcgagcgtatggggctttgtcacccgctctg

taggcccggccggcgccagccgacaaccaatcatccttttttcaggttgacctcggatcaggtagggatacccgctgaacttaag
catatcaataagcggagga

>BRD2-149B
ttgtaccctcaagctttgcttgnnnttgggcgtcttgtctctagctttgctggagaskmkcmttamaswaatwkgmagscggs
ctggaaccnntngggkttmsgagcsywgcwsaaktaykcacyctctrtcnnnnnngntytwgcrtmcattaagcctttkttt
caacttgktgrsytcgccsaycagstagggayaaacgmtraaccttwwgtmatwkcaatmagcgtsagkaa

>BRD2-113Y
atcatgaacacctcactactataacttttttgttgtaatagtggacttggaatgcgctggtgtcaatcnnnccttctaaaatggagttct
tgcactgttgatttgttttaacagcnnnttggcctaataagtttcggcagtannnttgaaatcactggctttttwttrcwtagaawctr
ayartcttkwwatacagamtttnnnnggtgacttgtctcctcttcaaattttgatctgaaatcaggtaggattacccgc

>BRD3-078B
tagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacacaaatatgaaggcgggctggaac
ctctcggggttacagccttgctgaattattcacccttgtcttttgcgtacttcttgtttccttggtgggttcgcccaccactaggacaaa
cataaacctttgtaattgcaatcagcgtcagtaacaaattaataattacaactttcaacaacggatctcttggttctggcatcgatga
agaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccttggtat
tccaaagggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtcttgtctctagctttgctggagactcg
ccttaaagtaattggcagccggcctactggtttcggagcgcagcacaagtcgcactctctatcagcaaaggtctagcatccattaa
gccttttttttcaacttttgacctcggatcaggtagggatacccgctgaacttaagcatat

>BRD3-078C
tggtcatttagaggaagtaaaagtcgtaacaaggttgctgtaggtgaacctgcagcaggatcattaaaaaacgagcctaaacgct
caaaaacaccgcgaacacacctctgttgcctcggtggtgacggtccggggaagtcccggacttcagccgccggcggcccgaa
aactactatttgatggcactctgaggcactctaaacgagtaaaaactttcagcaacggatctcttggttctggcatcgatgaagaac
gcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccggcattcc
ggcgggcatgcctgttcgagcgtcatttcaaccctcaggctccgcctggtgttggggccttttcgcatgaaagccccgaaaagca
gtggcgggctcgcagcggacccgggcgtagtagcacacctcgctcagggccccgccgcgggcgccggccgtgaaacctctt
tttacccaaggttgannnnngatcaggtaggaatacccgctgaacttaagcatatcaataagcggag

>BRD3-078E
ttagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacacaaatatgaaggcgggctggaa
cctctcggggttacagccttgctgaattattcacccttgtcttttgcgtacttcttgtttccttggtgggttcgcccaccactaggacaa
acataaacctttgtaattgcaatcagcgtcagtaacaaattaataattacaactttcaacaacggatctcttggttctggcatcgatg
aagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccttggt
attccaaagggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtcttgtctctagctttgctggagact
cgccttaaagtaattggcagccggcctactggtttcggagcgcagcacaagtcgcactctctatcagcaaaggtctagcatccatt
aagccttttttttcaacttttgacctcggatcaggtagggatacccgctgaacttaagcatat

>BRD3-078F
agaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacacaaatatgaaggcgggctggaacc
tctcggggttacagccttgctgaattattcacccttgtcttttgcgtacttcttgtttccttggtgggttcgcccaccactaggacaaac
ataaacctttgtaattgcaatcagcgtcagtaacaaattaataattacaactttcaacaacggatctcttggttctggcatcgatgaa
gaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccttggtatt

ccaaagggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtcttgtctctagctttgctggagactcg
ccttaaagtaattggcagccggcctactggtttcggagcgcagcacaagtcgcactctctatcagcaaaggtctagcatccattaa
gccttttttttcaacttttgacctcggatcaggtagggatacccgctgaacttaagcatat

>BRD3-078I
aggaagtaaaagtcgtaacaaggttgctgtaggtgaacctgcagcaggatcattaaaaaacgagcctgaacgctcaaaaacac
cgcgaacacacctctgttgcctcggtggtgacggtccggggaagtcccggacttcagccgccggcggcccgaaaactactatt
tgatggcactctgaggcactctaaacgagtaaaaactttcagcaacggatctcttggttctggcatcgatgaagaacgcagcgaa
atgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccggcattccggcgggcat
gcctgttcgagcgtcatttcaaccctcaggctccgcctggtgttggggccttttcgcatgaaagccccgaaaagcagtggcggg
ctcgcagcggacccgggcgtagtagcacacctcgctcagggccccgccgcgggcgccggccgtgaaacctctttttacccaa
ggttgacctcggatcaggtaggaatacccgctgaacttaagcatat

>BRD3-078J
aggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacaagtgaccccggtctaaccaccgggat
gttcataacccttttgttgtccgactctgttgcctccggggcgaccctgccttcgggcgggggctccgggtggacacttcaaactct
tgcgtaactttgcagtctgagtaaacttaattaataaattaaaacttttaacaacggatctcttggttctggcatcgatgaagaacgca
gcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccctggtattccgggg
ggcatgcctgttcgagcgtcatttcaccactcaagcctcgcttggtattgggcaacgcggtccgccgcgtgcctcaaatcgaccg
gctgggtcttctgtcccctaagcgttgtggaaactattcgctaaagggtgttcgggaggctacgccgtaaaacaaccccatttcta
aggttgacctcggatcaggtagggatacccgctgaacttaagcatat

>BRD3-078K
gaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaggagtatttgcgggaaatcgaaagaaa
gtaccgctctcccacccgtgactatatactatgttgctttccgggcttcaacccccggagaggaccaaactcttgaatttattactgt
ctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaa
tgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccctggtattccgggggcatgcctgttcgagcgt
cattaataccaatcccttcgggggtcttggggcctgggatctcccagctcttaaaatcagtggcggtgcctctcggctctaagcgt
agtaattcttctcgctatagtccccgggagaacacttgccataacccccacacttttaaggttgacctcggatcaggtagggatacc
cgctgaacttaagcatatc

>BRD3-078L
ggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacacaaatatgaaggcgggctggaacctct
cggggttacagccttgctgaattattcacccttgtcttttgcgtacttcttgtttccttggtgggttcgcccaccactaggacaaacata
aacctttttgtaattgcaatcagcgtcagtaacaaattaataattacaactttcaacaacggatctcttggttctggcatcgatgaagaa
cgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtattcca
aagggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtcttgtctctagctttgctggagactcgcctt
aaagtaattggcagccggcctactggtttcggagcgcagcacaagtcgcactctctatcagcaaaggtctagcatccattaagcc
ttttttttcaacttttgacctcggatcaggtagggatacccgctgaacttaagca

>BRD3-078M
agaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaagagtaagggtctctggcccgaacct
ccaaccctctgtcgttataactacttcgttgctttggcgggaccgcgagggtcctcccgagcgcgccggtctccggacaggcga
gcgcccgccagagtctaaccaaactcttgtttttaaaccagtcgtctgagtataaaattttaattaaattaaaactttcaacaacggat
ctcttggttctcgcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaa

cgcacattgcgccccttggtattccgaggggcatgcctgttcgagcgtcattacaccactcaagcactgcttggtattgggcatcg
tccgtcgtaaaggcgggcgtgcctcgaagacctcggcggggtttctccaacttcgggcgtagtagagttaaatcaaaacgtctta
taagcttggtgagatctcattgccgttaaacctttatatattttcaggttgacctcggatcaggtagggatacccgctgaacttaagc
atatc

>BRD3-079A
tagaggaagtaaaagtcgtaacaaggttgctgtaggtgaacctgcagcaggatcattaaaaaacgagcctaaacgctcaaaaac
accgcgaacacacctctgttgcctcggtggtgacggtccggggaagtcccggacttcagccgccggcggcccgaaaactact
atttgatggcactctgaggcactctaaacgagtaaaaactttcagcaacggatctcttggttctggcatcgatgaagaacgcagcg
aaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccggcattccggcggg
catgcctgttcgagcgtcatttcaaccctcaggctccgcctggtgttggggccttttcgcatgaaagccccgaaaagcagtggcg
ggctcgcagcggacccgggcgtagtagcacacctcgctcagggccccgccgcgggcgccggccgtgaaacctctttttaccc
aaggttgacctcggatcaggtaggaatacccgctgaacttaagcatat

>BRD3-079B
tagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacacaaatatgaaggcgggctggaac
ctctcggggttacagccttgctgaattattcacccttgtcttttgcgtacttcttgtttccttggtgggttcgcccaccactaggacaaa
cataaacctttgtaattgcaatcagcgtcagtaacaaattaataattacaactttcaacaacggatctcttggttctggcatcgatga
agaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtat
tccaaagggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtcttgtctctagctttgctggagactcg
ccttaaagtaattggcagccggcctactggtttcggagcgcagcacaagtcgcactctctatcagcaaaggtctagcatccattaa
gccttttttttcaacttttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>BRD3-079G
tggtcatttagaggaagtaaaagtcgtnncttggtcnnntagaggaagtaaaagtcgtaacaaggttgctgtaggtgaacctgca
gcaggatcattaaaaaacgagcctgaacgctcaaaaacaccgcgaacacacctctgttgcctcggtggtgacggtccggggaa
gtcccggacttcagccgccggcggcccgaaaactactatttgatggcactctgaggcactctaaacgagtaaaaactttcagcaa
cggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatc
tttgaacgcacattgcgcccgccggcattccggcgggcatgcctgttcgagcgtcatttcaaccctcaggctccgcctggtgttg
gggccttttcgcatgaaagccccgaaaagcagtggcgggctcgcagcggacccgggcgtagtagcacacctcgctcagngc
cccgccgcgggcgccggccgtgaaacctctttttacccaaggttgacctcggatcaggtaggaatacccgctgaacttaagcat
atca

>BRD3-081C
gaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaggagtatttgcgggaaatcgaaagaaa
gtaccgctctcccacccgtgactatatactatgttgctttccgggcttcaaccccggagaggatcaaactcttgaatttattactgtct
gagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatg
tgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggcatgcctgttcgagcgtca
ttaataccaatcccttcgggggtcttggggcttgggatctcccagctcttaaaatcagtggcggtgcctttcggctctaagcgtagt
aattcttctcgctatagtccccgggagaacacttgccataacccccatacttttaaggttgacctcggatcaggtagggatacccgc
tgaacttaagcatat

>BRD3-081D
tcctccgcttattgatatgcttaagttcagcgggtatccctacctgatccgaggtcaaacttagaaagttgggggttgatggctagc
gtccgccgagtccctatagcgaggatatttactgcgctcagggcctcgacggcaccgccactgattttaaggcccgccaggtga

gtggcgaagcccaataccaagccaggcttgagggttgtaatgacgctcgaacaggcatgcccctcggaatrccaaggggcgc
aatgtgcgttcaaagattcgatgattcactgaattctgcaattcacattacttatcgcatttcgctgcgttcttcatcgatgccagaacc
aagagatccgttgttgaaagttttaactattatatagtactcagacgacactaataatcagggttttgggtcctctggcgggcgctca
ccagccggggccggtggtcgggcgtgagcctgacggcccgccaaagcaacaaagtgataataacacagggtgggagatcta
ccnntaagggcatgaactctgtaatgatccctccgcaggttcacctacggagaccttgttacgacttttacttcctctaaatgacca
ag

>BRD3-081G
agaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattacagtattctattgcctgcgcttaattgcg
cggcgataaaccttacacaacgtgtttttttaatataaactattactttggtttggctaagaaattagttgagccagaggtgatttaaac
ttcaattttattgaattgttattttaattttatgtcaatttgttgattaaattcaaaacaatcttcaaaactttcaacaacggatctcttggttct
cgcatcgatgaagaacgcagcgaaatgcgataagtaatatgaattgcagattttcgtgaatcatcgaatctttgaacgcacattgc
accctctggtattccagagggtatgcctgtttgagcgtcatttctctctcaaacctttgggtttggtattgagtgatactcttagtcgaa
ctaggcgtttgcttgaaatatattggcacgagtagtgttgaacagtgttgtctgaacatcaatgtattaggtttatccaactcgttgaa
gcgtttaggtattactattcttcattaggctttgccttataaaacacaaacaagtttgacctcaaatcaggtaggattacccgctgaact
taagc

>BRD3-081H
cttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacgcgccccgg
cgcacccagaaaccctttgtgaacttatacccataccgttgcctcggcgcaggccggcctctgctgcagaggccccctggagac
agggagcagccagccggtggccaaccaaactctgtttctacagtgaatctctgagtaaaaaaacataaatgaatcaaaactttca
acaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcg
aatctttgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcgtggcttggtga
tggggcactgcccgtaaaagggcaggccctgaaatctagtggcgagctcgccaggaccccgagcgtagtagttatatctcgct
ctggaaggccctggcggtgccctgccgttaaaccccccaacttctgaaaatttgacctcggatcaggtaggaatacccgctgaact
taagcatatcaataagcggag

>BRD3-081I
ctccgcttattgatatgcttaagttcagcgggtattcctacctaatccgaggtcaaccactagaaggtataaggggtttcacggcag
gccctaggtacactgcaggagcgagaaaagatttactacgctacgagtgcacctaacaccgccactacctttggggagctacag
ggcccttacgggcctgtagggtcccaacactacgcaacagggcgtaagggttgaaatgacgctcgaataggcatgcccacta
gaatactaatgggcgcaatatgcgttcaaagattcgatgattcactgaattctgcaattcacattacgtatcgcatttcgctgcgttctt
catcgatgccagaaccaagagatccgttgttgaaagttttaacgtattaagttataggtttcagagatacactatacaaaacagagt
gtagaggtccgccggcggagcgtacagtgcaaactactacagggtagctccagggtagccacctacagggtaggtggctccc
ggaaaagctctccagggtagccgccgctcagcgccgaggcaacgacggtaaggttcacaaagggttgggagtttttgtagaac
tcagtaatgatccctccgctggttcaccaacggagaccttgttacgactttacttcctctaantgacca

>BRD3-081J
ctccgcttattgatatgcttaagttcagcgggtattcctacctaatccgaggtcaaccactagaaggtataaggggtttcacggcag
gccctaggtacactgcaggagcgagaaaagatttactacgctacgagtgtacctaacaccgccactacctttggggagctacag
ggcccttacgggcctgtagggtcccaacactacgcaacagggcgtaagggttgaaatgacgctcgaataggcatgcccacta
gaatactaatgggcgcaatatgcgttcaaagattcgatgattcactgaattctgcaattcacattacgtatcgcatttcgctgcgttctt
catcgatgccagaaccaagagatccgttgttgaaagttttaacgtattaagttataggtttcagagatacactatacaaaacagagt
gtagaggtccgccggcggagcgtacagtgcaaactactacagggtagctccagggtagccacctacagggtaggtggctccc

ggaaaagctmtccagggcagccgccgctcagcgccgaggcaacgacggtaaggttcacaaagggttgggagtttttgtagaa
ctcagtaatgatccctccgctggttcaccaacggagaccttgttacgacttttacttcctctaaatgacca

>BRD3-082B
aggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaggagtatttgcgggaaatcgaaagaaag
taccgctctcccacccgtgactatatactatgttgctttccgggcttcaacccccggagaggaccaaactcttgaatttattactgtct
gagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatg
tgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggcatgcctgttcgagcgtca
ttaataccaatcccttcgggggtcttggggcctgggatctcccagctcttaaaatcagtggcggtgcctctcggctctaagcgtagt
aattcttctcgctatagtccccgggagaacacttgccataacccccacacttttaaggttgacctcggatcaggtagggatacccg
ctgaacttaagcatatca

>BRD3-082C
tggtcanttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagttctacaaaaactc
ccaaccctatgtgaaccttaccgtcgttgcctcggcgctgagcggcggctaccctggagagcttttccgggagccacctaccct
gtaggtggctaccctggagctaccctgtagtagtttgcactgtacgctccgccggcggacctctacactctgttttgtatagtgtatc
tctgaaacctataacttaatacgttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgata
cgtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcatattgcgcccattagtattctagtgggcatgcctattcgag
cgtcatttcaacccttacgccctgttgcgtagtgttgggaccctacaggcccgtaaagggcctgtagctccccaaaggtagtggc
ggtgttaggtgcactcgtagcgtagtaaatcttttctcgctcctgcagtgtacctagggcctgccgtgaaaccccttataccttctag
tggttgacctcggattaggtaggaatacccgctgaacttaagcatatcaataagcggag

>BRD3-082M
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagtttacgctctacaaccctttgtg
aacatacctataactgttgcttcggcgggcagggtctccgtgaccctcccggcctcccgcccccgggcgggtcggcgcccgcc
ggaggataaccaaactctgatttaacgacgtttcttctgagtggtacaagcaaataatcaaaacttttaacaacggatctcttggttct
ggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattg
cgcccgccagcattctggcgggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtgttggggccctacggctga
cgtaggccctcaaaggtagtggcggaccctcccggagcctcctttgcgtagtaactttacgtctcgcactgggatccggaggga
ctcttgccgtaaaacccccaattttccaaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatat

>BRD3-082O
tcttggtcatttagaggaagtaaaagtcgtaacaaggttgctgtaggtgaacctgcagcaggatcattaaaaaacgagcctaaac
gctcaaaaacaccgcgaacacacctctgttgcctcggtggtgacggtccggggaagtcccggacttcagccgccggcggccc
gaaaactactatttgatggcactctgaggcactctaaacgagtaaaaactttcagcaacggatctcttggttctggcatcgatgaag
aacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccggcatt
ccggcgggcatgcctgttcgagcgtcatttcaaccctcaggctccgcctggtgttggggccttttcgcatgaaagccccgaaaag
cagtggcgggctcgcagcggacccgggcgtagtagcacacctcgctcagggccccgccgcgggcgccggccgtgaaacct
ctttttacccaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatca

>BRD3-083A
gaggaagtaaaagtcgtaacaaggttgctgtaggtgaacctgcagcaggatcattaaaaaacgagcctaaacgctcaaaaaca
ccgcgaacacacctctgttgcctcggtggtgacggtccggggaagtcccggacttcagccgccggcggcccgaaaactactat
ttgatggcactctgaggcactctaaacgagtaaaaactttcagcaacggatctcttggttctggcatcgatgaagaacgcagcga
aatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccggcattccggcgggc

atgcctgttcgagcgtcatttcaaccctcaggctccgcctggtgttggggccttttcgcatgaaagccccgaaaagcagtggcgg
gctcgcagcggacccgggcgtagtagcacacctcgctcagggccccgccgcgggcgccggccgtgaaacctctttttaccca
aggttgacctcggatcaggtaggaatacccgctgaacttaagcatatc

>BRD3-083D
tagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattactgagtgagggctccggcccaacct
ccaaccccatgtgaatccgacctctgttgcctcgggggcgacccggccctgcgccggggcccccggtggaccactcaactct
gcatctttgcgtctgagtatgatatttgaatcaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcg
aaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggc
atgcctgttcgagcgtcattacaccactcaagcctggcttggtattgggcacggcggcttcgcggccgcccgcctcaaagtctcc
ggctggaccgaccgtctctaagcgttgtgacttcattggaccgcttgcgagtacgggacagtccgtggccgttaaacccccccatg
aaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>BRD3-083K
gaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattactgagtgagggcccccgggcccgacc
tccaaccctatgtctaccgaccctgttgcctcgggggtgacccggcgcccgttccccgggcgccccgggggcccccggcggac
ccctcaacgctgcatctgtgcgtccgagtcatatttgaatcaatcaaaactttcaacaacggatctcttggttctggcatcgatgaag
aacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccgtggtatt
ccgcggggcatgcctgttcgagcgtcatttcaccaatcaagcctcgcttggtattgggcgtccgcggcggcaacgccgcgcgc
cccaatgtctccggctgagccgcctgtctctaagcgttgtggtctaactgttcgctttggagtatgggctggatcatcgccgttaaa
caccccatcaaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaat

