Ancient DNA in *Lycoptera* Fossils: Key Discoveries for Evolution and Methodological Reassessment

Wan-Qian Zhao[1,11*], Zhan-Yong Guo[2], Zeng-Yuan Tian[1], Tong-Fu Su[3], Gang-Qiang Cao[1], Zi-Xin Qi[4], Tian-Cang Qin[5], Wei Zhou[6], Jin-Yu Yang[7], Ming-Jie Chen[8], Xin-Ge Zhang[9], Chun-Yan Zhou[10,15], Chuan-Jia Zhu[11], Meng-Fei Tang[1], Di Wu[1], Mei-Rong Song[3], Yu-Qi Guo[11*], Li-You Qiu[12*], Fei Liang[13], Mei-Jun Li[14*], Jun-Hui Geng[1], Li-Juan Zhao[11], Shu-Jie Zhang[11]

Affiliations:

[01] School of Agricultural Sciences, Zhengzhou University, Zhengzhou, China

[02] College of Agronomy, Henan Agricultural University, Zhengzhou, China

[03] College of Sciences, Henan Agricultural University, Zhengzhou, China

[04] College of Animal Science and Technology, Henan University of Animal Husbandry and Economy, Zhengzhou, China

[05] Zhengzhou ToYou Bioengineering Co., Ltd, Zhengzhou, China

[06] Jiangsu Institute of Product Quality Supervision and Inspection, Nanjing, China

[07] Haining BoShang Biotechnology Co., Ltd, Haining, China

[08] NewCore Biotechnology Co., Ltd, Shanghai, China

[09] College of Tobacco, Henan Agricultural University, Zhengzhou, China

[10] Jiesai Biotechnology Co., Ltd, Haining, China

[11] School of Life Sciences, Zhengzhou University, Zhengzhou, China

[12] College of Life Sciences, Henan Agricultural University, Zhengzhou, China

[13] College of Paleontology, Shenyang Normal University, Shenyang, China

[14] National Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), China

[15] Jiangsu Jiayu Biomedical Technology Co., Ltd, Jiangyin, China

*Corresponding authors and emails:

wqzhao@zzu.edu.cn; guoyuqi@zzu.edu.cn; qliyou@henau.edu.cn; meijunli@cup.edu.cn

Abstract

Obtaining high-quality ancient DNA (aDNA) is a crucial process for advancements in molecular paleontology. However, current research is restricted to a limited range of fossil types due to DNA degradation and contamination from environmental DNA (eDNA). Utilizing the technology of nanoparticle affinity beads, we successfully extracted DNA from *Lycoptera davidi* fossils and obtained 1,258,901 DNA sequences. To ascertain their origins, we developed and implemented a novel protocol, the "mega screen method", designed to categorize aDNA fragments. This approach successfully identified 243 Actinopterygii species-specific fragments (SSFs), which subsequent analysis confirmed as original *in situ* DNA (oriDNA) sequences. *Lycoptera davidi,* a ray-finned fish from the Early Cretaceous Jehol Biota, dates back approximately 120 million years (Ma). Notably, the absence of deamination in certain oriDNA and paleoenvironmental DNA (paeDNA) fragments calls into question its reliability as a definitive marker for aDNA identification, a phenomenon potentially attributable to the protective effects of fossil diagenesis. Additionally, we identified a novel mechanism, "coding region sliding replication and recombination", responsible for generating new transposase. This mechanism, distinct from conventional processes such as horizontal gene transfer or vertical inheritance from parents, can facilitate a rapid genomic diversification. The transposase CDS insertion into the GAS2 gene of *Anabarilius grahami* offers molecular insights into the genetic interactions between ancestral carp and *Lycoptera* and the swift emergence of the expansion of early fish species starting from the early Cretaceous. These findings enhance our understanding of fossil DNA preservation, lay the groundwork for an ancient species DNA database, and introduce innovative tools for exploring evolutionary relationships across geological timescales.

Keywords: nanoparticle affinity beads, DNA, *Lycoptera*, Cretaceous, deamination, transposase, evolution

Introduction

The acquisition of high-quality DNA data is fundamental to molecular paleontological research. Although scientists declare that they have successfully retrieved oriDNA from some fossils dating back approximately 1 Ma, extracting amplifiable oriDNA from most fossils continues to pose a significant challenge [1]. A majority of the oriDNA has been lost, degraded, or crosslinked with macromolecules, making it unsuitable for use as an amplification template. Fossils embedded in lithological matrices can adsorb exogenous eDNA, with some potentially becoming embedded or subsequently replaced by newly introduced eDNA [2,3,4]. eDNA is integrated into fossils at distinct temporal stages: during pre- or post-lithification phases (paleoenvironmental DNA, or paeDNA and in recent or contemporary periods (present-day DNA, or preDNA). Predominantly, the eDNA found in most fossils originates from bacteria, often overshadowing the signals of oriDNA [5,6]. Additionally, studies have highlighted that oriDNA fragments frequently undergo deamination, which contributes to the degradation of the DNA strand [7,8].

The conditions under which fossils are buried and the mechanisms of their formation are crucial for the preservation of oriDNA. Research has shown that certain fossilized volcanic tuffs, rapidly encased by volcanic ash and formed from living organisms, still contain intact nuclei of *in situ* organisms preserved at the outlines of their remains. For instance, intact nuclear structures are commonly found in fossils such as Jurassic plants and the Jehol Biota; nucleic acid stains reveal the presence of nucleic acid material in the nucleus location [9,10,11,12,13]. While these findings are fascinating, they also raise important questions. For instance, if the nucleus tests positive for nucleic acids, are these derived from oriDNA? Do they include long-stranded DNA sequences (>100 bp) suitable for amplification templates? Addressing these questions is key to driving the most critical technological advancements in molecular paleontology [14].

Lycoptera davidi is an extinct freshwater fish of the order Osteoglossiformes. As prominent fossils of the Jehol Biota, they hold significant biostratigraphic importance for the early Cretaceous period (J3-K1), dating back approximately 120 Ma. The fossils were preserved in sedimentary rock and formed under hydrostatic conditions when the fish bodies were quickly encapsulated by volcanic ash [15,16].

In this study, fossil DNA extraction was performed using nanoparticle affinity beads and sequencing of 1,258,901 DNA fragments [17]. The mega screen method was employed to distinguish oriDNA fragments from others, successfully identifying 243 *Lycoptera* fragments (oriDNAs). Additionally, numerous DNA fragments from ancient hominids were detected, likely originating from environmental sources and deposited within fossil voids in antiquity. Intriguingly, most of these fragments exhibit no evidence of deamination, suggesting exceptional preservation. Further analysis using amino acid sequence data revealed a self-renewal mechanism within transposase genes, termed "coding region sliding replication and recombination". This mechanism appears to underpin the rapid diversification of Cretaceous fish, offering significant insights into fish evolution. Moreover, the DNA evidence derived from non-skeletal remains introduces novel methodologies for evolutionary studies, fundamentally advancing the field of molecular paleontology.

Results and Discussion

I. Extracting DNA information using the mega screen method

Utilizing the mega screen method, 1,258,901 DNA fragments of diverse origins were classified; the relevant biological information was extracted. This methodology involves two steps outlined in Materials and Methods and Figure S1.

All sequences were divided into subsets based on the lineage

Initially, as the *Lycoptera* genome data was unavailable, we conducted a BLAST search to align the fossil DNA sequences with all sequences in the NCBI database (Version 5, Nucleotide sequence Database) using the Best E-value mode (BE mode). We processed data from the 1,258,901 reads, creating subsets based on genealogical lineages and determining the total number of sequences (TS) within each subset. Among these, 674,472 total sequences (TS) were classified as bacterial sequences, 131,312 TS as primate sequences, 30,211 TS as angiosperm sequences, and 26,852 TS as fungal sequences. Additionally, 19,747 TS were categorized in the arthropod sequence subset. With the 10,000 Fish Genomes Project (F10K) nearing completion, the genomes of representative species have all been sequenced, providing a more comprehensive reference database for fish sequence matching. Leveraging these advancements, we have identified 11,313 sequences aligning with the subset of ray-finned fish sequences (Table S1).

The Actinopterygii subset was dominated by aDNA

To establish the sequence selective threshold, we initially set the E-value below 1E-07 and identified the sequences that met this criterion as qualified sequences (QS). Among the screened sequences, 693 met the threshold, while 10,620 did not. We speculate that the inability of these *Lycoptera* sequences to meet the threshold could be attributed to the absence of *Lycoptera* genome data, inadequate matches, and limited similarity to contemporary fish genomes. Some unaligned sequences may originate from non-fish species that have not yet been sequenced. Due to these limitations, we can't analyze these sequences in more detail in this work.

To assess the closeness of a subset to the respective modern genomes, we introduced the metrics, Affinity Index, and set the thresholds at 90% and 97.5%, respectively (see Materials and Methods). The percentage of the ray-finned fish subset passing the two threshold sequences was 8.51% and 4.91%, respectively. Compare this to other subsets; for example, the two percentage values for the human subset were 94.78% and 91.56%, respectively, while the percentages for amphibians and reptiles were 15.15% and 12.12%, respectively. In addition, our analysis should introduce some other factors, such as geographical factors; the local geography is as follows: The cold climate and hilly topography of the region result in a sparse distribution of amphibians and reptiles, which limits their contribution to the local eDNA; therefore, the Affinity Index of this subset can be used as a background value to screen the corresponding lineages. From the above, it can be assumed that the ray-finned fish subset meets our criteria, and most of the sequences in this subset do not come from the local environment, which are aDNA or oriDNA. The Affinity Index for the human subset is high (Table S1), and most of the DNA fragments in this subset are recently contaminated preDNA sequences. As many extant species have yet to be sequenced, the values of the Affinity index for some subsets are correlational and may fluctuate in future studies.

Some features of the ray-finned fish DNA sequences

When comparing the DNA sequences of 243 ray-finned fish (Tables S3A, S4) to the sequences in the human subset, it was observed that they are similar in length and GC content. However, there were notable differences in metrics such as Per Identity, Query Cover, average number of base deletions, and QS/TS (%). The human subset displayed a lower frequency of insertions/deletions when compared to that of ray-finned fish due to the sequence differences between preDNA and aDNA. The DNA sequences within the fish subset demonstrated aDNA characteristics overall, as depicted in Table S2, providing a sound basis for effectively distinguishing the ray-finned fish subset (Subset Affinity: 52.87%; see Materials and Methods) from the preDNA-dominated human subset (Subset Affinity: 97.15%).

Identification of each sequence in the fish subset

In the subset of ray-finned fish, we initially utilized the Best E-value mode to identify 693 high-quality matches. However, some of these matches had the same E-value when hitting genomes of different species, which made it difficult to determine their genealogical affiliation. Therefore, we concluded that these sequences exclusively trace their genealogical origin to ray-finned fish. Consequently, we used the MS mode to ensure the reliability of the uniqueness hit, thereby enabling the identification of species-specific fragments (SSFs; Materials and Methods). Consequently, we confirmed 243 sequences as Actinopterygii SSFs.

It is reasonable to deduce that the sequences of the ray-finned fish subset can be divided into three parts: oriDNA sequences, paeDNA sequences, and preDNA sequences. By excluding paeDNA and preDNA sequences, oriDNA sequences can be recognized. These 243 sequences were then categorized into three groups: the "local fish group", consisting of 180 sequences (Table S4); and the "non-local fish group (including marine fish)," which contains 49 sequences. The third group, the "other ray-finned group," comprises 14 sequences that could not be classified at the order level, and its Affinity Index value is only 77.51%, indicating a notable distinction from the known genome.

The *Lycoptera* oriDNA fragments: We further determined the source of the sequences of the Actinopterygii subset by combining the influence of geographical factors, geological changes, and evolutionary laws. Since there has been no marine transgression since the Cretaceous, the 49 sequences of the non-local fish group should be oriDNA. This result also shows a genomic connection between the *Lycoptera* and modern marine fish. In the local fish group, the 180 sequences mainly belong to Cypriniform genomes. The average values of the Affinity Index of the above two groups of sequences are 80.34% and 83.85%, respectively, and their conservation levels (average values) are almost equal. In addition, the fossil site is currently a hilly area without lakes and rivers. These conditions indicate that most of the sequences in the local fish group cannot be preDNA. At the same time, considering freshwater fish eventually become prey to other species, it is difficult for their DNA to enter the environment and become eDNA. Therefore, we can roughly infer that most of the 180 sequences in the local fish group and the 14 sequences in the non-order group could be oriDNA (Tables S3, S4).

II. Fossil DNA: The multiple resources hypothesis

The *Lycoptera* fossil can be defined as a rock that contains internal voids. We have developed a method to quantify the percentage of these internal spaces relative to the total volume of the rock, which ranges from 10.8% to 11.3% (see Materials and Methods). Due to diffusion and osmosis, various molecules not embedded in the fossil may be lost as internal liquid water drains away when the surrounding environment becomes arid. Subsequently, when water returns to the environment, DNA molecules may be transported into the dry voids within the fossil and become trapped. Furthermore, some molecules might be encased by minerals in the water and deposited onto the inner surfaces of the rock. However, if the environment dries out again, free molecules could escape.

Given that advanced primates exhibit the social behavior of burying their dead, DNA from the genomes of these primates and associated species, such as raising animals, prey, and surrounding vegetation, can continuously be introduced into the environment of long-term settlements, contributing to eDNA. Consequently, eDNA can infiltrate fossil rocks, and some of this DNA may also become buried in minerals within the water and permanently retained in the fossil. Thus, it is reasonable to conclude that eDNA from later periods can enter the internal voids of fossils formed in the Cretaceous period.

The mechanism of DNA preservation in *Lycoptera* fossils remains unclear. Water is the primary factor in DNA deamination[18]. Fossils in sedimentary rocks are formed when organisms are rapidly covered by volcanic ash in a calm water environment (Figure 2). This forms a protective shell around the organism, which prevents the chemical groups on the DNA molecules from coming into contact with water, allowing nucleic acids to be preserved for a long time and possibly protecting them from external elements such as oxygen and chemicals. Although some oriDNA fragments may

be lost or degraded due to geological processes, remnants may still exist. The eDNA in fossils comes from the environment, adsorbed on rock surfaces, cavities, and cracks; part of it has become permanent sediment [2,3,4,5].