>BRD3-083P
tagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattactgagtgagggctccggtccaacctc
caaccccatgtgaatccgacctctgttgcctcgggggcgacccggccctgcgccggggcccccggtggaccactcaactctg
catctttgcgtctgagtatgatatttgaatcaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcga
aatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggcat
gcctgttcgagcgtcattacaccactcaagcctggcttggtattgggcgaggcggctttgcggccgcccgcctcaaagtcttcgg
ctggactgaccgtctctaagcgttgtgacttcattggaccgcttgcgagtacgggacagcccgtggccgttaaacccccccatgaa
aggttgacctcggatcaggtagggatacccgctgaacttaagcatatca

>BRD3-083W
aagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaaaagtttacgtcctacttgcctccgtgcaaaag
gataagattccacgcgtgaaacagcgagctcttcggagcggtgctgggcctgcccctgtgaaccgaacccacctgcttctcgga
gcggtgtgccattagggatcaggatcagtgcagcttgcggtgcgctctagtaatagagccttcgggatggatatacttaacccgt
gtttaccttctttgttgctttggcgggtcgcctctaaggcgttggcctcggctaaaccgtacccgccagagaactcaaaactcttttg
ttttagtgacgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaat
gcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgaggggcatgc
ctgttcgagcgtcattatgaccaatcaagcctggcttggtgttggggtgcgcggtcccgcggcccttaaaatcagtggcggtgcc
gtccggctctaagcgtagtaaatttcctcgctatagtgtccgggtggttacctgccagaaccccccaatctcacggttgacctcgg
atcaggtagggatacccgctgaacttaagca

>BRD3-083Y
agaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattacctttctatgcagaagtattgtgtagggc
aacctcacaatactccctgtataccacccttgtcttttgcgcactcatatttcctcggcaggcttgcctgccgattggacaatttatact

atttttaattttgaatcagcgtctgaaataattaataattacaactttcaacaacggatctcttggttctggcatcgatgaagaacgcag
cgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccatgggg
catgcctgttcgagcgtcatttgtaccttcaagctctgcttggtgttgggtgtttgtccttctcatgtggttggactcgccttaaagtaat
tggcagccagtgttttggttttgaagcgcagcacaagtcgcgattctagctaacatcaccagcttccagcaagcctttattacctttg
acctcggatcaggtagggatacccgctgaacttaagcat

>BRD3-084B
agaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacacaaatatgaaggcgggctggaacc
tctcggggttacagccttgctgaattattcacccttgtcttttgcgtacttcttgtttccttggtgggttcgcccaccactaggacaaac
ataaacctttgtaattgcaatcagcgtcagtaacaaattaataattacaactttcaacaacggatctcttggttctggcatcgatgaa
gaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtatt
ccaaagggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtcttgtctctagctttgctggagactcg
ccttaaagtaattggcagccggcctactggtttcggagcgcagcacaagtcgcactctctatcagcaaaggtctagcatccattaa
gccttttttttcaacttttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>BRD3-084F
agaggaagtaaaagncgtaacaaggtctccgtaggtgaacctgcggagggatcattactgagtgagggctccggcccaacct
ccaaccccatgtgaatccgacctctgttgcctcgggggcgacccggccctgcgccggggccccccggtggaccactcaactct
gcatctttgcgtctgagtatgatatttgaatcaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcg
aaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggc
atgcctgttcgagcgtcattacaccactcaagcctggcttggtattgggcacggcggcttcgcggccgcccgcctcaaagtctcc
ggctggaccgaccgtctctaagcgttgtgacttcattggaccgcttgcgagtacgggacagtccgtggccg

>BRD3-084I
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaggagtatttgcgggaaatcgaaaga
aagtaccgctctcccacccgtgactatatactatgttgctttccgggcttcaacccccggagaggaccaaactcttgaatttattact
gtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagt
aatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggcatgcctgttcgagc
gtcattaataccaatcccttcgggggtcttggggcctgggatctcccagctcttaaaatcagtggcggtgcctctcggctctaagc
gtagtaattcttctcgctatagtccccgggagaacacttgccataacccccacacttttaaggttgacctcggatcaggtagggata
cccgctgaacttaagcatatcaat

>BRD3-085E
ttggtcatttagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacacaaatatgaaggcgg
gctggaacctctcggggttacagccttgctgaattattcacccttgtcttttgcgtacttcttgtttccttggtgggttcgcccaccact
aggacaaacataaaccttttgtaattgcaatcagcgtcagtaacaaattaataattacaactttcaacaacggatctcttggttctggc
atcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgc
cctttggtattccaaagggcatgcctgttcgagcgtcatttgtaccctcaagctttgcttggtgttgggcgtcttgtctctagctttgct
ggagactcgccttaaagtaattggcagccggcctactggtttcggagcgcagcacaagtcgcactctctatcagcaaaggtcta
gcatccattaagccttttttttcaacttttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>BRD3-085G
ggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattactgagtgagggctccggtccaacctccaa
ccccatgtgaatccgacctctgttgcctcgggggcgacccggccctgcgccggggccccccggtggaccactcaactctgcatc
tttgcgtctgagtatgatatttgaatcaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatg

cgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggcatgcct
gttcgagcgtcattacaccactcaagcctggcttggtattgggcgaggcggctttgcggccgcccgcctcaaagtcttcggctgg
actgaccgtctctaagcgttgtgacttcattggaccgcttgcgagtacgggacagcccgtggccgttaaacccccccatgaaaggt
tgacctcggatcaggtagggatacccgctgaacttaagcatatc

>BRD3-085R
gtttcctcggcgtgcgcannnggatgcccgcgggggatccgggcctgccctgcagctaccccgtagccgcgcggcaaggcc
ggatcctcagcacacggcgccgctcaaggtcccgccgaagtaccccaaactctgttttacgtggaattctgaatgcttcaactaa
ataagttaaaactttcaacaacggatctcttggttctggcatcgwtgrwgaacgcagcgraatgsgatmaktarwgtgaatatsc
araaytcwsysaatcatcnnatctttgaacgcacattgcgcccattagtattctagtgggcatgcctattcgagcgtcatttcaacc
cttaagcccgcgttgcttagcgttgggagtctgcgggtcaccccgcagctcctgaaaaccattggcggagtcagggagcactct
aagcgtattacacttttttatttttgtctcgctttggacgttgccccggccccacgccgtaaaacaccccataccaaatgttgacctcg
aattaggtaggaatacccgctgaacttaagcatatcaa

>BRD3-085T
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaaaaatcagtagcttgctactgtgctg
ggggagatgggaaagtcctggctgaaaagtcggcgtaccggtacctgcacccctacgatacccttgtcttttgagcacttatgttt
ccttggtaggcttgcctgccaacaggacaaactttaaaccttttaaatatcaatcagcgtctgaattataattaataattacaactttc
aacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatc
gaatctttgaacgcacattgcgccccttggtattccatggggcatgcctgttcgagcgtcatttgtaccctcaagctatgcttggtgtt
gggtgattgtccagcctgtatggccaggactcgccttaaagtaattggcagccagtgttttggtagaaagcgcagcacattttgcg
atttactccaataatattagcaaccataaagctcatttatcacttttgacctcggatcaggtagggatacccgctgaacttaagcatat
caataagcggagga

>BRD3-090A
tagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattaccgagtgagggctccggcccaacct
ccaaccccatgtgaatccgacctctgttgcctcgggggcgacccggccctgcgccggggccccccggtggaccactcaactct
gcatctttgcgtctgagtatgatatttgaatcaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcg
aaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggc
atgcctgttcgagcgtcattacaccactcaagcctggcttggtattgggcgaggcggctttgcggccgcccgcctcaaagtcttc
ggctggactgaccgtctctaagcgttgtgacttcattggaccgcttgcgagtacgggacagcccgtggccgttaaacccccccatg
aaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaata

>BRD3-105A
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagttaccgctctataaccctttgtga
acgtacctaaccgttgcttcggcgggcaggggaagcctctcgcgggcctcccctcccggcgccggccccccaccacggggac
ggggcgcccgccggaggaaaccaaactctatttacacgacgtctcttctgagtggcacaagcaaataattaaaacttttaacaac
ggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctt
tgaacgcacattgcgctcgccagcattctggcgagcatgcctgttcgagcgtcatttcaaccctcaagcaccgcttggttttgggg
ccccacggccgacgtgggcccttaaaggtagtggcggaccctcccggagcctcctttgcgtagtaactaacgtctcgcactgg
gatccggagggactcttgccgttaaacccccaaattctttacaggttgacctcggatcaggtaggaatacccgctgaacttaagca
tatcaat

>BRD3-105B
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagttaccgctctataaccctttgtga
acgtacctaaccgttgcttcggcgggcaggggaagcctctcgcgggcctcccctcccggcgccggcccccaccacggggac
ggggcgcccgccggaggaaaccaaactctatttacacgacgtctcttctgagtggcacaagcaaataattaaaacttttaacaac
ggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctt
tgaacgcacattgcgctcgccagcattctggcgagcatgcctgttcgagcgtcatttcaaccctcaagcaccgcttggttttgggg
ccccacggccgacgtgggcccttaaaggtagtggcggaccctcccggagcctcctttgcgtagtaactaacgtctcgcactgg
gatccggagggactcttgccgttaaaccccccaaattctttacaggttgacctcggatcaggtaggaatacccgctgaacttaagca
ta

>BRD3-105M
aggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattactgaactagtaattctctgaaaggtcgccg
gtacccggtcccccccctaaacaaggggggcccggggaaggtcctctcacacccttgtgtaccttaccatgttgctttggcgggc
cgacccggtttcgacccgggcggccggcgcccccagccctcaccggccaggacgtcaggctaagcgcccgccagtatacaa
aactcaagcgattatttcgtgtagttctgataaattattcaattaattaaaactttcaacaacggatctcttggttctggcatcgatgaag
aacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggtattc
cggagggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtattgggcgacgtccgctgccggacgcgcctcga
agacctcggcgacggcgtcttagcctcgagcgtagtagtaacatctcgctttggagtgctaggcgtcggccgccggacaatcga
ccttcggtctattacttccaaggttgacctcggatcaggtagggatacccgctgaacttaagcatat

>BRD3-107A
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattaccgagttatctaaactcccaaccctatg
tgaaccttaccgtcgttgcctcggcgggccgcgcttacccggtagctaccctgtagctacccggtaggtgcgctacaggcccgc
cggtggactactaaactctgttttaatactgtatctctgaatgcttcaacttaataagttaaaactttcaacaacggatctcttggttctg
gcatcgatgaagaacgcagcgaaatgcgataactaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcg
cccattagtattctagtgggcatgcctattcgagcgtcatttcaacccttaagcctagctgcttagtgttgggaatctgccccgtaact
tacggggcagctccctaaagtcatcggcggagttagggcatactctaagcgtagtaatattcttctcgcttctgtagttgtcctggc
ggcttgccgttaaaccccctatattttctagtggttgacctcggattaggtaggaatacccgctgaacttaagcatatc

>BRD3-111B
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagttaccgctctacaacccttttgtg
aacatacctaactgttgcttcggcgggcagaggttccctcgcggaacccctcccggtgacgccctcacgggcgtcgcgcccgc
cggaggataccaaactctattttaacgacgtttcttctgagtggcacaagcaaataattaaaacttttaacaacggatctcttggttct
ggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattg
cgcccgccagcattctggcgggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtgttggggccctacggttaaa
cgtaggcccttaaaggtagtggcggaccctctcggagcctcctttgcgtagtaactaacgtctcgcactgggattcggagggatt
ctagccgtaaaaccccccaatttttaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatc

>BRD3-111E
gaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattactgaactagtaattctctgaaaggtcgcc
ggtacccggtcccccccctaaacaaggggcccggggaaggtcctctcacacccttgtgtaccttaccatgttgctttggcggg
ccgacccggtttcgacccgggcggccggcgcccccagccctcaccggccaggacgtcaggctaagcgcccgccagtataca
aaactcaagcgattatttcgtgtagttctgataaattattcaattaattaaaactttcaacaacggatctcttggttctggcatcgatgaa
gaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggtatt
ccggagggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtattgggcgacgtccgctgccggacgcgcctcg

aagacctcggcgacggcgtcttagcctcgagcgtagtagtaacatctcgctttggagtgctaggcgtcggccgccggacaatcg
accttcggtctattacttccaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatca

>BRD3-111J
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattaccgagtttacaactcccaaacccctgt
gaacataccttaatgttgcctcggcggatcagcccgcgccccgtaaaacgggacggcccgccagaggacccaaactctaatgt
ttcttattgtaacttctgagtaaaacaaacaaataaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgca
gcaaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgctggtattccggcg
ggcatgcctgttcgagcgtcatttcaaccctcaagcccccgggtttggtgttggggatcggctctgcccttctgggcggtgccgc
ccccgaaatacattggcggtctcgctgcagcctccattgcgtagtagctaacacctcgcaactggaacgcggcgcggccatgc
cgtaaaaccccaacttctgaatgttgacctcggatcaggtaggaatacccgctgaacttaagcatatc

>BRD3-111K
aggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagttaccgctctataaccctttgtgaac
gtacctaaccgttgcttcggcgggcaggggaagcctctcgcgggcctcccctcccggcgccggcccccaccacggggacgg
ggcgcccgccggaggaaaccaaactctatttacacgacgtctcttctgagtggcacaagcaaataattaaaactttaacaacgg
atctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttg
aacgcacattgcgctcgccagcattctggcgagcatgcctgttcgagcgtcatttcaaccctcaagcaccgcttggttttggggcc
ccacggccgacgtgggcccttaaaggtagtggcggaccctcccggagcctcctttgcgtagtaactaacgtctcgcactgggat
ccggagggactcttgccgttaaaccccccaaattctttacaggttgacctcggatcaggtaggaatacccgctgaacttaagcatat

>BRD3-111N
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagttcttacaaactccaaccctttgt
gaaccataccaactgttgcctcggcgcgagctgcggctgcttggtagctacccggtagtcacctacccgggagctacccggta
gcgtctgcgtacaggcccgccgaaggaccaccaaactctgtttgacagtgtattctgaatgcttcaactaaatagttaaaactttca
acaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcg
aatctttgaacgcacattgcgcccattagtattctagtgggcatgcctattcgagcgtcatttcgaccctgaagccctagttgcttcg
cgttgggactctactggctaccctgtagttccctaatgacagtggcggagttcaggtgtactctcagcgtagtaatttcatctcgcttt
tgcagtagcctggtcgccggccgtaaaaccccctattttctagtggttgacctcggattaggtaggaatacccgctgaacttaagc
atatc

>BRD3-111T
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattacagagttgcaaaactccctaaaccatt
gtgaacgttacctataccgttgcttcggcgggcggccccggggtttacccccgggcgcccctgggccccaccgcgggcgcc
cgccggaggtcaccaaactcttgataatttatggcctctctgagtcttctgtactgaataagtcaaaactttcaacaacggatctcttg
gttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcac
attgcgcccgccagcattctggcgggcatgcctgttcgagcgtcatttcaaccatcaagcccccgggcttgtgttggggacctgc
ggctgccgcaggccctgaaaagcagtggcgggctcgctgtcgcaccgagcgtagtagcatacatctcgctctggtcgcgccg
cgggttccggccgttaaaccacctttaacccaaggttgacctcggatcaggtaggaagacccgctgaacttaagcatat

>BRD3-113C
ggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacttgcggaaggatcattacaagtgaggctaccgaacgc
tggaaacagcggttaggagcttacacccacccgtgtttacatactattgttgctttggcgggccgtggcctccactgcgggctctg
ctcgtgtgtgcccgccagagaaccaaactctgaatgttagtgatgtctgagtactatctaatagttaaaactttcaacaacggatctc
ttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacg

cacattgcgccctgtggtattccgcagggcatgcctgttcgagcgtcatttaaccactcacgcctagcgtggtattggggcacgc
ggtctccgcggccctcaaaattagtggcggcgccggtgggctctaagcgtagtacatactcccgctatagagttcccccggtgg
ctcgccagaacccccaatttttacaggttgacctnnnatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>BRD3-113P
gaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaagagatcatgccctcacgggtagacct
cccacccttttgtttacaatgcctttgttgctttggcgggcccgtttggccccgcgctgaacaaccggccccggctggtcagtgccc
gccagagaaccgaaaactctgaattaaatgtcgtctgagtactatgtaatagttaaaactttcaacaacggatctcttggttctggca
tcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccc
tctggtattccgggggggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggtattgggctacacccgactgggtggg
ccttaaaatcagtggcggtgccatctggctctaagcgtagtaattcttctcgctctggagatctaggtgtttgcttgccagcaacccc
caatttatcaaaggttgacctcggatcaggtagggatacccgctgaacttaagcata

>BRD3-113Y
tggtcatttagaggaagtaaaagtcgtaacaaggnnnccgttggtgaaccagcggagggatcattacagagttgcaaaactccc
taaaccattgtgaacgttacctataccgttgyttcggcgggcggccccggggtttaccccccgggcgcccctgggccccaccg
cgggcgcccgccggaggtcaccaaactcttgataatttatggcctctctgagtcttctgtactgaataagtcaaaactttcaacaac
ggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctt
tgaacgcacattgcgcccgccagcattctggcgggcatgcctgttcgagcgtcatttcaaccatcaagcccccgggcttgtgttg
gggacctgcggctgccgcaggccctgaaaagcagtggcgggctcgctgtcgcaccgagcgtagtagcatacatctcgctctg
gtcgcgccgcgggttccggccgttaaaccacctttaacccaaggttgacctcggatcaggtaggaagacccgctgaacttaag
catatc

>BRD3-121F
agaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaaaagttatgtcctacttgcctccgtgca
aaaggataagattccacgtgtgaaacagtgagctctttggagcgaagctgggccagcccctgtgaaccgaaaccgcctgcttct
cggagcgtggtgctattaggcgatcaggatcagtgcagcttgcggtgcgctctagtaatagggccttcgggatgaatatacttaa
cccgtgtctaccttctttgttgctttggcgggtcgcctctaaggcgttggcttcggctaaccgtacccgccagaggattcaaactctt
ttgttttagtgacgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaa
atgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgaggggcatg
cctgttcgagcgtcattatgaccaatcaagcctggcttggtgttggggttcgcggtcccgcggcccttaaactcagtggcggtgc
cgcccggctctaagcgtagtaaatttcctcgctatagcgtccgggtggttacccgccagaaccccccaatcttacggttgacctcg
gatcaggtagggatacccgctgaacttaagcatatca

>BRD3-121J
gaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattaccgagtgagggctccggcccaacctcc
aaccccatgtgaatccgacctctgttgcctcgggggcgacccggccctgcgccggggccccccggtggaccactcaactctgc
atctttgcgtctgagtatgatatttgaatcaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaa
atgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggcatg
cctgttcgagcgtcattacaccactcaagcctggcttggtattgggcgaggcggctttgcggccgcccgcctcaaagtcttcggc
tggactgaccgtctctaagcgttgtgacttcattggaccgcttgcgagtacgggacagcccgtggccgttaaacccccccatgaaa
ggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>BRD3-124A
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagttaccgctctacaacccttttgtg
aacatacctaaccgttgcttcggcgggcaggggaagcctctcgtgggcggaccctcccggcgccggccccgtcacgggggc
ggagcgcccgccggaggaaaccaaactctatttacacgacgtctcttctgagtggcacaagcaaataattaaaacttttaacaac
ggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctt
tgaacgcacattgcgctcgccagcattctggcgagcatgcctgttcgagcgtcatttcaaccctcaagcaccgcttggttttgggg
ccccacggccgacgtgggcccttaaaggtagtggcggaccctcccggagcctcctttgcgtagtaactaacgtctcgcactgg
gatccggagggactcttgccgttaaaccccccaaatttttttcaggttgacctcggatcaggtaggactacccgctgaacttaagcat
at