Our proposal delineates three sources of fossil DNA, as illustrated in Figure 2: the first being oriDNA before fossil formation, including the fish DNA and DNA from *in situ* species, such as food found in the stomachs of fishes, minute aquatic organisms, and fish parasites. The second source is paeDNA, and the third is preDNA. preDNA primarily originates from human activities, the surrounding plants, raising animals, parasites, bacteria, viruses, etc. The Beipiao area was a center of agriculture and animal husbandry in China. Ancient eDNA has infiltrated fossils since the remains were buried in mud under hydrostatic conditions, particularly during periods of considerable lacunae before the fossilization of dense formations. Following full petrification, the impermeability of shale presents challenges for water-soluble compounds to permeate the pores. Despite these, the possibility remains that minuscule quantities of matter continue to seep into the fossils.

To elucidate the origin of DNA in fossils, we developed a model diagram based on our hypotheses and speculations (Figure 1, right): (1) Before the fossil forms a dense structure, the internal DNA fragments are primarily composed of oriDNA. At the same time, paeDNA molecules gradually enter the pores and remain inside. (2) As the fossilization process progresses, the interior becomes denser, causing the voids that can accommodate foreign molecules to decrease and sometimes disappear entirely. Concurrently, the ingress and egress of internal DNA molecules are obstructed, embedding the DNA within the rock formation and rendering it immobile. (3) Following this process, the surface area of the internal pores continuously decreases, reducing their capacity to accommodate foreign molecules. The remaining open gaps are relatively robust and resistant to collapse or blockage. Subsequent generations of external paeDNA can still enter and adsorb to the surface of the pores, but are rarely embedded and remain predominantly mobile. (4) These mobile DNA fragments are displaced over time by later generations of external molecules. This replacement process persists over time, with older molecules being supplanted by newer ones up to the present day. (5) The oriDNA and paeDNA embedded within the fossils undergo continuous degradation. Consequently, this leads to a decrease in oriDNA and paeDNA and an increase in preDNA. In this study, a limited number of fish fragments (11,313 TS) were detected, whereas a larger quantity of human DNA fragments (127,977 TS) was identified, predominantly originating from preDNA (Table S1). This exemplifies the phenomenon of so-called "fewer old DNA molecules and more new ones" in fossils (Figure 1).

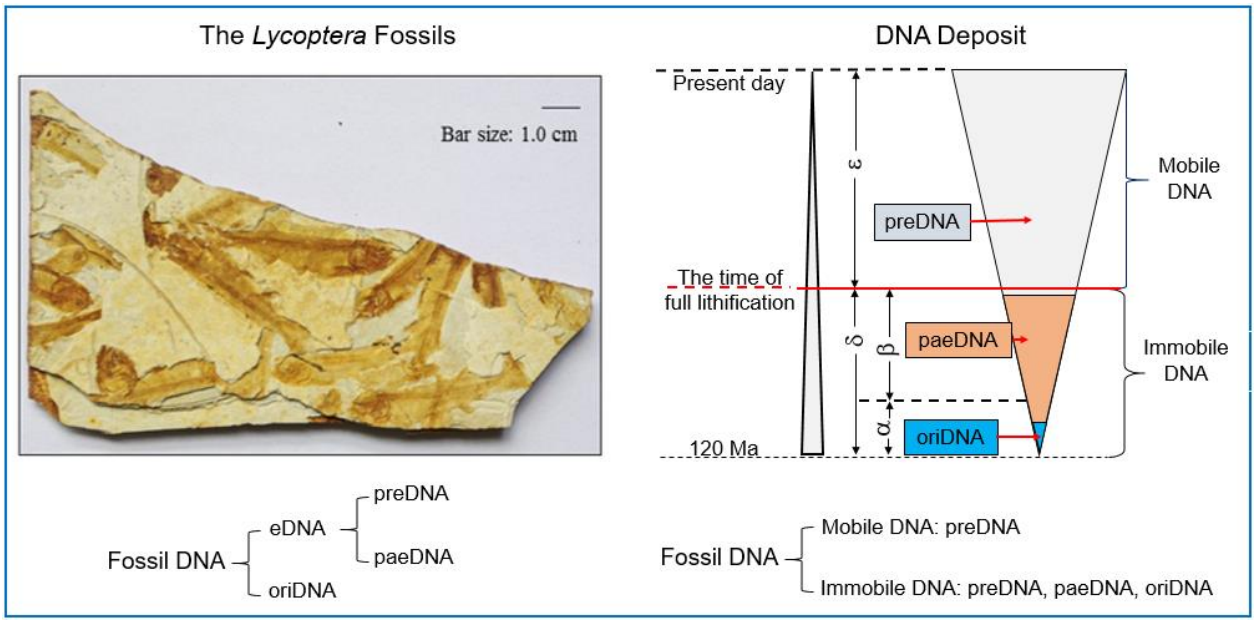


Figure 1. Formation of Fossils and Their DNA Contents

The fossils were found in Beipiao City, Chaoyang City, Liaoning Province. The longitude is 120.84, and the latitude is 41.60. The process of fossil formation and preservation of DNA**:** During early lithification (α), a significant amount of volcanic ash descends into the water, rapidly forming a thick, viscous paste that envelops and buries the fish within a tranquil, undisturbed water environment. This results in enclosed inclusions isolated from their surroundings, allowing some paeDNA to become trapped within them. As lithification progressed (β), these inclusions gradually desiccated and were compressed into a thin, dense layer known as the "fish layer", each approximately 1 mm thick. The brown sections in the shape of fish within a plate represent the fish remains, with the surrounding whitened sections acting as the cofferdam area. Additionally, there is the "non-residual layer" without any fish remains. After complete lithification (δ), the "fish layer" and the "non-residual layer" are mixed and overlapped in multiple layers to form dense tuff plates and are further petrified into lithological fossils. After the soft tissue has decayed, the area between the skeletal remains becomes very porous, with many tiny cavities connected to surface cracks in the fossil rock. This allows for the transfer and exchange of internal and external molecules, providing a pathway for eDNA to enter the fossil.

III. Absence of deamination in select oriDNA and paeDNA fragments

Although deamination is a type of DNA damage and a criterion for assessing aDNA in many studies, non-deamination was observed across the 243 oriDNA sequences (Tables S3A, S4). It may be attributed to the sealed structure of fossils following diagenesis, which likely restricts exposure to external environmental conditions, thereby shielding oriDNA from outside factors. Nevertheless, some DNA fragments located on fossil surfaces, such as paeDNA or preDNA, may remain vulnerable to environmental influences, lacking the protective rock encasement and resulting in deamination. These heavily altered sequences were identified and excluded through analysis of the MS mode.

We identified numerous fragments that aligned with non-human primates' genomes (NHP genomes) in our DNA extracts. The affinity values for these fragments range from 38% to 90%, and the mega screen method suggests a "unique" relationship between some sequences (SSFs) and the NHP genomes (Table S3B). Additionally, several facts support an opinion: these fragments likely originate from unclassified genomes of ancient Hominoidea; e.g. (1) The *Pan* lineage is not distributed throughout the East Asia region, and there are no fossil records of *Pan* species in the fossil production area, indicating these fragments are not a result of eDNA contamination; (2) our laboratory has never been exposed to samples containing *Pan* DNA. Furthermore, these fragments do not fully align with known *Pan* genomes, and there is no existing sequencing record, ruling out experimental contamination; (3) *Pan* genomes have been extensively sequenced, but the observed affinity gaps between the known ape and the sequencing data should not be so pronounced. (4) No discernible deamination was observed in these sequences.

IV. Deamination is not suitable as the primary criterion for the authenticity of aDNA fragments

In the method developed by the Pääbo team, deamination is the main criterion for identifying aDNA fragments [19,20,21,22,23,24]. However, several points challenge this perspective: (1) The assumption that the fossil oriDNA originates from a single individual of the corresponding species overlooks the potential for eDNA contamination from either the same or different species over time following the formation of the fossils; (2) The selection process involves only a limited number of closely related genomes as BLAST references, failing to consider the broader influence of numerous environmental species and neglecting to utilize all available genome sequencing databases; (3) There is insufficient consideration of the unreliability of BLAST results, mainly when the E-value is excessively high, rendering them outside a credible interval; (4) In this work, many DNA fragments with non-deamination match to the genomes of ancient organisms, including *Pan troglodytes*, suggesting that this characteristic does not serve as a reliable logical criterion for oriDNA identification (Table S3A, S3B). Consequently, the genomes of Neanderthals, Denisovans, and several species assembled by Pääbo's team likely represent "mixed genomes". The belief that deamination is necessary for identifying aDNA/oriDNA is incorrect, which has hindered the search for long-chain aDNA or oriDNA, which could offer more comprehensive genetic information. As a result, several significant flaws in the literature concerning Pääbo's research on ancient species should be addressed, and these works should be re-evaluated.

V. Genomic links between the *Lycoptera* and other fishes

The relationship between the genomes of ray-finned fish and their morphological development is not fully understood. The classification of these fish is mainly based on their physical characteristics. A study by Zhang concluded that *Lycoptera* belongs to the order Osteoglossiformes [25]. Early researchers, such as Cockerell and Berg, suggested that species of the Cypriniformes order originated from Jurassic Lycoptera [26,27], while Rosen proposed that they may have originated from Gonorynchiformes[28]. The oldest known carp fossils date back about 60 million years. Carp fishes are found in freshwater environments worldwide, except Antarctica, South America, and Australia, indicating that their origins predate the breakup of the Pangaea continent. Molecular paleontological studies have provided valuable insights into the origins of carp. Tao's research suggests that the genomes of carp fishes can be traced back to the early Jurassic period, around 193 Ma [29]. Over 75% of the more than 180 sequences of *Lycoptera* oriDNA in the present study align with the carp genome, indicating a genomic connection between them. Furthermore, the genetic relationship of *Lycoptera* with carp and other native fishes requires further investigation. Although the total length of the oriDNA obtained in this study is limited, further exploration is necessary for a comprehensive phylogenetic analysis.

VI. Decoding fossil DNA: Revealing parasites and prey

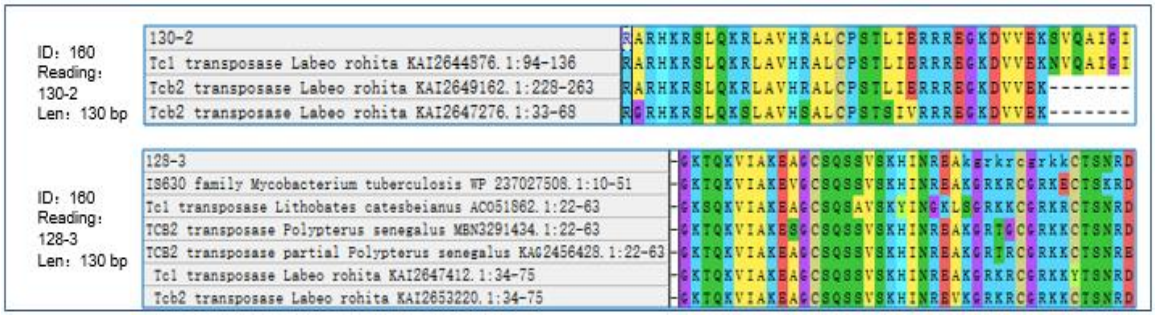
In the fossil DNA, the ID: 282 sequence (Table S3) corresponds to the 28S rDNA sequence of the class *Ichthyosporea*. These are flagellated microorganisms that belong to the Opisthokonta (unranked). This small fish parasite's taxonomic position lies between animals and fungi [30]. Additionally, we have identified 4 DNA sequences corresponding to the *Macrobrachium nipponense* genome (IDs: 285-287) with a Subset Affinity value of 83.24%, of which 3 DNA sequences were highly divergent from the genomes of existing species. Two other sequences corresponding to *Penaeus vannamei* and *Rhynchothorax monnioti* (sea spider) genomes, both non-native species, were hypothesized to be derived from close relatives in their ancient decapod relatives (Table S3A, ID: 283, ID: 289). Upon using 3D X-ray computed tomography to scan the fossils, we found no remains of other animals. This result suggests that these fragments originated from the parasites and prey of the *Lycoptera* and that they are oriDNA rather than eDNA (Figure S1). It also brings new insights for studying their evolution.

VII. New mechanism for generating transposase-encoding sequences

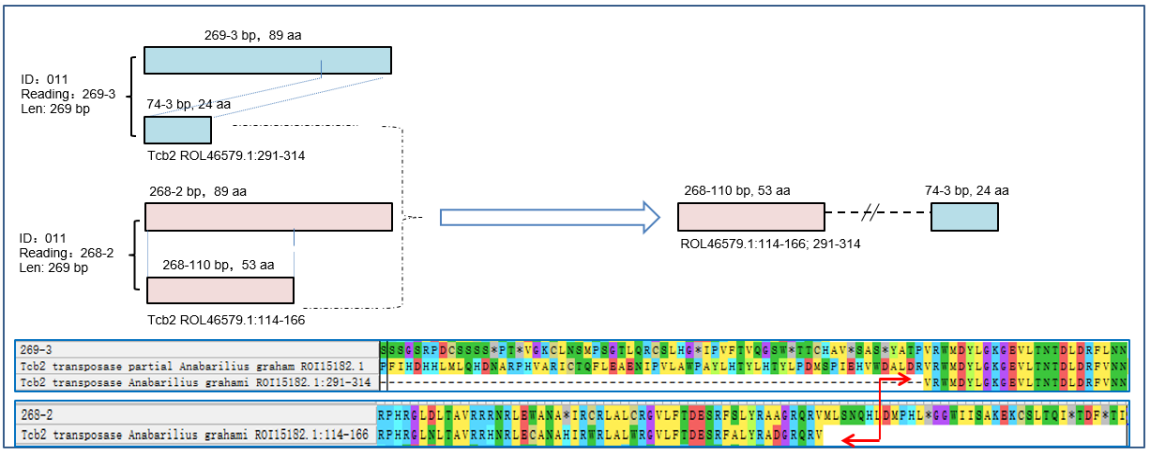
Transposons constitute over 40% of the fish genome, with transposase as a critical component in the transposition process [31,32]. According to Frith's research, transposon sequences in modern genomes can be regarded as "molecular fossils," exhibiting substantial sequence similarity across diverse organisms, including prokaryotes, arthropods, and mammals [33]. Transposons play a key role in introducing new DNA sequences into the genome. Our analysis of 10 fish aDNA fragments has revealed a novel mechanism for transposase gene formation, designated "coding region sliding replication and recombination". This mechanism diverges from established conventional processes such as horizontal gene transfer, vertical inheritance, and gene or chromosomal duplication [34].

This mechanism comprises three distinct types of coding: (1) The initial sliding duplication takes place within a CDS to generate a segment, followed by a shift in the reading frame that produces a different segment. Subsequently, the two distinct segments can be integrated with the segments derived from other transposase genes to construct a novel CDS (Figure 2A). (2) After sliding replication, different segments are generated, which are then assembled in a tandem manner and then connected to the CDS segments from the other transposase to form a new enzyme (Figures 2B, 2C). (3) Indel events occurring at a single locus: when an insertion or deletion of a base in a CDS results in a shift in the reading frame. Figure 2D suggests a "deletion-correction mechanism" that can utilize the mutant CDS where the Indel has occurred to generate a novel enzyme.

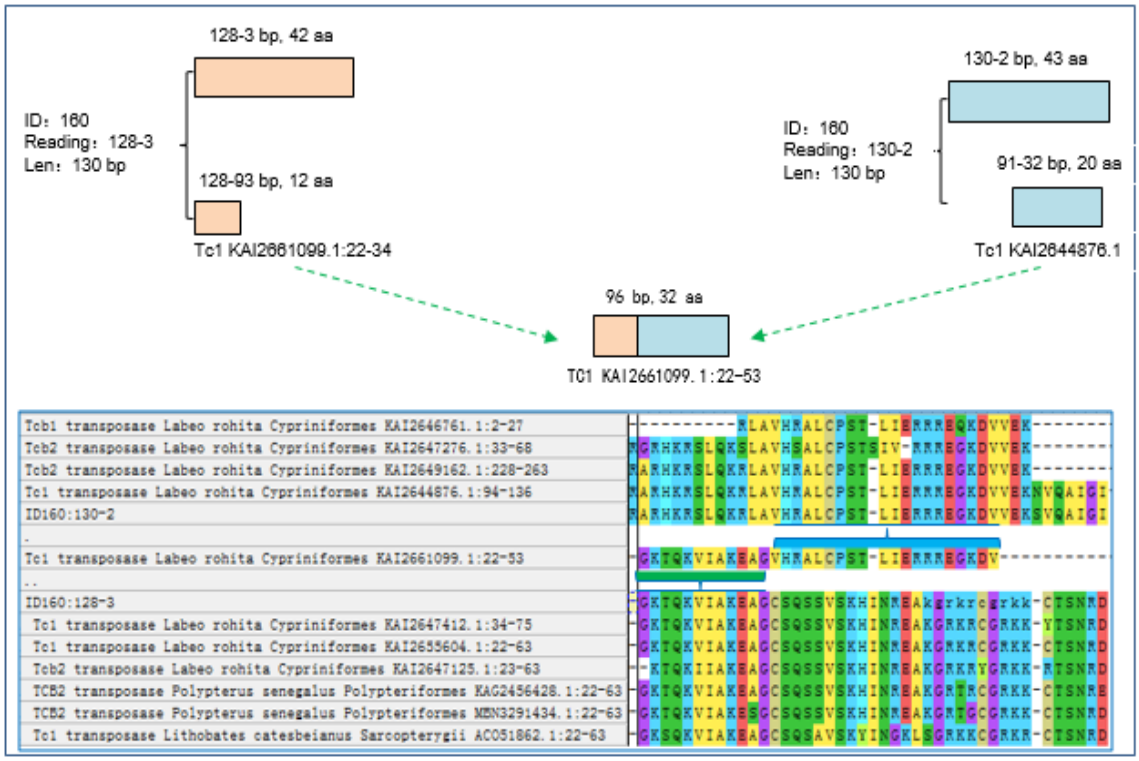
A


ID: 160
Reading: 130-2
Len: 130 bp
130-2
Tc1 transposase Labeo rohita KAI2644876.1:94-136
Tcb2 transposase Labeo rohita KAI2649162.1:228-263
Tcb2 transposase Labeo rohita KAI2647276.1:33-68
ID: 160
Reading: 128-3
Len: 130 bp
128-3
IS630 family Mycobacterium tuberculosis WP 237027508.1:10-51
Tc1 transposase Lithobates catesbeianus AC051862.1:22-63
TCB2 transposase Polypterus senegalus MBN3291434.1:22-63
TCB2 transposase partial Polypterus senegalus KAG2456428.1:22-63
Tc1 transposase Labeo rohita KAI2647412.1:34-75
Tcb2 transposase Labeo rohita KAI2653220.1:34-75

B


269-3 bp, 89 aa
ID: 011
Reading: 269-3
Len: 269 bp
74-3 bp, 24 aa
Tcb2 ROL46579.1:291-314
268-2 bp, 89 aa
ID: 011
Reading: 268-2
Len: 269 bp
268-110 bp, 53 aa
Tcb2 ROL46579.1:114-166
268-110 bp, 53 aa
74-3 bp, 24 aa
ROL46579.1:114-166; 291-314
269-3
Tcb2 transposase partial Anabarilius graham ROI15182.1
Tcb2 transposase Anabarilius grahami ROI15182.1:291-314
268-2
Tcb2 transposase Anabarilius grahami ROI15182.1:114-166

C


128-3 bp, 42 aa
ID: 160
Reading: 128-3
Len: 130 bp
128-93 bp, 12 aa
Tc1 KAI2661099.1:22-34
130-2 bp, 43 aa
ID: 160
Reading: 130-2
Len: 130 bp
91-32 bp, 20 aa
Tc1 KAI2644876.1
96 bp, 32 aa
TC1 KAI2661099.1:22-53
Tcb1 transposase Labeo rohita Cypriniformes KAI2646761.1:2-27
Tcb2 transposase Labeo rohita Cypriniformes KAI2647276.1:33-68
Tcb2 transposase Labeo rohita Cypriniformes KAI2649162.1:228-263
Tc1 transposase Labeo rohita Cypriniformes KAI2644876.1:94-136
ID160:130-2
Tc1 transposase Labeo rohita Cypriniformes KAI2661099.1:22-53
ID160:128-3
Tc1 transposase Labeo rohita Cypriniformes KAI2647412.1:34-75
Tc1 transposase Labeo rohita Cypriniformes KAI2655604.1:22-63
Tcb2 transposase Labeo rohita Cypriniformes KAI2647125.1:23-63
TCB2 transposase Polypterus senegalus Polypteriformes KAG2456428.1:22-63
TCB2 transposase Polypterus senegalus Polypteriformes MBN3291434.1:22-63
Tc1 transposase Lithobates catesbeianus Sarcopterygii AC051862.1:22-63

D

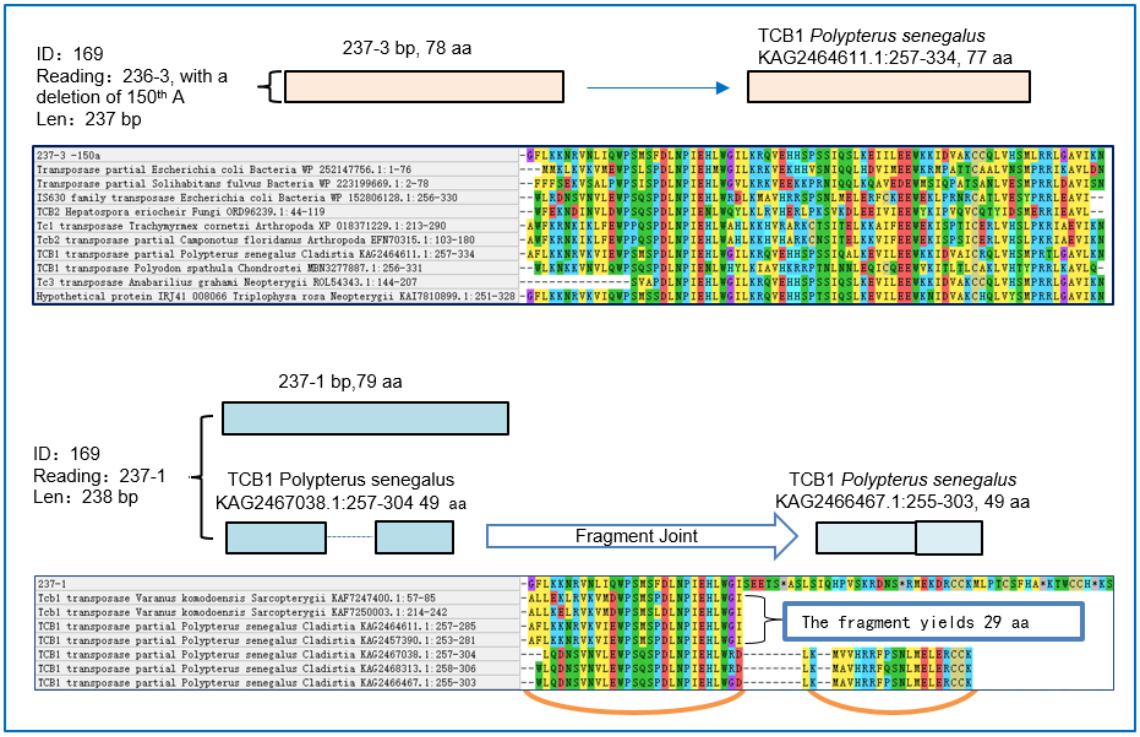


Figure 2. Proposed model for a new mechanism of generating transposase-encoding sequences

A. Multiple fragments produced by sliding replication are spliced and fused with other gene fragments separately to form a new enzyme. **B**. The two fragments generated by sliding duplication were spliced together with other gene fragments to constitute the coding region of a new enzyme. **C**, the two fragments generated by sliding duplication are joined together and spliced with other gene fragments to form the coding region of a new enzyme. **D**, A site in the coding region of the transposase undergoes an Indel, but still passes clips to constitute the coding region of the new enzyme.

During the Mesozoic period, ray-finned fish underwent significant evolutionary developments, resulting in the emergence of numerous new Orders. This diversification was facilitated by the ability of their genomes to generate a substantial number of novel transposases, through the mechanism of "coding region sliding replication and recombination", independent of external DNA contributions. This intrinsic capability likely drove a rapid expansion and divergence of their genomes, accelerating the appearance of new species. Furthermore, transposases were found to integrate into the coding regions of functionally regulated proteins. For instance, the N-terminal portion of an oriDNA fragment (sequence ID: 011) shares 52 amino acids with both the Tcb2 protein and the growth arrest-specific protein 2 (GAS2) of *Anabarilius grahami*, as its C-terminal portion corresponds to the conserved region of GAS2 (Figure 3). These findings indicate that the coding DNA sequence (CDS) of the transposase was inserted into the GAS2 gene, with this translocation likely occurring before fossilization in the Early Cretaceous. This event underscores the early evolution of fish genomes and elucidates a pathway for genome content expansion. Furthermore, it provides molecular evidence supporting genetic exchange between ancestral carp and *Lycoptera.*

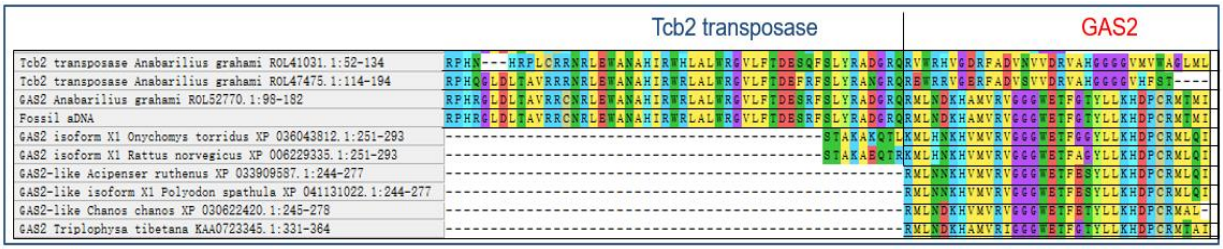


Figure 3. The transposase sequence is fused in the GAS2 coding region

VIII. Summary

We developed a novel protocol, the mega screen method, employing SSF (rather than assembled genome approaches) to identify and extract information from aDNA fragments of extinct species preserved in fossils. Using this approach, we successfully identified 243 oriDNA fragments from *Lycoptera davidi*, dating to the Early Cretaceous period, approximately 120 Ma. By leveraging aDNA sequence characteristics, we traced the phylogenetic descendants of the *Lycoptera* genome, which elucidated the molecular mechanisms underlying its rapid evolution, including a novel mechanism of transposase gene formation. This intrinsic dynamic process is closely linked to the rapid emergence of new ray-finned fish species during the Cretaceous period. These findings deepen our understanding of fossil DNA composition, lay the foundation for establishing a DNA database of ancient species, and provide valuable reference points for investigating evolutionary relationships between ancient and modern species. Moreover, they enable phylogenetic tracing and exploration of molecular mechanisms driving evolution or extinction in all extinct species, significantly enriching the paleontological information landscape and opening potential pathways for the resurrection of extinct organisms.

Surprisingly, this research provides a new perspective on "Panspermia," the theory for exploring the origin of Earth's life. Inspired by the study of extraterrestrial nucleobases in the Murchison meteorite [36], we propose the possibility that if rock fossils, such as the *Lycoptera* fossils, are transported into space, the DNA inside could potentially remain dormant, effectively and uniquely preserving life. Until these fossils land in a suitable extraterrestrial environment and release the contents to become the seeds of life, catalyzing the colonization of the planet.