>BRD3-125A
aggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattactgagttagggagcaatcccgacctccaa
ccctttgtgaacgcatcatgttgcttcgggggcgaccctgccgttcgcggcattcccccggaggtcatcaaaacactgcattctt
acgtcggagtaaaaagttaatttaataaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcga
taagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgtt
cgagcgtcatttcaccactcaagcctcgcttggtattgggcgtcgcgagtctctcgcgtgcctcaaagtctccggctgtcttgattc
aattcccagcgttgtggcaactatttcgcagtggaaacgagtcgatgtggccgttaaatctttcaaaggttgacctcggatcaggta
gggatacccgctgaacttaagcatatcaata

>BRD3-126C
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattactgagtgagggtctccgg
gcccgacctccaacccttttgtgaaccacctcgttgcttcgggggcgaccctgccgttcgcggcgtggggccccggagcccat
caaaccctgcgtaactaagtcgtcggagtttaaacaaattaaacaaaactttcaacaacggatctcttggttcgggcatcgatgaa
gaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccctggtat
tccgggggggcatgcctgtccgagcgtcattacaccaatcaagcctggcttggtattgggcgtcgcgggtctgacccgcgcgcct
taaagtctcaccggctgagcggcgtcgtctctaagcgttgtggaaactattcgcggagaggaggcgcggcgtggccgttaaac
accccatcaaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>BRD3-154C
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtagagttttgcctactacagcacgcctacccggcgcttgcta
gtctcgcgagaggcgacacgctaaatgtgcggagagtcctggctcgccaggcaatccgcagtcactttcgggtgattcagaga
ctcaatgagcgtgggttctttttaaagagcttgagatagagcccgtcctcgctggtcacagcggggattcaacggaacctgcgga
aggatcattaaagaataaacgggcctccgggccccctattctcaccctatgtttaccacactttgttgccttggcgcactgcgccag
cggatcaaaacccttgaatcattgctgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatga
agaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggta
ttccgggggggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggtattgggctcgccctgtagggcttgcctcaaag
tcagtggcggctaccgtccgaccttcagcgcagtactatccgtcgctgttagggaaggcctataacccgtcatcacaaccccccac
tttacaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataa

>BRD3-154E
agaggaagtaaaagtcgtaacaaggtttccgtaggtcaagttttgcctaccacagcagcccccgctgctagtctcctagtgaggc
gacacgctcaatttgcggagaggtcctcttctgaggaaaatccgcagcccatctcccttacccggaggtcggttcagaggctcaa
tgagtgtgggttcagcgcagctgagcttaagatagagtccgtcttggctggcaacagcgaagattcaacggaacctgcggaag
gatcattaaagaatacatggccttcgggtcctattctcacccttgtttaccaaaactcttgttgccttggcgcatttatgcgccaaag
gaatcaaacccttgaatctctgctgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaag

aacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcctcctggtattc
cgggaggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggtattgggctcgccttctttggcctgcctcaaaatcag
tggcggcacagtccgatcctcaagcgcagtaatacacgacgcttgccggtgaaggttgctgctccagaaaccccccacaaact
aaaggttgacctcgaatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>BRD3-154F
tagaggaagtaaaagtcgtaacaaggtttccgtaggtcaagtcttgcctaccacagcagccccggctgctagtctctttagtgagg
cgacacgctcaatttgcggagaggtcctcttctgaggaaaatccgcagcccatctcccttacccggaggtcggttcagaggctca
atgagtgtgggttcagcgcagctgagcttaagatagagtccgtcttggctggcaacagcaaagattcaacggaacctgcggaag
gatcattaaagaatacatggccttcgggtcctattctcacccttttgtttaccaaaactcttgttgccttggcgcattcgtgcgccaaag
gaatcaaacccttgaatctctgctgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaag
aacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcctcctggtattc
cgggaggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggtattgggctcgcctcctttggcctgcctcaaaatca
gtggcggcacagtccgatcctcaagcgcagtaatacacgacgcttgccggtgaaggttgctgctccagaaaccccccacaaac
taaaggttgacctcggatcaggtagggatacccgctgaacttaagcata

>BRD3-154J
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaannngcggaaggatcattaatgattaactgtctgtcgag
cttgctcacagacttatcatatccataacacctgtgcacttgtcggatggcttagtgaagaccgcaaggttggatctatccatctactt
tacataacaatttagtaacaaatgtagtcttattataacataataaaactttcaacaacggatctcttggctctcgcatckatgaagaa
cgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaccttgcgctccctggtattccg
gggagcatgcctgtttgagtgtcatgaaaaccctcaaccttagattggttaacacctttctttggcttggatttggacgtttgccgatg
ataagtcggctcgtcttaaaagtaatagctggatctgtctcgcgacatggtttgacttggcgtaataagtatttcgctaaggacatctt
cggatggccgcgttgcaagactaaagaccgctttctaatccattgatcttcggattaatattcttgacatctggcctcaaatcaggta
ggactacccgctgaacttaagcatatca

>BRD3-155A
aggaagtaaaagtcgtaacaaggtttccgtaggtcaagtcttgcctaccacagcagccccggctgctagtctcctagtgaggcg
acacgctcaatttgcggagaggtcctccttctgaggaaaatccgcagcccacctcccctacccggaggtcggttcagaggctca
atgagtgtgggttcagcgcagctgagcttaagatagagtccgtcttggctggcaacagcgaagattcaacggaacctgcggaa
ggatcattaaagaatacatggccttcgggccctattctcacccttgtttaccaaaactcttgttgccttggcgcattcgtgcgccaaa
ggaatcaaacccttgaatctctgctgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaa
gaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcctcctggtatt
ccgggaggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggtattgggctcgcctcctttggcctgcctcaaaatca
gtggcggcacagtccgatcctcaagcgcagtaatacacgacgcttgctggtgaaggttgctgctccagaaaccccccacaaacta
aaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcat

>JAJ1-001
cggaaggatcattatccatctcaaccaggagtggtggcggccaggcattgcgtcctcggttcagccccgctttcctgcatccttttt
ttacgagcacctttcgttctccttcggcggggcaacctgccgctggaacttaacaaaaccttttttgcatctagcattacctgttctg
atacaaacaatcgttacaactttcaacaatggatctcttggctctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtg
aattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccatggggcatgcctgttcgagcgtcatct
acaccctcaagctctgcttggtgttgggcgtctgtcccgcctctgcgcgcggactcgccccaaattcattggcagcggtctttgcc
tcctctcgcgcagcacaattgcgtctgcgaggggcgtggcccgcgtccacgaagcaacattaccgtctttgacctcggatcag
gtagggatacccgctgaacttaagcata

>JAJ1-005
cggagggatcattacagagactctgccctttgggtagacctcccaccctgtgtcgttataccttcgttgctttggcgggccgcggg
gctccggccctgcccctggctccggctagggcgcgcccgccagaggactcccaaacctgaatgttagtgtcgtctgagtactat
ataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgca
gaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggcatgcctgttcgagcgtcattacaaccc
tcaagctctgcttggtcttgggcgtcaccggtccccggtgtgcctcaaaatcagtggcggcgccgtctggctctaagcgtagtac
atactctcgctacagacgcccggcggatgctggccagcaacccccaatctatcaaggttgacctcggatcaggtagggataccc
gctgaacttaagcatat

>JAJ1-009
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattatagagttttctaaactcccaacccatgt
gaacttaccattgttgcctcggcagaagctacctggttaccctaccttggaacggcctaccctgtagcgccttaccctggaacggc
ctaccctgtaacggctgccggtggactaccaaactcttgttattttattgtaatctgagcgtcttattttaataagtcaaaactttcaaca
acggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaat
ctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccttaagcctagcttagtgttggg
agcctaccgcttttgctagcggtagctcctgaaatacaacggcggatctgcgatatcctctgagcgtagtaatttttatctcgcttttg
actggagwwgcagcgtctttagccgctaaaccccccaaatttttaatggttgacctcggatcaggtaggaatacccgctgaact
taagcatatcaataagcggag

>JAJ1-013
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattatagagtttttaaactcccaa
cccatgtgaacttaccattgttgcctcggcagaagctacctggttaccttaccttggaacggcctaccctgtagcgccctaccctg
gaacggcctaccctgtaacggctgccggtggactaccaaactcttgttattttattgtaatctgagcgtcttattttaataagtcaaaa
ctttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaat
catcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccttaagcctagctta
gtgttgggagcctactgcttttgctagcggtagctcctgaaatacaacggcggatctgcgatatcctctgagcgtagtaatttttatct
cgcttttgactggagttgcagcgtctttagccgctaaaccccccaatttttaatggttgacctcggatcaggtaggaatacccgctg
aacttaagcatatca

>JAJ1-021
cggagggatcattacagagactctgccctttgggtagacctcccaccctgtgtcgttataccttcgttgctttggcgggccgcggg
gctccggccctgcccctggctccggctagggcgcgcccgccagaggactcccaaacctgaatgttagtgtcgtctgagtactat
ataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgca
gaattcagtgaatcatcgaatctttgaacgcacattgcgcccctggtattccgggggggcatgcctgttcgagcgtcattacaaccc
tcaagctctgcttggtcttgggcgtcaccggtccccggtgtgcctcaaaatcagtggcggcgccgtctggctctaagcgtagtac
atactctcgctacagacgcccggcggatgctggccagcaacccccaatctatcaaggttgacctcggatcaggtaggaatacccc
gctgaacttaagcata

>JAJ1-025
cggagggatcattactgagttatctaaactccaaccctatgtgaaccttaccgtcgttgcctcggcgggctgcgtctaccctgtagc
taccctgtagctacccggtaggcgtgctacaagcccgccggtggaccactaaactctgttataaatactgtatctctgaatgcttca
acttaataagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaatt
gcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctattcgagcgtcatttcaac
ccttaagcctaagttgcttagcgttgggaatctgccctgtattacggggcagttccctaaagttatcggcggagttagggcatactc

taagcgtagtactattattctcgcttctgcagttgtcccgacggcttgccgctaaacccctatattttctagtggttgacctcggattag
gtaggaatacccgctgaacttaagca

>JAJ1-029
gaacctgcggagggatcattacaagtgaccccggtctaaccaccgggatgttcataacccttttgttgtccgactctgttgcctccg
gggcgaccctgccttcgggcgggggctccgggtggacacttcaaactcttgcgtaactttgcagtctgagtaaacttaattaataa
attaaaacttttaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattc
agtgaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgttcgagcgtcatttcaccactcaag
cctcgcttggtattgggcaacgcggtccgccgcgtgcctcaaatcgaccggctgggtcttctgtcccctaagcgttgtggaaact
attcgctaaagggtgttcgggaggctacgccgtaaaacaaccccatttctaaggttgacctcggatcaggtagggatacccgctg
aacttaagcata

>JAJ1-033
agcattacctgttctgatacaaacaatcgttacaactttcaacaatggatcnnntggctcnngcatcgatgaagaacgcagcgaa
atgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcnnattncgccccttggtattccatggggcatg
cctgttcgagcgtcatctacaccctcaagctctgcnnngtgttgggcgtctgtcccgcctctgcgcgnnnactcgccccaaattc

>JAJ1-041
cggagggatcattactgagttatctaaactccaaccctatgtgaaccttaccgtcgttgcctcggcgggctgcgtctaccctgtagc
taccctgtagctacccggtaggcgtgctacaagcccgccggtggaccactaaactctgttataaatactgtatctctgaatgcttca
acttaataagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaatt
gcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctattcgagcgtcatttcaac
ccttaagcctaagttgcttagcgttgggaatctgccctgtattacggggcagttccctaaagttatcggcggagttagggcatactc
taagcgtagtactattattctcgcttctgcagttgtcccgacggcttgccgctaaacccctatattttctagtggttgacctcggattag
gtaggaatacccgctgaacttaagcata

>JAJ1-045
cggaaggatcattatccatctcaaccaggagtggtggcggccaggcattgcgtcctcggttcagccccgctttcctgcatccttttt
ttacgagcacctttcgttctccttcggcggggcaacctgccgctggaacttaacaaaaccttttttgcatctagcattacctgttctg
atacaaacaatcgttacaactttcaacaatggatctcttggctctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtg
aattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccttggtattccatggggcatgcctgttcgagcgtcatct
acaccctcaagctctgcttggtgttgggcgtctgtcccgcctctgcgcgcggactcgccccaaattcattggcagcggtctttgcc
tcctctcgcgcagcacaattgcgtctgcgagggggcgtggcccgcgtccacgaagcaacattaccgtctttgacctcggatcag
gtagggatacccgctgaacttaagcatat

>JAJ1-049
acctgcggaaggatcattatccatctcaaaccaggtgcggtcgcggcccccgggggcttgctcccgggtggtaggggtaacac
cctcacgcgccgcctgcctgtaccctctttttacgagcacctttcgttctccttcggcggggcaacctgccgctggaaccaaaaca
aaaccttttttgcatctagcattacctgttctgatacaaacaatcgttacaactttcaacaatggatctcttggctctggcatcgatgaa
gaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccttggtatt
ccatggggcatgcctgttcgagcgtcatctacaccctcaagctctgcttggtgttgggcgtctgtcccgcctctgcgcgcggactc
gccccaaattcattggcagcggtctttgcctcctctcgcgcagcacaattgcgtctgcgggggggcgtggcccgcgtccacgaa
gcaacattaccgtctttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>JAJ1-057
agcatttcctgttcngatacaaacaatcgttacaactttcaacaatggatnnnntggctctggcatcgatgaagaacgcagcgaa
atgcgataagtagtgtgaattgcagaattcagtgaatcannnaatcnnngaacgcacattgcgccccttggtattccatggggca
tgcctgtnngagcgtcatctacaccctcaagctctgcnnggtgttgggcgtctgtcccgcc

>JAJ1-061
cgtaggtgaacctgcggaaggatcattactgaaaatgtaataaacccttcaggttttggaaggggagccgtcaaaagcttccct
ggtacatgcctcacccctttgtatatctaccatgttgctttggcgggccgacccggtttcgacccgggcggccggcgccccagc
ctgcttgccaggccaggacgcccggccaagtgcccgccagtatacaaaactccagcgattattttgtgtagtcctgagaatttatt
caataaattaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgca
gaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggcattccggagggcatgcctgttcgagcgtcatttcaaccc
tcaagctctgcttggtattgggcgacgtctgctgtcagacgcgcctggaagacctcggcgacggcattccagcctcgagcgtag
tagtaaaatatctcgctttggaggatggggtgacggcttgccggacaaccgacctctggtcatttttttccaaggttgacctcggatc
aggtagggatacccgctgaacttaagcatatcaataagcggag

>JAJ1-065
ccgtaggtgaacctgcggagggatcattactgagtgagggccccggnccgacctccaacccttttgtgaaccgaacttgttgcttc
gggggcgaccctgccgctcgcggccgcggcgcccccggaggccttcaaaaccctgcgttctgtacgtcggagtctcaagtaa
attgaacaaaactttcaacaacggatctcttggtactggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcag
aattcagtgaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgttcgagcgtcatttcaccact
caagccccgcttggtattgggcgtcgcgggtctccgcgcgcctcaaagtctccggctgagctgttgtaaaccctagcgttgtgaa
acatattcgcttgtggtctgcgcagccgcggccgttaaacttttcaaaggttgacctcggatcaggtagggatacccgctgaactt
aagcatatcaataagcggagga

>JAJ1-073
cggaaggatcattaaagagtaagggtgctcagcgcccgacctccaacccttttgttgttaaaactaccttgttgctttggcgggacc
gctcggtctcgagccgctggggattcgtcccaggcgagcgcccgccagagttaaaccaaactcttgttatttaaccggtcgtctg
agttaaaattttgaataaatcaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatgcgataagta
atgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgaggggcatgcctgttcgagcg
tcattacaccactcaagctatgcttggtattgggcgtcgtccttagttgggcgcgccttaaagacctcggcgaggccactccggct
ttaggcgtagtagaatttattcgaacgtctgtcaaaggagaggaactctgccgactgaaacctttatttttctaggttgacctcggat
caggtagggatacccgctgaacttaagcatat

>JAJ1-093
ccgtaggtgaacctgcggaaggatcattatccatctcaaaccaggtgcggtcgcggcccccgggggcttgctcccgggtggta
ggggtaacaccctcacgcgccgcctgcctgtaccctcttttacgagcacctttcgttctccttcggcggggcaacctgccgctgg
aaccaaaacaaaaccttttttgcatctagcattacctgttcngatacaaacaatcgttacaactttcaacaatggatcncttggctctg
gcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgc
gccccttggtattccatggggcatgcctgttcgagcgtcatctacaccctcaagctctgcttggtgttgggcgtctgtcccgcctct
gcgcgcggactcgccccaaattcattggcagcggtctttgcctcctctcgcgcagcacaattgcgtctgcgggggggcgtggc
ccgcgtccacgaagcaacattaccgtctttgacctcggatcaggtagggatacccgctgaacttaagcat

>JAJ1-097
cggagggatcattacagagttatccaactcccaaacccatgtgaacttatctctttgttgcctcggcgcaagctacccgggacctc
gcgccccgggcggcccgccggcggacaaaccaaaactcttgttatcttcgttgattatctgagtgtcttatttaataagtcaaaactt

tcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatca
tcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccctaagcacagcttatt
gttgggaacctacggcttcgtagttcctcaaagacattggcggagtggcagtggtcctctgagcgtagtaatcttttatctcgcttct
gttaggtgctgcccccccggccgtaaaaccccccaatttttctggttgacctcggatcaggtaggaatacccgctgaacttaagca
ta

>JAJ1-105
cttccgtaggtgwwcctgaykgaarratcattmtcmatgctsaaaccaggtgcggtcgcggcccccgggggcttgctcccg
ggtggtaggggtaacaccctcacgcgccgcctgcctgtaccctctttttacgagcacctt

>JAJ1-113
gcggaaggatcattatccatctcaaaccaggtgcggtcgcggcccccgggggcttgctcccgggtggtaggggtaacaccctc
acgcgccgcctgcctgtaccctctttttacgagcacctttcgttctccttcggcggggcaacctgccgctggaaccaaaacaaaa
cctttttgcatctagcattacctgttctgatacaaacaatcgttacaactttcaacaatggatctcttggctctggcatcgatgaagaa
cgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccctggtattcca
tggggcatgcctgttcgagcgtcatctacaccctcaagctctgcttggtgttgggcgtctgtcccgcctctgcgcgcggactcgcc
ccaaattcattggcagcggtctttgcctcctctcgcgcagcacaattgcgtctgcgggggggcgtggcccgcgtccacgaagca
acattaccgtctttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>JAJ1-117
cggagggatcattactgagtgagggccccggcccgacctccaacccttgtgaaccgaacttgttgcttcgggggcgaccctgc
cgctcgcggccgcggcgcccccggaggccttcaaaaccctgcgttctgtacgtcggagtctcaagtaaattgaacaaaactttc
aacaacggatctcttggtactggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcat
cgaatctttgaacgcacattgcgcccctggtattccgggggggcatgcctgttcgagcgtcatttcaccactcaagccccgcttgg
tattgggcgtcgcgggtctccgcgcgcctcaaagtctccggctgagctgttgtaaaccctagcgttgtgaaacatattcgcttgtg
gtctgcgcagccgcggccgttaaacttttcaaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataa
gc

>JAJ1-121
ttccgtaggtgaacctgcggaaggatcattactgaaaatgtaataaacccttcaggttttggaaaggggagccgtcaaaagcttcc
ctsgkacatrcctsaccyctkgtcatwtytwccaakgttgacytyggatcrggccgacccggtttcgacccgggcggccggc
gcccccagcctgcttgccaggccaggacgcccggccaagtgcccgccagtatacaaaactccagcgattattttgtgtagtcct
gagaatttattcaataaattaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaa
tgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggcattccggagggcatgcctgttcgagcgt
catttcaaccctcaagcnnngcttggtattgggcgacgtctgctgtcagacgcgcctggaagacctcggcgacggcattcc

>JAJ1-125
cggagggatcattacaagtgaccccggtctaaccaccgggatgttcataacccttttgttgtccgactctgttgcctccggggcgac
cctgccttcgggcgggggctccgggtggacacttcaaactcttgcgtaactttgcagtctgagtaaacttaattaataaattaaaac
ttttaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatc
atcgaatctttgaacgcacattgcgcccctggtattccgggggggcatgcctgttcgagcgtcatttcaccactcaagcctcgcttg
gtattgggcaacgcggtccgccgcgtgcctcaaatcgaccggctgggtcttctgtcccctaagcgttgtggaaactattcgctaa
agggtgttcgggaggctacgccgtaaaacaaccccatttctaaggttgacctcggatcaggtagggatacccgctgaacttaag
catat

>JAJ1-129
cggaaggatcattaaaagtaaatatgtattctctggaatatacgtttcaaacccatctgtgtatcttacctgttgcttccgtaccgcgtg
gacttcggtcccctcatgctgtacttcggtacggttccagagtgccggtgcggaaggtatacatcaaactcttgcattccatgtcttc
tgtctgaactgttaataacaaaagtttaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatgcga
taagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcctcctggtattccgggaggcatgcctgttc
gagcgtcattaaaaaccactcaagctcttttgcttggtcatggaagaagagtccgcttgccggtctccctttcgaaatgcaatggcg
gaaggtctcttgtacccagcgtagtaataacttttcgctttgtgcattgattcattcctgccgtcaaaccccccaattttcttaggttgacc
tcggatcaggtagggatacccgctgaacttaagcatat

>JAJ1-133
acctgcggaaggatcattacctagagttgcgggctttgcctgccatctcttacccatgtcttttgagtaccttacgtttcctcggcgg
gtccgcccgccgactggacaatttaaaccacttgcagttgcaatcagcgtctgaaaaaacttaatagttacaactttcaacaacgg
atctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttg
aacgcacattgcgcccttggtattccatggggcatgcctgttcgagcgtcatttgtaccttcaagccttgcttggtgttgggtgtttg
tctcgcctctgcgcgtagactcgcctcaaaacaattggcagccggcgtattgatttcggagcgcagtacatctcgcgctttgcact
cagaacgacgacgtccaaaagtacatttttacactcttgacctcggatcaggtagggatacccgctgaacttaagcata

>JAJ1-137
tgcggaaggatcattacctagagtttgtagacttcggtctgctacctcttacccatgtcttttgagtaccttcgtttcctcggcgggtcc
gcccgccgattggacaacattcaaacccttgcagttgcaatcagcgtctgaaaaaacataatagttacaactttcaacaacggat
ctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaa
cgcacattgcgcccttggtattccatggggcatgcctgttcgagcgtcatttgtaccttcaagctctgcttggtgttgggtgtttgtc
tcgcctctgcgtgtagactcgccttaaaacaattggcagccggcgtattgatttcggagcgcagtacatctcgcgctttgcactcac
aacgacgacgtccaaaagtacatttttacactcttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>JAJ1-141
cggaaggatcattatccatctcaaaccaggtgcggtcgcggcccccgggggcttgctcccgggtggtaggggtaacaccctca
cgcgccgcctgcctgtaccctcttttacgagcacctttcgttctccttcggcggggcaacctgccgctggaaccaaaacaaaac
ctttttgcatctagcattacctgttctgatacaaacaatcgttacaactttcaacaatggatctcttggctctggcatcgatgaagaac
gcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccttggtattccat
ggggcatgcctgttcgagcgtcatctacaccctcaagctctgcttggtgttgggcgtctgtcccgcctctgcgcgcggactcgcc
ccaaattcattggcagcggtctttgcctcctctcgcgcagcacaattgcgtctgcgggggggcgtggcccgcgtccacgaagca
acattaccgtctttgacctcggatcaggtagggatacccgctgaacttaagcatat

>JAJ1-145
ctgcggaaggatcattactgaaaatgtaataaacccttcaggttttggaaggggggagccgtcaaaagcttccctggtacatgcctc
acccccttgtatatctaccatgttgctttggcgggccgacccggtttcgacccgggcggccggcgcccccagcctgcttgccagg
ccaggacgcccggccaagtgcccgccagtatacaaaactccagcgattattttgtgtagtcctgagaatttattcaataaattaaaa
ctttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaat
catcgaatctttgaacgcacattgcgccctctggcattccggagggcatgcctgttcgagcgtcatttcaaccctcaagctctgctt
ggtattgggcgacgtctgctgtcagacgcgcctggaagacctcggcgacggcattccagcctcgagcgtagtagtaaaatatct
cgctttggaggat

>JAJ1-149
acctgcggaaggatcattaaagagttagggtcccagtggcccaacctccaaccctctgttgttaaaactaccttgttgctttggcgg
gaccgttcggtctccgagcgcgtcaacccccgggtcggcgagcgcccgccagagtccaaccaaactcttgtattaaaccagtc
gtctgagtataaaattttaattaaattaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatgcgat
aagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgaggggcatgcctgttc
gagcgtcattacaccactcaagcactgcttggtattgggcatcgtccgtcgaaaggcgggcgtgcctcgaagacctcggcggg
gtttcaccgacttcgggcgtagtagagttaaatcaaaacgtctcataagcttgatggagcctcattgccgttaaacctttttatttttata
ggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>JAJ1-157
cggagggatcattactgagttctacaaaaactcccaacccttttgtgaattataccttagttgcctcggcgtcgagcgagtcaaccta
cccgggagctaccctggagtcacctaccctgtagaacctaccctgtaggacctaccctgtagacggctaccctggagctaccct
gtagttgcactttcgcgtcgccggtggactaccaaactcttatatgtatagtgtatctctgaattcttaacaaaattagttaaaactttca
acaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcg
aatctttgaacgcatattgcgcccagtagtattctactgggcatgcctattcgagcgtcatttcaacccttacgccctgtagcgtagt
gttaggactctactctttagagagcagttccctaaaaccagtggcagtgttcggtacactcatagcgtagtaattcttctcgcttctgc
agtggcccgtactactcgccgtaaaaccccctaattttctaatggttgacctcggattaggtaggaatacccgctgaacttaagcata
tc

>JAJ1-161
ccgtaggtgaacctgcggagggatcattacagagttcatgcccttcggggtagatctcccacccgtgttattatactattgttgcttt
ggcgggcccgcctcggccaccggctccggctggtgagcgcccgccagaggaccccaaactctgaaatttagtgtcgtctgag
tactattttaatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtga
attgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggcatgcctgttcgagcgtcatta
caaccctcaagctctgcttggtattgggcgtccccggcagcggggtgccctaaaatcagtggcggtgccgtctggctctaagcg
tagtaaatctctcgctctggagacccggcggatgcttgccagcaacccccaatttttaaggttgacctcggatcaggtagggata
cccgctgaacttaagcatatcaataagcggag

>JAJ1-163
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattatagagttttctaaactcccaacccatgt
gaacttaccattgttgcctcggcagaagctacctggttaccctaccttggaacggcctaccctgtagcgccttaccctggaacggc
ctaccctgtaacggctgccggtggactaccaaactcttgttattttattgtaatctgagcgtcttattttaataagtcaaaactttcaaca
acggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaat
ctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccttaagcctagcttagtgttggg
agcctaccgcttttgctagcggtagctcctgaaatacaacggcggatctgcgatatcctctgagcgtagtaatttttatctcgcttttg
actggagttgcagcgtctttagccgctaaacccccccaaatttttaatggttgacctcggatcaggtaggaatacccgctgaactta
agcatatca

>KP1-013B
cggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtttatcgtaccttgttgcttcggcgggcccgcc
tcacggccgccgggggcttctgccctctggcccgcgcccgccgaagacaccattgaacgctgtctgaagattgcagtctgag
caattagctaaataagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgatacgtaat
gtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccttggtattccgggggggcatgcctgtccgagcgt
cattgctgccctcaagcacggcttgtgtgttgggctccgtcctccttcccgggggacgggcccgaaaggcagcggcggcaccg

cgtccggtcctcgagcgtatggggcttcgtcttccgctcttgtaggcccggccggcgcttgccgacaacaatcaatctttttttcagg
ttgacctcggatcaggtagggatacccgctgaacttaagcatat

>KP1-017A
ctgcggaaggatcattacagagttcatgcccttcggggtagacctcccacccttgtgtattattaccttgttgctttggcggtgccgc
gtggcttcggccgcgccttgggctctcgagcccgagcgtgcccgccagaggaaacccaaactctgaatatttttgtcgtctgagt
actatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaa
ttgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccgggggggcatgcctgttcgagcgtcatttca
accctcaagctctgcttggtattgagccccgccagcgatggcgggccctaaaatcagtggcggcgccgctgggtcctgagcgt
agtaattctctcgctacagggtccccgcgtgcttctgccaacaaccccaaattttctatggttgacctcggatcaggtagggatacc
cgctgaacttaagcatat

>KP1-017C
gatatgcttaagttcagcgggtagtcctacctgatttgaggtcaaattgtcaaaggtattgtccttgcggacggttagaagcgagtct
aaaccctgtccacggcgtagataattatcacaccaatagacggagctcagtacgaactcgctaatgcatttcaggggagcagac
cgcgctgaggcagcctgcacaaaccccccacatccaagcctcggagaaccgttcggaaaacgggtgaggttgagaatttaatga
cactcaaacaggcatgctcctcggaataccaaggagcgcaaggtgcgttcaaagattcgatgattcactgaattctgcaattcac
attacttatcgcatttcgctgcgttcttcatcgatgcgagagccaagagatccgttgctgaaagttgtatagtgttttataggcgatca
agcccattgactacattctatatcatacttgtggggtgtgtaaaaagacgtagagcctggaaattcgaggagagacacctcaaga
gggcaatcctcgcatccgcactcagagagcacgagagtcatccagacctacagtcggtgcacaggtggatagataaaaatggc
gggcgtgcacaatgctccgaggagccagctacaaccaagacaccatagttattcgttaatgatccttccgc

>KP1-017E
tgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtttatcgtaccttgttgcttcggcgggcccg
cctcacggccgccgggggcacctgcccccgggcccgcgcccgccgaagacaccattgaactctgtctgaagattgcagtct
gagcgattaactaaatcagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgataagt
aatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggcatgcctgtccgag
cgtcattgctgccctcaagcacggcttgtgtgttgggctccgcccccctcccgggggggcgggcccgaaaggcagcggcggca
ccgcgtccggtcctcgagcgtatggggctttgtcacccgctctgtaggcccggccggcgcccgccggcgaccccaatcaatct
ttccaggttgacctcggatcaggtagggatacccgctgaacttaagcatat

>KP1-045A
tgcggaaggatcattactgagtgagggcccctcggggtccaacctcccacccgtgtttaacgaaccttgttgctttggcgggccc
gcctcacggccgccgggggcatctgcccccgggcccgcgcccgccgaagccacctgtgaactctgtctgaagtatgcagtc
tgagacaattattaaattaattaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgataact
aatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccctctggtattccggagggcatgcctgtccgag
cgtcattgctgccctccagcccggctggtgtgttgggccccgccccccttcccgggggggcgggcccgaaaggcagcggcg
gcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctcttgtaggcccggccggcgccagccgacccccctcaatcta
ttttttcaggttgacctcggatcaggtagggatacccgctgaacttaagcatat

>KP1-045C
ttccgtaggtgaacctgcggaaggatcattactgagtgagggcccctcggggtccaacctcccacccgtgtttaacgaaccttgtt
gctttggcgggcccgcctcacggccgccgggggcatctgcccccgggcccgcgcccgccgaagccacctgtgaactctgt
ctgaagtatgcagtctgagacaattattaaattaattaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcag
cgaaatgcgataactaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccctctggtattccggagg

gcatgcctgtccgagcgtcattgctgccctccagcccggctggtgtgttgggccccgccccccttcccgggggggcgggccc
gaaaggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctcttgtaggcccggccggcgccagc
cgacccccctcaatctatttttttcaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagc

>KP1-045K
tattgatatgcttaagttcagcgggtatccctacctgatccgaggtcaaaagtgagaaaaatgtggtcttgatggatgctcaaccat
ggctgatcagaagtgcaagattgtgctgcgctccgaaaccagtaggccggctgccaatcattttaaggcgagtctcgtgagaga
caaagacgcccaacaccaagcaaagcttgagggtacaaatgacgctcgaacaggcatgccctttggaataccaaagggcgca
atgtgcgttcaaagattcgatgattcactgaattctgcaattcacactacgtatcgcatttcgctgcgttcttcatcgatgccagaacc
aagagatccgttgttgaaagttgtaataattacattgtttactgacgctgattgcaattacaaaaaaaggtttatggttgggtcctggt
ggcgggcgaacccgcccaggaaacaagaagtgcgcaaaagacatgggtgaataattcagacaagctggagcccccaccga
gatgaggtcccaacccgctttcatattgtgtaatgatccctccgcag

>KP1-063J
ttccgtaggtgaacctgcggaaggatcattactgagtgagggcccctcggggtccaacctcccacccgtgtttaacgaaccttgtt
gctttggcgggcccgcctcacggccgccgggggggcatctgccccgggcccgcgcccgccgaagccacctgtgaactctgt
ctgaagtatgcagtctgagacaattattaaattaattaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcag
cgaaatgcgataactaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccctctggtattccggagg
gcatgcctgtccgagcgtcattgctgccctccagcccggctggtgtgttgggccccgccccccttcccgggggggcgggccc
gaaaggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctcttgtaggcccggccggcgccagc
cgacccccctcaatctatttttttcaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>KP1-063N
ttccgtaggtgaacctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtctatcgtaccttgttg
cttcggcgggcccgccgtttcgacggccgccggggaggccttgcgccccgggcccgcgcccgccgaagaccccaacatg
aacgctgttctgaaagtatgcagtctgagttgattatcgtaatcagttaaaactttcaacaacggatctcttggttccggcatcgatga
agaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggt
attccgggggggcatgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggcccccgtcccccctctcccgggg
gacgggcccgaaaggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacctgctctgtaggcccggccg
gcgccagccgacacccaactttatttttctaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagc
ggagga

>KP1-075B
acctgcggaaggatcattactgagtgagggcccctcggggtccaacctcccacccgtgtttaacgaaccttgttgctttggcggg
cccgcctcacggccgccgggggggcatctgccccgggcccgcgcccgccgaagccacctgtgaactctgtctgaagtatgca
gtctgagacaattattaaattaattaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgata
actaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccctctggtattccggagggcatgcctgtccg
agcgtcattgctgccctccagcccggctggtgtgttgggccccgccccccttcccgggggggcgggcccgaaaggcagcgg
cggcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctcttgtaggcccggccggcgccagccgacccccctcaat
ctatttttcaggttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>KP1-089A
tgtattattaccttgttgctttggcggtgccgcgtggcttcggccgcgccttgggctctcgagcccgagtgtgcccgccagaggaa
acccaaactctgaatatttttgtcgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaaga
acgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattcc

gggggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtattgagccccgccagcgatggcgggctctaaaatc
agtggcggcgccgctgggtcctgagcgtagtaattctctcgctacagggtccccgcgtgcttctgccaacaaccccaaattttcta
tggttgacctcggatcaggtagggatacccgctgaacttaagcatat

>KP1-091A
ttccgtaggtgaacctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtttatcgtaccttgttg
cttcggcgggcccgcctcacggccgccgggggggcacctgcccccgggcccgcgcccgccgaagacaccattgaactctgtc
tgaagattgcagtctgagcgattaactaaatcagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcag
cgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggg
gcatgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggctccgcccccctcccgggggggcgggcccgaa
aggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctctgtaggcccggccggcgcccgccgg
cgaccccaatcaatctttccaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>KP1-119C
gcggaaggatcattatccatctcaaaccaggtgcggtcgcggcccccgggggcttgctcccgggtggtaggggtaacaccctc
acgcgccgcctgcctgtaccctctttttacgagcacctttcgttctccttcggcggggcaacctgccgctggaaccaaaataaaac
ctttttgcatctagcattacctgttctgatacaaacaatcgttacaactttcaacaatggatctcttggctctggcatcgatgaagaac
gcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccat
ggggcatgcctgttcgagcgtcatctacaccctcaagctctgcttggtgttgggcgtctgtcccgcctctgcgcgcggactcgcc
ccaaattcattggcagcggtctttgcctcctctcgcgcagcacaattgcgtctgcgggggggcgcggcccgcgtccacgaagc
aacattaccgtctttgacctcggatcaggtagggatacccgctgaacttaagc

>KP1-119E
cggagggatcattattagaagccgaaaggctacttaaaaccatcgcgaactcgtccaagttgcttcggcggcgcggcctccctc
acgggggcgccgcagccccgcctctccggaggtgtggggcgcccgccggaggtacgaaactctgtattatagtggcatctct
gagtaaaaaacaaataagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagt
aatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgctagtactctagcgggcatgcctgttcgagc
gtcatttcaaccctcaagccctgcttggtgttggggccctacggctgccgtaggccctgaaaggaagtggcgggctcgctacaa
ctccgagcgtagtaattcattatctcgctagggacgttgcggcgcgctcctgccgttaaagaccatctttaactcaaggttgacctc
ggatcaggtaggaatacccgctgaacttaagcatatca

>KP1-123A
acctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtttatcgtaccttgttgcttcggcgggc
ccgcctcacggccgccgggggggcacctgcccccgggcccgcgcccgccgaagacaccattgaactctgtctgaagattgca
gtctgagcgattaactaaatcagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgat
aagtaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggcatgcctgtc
cgagcgtcattgctgccctcaagcacggcttgtgtgttgggctccgcccccctcccgggggggcgggcccgaaaggcagcggc
gg

>KP1-123B
attgatatgcttaagttcagcgggtatccctacctgatccgaggtcaacctggaaagattgattggggtcgccggcgggcgccgg
ccgggcctacagagcgggtgacaaagccccatacgctcgaggaccggacgcggtgccgccgctgcctttcgggcccgcccc
ccgggaggggggcggagcccaacacacaagccgtgcttgagggcagcaatgacgctcggacaggcatgcccccggaata
ccaggggcgcaatgtgcgttcaaagactcgatgattcactgaattctgcaattcacattacttatcgcatttcgctgcgttcttcatc
gatgccggaaccaagagatccgttgttgaaagttttaactgatttagttaatcgctcagactgcaatcttcagacagagttcaatggt

gtcttcggcgggcgcgggcccgggggcaggtgccccccggcggccgtgaggcgggcccgccgaagcaacaaggtacgat
aaacacgggtgggaggttggacccagagggccctcactcggtaatgatccttccgca

>KP1-123C
gaaggatcattatccatctcaactgagccggcgcggccccatttggtggtaacgcgaaagcgtgtacgcgccgtcactcttaccc
ttttttacgagtaccttcgttctccttcggtggggcaacctgccgctggaatcaacaaaaccttttttgcatctagcattacctgttctg
atacaaataatcgttacaactttcaacaatggatctcttggctctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtg
aattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccttggtattccatggggcatgcctgttcgagcgtcatct
acaccctcaagctctgcttggtgttgggcgtctgtcccgcctctgcgcgtggactcgccccaaattcattggcagcggtctttgcct
cctctcgcgcagcacaattgcgtttcttgggggggtgggtcgcatccacgaagcaacattaccgtctttgacctcggatcaggta
gggatacccgctgaacttaagcatatca