References and notes

1. Orlando, L. et al. Ancient DNA analysis. *Nat Rev Methods Primers* 1, 14 (2021).
2. Pedersen, M. W. et al. Ancient and modern environmental DNA. *Phil. Trans. R. Soc*. B370:20130383 (2015)
3. Foucher, A. et al. Persistence of environmental DNA in cultivated soils: implication of this memory effect for reconstructing the dynamics of land use and cover changes. *Sci Rep* 10, 10502 (2020).
4. Beng, K. C. & Corlett, R. T. Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. *Biodiversity and Conservation* 29, 2089-2121 (2020).
5. Miao, B. et al. Assessment of contaminants associated with gold-standard ancient DNA protocols. *Science Bulletin* 68:5-9 (2023).
6. Liang, R. et al. Genome-centric resolution of novel microbial lineages in an excavated Centrosaurus dinosaur fossil bone from the Late Cretaceous of North America. *Environmental Microbiome* 15, 8 (2020).
7. Briggs, A. W. et al. Removal of deaminated cytosines and detection of *in vivo* methylation in ancient DNA. *Nucleic Acids Res*. 38(6):e87 (2010).
8. Kjær, K.H. et al. A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA. *Nature*. 612(7939):283-291 (2022).
9. Pang, K. et al. The nature and origin of nucleus-like intracellular inclusions in Paleoproterozoic eukaryote microfossils. *Geobiology* 11, 499-510 (2013).
10. Bomfleur, B., McLoughlin, S. & Vajda, V. Fossilized nuclei and chromosomes reveal 180 million years of genomic stasis in royal ferns. *Science* 343, 1376-1377 (2014).
11. Yin, Z. et al. Nuclei and nucleoli in embryo-like fossils from the Ediacaran Wengan Biota. *Precambrian Research* 301, 145-151 (2017).
12. Bailleul, A. M. & Li, Z. DNA staining in fossil cells beyond the Quaternary: Reassessment of the evidence and prospects for an improved understanding of DNA preservation in deep time. *Earth Science-Reviews* 216 (2021).
13. Zheng, X. et al, Z. Nuclear preservation in the cartilage of the Jehol dinosaur *audipteryx*. *Commun Biol* 4, 1125 (2021).
14. Black, R. Possible Dinosaur DNA Has Been Found. *Scientific American* April 17 (2020).
15. Grabau, A. Stratigraphy of China. Part 2 Mesozoic. Beijing: Geological Survey of China 1-774 (1928).
16. Gu, Z. Jurassic and Cretaceous of China. Beijing: Science Press 1–84 (1962).
17. Su, T. et al. Development of nucleic acid isolation by non-silica-based nanoparticles and real-time PCR kit for edible vegetable oil traceability. *Food Chem* 300, 125205 (2019).
18. Lindalh, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709-715 (1993).
19. Fu, Q. et al., DNA analysis of an early modern human from Tianyuan Cave, China. P*roceedings of the National Academy of Sciences* 110, 2223-2227 (2013).
20. Wang, H. et al., 12 Human genetic history on the Tibetan Plateau in the past 5100 years. *Science Advances* 9, eadd5582 (2023).
21. Dabney, J. et al., Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proceedings of the National Academy of Sciences* 110, 15758-15763 (2013).
22. Mao, X. et al., The deep population history of northern East Asia from the Late Pleistocene to the Holocene. *Cell* 184, 3256-3266.e3213 (2021).
23. Bai, F. et al., Ancient genomes revealed the complex human interactions of the ancient western Tibetans. *Current Biology* 34, 2594-3605.e2597 (2024).
24. Zhang, D. *et al.*, Denisovan DNA in Late Pleistocene sediments from Baishiya Karst Cave on the Tibetan Plateau *Science* **370**, 584-587 (2020).
25. Zhang, J.Y. Phylogeny of Osteoglossomorpha. *Vertebrata PalAsiatica*. 44(1): 43-59. (2006).
26. Cockerell, T. D. A. The Affinities of the Fish Lycoptera Middendorffi. *Bull Amer Mus Nat Hist* 51, 313–318 (1925).
27. Berg LS. Classification of fishes both recent and fossil. *Trudy Zoologischeskogo Instituta* 5, 85–517 (1940).

28. Rosen, D. G., Anderson, S. & Weitzman S. Origin of the Weberian apparatus and the relationships of the ostariophysan and gonorynchiform Fishes. *American Museum Novitates. Environmental Science* 2428, 1-25 (1970).
29. Tao, W. et al. Phylogenetic relationships of Cypriniformes and plasticity of pharyngeal teeth in the adaptive radiation of cyprinids. *Sci China Life Sci* 62, 553-565 (2019).
30. Mendoza, L., Taylor, J. W. & Ajello L. The class mesomycetozoea: a heterogeneous group of microorganisms at the animal-fungal boundary. *Annu Rev Microbiol*. 56:315-44 (2002).
31. Etchegaray, E. et al. Transposable element-derived sequences in vertebrate development. *Mob DNA* 12, 1 (2021).
32. Tafalla, C., Estepa, A. & Coll, J. M. Fish transposons and their potential use in aquaculture. *J Biotechnol* 123, 397-412 (2006).
33. Frith, M. C. Paleozoic Protein Fossils Illuminate the Evolution of Vertebrate Genomes and Transposable Elements. *Mol Biol Evol* 39 (2022).
34. Gilbert, C. & Feschotte, C. Horizontal acquisition of transposable elements and viral sequences: patterns and consequences. *Curr Opin Genet Dev* 49, 15-24 (2018).
35. Hughes, L.C. et al. Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *Proc Natl Acad Sci USA*. 115:6249-6254 (2018).
36. Martins, Z. et al. Extraterrestrial nucleobases in the Murchison meteorite. *Earth and Planetary Science Letters.* 270(1–2):130-136 (2008).
37. Gilbert, M. T. P., et al, Assessing ancient DNA studies. *TRENDS in Ecology and Evolution*. 20(10):541-544 (2005).

Acknowledgments: We extend our gratitude to the "Chao-Xiang Talent Program" foundation of Haining City, the support that covers the costs of fossil collection, DNA extraction, and DNA sequencing from 2020 to 2021. We extend our heartfelt appreciation to the following individuals for their invaluable contributions to this project: Dr. H. J. Zheng from the Shanghai Institute for Biomedical and Pharmaceutical Technologies, Dr. J. Y. Ma from the Nanjing Institute of Geology and Paleontology, at the Chinese Academy of Sciences, Dr. A. M. Bailleul from the Institute of Vertebrate Paleontology and Paleoanthropology, at the Chinese Academy of Sciences, and Dr S. P. He from the Institute of Hydrobiology, also at Chinese Academy of Sciences. Their stimulating discussions regarding the manuscript have greatly enriched our work.

Author contributions: The study was conceptualized and designed by W.Q.Z. and Y.Q.G. The fossils were collected and overseen by F.L. and W.Q.Z. The DNA extraction from the fossil was carried out by W.Q.Z., C.Y.Z., and J.Y.Y. Alignment between the sequences and present genomes was completed by Z.Y.G. and M.J.C. DNA laboratory analysis, interpretations, taxonomic profiling, and annotation were conducted by Z.Y.G., Z.Y.T., L.J.Z., S.J.Z. and W.Q.Z. Statistical analyses were performed and completed by G.Q.C J.H.G, and X.G.Z. Phylogenetic analyses of mitogenomic DNA sequences were performed by Z.X.Q., W.Z., C.J.Z., M.F.T., and D. W. under the supervision of G.Q.C. and W.Q.Z. Figures were designed and finished by F.L., L.Y.Q., T.F.S., and M.R.S. Project coordination was managed by T.C.Q. W.Q.Z., Z.Y.T., M.J.L., and L.Y.Q wrote the manuscript.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: All data is available in the main text or the supplementary materials.

Supplementary Materials

Figure S1 to S2; Tables S1 to S5

Supplementary Materials for

Ancient DNA in *Lycoptera* Fossils: Key Discoveries for Evolution and Methodological Reassessment

Wan-Qian Zhao[1,11*], Zhan-Yong Guo[2], Zeng-Yuan Tian[1], Tong-Fu Su[3], Gang-Qiang Cao[1], Zi-Xin Qi[4], Tian-Cang Qin[5], Wei Zhou[6], Jin-Yu Yang[7], Ming-Jie Chen[8], Xin-Ge Zhang[9], Chun-Yan Zhou[10,15], Chuan-Jia Zhu[11], Meng-Fei Tang[1], Di Wu[1], Mei-Rong Song[3], Yu-Qi Guo[11*], Li-You Qiu[12*], Fei Liang[13], Mei-Jun Li[14*], Jun-Hui Geng[1], Li-Juan Zhao[11], Shu-Jie Zhang[11]

Corresponding authors and emails:
wqzhao@zzu.edu.cn; guoyuqi@zzu.edu.cn; qliyou@henau.edu.cn; meijunli@cup.edu.cn

The PDF file includes:
Materials and Methods
Figures. S1 to S2; Tables S1 to S5

Materials and Methods

I. The experimental procedure for the wet lab

1. X-ray 3D photography: The fossils originated from Beipiao City, Chaoyang City, Liaoning Province. They underwent a cleaning process involving tap water, deionized water, and precise cutting. Subsequently, the complete fish fossil underwent 3D X-ray scanning using a Zeiss Xradia 620 Versa scanner (Carl Zeiss AG, Germany) to ensure the absence of co-deposited species. The initial scan was conducted at a resolution of 45 μm to capture the overall structure. Specific areas such as the head, thorax, tail, and periphery were scanned at a higher resolution of 6 μm for detailed examination. The ZEISS 3D Viewer software was used to examine and reconstruct 2D orthogonal slices of the scans, helping to identify any additional species present in the scans and reconstructions.

2. Method for evaluating the internal volume ratio of fossils

After cleaning the fossils, dry them in a dryer (56°C, 72 hours) and weigh them (W1). Immerse the fossils in pure water for ultrasonic treatment (10 W, 2 hours). After taking them out, quickly wipe off the attached water and weigh them (W_2). Immediately put the wet fossils in a graduated container and immerse them in water. Calculate the total volume of the fossils (V_2) by reading the increased scale.

Internal volume of fossils (V_1) = (W_2 - W_1) × density of water

Internal volume ratio of fossils = V_1/ V_2 × 100%

The internal volume ratios of four *Lycoptera* fossil samples are 10.8%, 10.4%, 11.3%, and 10.8%.

3. DNA Extraction, DNA Library Construction, and Sequencing

The procedure was conducted in a BSL-2 industrial laboratory (Biosafety Level 2), following rigorous ancient DNA extraction protocols [37]. The process involved using UV light to disinfect all used items. We also applied Geobio® DNA Remover (from Jiaxing Jiesai Biotechnology Co., Ltd., Jiaxing, China) on clean benches and solid equipment. It is important to note that this laboratory has never been exposed to fish, non-human primates, or plant samples before this work.

The stone was cleaned with tap water and deionized water, and the surface was clean. Immerse it in Geobio® DNA Remover for a few seconds and then quickly remove it and soak it 3 times in deionized water. The stone was washed with deionized water to remove the additional detergent solution. After drying, the stone was cut into 10 × 10 × 5.0 (cm^3) raw materials, and the multilayered stone was divided into single layers approximately 0.8-1.2 mm thick. The textured portion of the fish was scraped off, and the dropped particles were collected in a clean tube. Scrape off the contoured portion at least 1.0 cm from the textured portion and collect the dropped particles in a separate tube. Harvest the same weight of fine powder from the non-textured (coffered) portion of the fossil at least 1.0 cm from the textured portion. Use a set of mortar and pestle and a mortar and pestle to grind the small particles into a fine powder. Fossil DNA was extracted from the above fine powdered material using the Geobio® DNA (Solid State Matter) Extraction Kit (Jiaxing Jiesai Biotechnology Co., Ltd., Jiaxing, China) [17].

During long-term geological evolution, some *in situ* DNA molecules within the fossils undergo cross-linking reactions with other macromolecules (e.g., proteins and polysaccharides) as a result of the action of physicochemical and biological factors, to assess the concentration of the extracts, we examined the optical density (OD) values of the DNA mix, which contains both DNA-linking complexes and free DNA molecules, at the peak of the UV 260 nm. Five replicate extractions were performed on the textured portion of the sample, and two replicate extractions were performed on the non-textured portion, the so-called cofferdam portion (Table S3).

The DNA extracted was repaired using NEB PreCR® Mix (M0309S). Next-generation sequencing (NGS) DNA libraries were prepared following the SSDLSP standard program (Sangon Biotech, CORP). Sequencing was conducted

on an Illumina NovaSeq 6000 (S-4 kit) using the "Dual Label Sequencing" standard procedure. The reads met quality control standards and exhibited a satisfactory quality distribution. Occasionally, it may be necessary to manually remove excess adapters from both ends of the reads (sequences) to obtain accurate information. Combining forward reads with their corresponding reverse reads resulted in 1,259,658 paired sequences for the fossil-textured portion (including 757 unnamed sequences) and 635 sequences for the fossil non-textured portions (matrix).

II. The experimental procedure for the dry lab

1. The mega screen method comprises two modes

Initially, we utilized the BE mode to cluster sequences by aligning fossil DNA with known sequences from the NCBI database (Version 5, Nucleotide Sequence Database). An E-value threshold of less than 1E-07 was applied to filter for qualified sequences (QS) and determine the best match at a specific taxonomic level.

Subsequently, we employed the MS mode to perform a search that excluded the initial result. If the quotient of the E-value of the first result divided by the subsequent one is ≤ 0.01 (1e-02), the fragment is considered specific to the first result (in a taxonomic unit), thereby confirming its "specificity". This indicates that the test sequence (query) corresponds to the subject identified in the initial search, suggesting a unique origin, such as Actinopterygii SSFs. Conversely, if the quotient surpasses this threshold, the sequence is presumed to be shared between two taxa, suggesting it does not derive from a single lineage. Consequently, we discarded the test sequence without deriving further conclusions.

In addition, the mega screen method can be applied to uncover relationships among similar sequences found across various genomic organisms, offering valuable insights into patterns of genome evolution.

The whole process involves the following steps:

1) Nucleotide Blast is employed to align sequences to the entire NCBI database without any limitation.

2) Creation of subsets for sorting the sequences: Using the BE mode, results are matched, and sequences are grouped into the appropriate subsets, determining the total number of sequences (TS). Subsequently, an E-value threshold of < 1E-07 is established to filter for qualified sequences (QS).

3) The subset primarily containing ancient DNA (aDNA) is selected.

4) Each sequence within this subset is screened individually using the MS mode to identify those originating from a single lineage, categorizing the “unique” sequences (SSFs).

5) Considering various factors, including local climate shifting, geological changes, and evolutionary principles of the host species, we can ascertain if the sequence is oriDNA.

2. Sequence affinity (Affinity): This metric measures the similarity between subject sequences and hit genomes. It is calculated by multiplying the values of Identity and Cover obtained from NCBI Blast and then converting the result to a percentage (Identity × Cover × 100%). Mutations in the genomic and mitogenomic sequences occur continuously during the evolutionary process, leading to significant sequence differences between ancient and modern species. However, this variation is less pronounced in conserved sequences compared to non-conserved sequences. A high-affinity value indicates that the sequence is highly coherent and related to the genome of the modern species, which is likely to be either a conserved sequence or a modern eDNA. A low-affinity value indicates a low similarity, and the sequence may be from a distantly related species either has or has not yet been sequenced. It is necessary to consider various factors, including the sequence composition and conservation, host species information, host genome sequencing data, etc., rather than relying on the Affinity value alone when determining whether a sequence is an oriDNA.

3. Subset affinity: This metric refers to the average of the sequence affinities of all sequences in a subset; it indicates whether the subset consists primarily of aDNA sequences. High values (close to 100%) indicate that the DNA sequences in the subset consist mainly of preDNA.