>KP1-131AA
acctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtctatcgtaccttgttgcttcggcgggc
ccgccstttcgacggccgccggggaggccttgcgccccgggcccgcgcccgccgaagaccccaacatgaacgctgttctg
aaagtatgcagtctgagttgattatcgtaatcagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagc
gaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggg
catgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggcccccgtccccctctcccgggggacgggcccga
aaggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacctgctctgtaggcccggccggcgccagccga
cacccaactttatttttctaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatca

>KP1-131B
tgcggaaggatcattaataaaaggatgcctccgggcataccccatccgtgtctatatactcttgttgctttggcaggccgtggtcttc
cactgtgggctctgcctgcatgtgcctgccagaggaccaaactctgaattttagtaatgtctgagtactatataatagttaaaactttc
aacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatc
gaatctttgaacgcacattgcacccggtggtattccgccgggtatgcctgttcgagcgtctgtagaacaacaaattaccaggtctt
gggtttcgacccaggcttgattctggggttgcggcatcgtctgcagccctaaagtaatgtggcggcaccgataggttctaagcgt
agtaatttctcctcgctacagagtctttcggcgcattgggtactcactcccggccataaaaccccccaattttagtttgacctcggatc
aagtagggataccgctgaacttaagcatatc

>KP1-131BB
gatatgcttaagttcagcgggtatccctacctgatccgaggtcaaacttagaatgtgggggtatatgacaggataccgcgagaca
cctatagcgaggaatatttactacgctcagggcctcgtggcaccgccactaattttaaggcccgccgagttaccggcgaagccc
aagaccaagcagttatgcttgagggttgtaatgacgctcgaacaggcatgccccccggaataccaaggggcgcaatgtgcgtt
caaagattcgatgattcactgaattctgcaattcacattacttatcgcatttcgctgcgttcttcatcgatgccagaaccaagagatcc
gttgttgaaagttttaactattatatagtactcagacgacaccaataatcagggtttaagatcctctggcggacgcgtaccagccga
aaccgatagccttgcggcggtccgccaaagcaacaatagta

>KP1-131C
tgaaaatagtttttgggttggtcggcgaagcgccggccgggcctacagagcgggtgacaaagccccatacgctcgaggaccg
gacgcggtactgccattgcctttcgggcccgtccctagggacgaggacccaacacacaagccgggcttgagggcagcaatga
cgctcggacaggcatgcccccccggaataccaggggggcgcaatgtgcgttcaaagactcgatgattcactgaattctgcaattca
cattatttatcgcatttcgctgcgttcttcatcgataccggaaccaagagatccgttgttgaaagttttaactaattaagttatgttctca
gactgcattataaaacagagttcaaatggcgtcctcggcggcacggatgccgccgaagcaacaagatgtaatagacacgggtg
ggaggttagctctgacgagcgtagactcagtaatgatccttccgcaggt

>KP1-131DA
aacctgcggaaggatcattactgagtctacgctcgtcagagctaacctcccacccgtgtctattacatcttgttgcttcggcggcat
ccgtgccgccgaggacgccatttgaactctgttttataatgcagtctgagaacataacttaattagttaaaactttcaacaacggatc
tcttggttccggtatcgatgaagaacgcagcgaaatgcgataaataatgtgaattgcagaattcagtgaatcatcgagtctttgaac
gcacattgcgccccctggtattccgggggggcatgcctgtccgagcgtcattgctgccctcaagcccggcttgtgtgttgggtcct
cgtccctagggacgggcccgaaaggcaatggcagtaccgcgtccggtcctcgagcgtatggggctttgtcacccgctctgtag
gcccggccggcgcttcgccgaccaacccaaaaactatttttttcaggttgacctcggatcaggtagggatacccgctgaacttaag
catatc

>KP1-131DB
ataacttgttgcctcggcattggttggcttcgaatgaagtcccttatacccttctgagtgtaaggagcagaccggccgacggcccc
tataaactcttgtttttgtaatatcatctgagtaaaacaactaaaatgaatcaaaactttcaacaacggatctcttggttctggcatcgat
gaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccggtg
gtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaaaatcttgtattattggtgttggaggaatacctgtaacagggt
accctctgaaatttagtggcgggctcgctagaattttgagcgtagtaattatacctcgtttttaaagactagtgggacttcttgccgta
aaaccccccaactttctgaaatttgacctcggatcaggtaggaatacccgctgaacttaagcatatca

>KP1-131F
tgcggaaggatcattaaccattggataccttcgggtatatcccatccgtgtctacatactcttgttgctttggcaggccgtggtcacc
cactgtgggctatgcctgcatgcgcctgccagaggaccaaactctgaatattagtgatgtctgagtactatataatagttaaaacttt
caacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcat
cgaatctttgaacgcacattgcacccggtggtattccgccgggtatgcctgttcgagcgtcattataaccactcaagcctgtcttgg
tgttggggttgcgaatcttttgcagccctcgagtctcgtagcgccacctgtgggttctaagcgtagtaatttctcctcgctatagaac
ctgctcggggaaaagtataattcgtagcctggttctatgtgcccgctataaaaccccccaatttttaaaggttgacctcggatcaagta
gggatacccgctgaacttaagcatat

>KP1-131H
ctccgcttattgatatgcttaagttcagcgggtattcctacctgatccgaggtcaaccattaaaaaagtgccccccgagagggtgc
ggtttatggctgttgtctgtacggcttgcagaagcgagataaaaaattactacgctcagagcacgaacagactccgccactggttt
tgaggagctgcgtattaggcagtctcccaacactaagctaggcttaagggttgaaatgacgctcgaacaggcatgcccactaga
atactaatgggcgcaatgtgcgttcaaagattcgatgattcactgaattctgcaattcacattacttatcgcatttcgctgcgttcttcat
cgatgccagaaccaagagatccgttgttgaaagttttaacttatttcttagttttgattcagatttgacaaaaattaacaagagtttagta
gtccaccggtggcagcgctgtttccagcaccgaccaccgaggcaacagtg

>KP1-131I
cggagggatcattacacaatatgaaagcgggttgggacctcatctcggtgggggctccagcttgtctgaattattcacccatgtctt
ttgcgcacttcttgtttcctgggcgggttcgcccgccaccaggacccaaccataaaccttttttttgtaattgcaatcagcgtcagtaa
acaatgtaattattacaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgatacgtagtgtgaa
ttgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtattccaaagggcatgcctgttcgagcgtcatttgta
ccctcaagctttgcttggtgttgggcgtctttgtctctcacgagactcgccttaaaatgattggcagccggcctactggtttcggagc
gcagcacaatcttgcacttctgatcagccatggttgagcatccatcaagaccacatttttctcacttttgacctcggatcaggtaggg
atacccgctgaacttaagcatat

>KP1-131J
acctgcggagggatcattaccgagtttacaactcccaaacccaatgtgaacatacctacatgttgcctcggcggaccgccccgg
cgccctcgggcaccggaaccaggcgcccgccgcaggacccaaacctctgtttttgttaagatttctcctctgagtggattttacaa
ataaatcaaaactttcaacaacggatctcttggctctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcag
aattcagtgaatcatcgaatctttgaacgcacattgcgcccgccagcattctggcgggcatgcctgtccgagcgtcatttcaaccc
tcaggctcccgcgcctggtgctggggatcggccttcaccggccggccccgaaatacagtggcggccccgcccgtgtacctct
gcgtagtagcatacacctcgcagctggaagcggcggcggccacgccggaaaaccccccgacttctgaaagttgacctcggatc
aggtaggaatacccgctgaacttaagcatat

>KP1-131K
cttccgtaggtgaacctgcggaaggatcattacctagagttgcgggctttgcctgccatctcttacccatgtcttttgagtaccttcgt
ttcctcggtgggttcgcccgccggttggacaacacttaaaccctttgtaattgaaatcagcgtctgaaaaaacttaatagttacaact
ttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatc
atcgaatctttgaacgcacattgcgccccttggtattccatggggcatgcctgttcgagcgtcatttgtaccttcaagctctgcttggt
gttgggtgtttgtctcgcctttgcgcgcagactcgcctcaaaacgattggcagccggcgtgttgacttcggagcgcagtacatctc
gcgctttgcactcataacgacgacgtccaaaaagtacatttttacactcttgacctcggatcaggtagggatacccgctgaacttaa
gcatatcaataagcggag

>KP1-131L
tgcggaaggatcattactgagtgagggccctctgggtccaacctcccacccgtgtttattgtaccttgttgcttcggtgcgcccgcc
tcacggccgccgggggggcttctgccccccgggtccgcgcgcaccggagacaccattgaactctgtctgaagattgcagtctgag
cataaactaaataagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgataactaatg
tgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgcccccctggtattccgggggggcatgcctgtccgagcgtc
attgctgccctcaagcacggcttgtgtgttgggctccgtccccccgggggacgggcccgaaaggcagcggcggcaccgagtcc
ggtcctcgagcgtatggggctttgtcacccgctctgtaggcccggccggcgccagccgacaaccaatcatccttttttcaggttg
acctcggatcaggtagggatacccgctgaacttaagcatat

>KP1-131M
acctgcggaaggatcattacagagtttatgctcgtcagagctaacctcccacccgtgtctattacatcttgttgcttcggcggcgca
ttcgtgtgccgccggggacaccatttgaactctgttttataatgcagtctgagaatataacttaattagttaaaactttcaacaacgga
tctcttggttccggtatcgatgaagaacgcagcgaaatgcgataaataatgtgaattgcagaattcagtgaatcatcgagtctttga
acgcacattgcgccccctggtattccgggggggcatgcctgtccgagcgtcattgctgccctcaagcccggcttgtgtgttgggtc
ctagtccttagggacaggcccgaaaggcaatggcagtaccgcgtccggtcctcgagcgtatggggctttgtcacccgctctgta
ggcccggccggcgctccgccgaccaaccaaaaaactatttttcaggttgacctcggatcaggtagggatacccgctgaacttaa
gcat

>KP1-131N
gatatgcttaagttcagcgggtatccctacctgatccgaggtcaaagacggtaatgttgcttcgtggacgcgggccgcgcacctc
gagaagcgcaatgtgctgcgcgagaggaggcaaggaccgctgccaatggatttggggcgagtccgcgcgcagaggcggga
cagacgcccaacaccaagcagagcttgagggtgtagatgacgctcgaacaggcatgccccatggaataccaaggggcgcaa
tgtgcgttcaaagattcgatgattcactgaattctgcaattcacactacttatcgcatttcgctgcgttcttcatcgatgccagagcca
agagatccattgttgaaagttgtaacgattgtttgtatcagaacaggtaatgctagatgcaaaaaaaggttttgataggttccaacgg
caggttgccccgccgaaggagaacgaaaggtgctcgtaaaaaaaggattcagacatgcggcgcgtgaaagtgttacctctacc
gcccgacggtagctgttgctcccgccgagggccgcgaccgcacctcatggaatagataatgatccTTCCGC

>KP1-131Q
tgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtctatcgtaccttgttgcttcggcgggcccg
ccgtttcgacggccgccggggaggccttgcgcccccgggcccgcgcccgccgaagaccccaacatgaacgctgttctgaaa
gtatgcagtctgagttgattatcgtaatcagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaa
atgcgataagtaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggcat
gcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggcccccgtccccctctcccgggggacgggcccgaaa
ggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacctgctctgtaggcccggccggcgccagccgaca
cccaactttatttttctaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>KP1-131R
cggaaggatcattaaccattggataccttcgggtatatcccatccgtgtctacatactcttgttgctttggcaggccgtggtcaccca
ctgtgggctatgcctgcatgcgcctgccagaggaccaaactctgaatattagtgatgtctgagtactatataatagttaaaactttca
acaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcg
aatctttgaacgcacattgcacccggtggtattccgccgggtatgcctgttcgagcgtcattataaccactcaagcctgtcttggtgt
tggggttgcgaatcttttgcagccctcgagtctcgtagcgccacctgtgggttctaagcgtagtaatttctcctcgctatagaacctg
ctcggggaaaagtataattcgtagcctggttctatgtgcccgctataaaaccccccaatttttaaaggttgacctcggatcaagtagg
gatacccgctgaacttaagcatatcat

>KP1-131S
tgcggagggatcattactgagtttaggtggaatccacccaactccaaccctttgtgaacacatcttgttgcttcgggggcgaccct
gccattcgtggcattccccccggaggtcatcaaaacactgcattcttacgtcggagtaaaaagttaatttaataaaactttcaacaa
cggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatc
tttgaacgcacattgcgcccccctggtattccgggggggcatgcctgttcgagcgtcattacaccactcaagcctcgcttggtattgg
gcgtcgcgagtctctcgcgcgcctcaaagtctccggctaggcagttcgtctcccagcgttgtggcaactatttcgcagtggagttc
gagtcgtcgcggccgttaaatctttcaaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>KP1-131T
ttccgtaggtgaacctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtctatcgtaccttgttg
cttcggcgggcccgccgtttcgacggccgccggggaggccttgcgcccccgggcccgcgcccgccgaagaccccaacatg
aacgctgttctgaaagtatgcagtctgagttgattatcgtaatcagttaaaactttcaacaacggatctcttggttccggcatcgatga
agaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgcccccctggt
attccgggggggcatgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggcccccgtccccctctcccgggg
gacgggcccgaaaggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacctgctctgtaggcccggccg
gcgccagcnancacccaactttatttttctaaggttgacctnnnatcaggtagggatacccgctg

>KP1-131V
tatgcttaagttcagcgggtatccctacctgatccgaggtcaaacttagaatgtgggggtggatgacaggataccacgagacacc
tatagcgagaagatttactacgctcagggcctcgtggcaccgccactgattttaaggcccgccgggtaagcggcgaagcccaa
gaccaagcagttatgcttgagggttgtaatgacgctcgaacaggcatgccccccggaataccaaggggcgcaatgtgcgttca
aagattcgatgattcactgaattctgcaattcacattacttatcgcatttcgctgcgttcttcatcgatgccagaaccaagagatccgtt
gttgaaagttttaactattatatagtactcagacgacactaatattcagggtttgagatcctctggcgaacgcgtaccagtcgaaacc
gatagccttgcggcggttcgccaaagcaacaatggtaacaataacatagg

>KP1-131W
tgcggagggatcattgctggaacgcgccccaggcgcacccagaaaccctttgtgaacttataccttactgttgcctcggcgcatg
ccggcccccccgggggcccctcccccggaggagcaggcacgccggcggccagcccaactcttgtttttacactgaaactctg
agaataaaacataaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagt
aatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagc
gtcatttcaaccctcaagcctggcttggtgatggggcactgcttcctacccagggagcaggccctgaaattcagtggcaagctcg
ccaggaccccgagcgcagtagttaaaccctcgctctggaaggccctggcggtgccctgccgttaaaccccccaacttctgaaaat
ttgacctcggatcaggtaggaatacccgctgaacttaagcatat

>KP1-131Y
ttccgtaggtgaacctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtctatcgtaccttgttg
cttcggcgggcccgccgtttcgacggccgccggggaggccttgcgcccccgggcccgcgcccgccgaagaccccaacatg
aacgctgttctgaaagtatgcagtctgagttgattatcgtaatcagttaaaactttcaacaacggatctcttggttccggcatcgatga
agaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgcccccctggt
attccgggggggcatgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggcccccgtccccctctcccgggg
gacgggcccgaaaggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacctgctctgtnnnnccggcc
ggcgccagccgacacccaactttatttttctaaggttgacctcggatcnnntagggatacccgctgaacttaagcatatcaataag
cgg

>KP1-131Z
tgcggaaggatcattagcgaagttcggaatgcgtgttcggtactgacgctgcccggcaacgggatgtgctcgtctggatgcgtgt
cccttctctattccaccccactgtgaaccaagcgtgcgagccgaagagagatcggaagctcgtatgcaaccctcaatataccccca
tcatgtatcagaatgtaccttgcgttaactcgcacaaatacaactttcaacaacggatctcttggctctcgcatcgatgaagaacgc
agcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaccttgcgcccccttggcattccgag
gggcacgcctgtttgagtgtcgtgaactcctccaccctctacctttttcggaaggcactgggctgggatttgggagcttgcgggtc
cctggccgatccgctctccttgaatacattagcgaagcccttgcggccttggtgtgatagtcatctacgcctcggcttagcgaacat
acggggacttgcttctaaccgtctcgcgagagacaactactaccaacttgacctcaaatcaggcgggactacccgctgaacttaa
gcatatc

>KP1-135B
aacctgcggagggatcattaccgagtttacaactcccaaacccaatgtgaacgttaccaaactgttgcctcggcggggtcacgc
cccgggtgcgtcgcagccccggaaccaggcgcccgccggaggaaccaaccaaactctttctgtagtcccctcgcggacgttat
ttcttacagctctgagcaaaaattcaaaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcg
aaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccgccagtattctggcgggc
atgcctgtccgagcgtcatttcaaccctcgaacccctccgggggatcggcgttggggatcgggacccctcacacgggtgccgg
cccctaaatacagtggcggtctcgccgcagcctctcctgcgcagtagtttgcacaactcgcaccgggagcgcggcgcgtccac
gtccgtaaaacacccaactttctgaaatgttgacctcggatcaggtaggaatacccgctgaacttaagcatatcaa

>KP1-135C
gcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtttaacgaaccttgttgcttcggcgggcccg
cctcacggccgccgggggcatccgcccccgggcccgcgcccgccgaagacacctgtgaacactgtctgaagttgcagtct
gagaaactagctaaattagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgataaat
aatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccctctggtattccggagggcatgcctgtccgag
cgtcattgctgccctcaagcacggcttgtgtgttgggcccccgtccccccccaccgggggggacgggcccgaaaggcagcggc
ggcaccgcgtccggtcctcgagcgtatggggctctgtcacccgctc

>KP1-135D
gatatgcttaagttcagcgggtatccctacctgatccgaggtcaaaagtgagaaaaatgtggtcttgatggatgctcaaccatggc
tgatcagaagtgcaagattgtgctgcgctccgaaaccagtaggccggctgccaatcattttaaggcgagtctcgtgagagacaa
agacgcccaacaccaagcaaagcttgagggtacaaatgacgctcgaacaggcatgccctttggaataccaaagggcgcaatg
tgcgttcaaagattcgatgattcactgaattctgcaattcacactacgtatcgcatttcgctgcgttcttcatcgatgccagaaccaag
agatccgttgttgaaagttgtaataattacattgtttactgacgctgattgcaattacaaaaaaaggtttatggttgggtcctggtggc
gggcgaacccgcccaggaaacaagaagtgcgcaaaagacatgggtgaataattcagacaagctggagcccccaccgagatg
aggtcccaacccgctttcatattgtgtaatgatccctccgca

>KP1-135F
tgcggagggatcattacagagttatccaactcccaaacccatgtgaacttatctctttgttgcctcggcgcaagctacccgggacc
tcgcgccccgggcggcccgccggcggacaaaccaaaactcttgttatcttcgttgattatctgagtgtcttatttaataagtcaaaa
ctttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaat
catcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccctaagcacagctta
ttgttgggaacctacggcttcgtagttcctcaaagacattggcggagtggcagtggtcctctgagcgtagtaatcttttatctcgctn
tctgttaggtgctgccccccggccgtaaaaccccccaatttttctggttgacctcggatcaggtaggaatacccgctgaacttaag
catatc

>KP1-143A
tttaaaaaaaaaaataaaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaattgcgatacgtaatat
gacttgcagacgtgaatcattgaatttttgaacgcacattgcgccttaaggtattcctcaaggcatgcgtggatgagcgatatttact
ctcaaaccacttggtttggtcttggcccccctttttttcatagggcctaaatatcaaatggtctctagaataagtttttagaaactcaaa
accttaactc

>KP1-143D
ggaaggatcattaacgagttttgaaacgagttgtagctggccttccgaggcatgtgcacgctctgctcatccactctacccctgtg
cacttactgtaggttggcgtgggctccttaacgggagcattctgccggcctatgtatactacaaacactttaaagtatcagaatgta
aacgcgtctaacgcatctataatacaacttttagcaacggatctcttggctctcgcatcgatgaagaacgcagcgaaatgcgataa
gtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaccttgcgctccttggtattccgaggagcatgcctgtttgag
tgtcatggaattctcaacttataaatccttgtgatctataagcttggacttggaggcttgctggcccttgttggtcggctcctcttgaat
gcattagctcgattccgtacggatcggctctcagtgtgataattgtctacgctgtgaccgtgaagtgttttggcgagcttctaaccgt
ccattaggacaacttttaacatctgacctcaaatcaggtaggactacccgctgaacttaagcatatc

>KP1-175G
ttccgtaggtgaacctgcggaaggatcattactgagtgcgggcccctcgcgggtccaacctcccacccgtgtctcttgaataccc
tgttgctttggcgggcccaccgggccacccccggtcgccgggggggcactgcgcccccgggcccgcgcccgccagagcgcc
tctgaaccctaatgaagaaggactgtctgagtctacgatataattatcaaaactttcaacaatggatctcttggttccggcatcgatg
aagaacgcagcgaaatgcgataagtaatgtgaattgcagaattccgtgaatcatcgaatctttgaacgcacattgcgcccccctgg
cattccgggggggcatgcctgtccgagcgtcatttctgccctcaagcccggcttgtgtgttgggcgtggtccccccggtgtcgggg
ggacctgccccaaaggcagcggcgacgttccgcctaggtcctcgagcgtatggggctttgtcacccgctcgggaggggcccta
cgggcgttggccacccaccaatttttttttacggttgacctcggatcaggtaggagttacccgctgaacttaagcatatcaataagcg
gag

>KP1-175L
tgcggaaggatcattactgagtgagggcccctcggggtccaacctcccacccgtgtttaacgaaccttgttgctttggcgggccc
gcctcacggccgccgggggggcatctgccccccgggcccgcgcccgccgaagccacctgtgaactctgtctgaagtatgcagtc
tgagacaattattaaattaattaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgataact
aatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccctctggtattccggagggcatgcctgtccgag
cgtcattgctgccctccagcccggctggtgtgttgggccccgcccccctttcccgggggggcgggcccgaaaggcagcggcg
gcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctcttgtaggcccggccggcgccagccgacccccctcaatcta
ttttttcaggttgacctcggatcaggtagggatacccgctgaacttaagcatat

>KP1-175M
tgcggaaggatcattactgagtgagggcccctcggggtccaacctcccacccgtgtttaacgaaccttgttgctttggcgggccc
gcctcacggccgccgggggggcatctgccccccgggcccgcgcccgccgaagccacctgtgaactctgtctgaagtatgcagtc
tgagacaattattaaattaattaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgataact
aatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccctctggtattccggagggcatgcctgtccgag
cgtcattgctgccctccagcccggctggtgtgttgggccccgcccccctttcccgggggggcgggcccgaaaggcagcggcg
gcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctcttgtaggcccggccggcgccagccgacccccctcaatcta
ttttttcaggttgacctcggatcaggtagggatacccgctgaacttaagcatat

>KP2-001C
tgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtctatcgtaccttgttgcttcggcgggcccg
ccgtttcgacggccgccggggaggccttgcgccccgggcccgcgcccgccgaagaccccaacatgaacgctgttctgaaa
gtatgcagtctgagttgattatcgtaatcagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaa
atgcgataagtaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggcat
gcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggcccccgtccccctctcccgggggacgggcccgaaa
ggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacctgctctgtaggcccggccggcgccagccgaca
cccaactttatttttctaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>KP2-001F
acctgcggaaggatcattactgagtgagggcccctcggggtccaacctcccacccgtgtttaacgaaccttgttgctttggcgggg
cccgcctcacggccgccgggggggcatctgccccccgggcccgcgcccgccgaagccacctgtgaactctgtctgaagtatgca
gtctgagacaattattaaattaattaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgata
actaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccctctggtattccggagggcatgcctgtccg
agcgtcattgctgccctccagcccggctggtgtgttgggccccgcccccctttcccgggggggcgggcccgaaaggcagcgg
cggcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctcttgtaggcccggccggcgccagccgacccccctcaat
ctatttttcaggttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>KP2-005A
cgttcttcatcgatgcgagagccaagagatccgttgcnnaaagttgtatatagttgcgttatacgcaggttacattcttaaannnaa
gcgtttgtggtaaaacataggaaggcttccaagcgagctgttttaaggctcgcccttctgccgacctacaacaagtgcacagagg
ttgaagagtggatgagccaggtgtgcacatgccccgagaggccagctacaacccgttcagttactcgttaatgatccttccgtag
gtgaacctgcggaaggatcattaacgagtaactgaacgggttgtagctggcctctcggggcatgtgcacacctggctcatccact
cttcaacctctgtgcacttgttgtaggtcggcagaagggcgagccttaaaacagctcgcttggaagccttcctatgttttaccacaa
acgcttcagtttaagaatgtaacctgcgtataacgcaactatatacaactttcagcaacggatctcttggctctcgcatcgatgaaga
acgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaccttgcgctccctggtattcc
ggggagcatgcctgtttgagtgtcatggtatcctcatccttcataacttttgttatcgaaggcatggacttggaggtcgtgctggttc

ctcgttgaatcggctcctcttaaatgtattagcgtgagtgtaacggatcgcttcggtgtgataattatctgcgccgtggtcgtgaagt
aacataagcttgcgcttctaaccgtccttaagctggacaacataactttgacatctgacctcaaatcaggtaggactacccgctgaa
cttaagcatatcaa

>KP2-009A
ttccgtaggtgaacctgcggaaggatcattaacgagttttgaaacgagttgtagctggccttccgaggcatgtgcacgctctgctc
atccactctacccctgtgcacttactgtaggttggcgtgggctccttaacgggagcattctgccggcctatgtatactacaaacactt
taaagtatcagaatgtaaacgcgtctaacgcatctataatacaacttttagcaacggatctcttggctctcgcatcgatgaagaacg
cagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaccttgcgctccttggtattccgag
gagcatgcctgtttgagtgtcatggaattctcaacttataaatccttgtgatctataagcttggacttggaggcttgctggcccttgcg
gtcggctcctcttgaatgcattagctcgattccgtacggatcggctctcagtgtgataattgtctacgctgtgaccgtgaagtgttttg
gcgagcttctaaccgtccattaggacaacttttaacatctgacctcaaatcaggtaggactacccgctgaacttaagcatatcaat
aagc

>KP2-009B
acctgcggaaggatcattaacgagttttgaaatgagttgtagctggccttccgaggcatgtgcacgctctgctcatccactctaccc
ctgtgcacttactgtaggttggcgtgggctccttaacgggagcattctgccggcctatgtatactacaaacactttaaagtatcaga
atgtaaacgcgtctaacgcatctataatacaacttttagcaacggatctcttggctctcgcatcgatgaagaacgcagcgaaatgc
gataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaccttgcgctccttggtattccgaggagcatgcctgt
ttgagtgtcatggaattctcaacttataaatccttgtgatctataagcttggacttggaggcttgctggcccttgcggtcggctcctct
cgaatgcattagctcgattccgtacggatcggctctcagtgtgataattgtctacgctgtgaccgtgaagtgttttggcgagcttcta
accgtccattaggacaacttttaacatctgacctcaaatcaggtaggactacccgctgaacttaagcatatcaa

>KP2-013F
cttccgtaggtgaacctgcggaaggatcattacagtattcttttgccagcgcttaactgcgcggcgaaaaaccttacacacagtgt
ctttttgatacagaactcttgctttggtttggcctagagataggttgggccagaggtttaacaaaacacaatttaattatttttattgata
gtcaaattttgaattaatcttcaaaactttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatgcgataagta
atatgaattgcagattttcgtgaatcatcgaatctttgaacgcacattgcgcccctctggtattccagagggcatgcctgtttgagcgt
catttctctctcaaacccccgggtttggtattgagtgatactcttagtcgaactaggcgtttgcttgaaaagtattggcatgggtagta
ctggatagtgctgtcgacctctcaatgtattaggtttatccaactcgttgaatggtgtggcgggatatttctggtattgttggcccggc
cttacaacaaccaaacaagtttgacctcaaatcaagtaggaatacccgctgaacttaagcatatcaataagcggag

>KP2-025D
acctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtctatcgtaccttgttgcttcggcgggc
ccgccgtttcgacggccgccggggaggccttgcgccccgggcccgcgcccgccgaagaccccaacatgaacgctgttctg
aaagtatgcagtctgagttgattatcgtaatcagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagc
gaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggg
catgcctgtccgagcgtcattgctgccctcaagcacggcttgtgtgttgggcccccgtccccctctcccgggggacgggcccga
aaggcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacctgctctgtaggcccggccggcgccagccga
cacccaactttatttttctaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>KP2-033B
ttccgtaggtgaacctgcggaaggatcattaccgagtgagggccctttgggtccaacctcccacccgtgtttatttacctcgttgctt
cggcgggcccgccttaactggccgccgggggctcacgcccccgggcccgcgcccgccgaagacacccccgaactctgcc
tgaagattgtcgtctgagtgaaaatataaattatttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcg

aaatgcgatacgtaatgtgaattgcaaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggca
tgcctgtccgagcgtcattgctgccctcaagcccggcttgtgtgttgggccccgtcctccgattccgggggacgggcccgaaag
gcagcggcggcaccgcgtccggtcctcgagcgtatggggctttgtcacccgctctgtaggcccggccggcgcttgccgatcaa
cccaaatttttatccaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>KP2-033D
gatccttccgtaggtgaacctgcggaaggatcattgnntnnnccgcgctcgtccgcgcccgcggtannnggggcccgcccctt
cggggccggccctgtctgcaccctctgccattgtcgcacctcgcgtttcctcggcgggcccgcccgccaatggggaccccaaa
ccaaacccatttgcagtgcctgcagtaaacgtctcaaaacaatggaaatcaaaactttcaacaacggatctcttggttctggcatcg
atgaagaacgcagcgaaatgcgataagtagtgtgaattgcakaattcmgtgaatcatcgaatctttgaacgcacattgcgcccy
ytggyattcckkrgggcatgcctgttcgagcgtcatctaacccctcaagcaccgcttgatgttgggcgcttgtccccgcccccgc
gcgcggactcgcctcgaagacattggcggcctgtgtattggctacgagcgcagcagacc

>KP2-033E
tcgcatttcgctgcgttcttcatcgatgccagaaccaagagatccgttgttgaaagttttgatttccattgttttgagacgtttactgca
ggcactgcaaatgggtttggtttggggtccccattggcgggcgggcccgccgaggaaacgcgaggtgcgacaatggcagag
ggtgcagacagggccggccccgaagggcgggcccccgataccgcgggcgcggacgagcgcggtgtagacaatgatccttc
cgtaggtgaacctgcggaaggatcattgtctacaccgcgctcgtccgcgcccgcggtatcgggggcccgcccttcggggccg
gccctgtctgcaccctctgccattgtcgcacctcgcgtttcctcggcgggcccgcccgccaatggggaccccaaaccaaaccc
atttgcagtgcctgcagtaaacgtctcaaaacaatggaaatcaaaactttcaacaacggatctcttggttctggcatcgatgaagaa
cgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtattcctt
agggcatgcctgttcgagcgtcatctaacccctcaagcaccgcttgatgttgggcgcttgtccccgcccccgcgcgcggactcg
cctcgaagacattggcggcctgtgtattggctacgagcgcagcagacc

>KP2-033H
ctgcggaaggatcattaccgagtgagggccctctgggtccaacctcccacccgtgtttatcgtaccttgttgcttcggcgggccc
gcctcacggccgccgggggcacctgccccgggcccgcgcccgccgaagacaccattgaactctgtctgaagattgcagtc
tgagcgattaactaaatcagttaaaactttcaacaacggatctcttggttccggcatcgatgaagaacgcagcgaaatgcgataag
taatgtgaattgcagaattcagtgaatcatcgagtctttgaacgcacattgcgccccctggtattccgggggggcatgcctgtccga
gcgtcattgctgccctcaagcacggcttgtgtgttgggctccgcccccctcccgggggggcgggcccgaaaggcagcggcggc
accgcgtccggtcctcgagcgtatggggctttgtcacccgctctgtaggcccggccggcgcccgccggcgaccccaatcaatc
tttccaggttgacctcggatcaggtagggatacccgctgaacttaagcatatc

>JD1-017A
cacccgagagggggggaggcgactataaacaagtcccagcgcaagtcagcactagctggcaacacaatcgaactgcgggga
cgctttaaagcccagtggtaccaacccccccggtgaccggaggggggcgggcatcgcccgaggtcagaacccactgagatgta
caataagcaatccgcagcggtccgatcggaccgttcacagactaagtggttgtggggagcctgctccttaagatatagtcgggc
ccggcgggcagtcgccgggggttttgcgtccgtaggtgaacctgcggaaggatcattacaagtgaggccgccgaacaccggt
caaacggtggggaggcggctaacacccacccgtgtctacatactcttgttgctttggcgggccgtggcctccactgcgggctct
gctctgctcgtgtgtgcccgccagaggaccaaactctgaatgttagtgatgtctgagtactatttaatagttaaaactttcaacaacg
gatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatcttt
gaacgcacattgcgccctgtggtattccgcagggcatgcctgttcgagcgtcatttaaccactcacgcctggcgtggtgttgggg
cacgcggtctcgcggccctcgaaagcagtggcggcgccggtgggctctaagcgtagtacatactcccgctatagagttcc

>JD1-017B
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacgcgccccaggc
gcacccagaaaccctttgtgaacttataccttactgttgcctcggcgctagctggtccttcggggcccctcaccctcgggtgttga
gacggcccgccggcggccaacctaactcttgtttttacactgaaactctgagaataaaacataaatgaatcaaaactttcaacaac
ggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctt
tgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcttggtgatgggg
cactgcttttacccaaaagcaggccctgaaattcagtggcgagctcgccaggaccccgagcgcagtagttaaaccctcgctttg
gaaggccctggcggtgccctgccgttaaaccccccaacttctgaaaatttgacctcggatcaggtaggaatacccgctgaactta
agcatatcaata

>JD1-017C
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattacagagttatcaaaactcccaaacccat
gtgaacatacctcgcgttgcctcggcaggtggcgtctcgccccgtaagaacctaccctgtaggaccttacccggtagacgaccc
tgccgacggcccccgaaactctgttttatagcattaaacttctgaaaacataactaaataagttaaaactttcaacaacggatctctt
ggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgca
cattgcgcccactagtattctggtgggcatgcctgttcgagcgtcatttcaacccttaagcccctgttgcttagcgttaggagcctac
cggaactctctggtagctccccaaagtcagtggcggagccggttcgcactccagacgtagtagcttttacacgtcgcctgtagcg
cgggccggtcccctgccataaaacaccccaatttttataggttgacctcggatcaggtaggaatacccgctgaacttaagcatatc
a

>JD1-017E
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattacagagttgcaaaactcccaaaccatt
gtgaacgtcaccgtatatcgttgcttcggcgggtggttcccactcgggtaccgcatctccccctcggggtaggtagcccgccgg
aggcgtctcaactctcgattttaattggcatctctgagtaacttataaaataagtcaaaactttcaacaacggatctcttggttctggca
tcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccc
gctagtattctggcgggcatgcctgttcgagcgtcatttcaaccatcaagccccaggcttgtgttggggccctgcggccgtccgc
agcccccgaaaagcagtggcgggctcgctgtcacaccgagcgcagtagtatatcttcgctgggacgtgattgcgggttccggc
cgtgaaacaccccaatttatcaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatca

>JD1-017F
aggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaattaaacgcgaatggccgggcctcgtcac
caccccttctgggggtggaatcctgggggcccggtgtccttttgattatctcccatctgtgaactcttccactgttgcttccgtgag
cctgtactctgctctctttgggagtcggagcacctgaaaagccctggttctgccaggggcgcacggcaatggtaccacaactctg
gttttatgttgtatcagtctgagtggttaacacacaaacaagttaaaactttcaacaacggatctcttggttctcgcatcgatgaagaa
cgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcctcctggtattccg
ggaggcatgcctgttcgagcgtcatcaaacacatcctcaagtatttttgcttggttgtggaggaagatagtctgttgattatcccctct
ggaattcctcagcagagaccacagcagcctcaggtgtgataatactacctcgcttgagtgttattgtgggctcccgctggagagc
agcccccagcttttatcatgtttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>JD1-017G
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaacgagttttaagtgggcttgatgctgg
ctttaccaaaagcatgtgctcagcctcgctctaatccactcaacccctgtgcactattcggagtgttgcaagctgagacttgaatag
gtcttggtcgcgtttctgatgcggctgggttctaagacagtcaaagcttgtcctccggtctatatacaaactcttataatgtcttgtaga
atgtgttagcctctcgttaggcgaaacttaaatacaactttcaacaacggatctcttggctctcgcatcgatgaagaacgcagcgaa
atgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcaccttgcgctccttggtattccgaggagcatg

cctgtttgagtgtcatgttaatatcaaccccgtcggtttgtaaatcagaccgttgggtgttggacttggaggttcttgctggctgcaaa
gtcggctcctcttgaatgtattagcttgggcctgttgtgcgtgctcgcggtgtaataagttttattcatcgtttgggtcgtacacttattg
ggcttgcttctaactgtcttcggacaaacataacctttgactctgacctcaaatcaggtaggactacccgctgaacttaagcatatca
at

>SS1-033A
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattacaagaagccgaaaggctacttaaaa
ccatcgcgaacttatccaagttgcttcggcggcgcggttcccctcacagggcgccgcagccccgcctctcaggaggtatgggg
cgcccgccggaggtacgaaactctgtattatagtggtatctctgagtaaaaaacaaataagttaaaactttcaacaacggatctctt
ggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgca
cattgcgcccgctagtactctagcgggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtgttggggccctacgg
ctgccgtaggccctgaaaggaagtggcgggctcgctacaactccgagcgtagtaattcattatctcgctagggaggtgcagcgt
gctcctgccgttaaagaccccatctttaaccaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatca

>SS1-033B
aggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattacaagaagccgaaaggctacttaaaacca
tcgtgaacttatccaagttgcttcggcggcgcggctcccctcacagggcgccgcagcccccgcctctcaggaggtatggggca
cccgccggaggtactaaactctgtattatagtggcatctctgagtaaaaaacaaataagttaaaactttcaacaacggatctcttggt
tctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacat
tgcgcccgctagtactctagcgggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtgttggggccctacggctg
ccgtaggccctgaaaagaagtggcgggctcgctacaactccgagcgtagtaattcattatctcgctagggaggtgcggcgtgct
cctgccgttaaagacccatctttaaccaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatc

>SS1-033C
ttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagttcttacaaactcc
aacccttttgtgaaccataccaactgttgcctcggcgcgagctgcggctgcttggtagctacccggtagtcacctacccgggagct
acccggtagcgtctgcgtacaggcccgccgaaggaccaccaaactctgtttgacagtgtattctgaatgcttcaactaaatagtta
aaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagt
gaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctattcgagcgtcatttcgaccctgaagcccta
gttgcttcgcgtcgggactctactggctaccctgtagttccctaatgacagtggcggagttcaggtgtactctcagcgtagtaatttc
atctcgcttttgcagtagcctggtcgccggccgtaaaaccccctattttctagtggttgacctcggattaggtaggaatacccgctg
aactnaagcatatcaa