4. The percentage threshold for subset affinity (Affinity Index): The number of sequences within a subset that exceed a certain threshold as a percentage of the total number of sequences. In this study, we set the thresholds at 90% (indicating limited similarity between the sequences and their hit genomes) and 97.5% (indicating high similarity between the sequences and their hit genomes), respectively. This metric reflects the closeness of the relationship between all sequences within a lineage subset and the modern genome. For a subset, a low value indicates that the sequences are mostly ancient (aDNA), while a high value indicates that the sequences are mostly recent (preDNA).

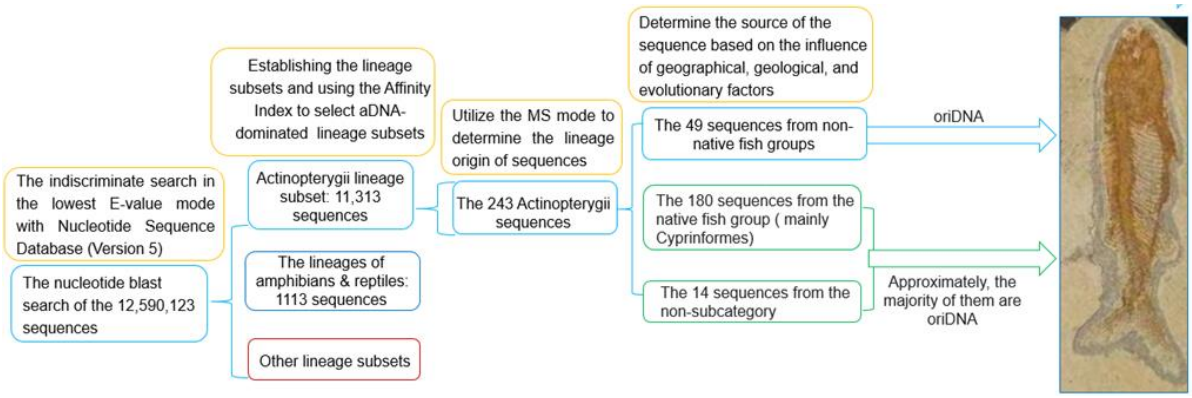


Figure S1. Selecting for the fossil oriDNA using the mega screen method

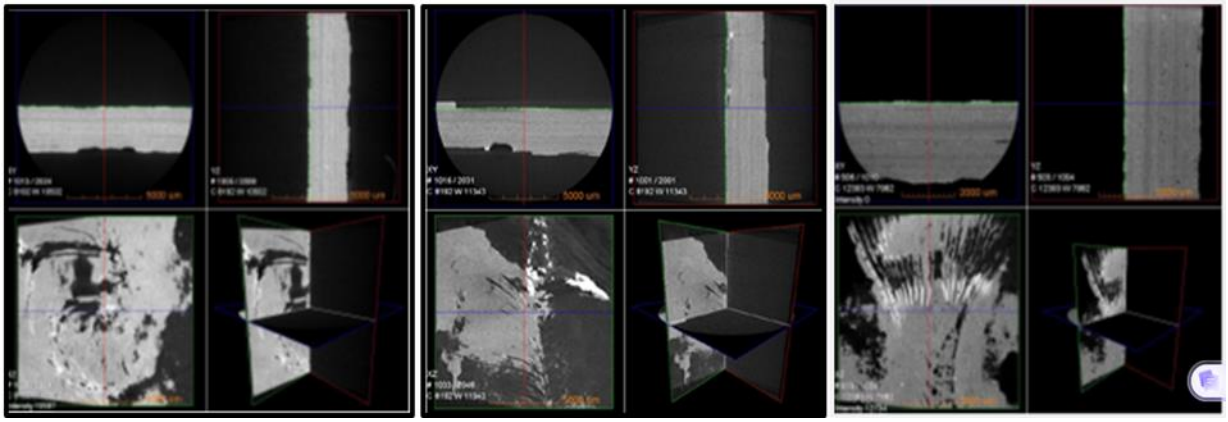


Figure S2. 3D X-ray scanning the Fossil Fish

Table S1. Sequencing reads and their matching lineage genomes

	Textured parts				Surrounding parts			
Organisms	TS	QS	Aff 90%	Aff 97%	TS	QS	Aff 90%	Aff 97%
Prokaryotes	674472	597813	68.48%	61.12%	333	259	54.83%	54.00%
Fungi	26852	19372	60.91%	51.96%	15	9	77.78%	66.67%
Algae	**1201**	**571**	**14.54%**	**8.41%**				
Bryophytes	366	186	90.32%	84.41%				
Gymnosperms	**513**	**442**	**11.54%**	**5.20%**				
Angiosperms	30211	20241	50.46%	39.79%	18	12	0.00%	0.00%
Protozoa	3913	2874	88.27%	78.84%				
Nematoda	1512	837	18.52%	8.12%				
Mollusca	**1248**	**51**	**3.92%**	**1.96%**				
Platyhelmintha	881	406	79.56%	73.40%				
Arthropoda	19747	2591	44.46%	41.14%	**11**	**11**	**18.18%**	**9.09%**
..Decapoda	**588**	**63**	**3.17%**	**3.17%**				
Ray-finned fishes	**11313**	**693**	**8.51%**	**4.91%**				
Amphibians & Reptiles	**1113**	**33**	**15.15%**	**12.12%**				
..Homo sapiens	127977	127656	94.78%	91.56%	187	187	89.30%	83.96%
..Livestock	2812	2167	59.07%	45.82%				

TS, each subset's total number of sequences; QS, the number of qualified sequences

Table S2. Some characteristics of DNA sequences in the ray-finned fish subset

Markers	*Homo sapiens*		oriDNA	
	Mean	SD	Mean	SD
Length	135.65	23.8	139	21.28
Per Identity	98.32%	6.46%	68.65%	21.05%
Query cover	98.82%	6.08%	77.02%	23.37%
Subset Affinity	97.15%	0.39%	52.87%	4.92%
Gap(s)/Sequence	0.00	0.27	2.25	2.70
GC%	42.05%	9.73%	44.68%	11.00%
QS/TS（%）	99.70%		6.14%	

Table S3A. DNA sequences of the ray-finned fishes and the food web symbiotic species

ID	Len	DNA Sequences
1	101	TGCATATTTCATTAAAGTCACCAGTGGAGGTGGGAAAGGAGTAGGTACAGCTCTCAGGGGAAG TAAAGGTTGTTTAGCTCAAGTGAATGTTACAGGAGCTG
2	111	GGGATCATTACAAGATTGTTCACTTAGCATAAGTTTTGGAGCTCATCTGAGAAGATGCCTCACTT TACCAAAGGTGTGTGACGGTAATAAAGGAAGGTGGCAAATGTAGCA
3	149	GACCAAGCAGAAAGAACCGACCAAAAACTGAACAGAACCGACCAAAACTGACCAAAACCGACA AAAAACGACCAAAACCGACCAAGAACCGACTAAGAACCGACTAAGAACTGACCGTATACAGAA CCGACCAAAATCAACCAGAACCG
4	133	TTTTATTCCTCAAAAATGAGTCTTTGTAAAAAAAAAAAAAAAACTTTACCAAGTATGATTTAGAAG CTTGTCCTTTGGCATGCAGAGAGGGATCTCCACAGAGCACATGCAAAATTAGTTTCATTGTTTCT AG
5	96	ATTGTGAGTCCTTCAGCAGTTCACAGATATCACTTCTCGTTCCTTCTCCATTGGGCAACCTGGC AGCACTGTCATGGACTGTTTCGGAATCAGAAG

7	74	GATTTTTTGGGGTTCATCCGCAGGTTGGACATCGGGATTTTTCGTTTCTTCCAAAGCGTGCGTA AATTGACCGG
8	97	GCCGTGCGCGATCTGCTCGAATTCAAGCCGCTCACGTCCGTTCCTGTGAGCCCTGTGACCCCT GTAACCCCTGTGACCTCTGTGACCCCAGTGAGCA
9	181	GTCAACACTGAAACACAAATCAGCTCGTCAGTTACTACAGCTCGTGAGCTACTACAGCTCATCA GTTACTACATCTCTTCAGTTACTACAGCTCGTTTGTTACTACAGCTCGTCAGTTACTACAGCTCG TCAGTCACTACATCTCGTCAGTCACTACATCTCGTCTGTTACTACAGCTGGT
10	125	CAAGGTTCCATCTGAGCTAAACACAAAACATGTATTCAAAGAAGAAAAAAAGGAATTTCCCTGA GCTATTTTTGCACATTCAACAAAAAGATGATCACACGTGTCATGCAAAAGAAGCACAATAT
11	269	ATATTGTTCAAAAATCTGTCTAAATCTGTGTTAGTGAGCACTTCTCCTTTGCCGAGATAATCCAT CCACCTCACAGGTGTGGCATATCAAGATGCTGATTAGACAGCATGACACGTTGTCTACCAGCTG CCCTGTACAGTGAAAACCGGGATTCATCCGTGAAGAGAACACCTCTGCAAAGTGCCAGACGGC ATCGAATTTAAGCATTTGCCCACTCAAGTCGGTTACGACGACGAACTGCAGTCAGGTCGAGAC CCCGATGAGGACGA
12	125	TGAACTAGTGACCTCATGAAAACATAGAAATGCTTGGTGGAATGACTACAAGAACATTGGCTTG TTGGGAACATGTTCTCTTTAACAGCAACACCTATAAAAGCCTCCTCCTTCAAGCCATTTTG
13	104	ATCGGTGACTGTATATGAAGACGATCACTGACTGTATATAAAGACGATCACTGATTGTATATAAA GACGATCGGTGACTGTATATGAAGACGATCACTGACTGT
14	68	TGTTGAATAAAAGTTGCCATACTACCAAAATGCCTGGACGCAATTAGATAGACCTACATGCAGT GAGA
15	164	AAAAGATGGTCCAACGAGTCAAAGCTGGATTTACAGGCCTGCATTGAGTGTTTTTGAAGTTGCA GCCACTGACCTGGACTGACACTGTGACTTCTTACATCAGTTTTTGTGAGGACGTGTGTGCTGAC TAAAACCTTACGCACATACAACAACCATAAACCCTG
16	163	CCTTCATCATTCCTCCTCTTCCTCTTCCCACACACGGTTTCAGCTGACAGTGGGGTGCTCTGGA CTCCACTTACAGCCTCTGTGCGGCGGGTAAATACGAGAGTTGTTAATAATAATTCTGACCTTAT GTTTTAATTTATTCCTCTATCACAGTGATCCCCAA
17	231	GTGTGTGAGTGTGTGTGCCCTCATGCTTCACTCTCACCTCGTCTTCAGGGTGGTGGTTGGTGA GAATCAGATGCAGTGATGCTGACTCGAGTATCACGTGATATACTTGCAGGTAGTTCGCAGCGG GATTCTCTAACAAGTGAGTAAAAAATTAGTAAAGCTTTAATTTGCTTTTCTGCAAATGGCTGAAAA GTGCAGCTCCCGACTCTCAATCAGTCCAAAGATCTGCTG
18	72	ATACAGCATGACTGGGCGGTTCTTCTGATTTCCAGTGTATAGCTATACATTTCTTTGCTGCTATC AGAGCCA
19	150	AAAGACCACACATGTGCAGTATATGTTGTCCTTTTTCATACTATTACAGCAGTATGTAAAGGGAC AAAAGAAAATAATCACGAGGATAGAAAGCAGCGACTGAAAAGAGCTTTTGCCGTAGATATTCTT GTTATAACATGTGTGGTCTTT
20	264	GCCTTCAGATCATCTGTATATTGTTGGATCGACTGTTTCTCATCTTTCTCTTGAAAATATCCCATA GATTCAGGTCAGGCATATTGGCTGGCCAATAAAGCACAGTAATATCATGGTCAGCAAACCACTT GGAAGTGGTTTTTGCACTGTGGGCAGGTGCTAAAGTCCTGCTGGAAAAGGAAATCAGCATCTC CATAAAGCTTGTCAGCAGATGGAAGCATAAAGTGCTTCAAAATCTCCTGGAAGATGGCTGCATT GACTTTG
21	96	GTGTGTGTGTGTGTGTGCTTCAGATATACCCTACATTATTGGGACAAAATATCCCTACTAATATC TGAAATTGTTCACCTTGTGGGAACATTTTTT
22	282	GTTAACTCCACATCAGCTACATAAATTCATCAACTAACCATTCAGAAACATCCTGTTGCAATCTA CATGTTGTCACTTCTTCTTAAATCTCTCCATCATCGTCCGATTCCGGTTTGAACATAAAAGGCTG AACAGTTTCTGACATTTTCAGTGAGCCTGTGTGAGGTAATCGGAGGTGCTAAAATGAGCTCTTG AAACTCTGTCCTCTTCTTGGGAGCAGCAGCTCATTTGCATTTAAAGGGACACACAAAATCTGAG TGTTTTTGCTCACCCTCAAAAAGT
23	128	CTACTAAAGACATATTTAAAACAGTTCATGTGACTGTTATGAAGTGACAAGAATACTTTTTGTGC GAAAAAAAAAATGGCTTTATTCAACAATATCTAGTAATGGGCGATTTCAAAACACTGCTTCAT
24	149	AAAGGAAAAGAATTGAAAGTGGGGGGAAAAATCACATTATGAAATAAATGTTTTTCTCCAAAACA CGTTGGCCACAATTATTGGCACCCTTTTATTCAATACTTTTTGCAACCTCCTTTTGCCAAGTTAAC AGATCTGAGTCTTCTCCT
25	276	GCAATTATTTAAGACAATTCTCCTAACCTTTTTGGCATAAGGAGGACCTATTATGCCCCCCATCA TATAAGTCTCTGATGTCCCCAGAATGTGTCTGTGAAGTTTCAGCTCAAAATACCCCACAGATCAT TTATTATAGCTTGTCAAATTTGCCCCTATTTGGATGTGAGCAAAAACACGGCGTTTTTGTGTGTG TCCCTTTAAATGCAAATGAGCTGCTGCTCCCGGCCCCCTTTCCAGAAGAGGGCGGAGCTTTAA CAGCTCAACAACAACAAA
26	264	TGTTTGTGTACACATACACACATATATAAAAATTCCTGTTTGGATTTAAATAGGCATATAATATAC TTAAGAGTCTAAGACTGGATCATTATTGCAGTGATTATTATGTTTCTAGTATGTTATATGGTTGAC AACAGTTATTCTAACCCTGACTGATGGAGTGTGTAGCTTTTCATTTCTTTAACAACCATGTAGGA

		AGTCAAATTGTTGCCATATTCCAGGATGACAATGTCAAGATTCCACAAATAGTGAAAGAATTGTTGG
27	240	GAATATCTTCTCCTCCCTCTTGCAGCATCTCTTCTCTTCTCTGATGACGAGGGCGGGGCAACCTGTCACTCACATGAGATCCACCAATGGCAAACCACAAACATCCAGACATCCAATCAATTCCCCATGGACAAAATCAAGTCCTGCCCTACATTTTTCTTGTTAGAGAAGGCATAGGGAAAAAAAGGCTATCACAACTTCGATTCCATGCCCACTTTAAAGGAACAACAGCATTTGTGG
28	103	TGTTCGAACCACCCACTCTCTAAGTGGGACTTGAACCCAGGTCCACCAGCATGGGAGTCAGGCACTCCAACAAGGACTCTAAAGACCGCAGTCTTTACACGCA
29	133	GATGAAATGCCTTTCCGTACCAGCAGATGGCAGTAGCTGCACAGCCAATCAGAATGATCAGATGGCCCGATGGACCGACGAGCTCCGACGCCAATTCAACATGTGGAATCGGCCGGATAAAAGCCGACAAGGA
30	276	CATCGACACAACAGCCAGTGGCACAGCTCCTCAACAAACCGTCCATACCGGTGTGATGAATACGATCCTCAACTGGACTGAACTGGAATAAATACTTTGAATGTTGCGATCCTATCAGACATAAGATAGCAACCTGAATCGTAACAAAGCACTGTTCGCCAGAGGAGAACTGGCCCCCCGACTAAGCCTGTTTTCTCCCAAGGTTTTTTTCTCCATTTTAACACAGAAGTCTATTTGCCACCTGTTTGCCACCTGATGTCACCTGTTGGAGTTTGG
31	100	TGACGATTCATCTGACCCAGCTGCCCAGTCCATACTAAAGGACCCTCTTGGCGCAACCGAATTTCAGCATCCTCCAGCCCAGCGTGGCACGAATACTGGA
32	179	CATCTTGGTTAAAGTTTTAGTATTTGTTTCCATGATTAATGTTTTTGACATTTTTAATAGTTTCTTTTATACATATCGTCGCTGTCAAAAACCTCAATTTGAAGCTTGATTTTGGTAATAAGTGTTAGTTTCTGTCATTAATTCCTCTCCCTCATGTCGTTCCACACCCGTAAGACCTT
33	189	ACATCCCATCCGCTGGGTTAAAACAATCCAATCCCTGGGTTAAAACAACTCTATCCCTGGGTTAAAACAACCCATTCGCTGGGTTAAAACAACCCATCCGCTGGGTTAAAACAATCCAATCCCTAGGTTAAACAACCCAATCGCTGGGTTAAAACAACCCAACCACTGAGTTAAAACCACCCAATCGCT
34	184	CAGAAAAAGTAACTAACAGTTTACACTTTCTTCGTAACTTCAATATGGAAGGCGGTCTGGCGGAAGCTACATATTTTACTTCATAACTTGTTAAATATGGATATTTTACACAAACGCATTGCTTTGCTTCAGAAGGCCTTTATTAACCCCCCGGAGCCGTGTGGAGCACATATTTATGATGGAT
35	146	TGTCCCAGGTGAAAAAAAAGTGCATTTCAATAATGTACTTAAAGTGCTCTATTTTCACACACTAATTTTGTACTTAGATGCTCTTAAGATTATCTTAAGAAGTACTAAAGAAGAATTTTTAGTATATTAAGTACAAAATTAGTGTG
36	169	TCATGATATTACGTATTAGTATTGTGCTCTCATCTAACTTGAAGAAGGGGCTATTTTACAGTACTTTTCAGTTCGTAATATTTAATGCTACAACATCTACGCACTGCACTATTCCAATTTTGCTTAGCTCAATAAACAAAAGAACATCGTTGTAAAATTACATTTGAGA
37	156	ACCAGCCACTTACTTGCATATATTTCGGGAGAAACTGATGAATTCACATGCTGTAAATCCGACTGACAGGACTCTGTGAACTGCAGCGCAAACTAAGATGGCGGCGCCCATCTCGCATTATAGATCAAGATAAAGATCGTTTATAAAGGTTTTTAA
38	194	CATGCCGTTCCACACCCGTAAAACCTTCGTTAATCTTCGGAACACAAATTACGATATTTTGGTTGAAATCCAATGACTCAGTGAGGCCTCCATAGCCAGCAATGACATTTCCTCTCTCAAGATCCATTAATGTACTAAAAACATATTTAAATCAGTTCATGTGAGTACAGTGGTTCAATATTAATATTATAAAG
39	149	GTTTCATCATATGCGTGCTCCTGGGAGCATTTTAGGCAGAATTGGTCACTTTTGCATGAAAAGTTACATTGCCACTAGAAATATGCTATTCTCACTCAGAAACGCTGTTCTGTGCTAGATATGCTTTAAAACAGCTAGTTTCATCATAT
40	225	TTTCTTCACTCGAAAAGAAATTAAGGTTTTTGATGAAAACATTCCAGGGTTATTCTCCTTATAGTGGATTTCAATTGGCCCCAAACGGTTGAAGGTCAAAATTACAGTTTCAGTGCAGCTTCAAATGTATATGCTTTATAGTCACAAATGATCGCATCACAAGTGCTTCCACCAGACCCCGATTCTGTATTCTTCAAAAAGCTTACACTGAATGTCCTACACCTT
41	203	GTTACTCAATTATGTTGTTAGAGAGTCCTAAATAGAATCACTCCGGTTCTATTGTCCTGAAGTCAATGGGGTTTTTTTTACTGGGTTTTTGGTTAAAGGTGGGGTAAGCCATTTTTCAAAAATGTTTTTGAACACTGTTGATATTTGAAATCAACCCAAACAAATCCACCCCTCTTTTCTCATTGCTCTGCCTTCAAAACTCA
42	225	TAAATGTTCATTTATTAAACATATTAATGTTAATTTCCAACATATTTGGGTTCATTTTAAGCAACAATACAGTAATTTTTAAACAATATTTGGGTTTTAAAAAAAACTACCCAGCACATTGGGGCAAATATTTAACCCAACCGCTGGGTTTGTCCATTTTCAACCAAACTTTTAGAGTGTACAAGTTCGATTAAACCAATTGAAATATGAGGAAAATATTTTTATA
43	269	CTAGCCCGCAAGGTTATGATTATCGACACAATCCTCGTCAGAAAGGCAAAGGGGGAGGTGTTGCTGTAATTTATAGTAATATTTACAGTATTAGTCAAAAGTCTTCCAAATATAATTCCTTCGAAGTGATGGTGCTTTACGTAACATTATGTAAGTTGACATTTGTGCTGGCTACTGTATACAGGCCACCAGGACACCATACAGACTTCATCAAAGAATTTGCTGCTTTTCTTTCAGAGTTAGTACTAACTGCAGATAAAGTCCTTGTT
44	207	CCTTTCCCTTCTGCTGGCCTCCGTAACAGACAAAGAGCTGATCTGAGGTCCTGAAGCTTCGGGTTCTGTCCACGTACAGGTGCAGAGCGCGGACGGGACAGAGCAAAGCTAGGGCTGGGTCTGCC

		TCCTCCGAAGGCAGCGCTTGCAGGTTCACTACCTGATCTCGAAAGGGAGTAGTAGGAACCTTG GGCACATATCCAGGCCGGG
45	239	TTCTCAGGCACTCTTTACTTAAAGCTGCAGTCCGGAACTTTTTTTGATTAAAAATTATCCAAAATC AATTTTTGAGCAAGTACATAACCAGCCAGTGTTCAAAACTATCGAATTATCTTAGCTCGATTCAC AACGGTAAGCTTGTAATAACGTTTTATAATAAGAGCGACCTGGTGGATTTCCGTGGGAAATTCA AGCATGCAGCAGTTCGTCTTTGTGTCATTACATCACGTCTGTTT
46	216	ATACTAATAAGCAAAAATGACTTTAGATTAGATTTAGAAAGAAGAGTGTTGAACTTCCTTCTACT GTATCCTAATCTTCTTTAATCCAGAATGGCAGCACAGCTGAGAGGTTTGTTTGAGCGGCGCCCT CTACTGAACAGGTGTAAATATGCATTTCCTTCAGCCTGAGGATTATTCATTTCACTTTTGGTGTG AAAGGGCCTTTAAAGCAATTAA
47	171	GGTTTGAGTTTTATGAACTTATTTCTTTTAATTTAGACAAACTTAAGTTTAACTGAAACTGGGTTG GGTTTTATATTTCCCAGCATGCTTTGCTCATGACACTCGAAAGGCCCATGCAACTGATTGCCTC AATTTTTTTGAGTTTTGCCAACTTATTCAGGTTTACAGTGT
48	91	ATAATCATGGACATGATTCGAAAAGGTTAAATGGGCCTGAAAACAAGTTACACTTGTGCTCCTC CCTGTCAAAAATGGCACATTGGAAATG
49	184	GAAAAGGGGGTTTCGGGCCAGAAAAATTGTTTCAATATATCTCCAAAAAGGACAGATGAGAGAA CCAGGAGTTTAAGCCCCGGGTATACTTCGGCCGTCTGCGTTTGTGAGTAATTTCGGGACCGCG CGGATGGTGCGCGTGCAACAGCGCATGCGCAAGCAGGACAGAAGAGTACACTAGGTG
50	186	TGGAGGCCTCACAGAGCCATCGGATTTCATCAAAAATATCTTAATTTGTGTTCTGAAGATGAAC GAAGGTCTTATGGGTGTAGAACGACATGAGGGTGAGTAATAAATGACAGAATTTTCATTTTTGG GTGAACTAACTCTTTAAAATTTTACTTGTCCAATGTTAAGAATGCATGCAGTTTTACT
51	131	CTGGAGTTGTGTTTACCTAACGTTAAAAATAAGCGGGGCAGCCAGAAATCTGCCCCAATGGCTC TCTTTGAGAGCGCGCTGCAGTTCCATTGTACACCACAGATTTACATAGGAAAGTTGTGGGGCG CTGT
52	176	ATGGCGTAATAATGTAACTTCTGTTGCTGTTCACGGTCAGGGACTAATTTTTTTCCGGCGGTAG GAATGTATTAGTGAAAATGTACTTCATGAAAGTTGCATTGATATATTTTTGGCTTTAATATTTGTA TTGTGTGGTAACCGCTTCGCATCGTGCCTAACGACAGGGAACGACA
53	210	CGAGCAGAGAAAAACAGCATTTTTCATGGACAAAGGAAAGGTACTCCTCTCAGAAAACATACAA ATAAAGATACTTTTAGATGTTTAATTGCTATTTGGAGCATCTGGATCTCTAGATGAGGGCTTAAT GAACTAATTTTTGTGAAAACCTCAAATTCGACCAAAACTGGCATTTCTCACTTGTTTTGCCTGTA GCTCATAATTAGCTGT
54	157	CTCGTCAACATCACTGACTGACCTCACAAATGCGCTTCTAGAAGAATGGTCAAAAATTCCCATA AATACACTCCTAAACCTTGTGGAAAGCCTTCCCAGAAGAGTTGAAGCTGTTATAGCTGCAAAGG GTGGGTCAACTCCATATTAAACCCTACGG
55	149	GGAGAAGAGTAAAATTTAGCTTCAAAATCAATTGAATGTTTTGTCGTGTGCGTGTGAGAGCACT GCTAGCATTCATTTCACTGTAAGACAGGAAAAGAGGAGCTGTCTGTGAGGATGTTGACTTTTTC CTTCGCTGATTTAATGACCTG
56	250	ATAGCAATGCAACTTATATATTTCAGGACATAATTAAAGCATCAAAACATTTAAATGCATTGAAAA TCAGAGGGAACTCATATGTTAGAAAGTTTATGTGAATATGAGTAACTCAAGAGCTTAAGCTTTGA GGCTTGGACGGTACTCATTGCTGTCCACAACAGCCTTCTGTGGTGTTAAAGAGTAAAAATACCA CTGAACTCAAATCCCTCTCAGAATTTCTGTATCCGGCATGTGTTTGATTTAACCA
57	140	CATATTTAATGTCCTGATGAACGCTTCCACCTGCTGACACAGGAAAAGTGACAGAACAATACAC CTCCCAACCCCCTCCATCCCTCCGTCCAGCAGCTTCAAAGCCCGTCTGCCCGGAACATCGCCA TCATCAATCCAGA
58	143	TAATGACTTCACCCTTATCAGAGTTCAGTGAGAGTTCAGGGAATGCAGAGTTATTCCCTAACATT CATCCACAACATAAAGCAGCATTCACAGCTTAATCATTCACTTTTGGTGCAGAGGAATCTTAAAA TGCAATTTACAAT
59	133	TATTTCAGTGTCACATGATCCTTCAGAAATCATTCTAATATGCTGATTTGATACTCAGTTATTATC AGTTTTGGAAACAGTTTTGCTGCTTAATATTTTTTTGGAACCTGTGATATTTTTTTCAGGATTCTTT
60	154	CATGCTGATACGGATGAGGTGACTAAACACATACAGTACTTATGTAGAAACACAAATGATGCCC GTATAATCTCCAAACACCAGTCTTATTCAGTTCAGGGTGTATTATCAAAGAAAATACCACCACTT TTTCATATTTCATGAACTTTAGCAT
61	153	GGATTGAGATCTGGAGAATTTGGAGGCCAAGTCAACATCTCAAACTTGTTGTTATGCTCCTTAA ACCATTCCTGAACCATTTTTTCTTTGTGGCAGGGCGCATTATCCTGCTGAAAGAGACCACAGCC ACCAAGGAATACCATTTCCATGAAA
62	93	ATTTAAATCAGAGTAGGTCACAAGGCCAAAGCCTCACCATCAGTTTGAAGAGCCCCTCTATTGA TCACAGATAAAATCAAGGTAAAACCATTG
63	166	GCAGGCCCATGACAGTGAACTGGGGCACAAAGATGCACTGAAGATCCTCCCCAACCTGATCAT TTGTTCCAGAAGTTTCCGTCTGGCAGATGTTTCAAGCAATCCAAACCAAGACCAACAGACAACC GCACTGCCAACACTGTACACATTCACATACATATATTCA