>SS1-033E
ttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattatagagtttttttaaactccca
acccatgtgaacttaccattgttgcctcggcagaagctacctggttaccttaccttggaacggcctaccctgtagcgccctaccctg
gaacggcctaccctgtaacggctgccggtggactaccaaactcttgttattttattgtaatctgagcgtcttattttaataagtcaaaa
ctttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaat
catcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccttaagcctagctta
gtgttgggagcctactgcttttgctagcggtagctcctgaaatacaacggcggatctgcgatatcctctgagcgtagtaatttttatct
cgcttttgactggagttgcagcgtctttagccgctaaaccccccaatttttaatggttgacctcggatcaggtaggaatacccgctg
aacttaagcatatca

>SS1-033H
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacagagttcatgcccttcgg
ggtagatctcccacccgtgttattatactattgttgctttggcgggcccgcctcggccaccggctccggctggtgagcgcccgcc
agaggaccccaaactctgaaatttagtgtcgtctgagtactattttaatagttaaaactttcaacaacggatctcttggttctggcatc
gatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccct
tggtattccgggggggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggtattgggcgtccccggcagcggggtgc
cctaaaatcagtggcggtgccgtctggctctaagcgtagtaaatctctcgctctggagacccggcggatgcttgccagcaaccc
ccaatttttaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>SS1-033I
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattatagagttttctaaactcccaacccatgt
gaacttaccattgttgcctcggcagaagctgctcggtgcaccttaccctggaacggcctaccctgtagcgccttaccctggaacg
gcctaccctgtagcggctgccggtggactaccaaactcttgttatttttattgtaatctgagcgtcttattttaataagtcaaaactttcaa
caacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcga
atctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccttaagcctagcttagtgttgg
gagcctactgcttttactagctgtagctcctgaaatacaacggcggatctgcgatatcctctgagcgtagtaaatttttatctcgctttt
gactggagttgcagcgtctttagccgctaaatcccccaatttttaatggttgacctcggatcaggtaggaatacccgctgaacttaa
gcatatcaat

>SS1-033L
gaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattactgagtgagggcccccgggctcgacct
cccacccttttgtgaacacaacttgttgcttcgggggcgaccctgccgtttacgcggcgagcgcccccggaggccttcaaacact
gcatcattgcgtcggagtttaagcaaattaaacaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcga
aatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggtattccggagggcat
gcctgttcgagcgtcatttcaccactcaagcctagcttggtattgggcgccgcggtgttccgcgcgcctcaaagtctccggctga
gctgtccgtctccaagcgttgtgatttcattaatcgctctggagcgcgggcggccgcggccgttaaatctttcacaaggttgacctc
ggatcaggtagggatacccgctgaacttaagcatatca

>SS1-033M
gaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaagagttagggtctctcaggcccgatctc
ccaacccattgtttaccgaaccactgttgcttcggcgggcccgtcttcacgaccgccggaggaccgtcaaaccgtcctctggtcc
gcgcccgccgatggccccaacttaacaaattcttaaatgaatcgtgtcattgtctgagtactatttgaaaaaataaattaaaactttca
acaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgcgaattgcagaattccagtgagtcat
cgaatctttgaacgcacattgcgccctttggtattccgaagggcatgcctgttcgagcgtcattatcacccctcaagccctctcttgg
gtttggtgttggacggcctggtcaggcaactgacccctcctaaagacaatgacggcgtcctcgtgggatccccggtacactgag
cttcttaactgagcacgtatcggatcaagggtctacgggacacggtctgaaaacctatatattttctcaaggttgacctcggatcag
gtagggatacccgctgaacttaagcatat

>SS1-033N
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaattaatcgcgaaaggggccgggttgc
cctacacccttctgggggtgatatggtgggcaactggccccgcttttgattatatcccatctgtgaacttttccactgttgcttccgt
gagcctgtactttgctctctttgagagttggagcacctgaaaagccctggttctaccaggggcgcacggcaatggtaccacaact
ctggttttatgttgtataagtctgagtggttaacacacaaacaagttaaaactttcaacaacggatctcttggttctcgcatcgatgaa
gaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcctcctggtatt
ccgggaggcatgcctgttcgagcgtcatcaaacacatcctcaagcatttttgcttggtcctggaggaagatggtctcttgactatcc

cctctggaattcctcagcagagaccacagcagcctcaggtgtgataattctacctcgcttgagtgttactgtgggctcccgctgga
gaacagccccccagcttttaccacgtttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>SS1-033P
gaactagtaattctctgaaaggtcgccggtacccggtccccctttaaacgaggggccggggaaggtcctctcacacccttgtgta
ccttaccaccgttgctttggcgggccgacccggtttcgacccgggcggccggcgcccccagccctcaccggccaggacgtca
ggctaagcgcccgccagtatacaaaactcaagcgattatttcatgtagtcctgatgaattattcaattaattaaaactttcaacaacg
gatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatcttt
gaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagctctgcttggtattgggcg
acgtccgctgccggacgcgcctcgaagacctcggcgacggcgtcctagcctcgagcgtagtagtaatatctcgctttggagtgc
tgggcgttggccgccggacaatcgacctttggtctattacttccaaggttgacctcggatcaggtagggatacccgctgaacttaa
gcatatcaataagcggag

>SS1-033R
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaattaatcgcgaaaggggccgggttgc
cctacacccttctgggggtgatatggtgggcaactggccccgcttttgattatatcccatctgtgaacttttccactgttgcttccgt
gagcctgtactttgctctctttgagagttggagtacctgaaaagccctggttctaccaggggcgcacggcaatggtaccacaactc
tggttttatgttgtataagtctgagtggtcaacacacaaacaagttaaaactttcaacaacggatctcttggttctcgcatcgatgaag
aacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcctcctggtattc
cgggaggcatgcctgttcgagcgtcatcaaacccatcctcaagcatttttgcttggtcctggaggaagatggtcttctgactatccc
ctctggaattcctcagcagagaccacagcagcctcaggtgtgataagtctacctcgcttgagtgttactgtgggctcccgctgga
gaacagccccccagcttttaccatgtttgacctcggatcaggtagggatacccgctgaacttaagcatatcaata

>SS1-033S
gtcatttagaggaagtaaaagtcgtaacaaggnnnccgttggtgaaccagcggagggatcattgctggaacaaacgccctcac
gggtgctacccagaaaccctttgtgaatactacctaaaaaatgttgcctcggcgcaggttggccctttacagggtcccttctcccc
taggggtttaggagcagaccggccggcggccctataaactcttgtttttgtaaaatcatctgagtaaaaaaaataaatgaatcaaa
actttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtga
atcatcgaatctttgaacgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagccta
gctttggtgttggaggaagacctgtaaaaaggtcccctcccaaatttagtggcgggctcgctagaattttgagcgtagtaattttac
ctcgtttttaaagactagtgggacttcttgccgtaaaaccccccaactttctgaaaattgacctcgaatcaggtaggaatacccgct
gaacttaagcatatcaataagcggag

>RAS1-015A
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattatagagttttctaaactcccaacccatgt
gaacttaccattgttgcctcggcagaagctacctggttaccctaccttggaacggcctaccctgtagcgccttaccctggaacggc
ctaccctgtaacggctgccggtggactaccaaactcttgttatttttattgtaatctgagcgtcttattttaataagtcaaaactttcaaca
acggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaat
ctttgaacgcacattgcgcccattagtattctagtgggcatgcctgttcgagcgtcatttcaacccttaagcctagcttagtgttggg
agcctaccgcttttgctagcggtagctcctgaaatacaacggcggatctgcgatatcctctgagcgtagtaatttttatctcgcttttg
actggagttgcagcgtctttagccgctaaaccccccaaatttttaatggttgacctcggatcaggtaggaatacccgctgaactta
agcatatca

>RAS1-015B
gttggtgaaccagcggagggatcattnctggaacgcgccccaggcgcacccagaaaccctttgtaaacttataccttactgttgn
nncggcgctagctggtcccncnnngcccctcaccnncgggtgttgagacagcccgtcggcggccaacctaactcttgttttta
cactgaaactctgagcacaaaacataaatgaatcaaaactttcaacaacggatctcttggttctggcatcnatgaaraacgcagcg
aaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggtattccggagggca
tgcctgttcgagcgtcatttcaaccctcaagcctggcttggtgatggggcactgcttcttacccaagaagcaggccctgaaattca
gtgg

>RAS1-015H
gaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattacagagactctgccctttgggtagacctc
ccaccctgtgtcgttataccttgttgctttggcgggccgcggggctccggccctgcccctggctccggctagggcgcgcccgc
cagaggactccaaaacctgaatgttagtgtcgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatc
gatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccct
tggtattccgggggggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggccttgggcgtcaccggtccccggtgtg
ccttaaaatcagtggcggcgccgtctggctctaagcgtagtaaacactctcgctacagacgcccggcggatgctggccagcaa
cccccaatctatcaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaa

>RAS1-015K
tcttggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaagaatatcaggcctca
cggccctattctcacccttgtctaccgaactttgttgctttggcaggcccctggtcactagctgactcgcgcagctaccagccatg
cctgccaaaggacccaaactctgaaatttttgctgtctgagtactatgtaatagttaaaactttcaacaacggatctcttggttctggc
atcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcc
ctctggtattccgggggggcatgcctgttcgagcgtcattacaaccctcacgcctgacgtggtattgggcagacccttggggcct
gcctcaaagtcagtggcggtaccgtgtgaccccgaagcgtagtactactcgtcgcttctgggcgtcaggcggcagcctgccag
caaacccccactttacaaggttgacctcgaatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>RAS1-015L
cttggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaagaatatcaggcctcac
ggccctattctcacccttgtataccgaactttgttgctttggcaggcccctggtcactagctgactcgcgcagctaccagctatgcc
tgccaaaagacccaaactctgaaatttttgctgtctgagtactatgtaatagttaaaactttcaacaacggatctcttggttctggcatc
gatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctc
tggtattccgggggggcatgcctgttcgagcgtcattacaaccctcacgcctggcgtggtattgggcagacccttggggcctgcc
tcaaagtcagtggcggcaccgtgtgaccccgaagcgtagtactactcgtc

>RAS1-015P
cttggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaagaataacacgcctcc
gggcgctattctcacccttgcgtaccatactttgttgccttggcgccgtgcgccagtggctcgaaacccgtgaatcattgctgtct
gagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatg
tgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccccctggcattccgggggggcatgcttgttcgagcgtc
attacaaccctcaagctccgcttggtgttgggctcgcctttctcggaaggcctgcctcaaagtcagtggcggccgccgtccgacc
ttcagcgcagtaatgctcgtcgctgcaagggaagggtagcaagcgccgtcatacaaccccccacacacaaggttgacctcggat
caagtagggatacccgctgaacttaagcatatcaataagcggag

>RAS1-015Q
gaggaagtaaaagtcgtaacaaggttttgctaagcaacagtcnnnctgtagcagagcttaacggagaccttgcacccgagagg
ggggaggcgactataaacaagcttccagcgcaagtcagcactactggcaacacaatcgaattgcggggacgctttaaagccca
ccggtaccaacagcccttctaagggctgggcgagcatagctcgaggttacaacccagtgggatggtacaataagcaatccgca
gcggccgagccgtccacagactaagtggttgtgggtaggagaaccctatctaagatatagtcgggcctgaggagagatccccg
gggcttctgcgtccgtaggtgaacctgcggaaggatcattaataaggataccgggaaaccggtagacccacccgtgtctatcta
ctcttgttgctttggcaggccgtggcctccaccgtgggctctgcctgcgcgtgcctgccagaggaccaaactctgaattttagtga
tgtctgagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagt
aatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccggtggtattccgccgggcatgcctgttcgag
cgtcattataaccactcaagcctagcttggtattggggttcgcggttccgcggctcctaaaatcggtggcggtgccggtgggctct
aagcgtagtaaatctcctcgctatagggtccctccggtcacttgccagaacccccaatttttacaggttgacctcggat

>RAS1-015R
cattgcgccctgtagsakagcttamcgsagrscwtgcctryycgagmgksrkkatrgcsamtataagccwagctwastmy
yggsgcwmgtcrgymycgctggcmwcwmaawcsartkgcggkgmcgctktaargctctgagcstagtacmwacts
ctcgctatmsagkscwgtsgstwscytagcyagarsycmcatayytkrtcaggwtgacctacaataagc

>RAS1-015U
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaaggagtataagacgaagct
tcggcttctgactccccatcctttgcttaccatcccccatacttcggcgggcgatatagagggaatgaacttctagcgatagaggat
cctccctctgtcccagctcgccggcattacaaaacctcaagctataacagtgctgtctgagcatttgcaagcaaaagaataaaaac
tttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatc
atcgaatctttgaacgcacattgcgccctctggtattccggagggcacacctgttcgagcgctagtacaacaatcaagctcagctt
ggtgttggatctggccctttcctcggaagggacctttccaaaactcatgggcgccgccgtccggctccaagcgtagcaagcgaa
agccctcgctttggatagtcgggtcgatggccgccggtaaacacttggggtaacccaaattactacagtttagcctcggatcagg
tggggatacccgctgaacttaagcatatcaataagcggag

>JD1-047D
cttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacgcgccccag
gcgcacccagaaaccctttgtgaacttataccttactgttgcctcggcgctagctggtccctcggggcccctcaccctcgggtgtt
gagacagcccgtcggcggccaacctaactcttgtttttacactgaaactctgagcacaaaacataaatgaatcaaaactttcaaca
acggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaat
ctttgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcttggtgatgg
ggcactgcttcttacccaagaagcaggccctgaaattcagtggcgagctcgccaggaccccgagcgcagtagttaaaccctcg
ctctggaaggccctggcggtgccctgccgttaaaccccccaacttctgaaaatttgacctcggatcaggtaggaatacccgctgaa
cttaagcatatcaataagcggagga

>JD1-047E
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattattagaagccgaaaggctacttaaaacc
atcgcgaactcgtccaagttgcttcggcggcgcggcctccctcacgggggcgccgcagccccgcctctccggaggtgtgggg
cgcccgccggaggtacgaaactctgtattatagtggcatctctgagtaaaaaacaaataagttaaaactttcaacaacggatctctt
ggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgca
cattgcgcccgctagtactctagcgggcatgcctgttcgagcgtcatttcaaccctcaagccctgcttggtgttggggccctacgg
ctgccgtaggccctgaaaggaagtggcgggctcgctacaactccgagcgtagtaattcattatctcgctagggacgttgcggcg
cgctcctgccgttaaagaccatctttaactcaaggttgacctcggatcaggtaggaatacccgctgaacttaagcatatcaa

>JD1-047F
aggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaattaatcgcgaaatgcgtcgtgccatgtaa
gactcacggcctcctgggccgtggggacgggtcccccgtgtggggataccgccgcatgggccttcgcctcccttcgattatatc
ccatctgtgaacacttccactgttgcttccgtgagcctgctgctgactccttctgggggtgagcgcctgaaaagccctggctctgct
aggggcgcacggcaatgctacaatcaccctggttttatgtatctccagtctgagtggtgaaacacatcaaacaagttaaaactttca
acaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcg
aatctttgaacgcacattgcgcctcctggtattccgggaggcatgcctgttcgagcgtcatgaacaacatcctcaagcattcttgctt
ggttgtggaggaagagggtccttggaccgtcccctctggaatttctcagcagagaccgcaacagcctcaggtgtgataacattac
ctcgcttgagggctgcagtgggctcccgctgcgttgatccccccagcaatactatgtttgacctcggatcaggtagggatacccg
ctgaacttaagcatatcaataagcggag

>JD1-047G
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattaccgagttataacaactccaaccctttgt
gaaccttaccgtcgtttcctcggcgcactgctgcgaccgccccgcaggcggccctcgcggccccgcgggcagctacttaccct
gtacggatctaccctgcagcagctgaaggcgcagcggcgccgaaaggaccgctacaaactcttttatccaagctacgtctgaac
aatttactatgcaaaatagtttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagt
aatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccattagtattctagtgggcatgcctattcgagcg
tcatttcgacccttaagccctgttgcttagcgttgggagtctgcggcttcggccgcagttccttaaagtcagtggcggagctgtggc
acactctaggcgtagtagttttaccgcctcgcctccagagtggccccggctgcctgccgtaaaaccccctaattttctagtggttgac
ctcggattaggtaggaatacccgctgaacttaagcatatcaa

>JD1-047K
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagttaccgctctataac
cctttgtgaacatacctaaccgttgcttcggcgggcaggggaagcctctcgcgggcctcccctcccggcgccggccccccacca
gggggacggggcgcccgccggaggaaaccaaactctatttacacgacgtctcttctgagtggcacaagcaaataattaaaacctt
tttaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcat
cgaatctttgaacgcacattgcgctcgccagcattctggcgagcatgcctgttcgagcgtcatttcaaccctcaagcaccgcttgg
ttttggggccccacggcacacgtgggcccttaaaggtagtggcggaccctcccggagcctcctttgcgtagtaactaacgtctcg
cactgggattcggagggactcttgccgtaaaaccccccaatttttttacaggttgacctcggatcaggtaggaatacccgctgaac
ttaagcatatcaa

>JD1-047L
cttggtcatttagaggaagtaaaagtcgtaacaaggtctnnnttggtgaaccagcggagggatcattgctggaacgcgcnccag
gcgcacccagaaaccctttgtgaacttataccttactgttgnnnnngcgctggctggtcnnnnggggcccctcaccnncgggt
gttgagacggcccgnnngnggccaacctgacnattgtttttacaccgaaactctgagaataaaacataaatgaatcaaaactttc
aacaacggatnnnntggttctggcatcgatgaagaacgcagnnaaatgcgataagtaatgtgaattgcagaattcagtgaatca
nnnaatctttgaacgcacattgancncnctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcnt
ggtgatggggcactgcttttacccaaaagcaggccctgaaatt

>JD1-047O
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacgcgccccaggc
gcacccagaaacnnnttgtgaacttatancttactgttgcctcggcgctagctggtccttcggggcccctcaccctcgggtgttga
gacggcccgccggcggccaacctaactcttgtttttacactgaaactctgagaataaaacataaatgaatcaaaactttcaacaac
ggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctt

tgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcttggtgatgggg
cactgcttttacccaaaagcanncctgaaattcag

>JD1-047P
agaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattacagagttaaacaaactccaaaccctttt
gtgaaccttaccataagttgcctcggcgtgcgctgcggctacccggtttccggccccgcagacgggtggttaccctgtagccgg
tcaacagcccgccgaaggacccctaaactctgtttaaaatggtgtattctgaattacttaaactaaaataagttaaaactttcaacaa
cggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatc
tttgaacgcacattgcgcccattagtactctagtgggcatgcctattcgagcgtcatttcgaccattaagccctgtagcttagcgttg
ggagcctacgccccttgcggcgtagctccttaaagtcagtggcggagttggagcacactctaagcgtagtaaacttatctcgcttct
gtagttgccccggctgcctgccgtaaaaccccctatatttctagtggttgacctcggattaggtaggaatacccgctgaacttaagc
atatcaata

>JD1-047Q
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattactgagttcttacaaactccaaccctttg
tgaaccataccaactgttgcctcggcgcgagctgcggctgcttggtagctacccggtagtcacctacccgggagctacccggta
gcgtctgcgtacaggcccgccgaaggaccaccaaactctgtttgacagtgtattctgaatgcttcaactaaatagttaaaactttca
acaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcg
aatctttgaacgcacattgcgcccattagtattctagtgggcatgcctattcgagcgtcatttcgaccctgaagccctagttgcttcg
cgttgggactctactggctaccctgtagttccctaatgacagtggcggagttcaggtgtactctcagcgtagtaatttcatctcgctt
tgcagtagcctggtcgccggccgtaaaaccccctattttctagtggttgacctcggattaggtaggaatacccgctgaacttaagc
atatca