64	78	AAACAGAGAACCTGCAGAATGGAAACGATCACGTAGTCTTCTTTCCTGGAGACATTCAGGTTGG TGCATGTCATCAAT
65	177	TGTAATTCTTCTTGCTCTTTTCTCTGCATCTCAAGCACATTATTTCATTCAGTCTGTGCTCCATGT GCTCAGCTTTTGCCTTTTCAGCTTTCTGGCCACAAAGTCTCTGCTGTTGCTTCTCCGTTCTTTCT TTATTTCACAAATTCTGCCATCTCATTCTCTTTGTCACTGAACCAA
66	150	CGACGCAAAATGCACTGCGCACCTAAAAATAAATATGTATAAAACGATGTTGATTTTTTTTTTTTT AAATAACGTAAATCTTTCATGGCTTTCTACTACATATTTCTGTAAGAGACATAATTTACGTTATATT AAGCCATACAAGTCTCA
67	165	ACACAACAGGAAATAGCTCCATGTTTTTTCCTGGGCGAGAGGGGCCAATAACTGTCACTTTAGT TTAGATGGCTCGAGCGCAGTGGTCCAATGTCTGTTTAACAGCTTGGGTTGGAAACTTCTGCAAT TAAAGAGTCCATGTGCACACGGAGATGCACTTCCCAG
68	171	AACCTATGTTTATAGACCCTTTTACAGCCCACGTAATCAAATAGTCACGCGTGCATTTTGGCAAC GGAAGTGGTGTTGTTCCCCAGTGGTTCTAACGTATTCATAAACATCCAGTGATTTATATGCTCGT AATTTCTCTCGGGTGTACACGCTCGGTTTCTCAACTAAATA
69	237	GATGACCGAAAATCACCTGGTGAAGGACTTTTATTGACAACTTGCGAGCAGCCAGCATTCCATG GGATGAGGCTGAAGTCGCCTCTGTGGATCGAAACTGATGGAGAACACCTGTCCTCCCATGTGT AACAGGCGTATGCTAGTCAGAGCTGTGCGAGTAAACCTCACTCCCCGATCTCAAGAGATGCTC TAGCGACTGACGCTAGAGGTTGCAGCCTTTAGCCTCTGTGTTAGAGT
70	167	AGTCATTTAGTCCACAGCATGTTAGTTCATTAGGGAAATGAACCTAAGAGGGGAGCTAAATTTC ACTAAATATATGCACTATATTAGCTTATATATTCATTATATTTACTTCACCTGTTAGTAATGTCTTA ATTATTTAATATATGTTTAAATAAATTTGTCATAAA
71	181	TGTCATGATCAGACTATCACATTGTTAGCTTAGCAGCAAGCTAACACCGAATTGGTGCAATTTTT TTCCTAAAAAAGGCAAGCACAAATCCCACCTGCAGCTTATTTGTTTGGTCAACATGACCACTCA GTCACCTGTCAAGAAAAGTTCAGACCTCATGGATTTCAGAAATATAATGGAT
72	123	ATAAAGTAGTTTTCATTTGCGCATTTCGGCAAATTTAGCAGTTCAGTCCGTACATACAGAAAACT TGCCTTCTTCATTACTACTCTTTTTGTAAATCAGTAGGTGGTATGGTATTCTGAACAT
73	115	TGAAGAGTTCAGATGCAAAAGCCTCTAAGTGCCGTCTGAAATTTTCTTCTAAAATGTACATTTTT ATCAAGCTTGTATGTTTATGTTAAGTTATTTAACTTTAATGGCAATTAAT
74	134	GACTATGTCTGCGCTTTCAAAAGTATGCAGTCTCAAAAATGAGGAGACCCCTCCTCGTTGCACT TTGGTGTGTGGACAGGAAGAGAGACAGCACAGCTGTCTTTACAAATATGTTCAGCTGAGTCAAC ACAGGA
75	260	ACTGTCCCTTAGAGCACACACACACACACACAACACACACAAAACCATATGGCCCCAGCATCAA GTGTCATGTTTACAAGCGTAGCTCATTGGCACTGCACTCGTCCTTCTTACCATTATCTTTCCATC TGAACTCATTCAAAGTAAGTCCAGCGTAGAAACATTAATCCTCAAGTTATTCCAAAACAGTTTCT TGCCATCCTGCTCATCATTTTAGCGCTTTGAAATGTGGGCTAGTTACACAAAGCCCCATCAAGC CT
76	134	TCATGTTGTTCCACACCCATAAAACTTTCGTTCTTCTTCGGAAGACAAATTAAGATATTTTTGATG ACAGAATGCTGTCAGATTTTCTTCCATTGACTGCCTTTGCAACTACCACTTTGACGCTTCAAGAA CTT
77	274	GGCGTTTAACAGGGGCATGGGCAGCTCGGGAGCTCAGGGCTTCGACTACGGCATCTCCAACA ACAAAGGTGAGAGTAAAGATCTGGGAGAGATAGAAATAAAATGTGATTTATTATTGAGTGACGC TTACATCTCTGTTTATATGGTTACAGTCCACCTGCCAAATGAACATTTTTTTTCATGCCATGCTCC TGCTAAATATCTGTCATATTTAATCAGATATATCTTGGTAGTTTATTGCCATATACATTCAATTGAT TTCAATCCAATGGTG
78	218	GTTGCCTTTCACTGAAGGGTCAGTTGAAGTCAAATCCTGTGTTCTGCTTAAGGTCACACTGATA GAAGAGACTCATTGTGCTCAAATAGAGAAATACTGAGAAGAACTGTCACTCACTCGACTAATAA AAAGCTCTCCCTTAAATAAAAACAAGTGCTTCACAAAACCTGTTTATTTCCTCACGCTTTAAGCA GACTGGTCAGGATGCTGTGGCTCAT
79	217	ATGAGCCACAGCATCCTGACCAGTCTGCTTAAAGCGTGAGGAAATAAACAGGTTTTGTGAAGCA CTTGTTTTTATTTAAGGGAGAGCTTTTTATTAGTCGAGTGAGTGACAGTTCTTCTCAGTATTTCTC TATTTGAGCACAATGAGTCTCTTCTATCAGTGTGACCTTAAGCAGAACACAGGATTTGACTTCAA CTGACCCTTCAGTGAAAGGCAA
80	249	CTGAGGCTGGGGTCAAGGCATCAAGAGCCACCACACACAGACGTGTCAAGGAATTTGGCTACA GTTGTGGTATTCCTCTTGTTAAGCCACTCCTGAACCACAGACAATGTCAGAGGCGTCTTACCTG GGCTAAGGAGAAGAATAACTGGACTGTTGCCCAGTGGTCCAAAATCCTCTTTTCAGATGAGATC AAGTTTCGGAAACCAAGGTCCTAGAGTCTGGAGGAAGGGTGGAGAAGTTCATCACCCA
81	163	TGAGTTATAAAGTCAGAATTCAAAGAAAAAACAGTCTTTTTTTTCCTCACAACTGGACATAACAC GCAATTGCAAGTTTATATCTCACAATTCTGACTTTATAACTCGCAATTGTGATTTTATATCTTGAA ATTTTGAGAAAAAAAGTCAGAATTGTGAGATA

82	144	GAAGCCCACTAACTCTCATTGATTTTGCTTCAGTGTCTCTTTATTGAAACTTTATCAGGTGGATC TGAGCGCATGGAGAGGGAATCCGGAACGAGTGATGAACTTTTCTGTCTCTCCTCCTCATGCGT ACACCCCGCTACATCG
83	218	TGGAATCGAGATGAGGGCATTTCAGGGTACATGAAAGAAAAGCTTTTCTCTTGCAGGGCTTTCT AAGGCAAAACAGCACTTCATAAATTAAAATTACTTGTCCACAAGCAACACTGGATCTTGAGTTCA GTTCCTGAGGCAATGTTTTAGACATTAATATAACAAAACCATTTGGGGATAAAGTGTAGTTCAGC ATGTGTGCGCGCATCTCGATTCCA
84	240	GAGTAATTGATTTAGAGAATGTTAAAATCAATTGGCAGACAGCTGGGGTCTGGGAGATTAGAAG TGAAGTATAGGTGGAGTGGAGAGCAGCAGACGGCCCCAGAGCAAACAGTAATCTCAGCATTAT AGACCAGGGGGAGATGTACCTGTTTTCTCTTATTTCACTTTCACTCTCTCTCTTGCTCCTTTTCT CCTGTTGGGGACGGTCCATAGTCTCTTACACCAACCTTGTCTCCATCT
85	192	TTAATCATTTTCATTTATGTCACTGCTATTGCTAGACGGCATAGCGGCAAAATGTCCGAAGGAAA TTATAGTGGAGCACAAAAAAACGGAAAATACAAAGGTTAAAAGAACAACGTGAACAATATATACT GAAAAAATACCCAAATTGACTTGATACTTCAAGGCATTGCATACATCAAGACAGAAGTGGCG
86	109	CCTATGCCAACATGTCCAGGTCACAGTCCAACGGGCACCATACCAACATTCCCTCACTGACACA TTAACACAATGAAACAGGTTGCAACCATTGTTAACAGCCTGCAAA
87	156	GGAATTTGTCTAAAATTAGCCCAAGAAGAAAGAAAGAGCTAATTATTGTCAACTTCACATGTCTT ATCCCACGCAAAGATGGTGAAACACTCACCCAACGTAAAGAAATTCTGATACGTCTACCACCAA TAACGCTTCCCTCTATGTGTTTCACAA
88	141	CTTAAATAACTTTTTAATTTTTTTTTTTTTTGAGTGATTCCAATATTCAGCAATAATTTGCCAAATG TCTTCTTTCAAAATAGACCTTAGATCTGTATTCCAAAGCATTCATGAATTACAACCATTTATTTCG CATGTCGT
89	276	AGCACTTACCGAATACAACAGACGAGTCAGTGCCATTGTGAGGGTACCGATAGTGACAGTAGT GGCAGCTGAAAATCTTGTCCAATGTATAAAATGAAACTCGGTGGACTGGACGAGATCATTTGTA GTCAGAGATGTTGCCTGTATTTAGCAAGAGCTGGTCTGAGGCATGTGGTTTACAAGCACCCAAT TTAAGGGTGTCCATGCCCCCTTTGCTATCCTATGGTCGTTGACACTCTACTTTTCTGTGAAGCC CATCAAGTTTTGATTGCCGTG
90	205	AACGATGGCAGGGGAAAAACGAACCGTTTTGAAAGAGTAATTACGTTGTAGCCTACTAGAATTG ATGCCGACCATGTTTATCATAAGTGCAACAAGTCATTTGCATGTGAAAGTAGAAACATGCAATTC AGACGAAACTTCTAAAACGTGCATATTAACTTTCCTCATTAAATGTGTCTGTGTGTCCGACACAG CCTCCACGACG
91	159	GATCGGAAGAGATGAAGAATATACGTCATAGATTGCTTTGGAAGTCAGGCATTTTGGTACATTA TCATGTGCCAGTTTGGTACATCCATACTACAACCAGAGGTCAGGAAAGATGCCACTATTGGCCA AAAATGATGACTAAACAGGTCAAAATTATGG
92	138	TAGAGTAAAGAGACGCAATTAAGCAGGTGGTTTTCCCCTCAATCTGGCGCCCCTGTGCGGTGG GCTGTAGGGGAGGGTTGATGTCGGAGCACTAGCCCCTCTCATCCTCAAGCACATGGTCCACAT TAGTGTGGAATG
93	218	CAAGCCAAAAACTGTTTTTAAAGTAAACTGGTGCCTTCCACTCCTCCCATGTGCTTCTCCTTTTC ACAGTCTGGCAAAAATAATGGGTAGTTCCAAAATATGTGGTCACATGGCAATAAAAGGGTACGG TTGCTAAGATACAGGGGGTGGGTCAACATACCCACTGGCTCATCTTACCCCACTCTCCCCTACA CATAGAAAACTCATAGGCCCGGTTT
94	178	GTGGACTATTTTAACAATGTCTTAGAGTCAGACAGCTCTCGGATTTCATCAAAAATATTTTAATTA ATTTGTGTTCCGAAGATGAACGAAGGTCTTTTAGGTGTGGAACGACATGAGGGTGAGTAATTAA TGACAGAATTTTCATTTTTGGGTGAACTAACCCTTTACGTTGCTTTGT
95	108	AGTGAATTGAATTTCACTATGTGTCTGTCTATGTGTCTCTGTGTGTGTGTGTGTGTCTGTCTATG TGTCTGTGTGTGTGTAGAGAAGTATTGGATATAGTGGTGGAAT
96	141	GCCACAAGCTGTTTTGTTCCTTGCCTTCAGGCAAGCGTTATTGCACTGTCAAAGCAAAGACTAG CAGACTAAATATTCCAACCCACTGCAGTTCATAGACTAAATAACACTTTTATAGGACATGTTCAA GGTTTTTTGCTG
97	173	GCGCTTATCCGGATTGTTCAAACCGACTTAAAACTAAAAGATTATGTTCTCTTGCGCGAACTCAT TCAGGAGTGGTGACTTTCCACGTATTCCCTATTAACTTCTGAGCTGCTCAGCTGAAGTTATGGTT GATTGATCTTCGGCTGGATATCAACACACCCATTAATGTACTA
98	287	GAAAAAGCCGAAGGGAAGCTGGAGGGGGAAAAAAAGTTCAGATTAGTTAAATAATGATTAAAAT TATTCAGAAAAATATATACTGATTATGTCTTACTTTCCCAACTGCTTTTGGAGGTCAACCATGTAC TGGTAACACTGGGCGTAATACGTCATCTGAGCCTCCACAAAATCATTTAGACAGCGGAGATGAT GGGCCTGGGAGAGAAATGATATGGTTTGTTATAGGTCAGACGTTGAGTAAATCGAGGGGGTTG TGAGATGTGGATTGGCAGGCGGCCTAGTGG
99	104	GACAAGTTGAAGAAAAAATGGAAGAGAAAAGATTATTATAGATATCATTGGGTTTTTTCAACAGA CATATATACAACATCAGTCCTTTAATCTGACTGGTGATA
100	209	CTTCACATTTGAGAGTCACAGATGTGCTGGAGGTCCATGAGACCACAATTTCTGTTGCAATTGG TCTGTCTTAAAATGAAATGTGTTAATGAGCAAACAGATAAAATAGATGAATTGTATATGAATAAAA