>JD1-047S
tggtcatttagaggaagtaaaagttgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacgcgccccggcg
cacccagaaaccctttgtgaacttatacccatactgttgcctcggcgcaggccggcctctgctgcagaggccccctggagacag
ggagcagcccgccggtggccaatcaaactctgtttctatagtgaatctctgagtaaaaaaaacataaatgaatcaaaactttcaac
aacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaa
tctttgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcttggtgatg
gggcactgcctgtgaaagggcaggccctgaaatctagtggcgagctcgccaggaccccgagcgtagtagttatatctcgctctg
gaaggccctggcggtgccctgccgttaaaccccccaacttctgaaaatttgacctcggatcaggtaggaatacccgctgaacttaa
gcatatcaataagcggag

>JD1-047W
tggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattaccaaaactttaggtttattcc
cccttcttagggggttgagccataccctttgtctacgtgtacctatcgttgtttcctcgggggggctatttgccccccgataggaatcc
cttaaaccctttgtattatagtattaaacctctgataactacctaaataatcacaactttcaacaatggatctcttggttctggcatcgat
gaagaacgcagcgaaatgcgaaaagtagtgtgaattgcagaattccgtgaatcatcgaatctttgaacgcacattgcgcccctcg
gtattccgtggggcatgcctgttcgagcgtcatttacaccctcaagctctgcttggtgttgggcgtctgtcctgcttcatgcatggac
tcgccccaaagtcattggcagcggtcgtgccagcttctcgcgcagcacatttgcgtttctcgaagctctggtggatcagcgtcca
gtaagctctctttatgacttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>JD1-047X
tagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacaaacgccctcgcgggcg
ctacccagaaaccctttgtgaacttatacctatcactgttgcctcggcatggttggcctcttatgaggtcccttctgtacaggaggag

caggccggccggtggccctataaactcttgttttttttgtaatatcatctgagcaaacaactataaatgaatcaaaactttcaacaacg
gatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatcttt
gaacgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagcctggctttggtgttgga
ggaccacttttcgttaagtggcctctgaaatttagtggcgggctcgctaggattttgagcgtagtaattatacctcgtttttaaagact
agtggttctcttgccgtaaaaccccccaacttttctgaaaattgacctcggattaggtaggaatacccgctgaacttaagcatatcaa
taagcggag

>JD1-047Y
tagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattatcgtggggcttcggccccgtcgagaa
aataccttgccttttatgagaaccatggtgtaggtcttccctgggattgtccgggatcggagcgtccatgtacctttttcctcggccg
gcttgccggccaatggggaacacaaaaacactgtcattgctgtaactgtctgataaaacaaacaaaaatcaaaactttcaacaac
ggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctt
tgaacgcacattgcgccctttggtattccttagggcatgcctgttcgagcgtcatttaaaccttcaagcactgcttggtgttgggtgc
ctgtcccgcctccgcgcgtggactcacctcaaattcattggcggccctcacgtcggctacgagcgcagcagaaacgcgaactc
gtgggcccgacggggaggctcccagaagctacaattcaccattttgacctcggatcaggtagggatacccgctgaacttaagca
tatcaa

>JD1-047Z
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacgcgcccccggc
gcacccagaaaccctttgtgaactcataccttactgttgcctcggcgcaggccggcccccctgggggcccctcaccccggtga
ggagcaggcccgccggcggccaagccaactcttgtttttacactgaagctctgagaaacaaacacaaatgaatcaaaactttca
acaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcg
aatctttgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcttggtga
tggggcgctgcttcctgcccagggagcaggccctgaaattcagtggcgggctcgccaggaccccgagcgcagtagttaaacc
ctcgctctggaaggccctggcggcgccctgccgttaaaccccccaactcctgaaaatttgacctcggatcaggtaggaatacccg
ctgaacttaagcatatcaataagcggag

>JD1-047AA
aggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattactgaaaatgtaataaacccttcaggttttgg
aaggggggagccgtcaaaagcttccctggtacatgcctcacccctttgtatatctaccatgttgctttggcgggccgacccggtttcg
acccgggcggccggcgcccccagcctgcttgccaggccaggacgcccggccaagtgcccgccagtatacaaaactccagc
gattattttgtgtagtcctgagaatttattcaataaattaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagc
gaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctctggcattccggaggg
catgcctgttcgagcgtcatttcaaccctcaagctctgcttggtattgggcgacgtctgctgtcagacgcgcctggaagacctcgg
cgacggcattccagcctcgagcgtagtagtaaaatatctcgctttggaggatggggtgacggcttgccggacaaccgacctctg
gtcatttttttccaaggttgacctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggag

>JD1-047AC
aaggatcattaattaatcgcgaaaggggccgggttgccctacacccctttctgggggtgatatggtgggcaactggccccgctttt
gattatatcccatctgtgaacttttccactgttgcttccgtgagcctgtactttgctctctttgagagttggagtacctgaaaagccctg
gttctaccaggggcgcacggcaatggtaccacaactctggttttatgttgtataagtctgagtggttaacacacaaacaagttaaaa
ctttcaacaacggatctcttggttctcgcatcgatgaagaacgcagcgaaatgcgataagtagtgtgaattgcagaattcagtgaat
catcgaatctttgaacgcacattgcgcctcctggtattccgggaggcatgcctgttcgagcgtcatcaaacacatcctcaagcattt
ttgcttggtcctggaggaagatggtctcttgactatcccctctggaattcctcagcagagaccacagcagcctcaggtgtgataatt
ctacctcgcttgagtgttactgtgggctcccgctggagaacagccccccagcttttaccatgtttgacc

>JD1-047AI
tggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacgcgccccaggc
gcacccagaaaccctttgtgaacttataccttactgttgcctcggcgctagctggtccttcggggcccctcaccctcgggtgttga
gacggcccgccggcggccaacctaactcttgtttttacactgaaactctgagaataaaacataaatgaatcaaaactttcaacaac
ggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctt
tgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttcaaccctcaagcctggcttggtgatgggg
cactgcttttacccaaaagcaggccctgaaattcagtggcgagctcgccaggaccccgagcgcagtagttaaaccctcgctttg
gaaggccctggcggtgccctgccgttaaaccccccaacttctgaaaatttgacctcggatcaggtaggaatacccgctgaactta
agcatatcaata

>JD1-047AJ
ggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacaaacgccctaac
gggtgctacccagaaaccctttgtgaatactacctaaaaaatgttgcctcggcgcaggttggccctttacagggtcccttctcccc
taggggtttaggagcagaccggccggcggccctataaactcttgtttttgtaaaatcatctgagtaaaaaaaataaatgaatcaaa
actttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtga
atcatcgaatctttgaacgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagccta
gctttggtgttggaggaagacctgtaaaaaggtcccctcccaaatttagtggcgggctcgctagaattttgagcgtagtaatttac
ctcgtttttaaagactagtgggacttcttgccgtaaaaccccccaactttctgaaaattgacctcgaatcaggtaggaatacccgct
gaacttaagcatatcaataagcggag

>JD1-047AN
aagcaacagtccttctgtagcagagcttaacggagaccttgcacccgagagnggggggaggcgactataaaagaactcccagc
gcaagtcagcccagctggcaacacaatcgaattgcggggacgctttaaagcccaccagtaccaactgcccccaagggcaggg
ccagctatgctggaggtcacaacctggtaggatattacaataagcaatccgcagcggctcagccgtccacagactaagtggttgt
gggtagggatcccctacttaagatatagtcgggccctgtgacaagtcataggggcttctgcgtccgtaggtgaacctgcggaag
gatcattaaccattggataccttcgggtatatcccatccgtgtctacatactcttgttgctttggcaggccgtggtcacccactgtgg
gctatgcctgcatgtgcctgccagaggaccaaactctgaattttagtgatgtctgagtactatataatagttaaaactttcaacaacg
gatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatcttt
gaacgcacattgcacccggtggtattccgccgggtatgcctgttcgagcgtcattataaccactcaagcctgtcttggtgttgggg
ttgcgaatattttgcagccctcgagtctcgtagcgccacctgtgggttctaagcgtagtaatttctcctcgctataggacctgctcgg
ggaaaagtataattcgtagcctggttctatgtgtccgct

>JD1-047AO
gaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattatcgtagggcttcggccctgtcgagatag
aacccttgccttttatgagtaccgtccgtttcctcggcaggtccgcctgccgatggacaaaaaaaaataaacccttttatcacctgtatc
agtctgaaaaaacaaataaaaattaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgata
agtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtattccttagggcatgcctgttcga
gcgtcatttaaaccttcaagctctgcttggtgttgggtgtctgtcccgcctcggcgcggactcgccttaaattcattggcggccggt
acattggcttcgagcgcagcagaaacgcgaactcggagactttggatcagctcccagaagctattttctatttttgacctcggatca
ggtagggatacccgctgaacttaagcatatca

>JD1-047AP
aggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattatcgagtttacaactcccaaacccttatgtg
aacatacctactgttgcttcggcggactcgccccggcgtccggacggcctcgcgccgcccgcggcccggacccaggcggcc
gccggagacccccaaactctgtattatcagcattttctgaatccgccgcaaggcaaaacaaatgaatcaaaactttcaacaacgg
atctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttg
aacgcacattgcgcccgccagcattctggcgggcatgcctgttcgagcgtcatttcaaccctcgacttccctttggggaaatcgg
cgttggggaacggcagcataccgccggccccgaaatgaagtggcggcccgtccgcggcgacctctgcgtagtaatccaactc
gcaccggaaccccgacgtggccacgccgtaaaacacccaacttctgaacgttgacctcggatcaggtaggaatacccgctgaa
cttaagcatatca

>JD1-047AQ
ttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacaaacgccctca
cgggtgctacccagaaaccctttgtgaatactacctcaaaaatgttgcctcggcgcaggttggccctttacagggtcccttctccc
ctaggggtttaggagcagaccggccggcggccctataaactcttgtttttgtaaaatcatctgagtaaaaaaaataaatgaatcaa
aactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtg
aatcatcgaatctttgaacgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagccta
gctttggtgttggaggaagacctgtaaaaaggtcccctcccaaatttagtggcgggctcgctagaattttgagcgtagtaattttac
ctcgtttttaaagactagtgggacttcttgccgtaaaaccccccaactttctgaaaattgacctcgaatcaggtaggaatacccgct
gaacttaagcatatcaata

>JD1-047AR
cccagaaacccttttgtgaatactacctaaaaatgttgcctcggcgcaggttggccctttacagggtcccttctcccctaggggttta
ggagcagaccggccggcggccctataaactcttgtttttgtaaaatcatctgagtaaaaaaaaaaaaatgaatcaaaactttcaac
aacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaa
tctttgaacgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagcctagctttggtgtt
ggaggaagacctgtaaaaaggtcccctcccaaatttagtggcgggctcgctagaattttgagcgtagtaattttacctcgttt

>JD1-047AX
cttggtcatttagaggaagtaaaagtcgtaacaaggtttccgtaggtgaacctgcggaaggatcattatcgtagggcttcggccct
gtcgagatagaacccttgcctttatgagtaccgtccgtttcctcggcaggtccgcctgccgatggacaaaaaaaataaaccctttat
cacctgtatcagtctgaaaaaacaaataaaaattaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcg
aaatgcgataagtagtgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgccctttggtattccttagggcat
gcctgttcgagcgtcatttaaaccttcaagctctgcttggtgttgggtgtctgtcccgcctcggcgcggactcgccttaaattcattg
gcggccggtacattggcttcgagcgcagcagaaacgcgaactcggagactttggatcagctcccagaagctattttctatttttga
cctcggatcaggtagggatacccgctgaacttaagcatatcaataagcggagga

>JD1-047AZ
ttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatcattgctggaacaaacgccctca
cgggtgctacccagaaaccctttgtgaatactacctaaaaatgttgcctcggcgcaggttggccctttacagggtcccttctcccc
taggggttaaggagcagaccggccggcggccctataaactcttgtttttgtaaaatcatctgagtaaaaaaaaataaatgaatca
aaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagt
gaatcatcgaatctttgaacgcacattgcgcccggtggtattccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagcc
tagctttggtgttggaggaagacccgtaaaaaggtcccctcccaaatttagtggcgggctcgctagaattttgagcgtagtaatttt
acctcgtttttaaagactagtgggacttcttgccgtaaaaccccccaactttctgaaaattgacctcgaatcaggtaggaatacccg
ctgaacttaagcatatcaataagcggag

>JD1-047BB
tagaggaagtaaaagtcgtaacaaggtctccgtaggtgaacctgcggagggatcattagaaaacgagttgcgagactcgaaaa
ccatcgcgaacatcacctcgtcgtcgttgcttcggcgggcggccccgcgaaggggcccgcaaggcccgcgagggcgcccg
ccggaggagcgtgaacactctgttttcttagcgcctctctgagtggatacagcgaaatgaatcaaaactttcaacaacggatctctt
ggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgca
cattgcgcccgccagtattctggcgggcatgcctgttcgagcgtcatttcaacccctcaagccccgcttggtgctggggccctac
gtgcgccgtaggccccgaattacagtggcgggctcgccgcgatcccgagcgtagtaacacacctcgctctggggctccgcgg
cgggtgccggccgtgaaaacgaccaaaacacaaacgcaaggttgacctcggatcaggtaggaatacccgctgaacttaagca
tatcaa

>JD1-047BH
aggaagtaaaagtcgtaacaaggttttgtgaagcaacagtccttctgtagcagagcttaacggagaccttgcacccgagagggg
ggaggcgactataaacaagcttccagcgcaagtcagcactgctggcaacacaatcgaattgcggggacgctttaaagcccacc
ggtaccaacggctccctttaggggccgggcgagcatgctcgaggtcacaatctggtgggatagcacaataagcaatccgcagc
ggcagagccgtccacagactaagtggttgtgggtaggagaaccctatctaagatatagtcgggcccgaggagagatccccgg
ggcttctgcgtccgtaggtgaacctgcggaaggatcattaaaaggatgccggaaacggtagaccccacccgtgtttacatactct
tgttgctttggcaggccgtggcctccaccgtgggttctgcctgcgtgtgcctgccagaggaccaaactctgaatgttagtgatgtct
gagtactatataatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatg
tgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcgcccggtggtattccgccgggcatgcctgttcgagcgtc
attataaccactcaagcctcgcttggtattggggtacgcggttccgcggctcctaaaatcagtggcggtgccggtgggctctaag
cgtagtacatctcctcgctacagggtccttccggttgtctgccagaaccccccattttacaggtg

>JD1-047BJ
cttggtcatttagaggaagtaaaagtcgtaacaaggtctccgttggtgaaccagcggagggatmattcrstggmrmgckcscc
aggcgcacccmgaarcscttwgtgaasttawaccctyrctgttgcswmggcsctrgctggtnntnnnggggcccctcacsct
cgggtgttgagamggcccgccggcggccaacctwactgawrwtttkacactcgrawcncnkagaakaawacaymrmt
gaactyaaramtwtcaayaagcggakctcttggtnnnggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaatt
gcagaattcagtgaatcannnaatctttgaacgcacattgcgccctctggtattccggagggcatgcctgttcgagcgtcatttca
accctcaagcctggcttggtgatggggcnnngcttttacccaa

>JD1-047BK
ccgttggtgaaccagnngagggatcattgcnggaacaaacgccctcacgggtgctacccagaaaccctttgtgaatactaccta
aaaatgttgcctcggcgcaggttggccctttacagggtcccttctcccctaggggtttaggagcagaccggccggcggccctat
aaactcttgtttttgtaaaatcatctgagtaaaaaaaaaaatgaatcaaaactttcaacaacggatctcttggttctggcatcgatgaa
raacgcagcgaaatgcgataagtaatgtgaattgcaraattcagtgaatcatcgaatctttgaacgcacattgcgcccggtggtatt
ccaccgggcatgcctgttcgagcgtcatttcaaccctcaaagcctagctttggtgttggaggaagacctgtaaaaaggtcccctc
ccaaatttagtggcgggctcgctagaattttgagcgtagtaattttacctcgtttttaaagactagtgggacttcttgccgtaaaaccc
cccaactttctgaaaattgacctnnnatcaggtaggaatacccgctgaacttaagcata

>JD1-047BM
cggggcgcttgccgaagnnnatgcaaccctcacggggtccgcgactataaacacaggatgcaagtcgcgatggcggcaaca
taatcgaattgcggggatctcctaaagccaagcagtaccaacaccgtccagaaatgtagggtgggccagcgttaatcccgctgg
aggttagaatctgctaggatgccacaatggatgatccgcagcgaagcccctaacggcgccagcctacggggaacgttcacag
actaagtggttgtgggtaggctcgtcctgcttaagatatagtcgggccctgtaggaaactacaggggctagccgttccgtaggtg

aacctgcggaaggatcattacagagttcatgcccttcggggtagacctcccacccttgtgtatcattatagaatgttgctttggcgg
gccgcgcctcgtgcgcctagattcgcgtctagcgtgcccgccagaggacccctaaactctgaatgttagtgtcgtctgagtactat
taaatagttaaaactttcaacaacggatctcttggttctggcatcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgca
gaattcagtgaatcatcgaatctttgaacgcacattgcgccccctggtattccgggggggcatgcctgttcgagcgtcatttatacca
atctagcctggctaggtgttgggcgccgccgcctggcggaccttaaaactagtggcggtgctctcaagctctacgcgtagtaatc
ttctcgctacagggtcttgggagatgctagccagcaaccccaaatttttctaggttgacctcgga

>JD1-047BN
tgtagcagagcttaacggagaccttgcacccgagagggggggaggcgactataagagaactcccagcgcaagtcagcctagct
ggcaacacaatcgaattgcggggacgctttaaagcccaccggtaccaactgctcctctggggcagggccagttatgctggagg
tcacaacccggtaggatgctacaataagcaatccgcagcggctccgccgtccacagactaagtggttgtgggtagggatcccct
acttaagatatagtcgggccctgtgataagtcatagggggcttctgcgtccgtaggtgaacctgcggaaggatcattaaccattgga
taccttcgggtatatcccatccgtgtctacatactcttgttgctttggcaggccgtggtcacccactgtgggctatgcctgcatgtgc
ctgccagaggaccaaactctgaatattagtgatgtctgagtactatacaatagttaaaactttcaacaacggatctcttggttctggc
atcgatgaagaacgcagcgaaatgcgataagtaatgtgaattgcagaattcagtgaatcatcgaatctttgaacgcacattgcacc
cggtggtattccgccgggtatgcctgttcgagcgtcattataaccactcaagcctgtcttggtgttggggttgcgaatcttttgcagc
cctcgagtctcgtagcgccacctgtgggttctaagcgtagttatttctcctcgctatagaatctgctgagggaaaagtataatnnnt
agcctggttctatgtgtccgctatcaaacccccaatttt

Curriculum Vitae

Candidate's full name: Ryleigh Anne Jean Stack

Universities attended:

University of New Brunswick, Saint John

2015-2019 Bachelor of Science with Honours

2019-2020 Master of Science

Conference Presentations:

Stack, R.A., J. Johnson and C. Gray. 2019. *Taxonomic analysis of a collection of endophytic fungi*. Presentation at the Maritime Natural Products Conference. Saint John, NB.

Debly, J.M., Lahey, M.E.B., Smit, S.E., **Stack, R.A.**¸Gray, C.A., Johnson, J.A. 2019. *Isolation of endophytes from medicinal plants*. Poster presentation at the Maritime Natural Products Conference. Saint John, NB.