		ATGACTCTGATATGTCTTGACTTACATTGAATAATTAATGTAAAGTTTTGGGATTTTATAGTTCCA TTTTTTCTAACTA
101	113	TCTCAAAATCTTACAGTCACTGCTGGAAAGGGTTCAAATATGCAGAAGATGCTGGAAAACTGAA GAATCTGCAGGACCTGGAGGATTTTTCTGAAGAACAGAGCTCAATTTAA
102	146	CATTGTTATGATTCCCTAGTCTTTACAGTAAAAGGACCTGGAGCCCTAGGGGTCAGTAGATGTT TTACAGGACTTCCTGTGTTGTGCGAAAGGGGGCTCTGGGGTCAGTACACATTATGCAACAAGG GGCCCGCAGGCCTAATTCT
103	213	ACTAGCGTTTCTGCTGAAGTTTTGCTTGTAAAAAAAGTTCATTGTCTATTGTCAATAGTGTAGTG ATGTATGAAGCACAGCTCACACAGAAGCCACTTTCAAGCAGCTTTTTGTTGGTCAATGCAATGA ATTGTGAAATCTTTCCCTTGTCTGAAATTCCCAATTGTGTGGATATCTATTGGAATTACTTGATTT TGCTTGTGAAAATTCATT
104	76	GCGCCAGATCGTGCCCACAGGACGTCATCGACCGGCATCAGCCCAAACTCTTCCTCTTGTCCC CCAAACATGAGCT
105	216	CTTATCCAAACGCTGCAGTGACTTGACTGAACTGTTTTTCTTTCATGTAACTGTCGCGTCCTTAC AAAACTTGCAGAAACCGATGTTAATATATCACCACCTTCAGTATTTTTCAGAATTTACTTGGAATG CACGTTCATCAGCGACAGCAGACAGTGTGTCTTTGTGTGGGACGTCCTGCAAAAGAACCCCAC AAACTCCGTAATAAGAAACGCG
106	165	TGAGAGCCATTTCCCCAAACACATTTACATCAAATTAAACATATATGCCAACATTTAAATCACGTA CTGTGAATTTCTCTAATGGAAATGTTATTTGACATGGCTAATTTATTTTATTAAAAAAATTAGACA AAAGGGTAACTTTTTAAACTTAGAAAACCCAAA
107	239	GCTGTGTGCCTGGAGCTGAGGGGCTGAGGTGAAGAGCATCCAGTGGCGTCTCACAGTTCCGC TAGGCAGCACACTAATTGAGAAGTCGCTCAGGGGTTGAGCGCAGGGGGCTCGCTGCACTGTG CTCCCTATCCAAATCAAGCCCCTGTAAGTGGGGCTGACATGGACCAATCACCATCTGGAGGTT CCTCCACTAAAACTCCACCATTCCTGACTACCATTAAAGGATCTGCTGCTAT
108	121	CACACGGACAGGTACTATTTTATGCACTGTGGCAAAATACAGTGTATTAAAAGGAATGAGTCAC TGCTGACCTTTCAGCAAGTTGTACTGCCATGTTTCCTTAAACATAGATTTGCTAAGA
109	136	TTTGGGTGGTTTCTTGGTGGGGGCTTATTGGCCCAAATGTTGTGAAATTTACAGGACCTTTATA CAGTGCATGGCATTGTAAAGACCTTAAGTCTATCCTTCACAGCCTCTCTTCATCTGTGTACAATT CAGTATG
110	251	AAAAATTGGATGACCAGTAGTTTCCTACTTCTAATTTCAGAAAAAAACAGAGATTCTAATTTTTGG ACCAAAAACTTCACGTAATAGCCTAGAATACTGTCTAACACTTGATGGCTGCTCTGTTAAGTCTT CGTCGTCAGTTAGGAACCTGGGTGTGCTCTTTGATACCAATCTTTCATTTGAAGGCCATGTTACT AGCATCTGTAAAACCGCATTCTTCCATCTTAAAAATATATCTAAACTACGACATA
111	135	GCATGGTTATACTTTTCTCTCTAATATTGTCAATCTTGCCAGTAAAGAAGTTCATAAAGTCATTAC TGTTGTGCTGTTTGGAAACATCAGAAGTTGAAGCTTTATTTCTTGTTAATTTTGTTGTTTAGATTT GTG
112	148	TTACCCTAACCAAAAGGACAAACTGTTGAGAAGGACAAAAAGGAGACTTTTTATCCAAAAAAGT GACCAGTTTCTGTTTTTTTTTTCTTTTCCAAAGAGAATGACCCTCATTGGACTGAATGTAAGAAAT ACTGTCTGAATCGGGCAG
113	97	TGTCTCTTTTGTCCACTTTCAACCACTTTTGTCCTGATTTCTTCGAGGGGATGGTAAATATATTG GCTTTTTCTGTTCTTTTGATCTAATGGACGAA
114	131	CACAGCGATCAGGTTATTGCCATGTCCACCTGCAAGAAAAACATAACAGTTGCATGAAAATCCA GTGTTGGGAACAGAGCAATCTAGCAGAAAGGTTAATTATTTATTATTGTAAGCAATCCCTGACAT AT
115	133	CACTTTCTCTAATAAAAAACATCAGTATGGTGAGGATGACGTTGCTGTTTTATCCACCGTTTGTG GCAGAATTGTGGTCAGGGGCTTGAAGGCTGACAGTGAGCCGTTAGTGTATATGTGCACTGAAA CACGG
116	133	ACAACAAAATGGAAATATTTGGATACTTTCTAACATTTACAGATGGAGAATATACCTGTGAAATG TAAGTTATGTACTAAGGAAAATGCCATATTTAAAATACATTAAACAGCGCAAGTATCTGTAAGAG CAT
117	193	AGGAAATGATATAGATAGAAAATAAAAACATGGAATATAAATCTAGTATGGATGATGCTAGGATA ATAAAAAACGTGCTGTGGATCCACTATTGCTCAAACAAGACTTCAGATGAAACTTTAGCAAGTTT TCTAGCAATCCATTATTAGCTTATACACCAATTCTGCAATGTTTCAGCACCATCTCACAATCA
118	221	GTTACATATGATTTCTATTTCAAATAAATGCTGTTCTTTTGAACTTTCTGTTCAGTAAAGTTTCCTG AAAAAAAAAAAAGTCACTGTTTCCACAAAGATTTTAAGCAGCACGACTGTTTTCGACATTGATAA TAATAACCATTATTAATAGGTGAGTACCAATATTAATTGATGATAATTGAGCACCAAATCAGCATA TTAAAATGATTTCTGAAGGATCA
119	189	CATATAGGTCCTATACCTATAAAACACACAAAAAAACATTTGAATGAAAACTTGAAATGTAATAGT TTGGACACTGTTGCAATAAATCTTGCCAAGAATTTCCTAACTTTCAACGTTTCGTTTGAATTGTGA AATACATGGGACTGAGACAATGAGACAATCAATCAATACAATCACTACACAAAACTA

120	126	GCCAATTGGGATATACTGTAGATAAATGAATGGGGAGGATAGGCTTATATAAGACTGTATAATTT GGTGTTTCGGTGCCAACTTGAGCCATGATGGTTTCAAGCCTTTGGACAAGTAGTAATTTAT
121	113	CATGGGTACCATGCAACGACATTCTGGACCCTAATCTACTGGACACCTTCCATTCTGCTCATCC CAACAGACCTGCCCCTCGTGGAAGAGGACAACCACCACGATGTCGGGGT
122	167	GTGTTGTGCCGAATTCTGACCCCTGTCAGATTATTAACCCCCTAGTATTGGTAGGTTTAGGATTA GATGTGGGGGAGAGGGGTTAGGATTAGGGGTTGGGTTAGGATTAGGCAATCAGGTAGCGACT TCAACGAGGGGGGTAATAATTTGGCAGCGAGTCGAAATTG
123	142	CTTTTTTAAAACCAGTTGGCAACATAATGACAGAATTTTCATTTTTGGGTGAACTGTCCCTTTAAA TAAATGTGGTGTAATGTTCTTAACTATGATTCAAGTGGTCTGTTGTGAATTGGGTCTTTGATTTG GATCTTTTATA
124	121	AAAGTATCATTCTGACTGGATTATAAGAGCAGACTGATCTCTTAGAATTCTGCTCAGAAACACTG CCATGAATAAAGAATGGTATAAAAACTTCCTGCAAGAGCAGCTTCTCCCAACGATA
125	278	ACTAGTGTGCGTTACTGTACAAGTTACGGGTTTACTACAGAAACGGAGAAGGTGAATTCAGGGC AGGATAAGGAAAAAAATAAATAAAAAAAATTGTGGTTTGGGAATTTGGATACAGGAGAGATGAA GTATCACAGAGCCAGGTAATTTGCAAGACCTGTCAAACTGTCGTTGCGGCAAAACAGAGTAGG GGCCATTCACACAGAATGTGTTTTTCTGGTTAAAAACATGAGACGCAGGTCTAGAGACTTGTTTT TTTAAAAATGTAACTTATTTTA
126	171	ATCTCTTTCAGTAGCTTAGTCGGTATAGGGTCTAACATACATGTTGTTGATTTTGATGATTTAACA AGTTTAGACAATTCTTCCTCTCCTATAGCAGAAAAGGAATGGAATTGTTCCTCAGGAGATCTGTA GCACGACATCTGTGATACTGTAGTAAATGGTTGCATGGTT
127	224	CCCCATGATCCTTTTCACAGGGAAGTTGTTTGGCATCACAGAATGGGTACAGGAGGCTCAGAG TTCGATTTTACCTGTAAATGTATGTATAGTATTTTCTATACTACTGTAATATTTATACAAGCTAGAT TTGGTACATTATTTTGGTAAAAACGACATTGTACTTCAACAAAAATACTTTACATTTACCTTTATC AGTCCATTCAAATGTTTCAAATACATAG
128	229	TTTTTCTTTTGTTGCTTGATTAACATTAATGACAGACAGCAGCAGGTTTATTAGGCTGCTGTCAC TTTAAGGACAAATGCACAGACCGAATATACTTATGGTTATTCAAAACTTTTTAACTTATTTAAAAC TGTGTTCACGAGGTCACTGGCCAAAAAAAGCATTTTGACATAATTTTGTGTTTTTGTCCATTAAA GTGCGACCAGACAGGCGCTAAACAGACAGTGTG
129	167	CACAATATTCTATTTCGTTCCCTCGATTTGCTAAATCGCGTACACAATTTACTATTTCGTTCCCTC GATTAGCTAAATTGCACGCATTATTTACTATTTCGTTCCCTCGATTTGCTAAATCTTGACGACGAT TTACTATTTCGCTCCCTCGATTTGCTAAATCTTGA
130	140	TTGAGCTTCAAGAAAAGACGTGCTGGAATTATAGTAAACAAAATAATAACATGAAAACGCTTTGC TGAAGTTCTGCTTGTCAGTGATGAAGAAATAAAAAATACTGCTCGCAGCCTCTTTCTTGTTCTGC CCCACATTTT
131	101	ACTGTACTGCTGGTTGCTATTAGGCATTTCACATCAAGCACGATAACTATAACAATAAAGTTTTA ATAAACATTGTTTCTTTGAGAATAGGGAAGTCCACA
132	141	CCCTCAACTGTATCGTAAAAATGTTCATTACACTAATGCTGTGTTCGAAATCGCCCCCTATACCC TCATTCACTATTCCCTACATTACTCCACTAATATAGTCCACTTGGAGTGAATGAAACGAGTGAGT GAATTCGGACA
133	90	CTATCCTGACAAAAAAAAAAATTATCCTAGTTTCCACAACTGTTTTCAACACTGATAATAATCAGA AATGTTTCTTGAGCAGCAAATCAT
134	105	ATATAAGACTGGTTTTGTGCTTCAGGGTCACATATAACATTAATTTACTTTGTCACTTGTGTTTTA GTCTCGCCCACAGGTGTTTGCTGGCATGTCTGATGGTCT
135	105	ATATAAGACTGGTTTTGTGCTCCAGGGTCACATATAACATTAATTTACTTTGTCACTTGTGTTTTA GTCTCGCCCACAGGTGTTTGCTGGCATGTCTGATGGTCT
136	112	TGTGCGATTTGGTTGATATCCACAATTAAAACTTATTTAATACAGACTTCAGCACATCGTTTGAC ATGTTTTTGCATTAGGATTTGAATGAATCGATTACCACTTGGCTGTT
137	178	TTGACTTCTGAGAGACACTGCCACTCTGAGAGGCCCCTTTTATACCCAACCATGTTGCCAATTG ACCTAATAAGTTGCAAATTGGTCCTCCAGCTGTACCTTATATGTACATTTAACTTTTCCGGCCTC TTATTGCTACCTGTCCCAACTTGTTTGGGATGTGTAGCTATCATGAAAT
138	111	TTTTTTTCTTGTTCAAGTTAAGCATAAAAGGCAGTGCTAAATCTTTTGTCCAGTGCAGATGTTAAA TTTAGGTAGGAATATGATGGGTAATTTAACACATTATTGCAGGAA
139	152	TAATTTGGACTTCTCTGTGCATGCTCTTTGAGGTGGTGACCATAGACCTGGATTATATGAGTCAC AGACAAGAACAGTTTGACGTAAAAATATAAAAATGGAAACTACTGCGGAAACAAAGACATCTAC ATCTTGGATGCCCTTGAGGTAAG
140	116	TAAAATGGTGTAAGATTTAGGAATTTGATAATGTACTTATCCATTTTGTTGCGAATGCAGTGCAC TTGCAGAGAGAGACCGCCCACACACTCTTCTGATTGGCTGTGATTTTTGAT
141	249	TTTGCGTTCTCTCACAAAACTTTTGTGTTCCGCTAAGAAACTTTGTGTTCCCTCGCAAATCTTTTG CATTTCCCCCGACAAACTTGATTTGCAGGCAACTAAAAAGTTTTGCAAGCGAACGCAGTTTCTC AGGGGAACGCAAACATTTTGTGAGTGAACACAAAGTTTCCTGGGGAATGCAAAAAAAAATTGCG AAAGAATGCAAAAGTATTAACATATAATCTTTGCGTTCTCTCGCAAATTTTTTTG

142	233	AAAATCTTGTAATAAAAGATGATGTTGTATTGGCTATGACTAGACAGCAAATACAGACTCTCCAT GTTTTTTGAGTATTATTAAAGTACAGGCAAAAAAAAAAGTAAAAAAGTTATTAACCAGTATTTTTT CTAATCAAATCGGTTTCAGATTGGTAATAATACAGAAGGAACACATGAACAGAAACGTTAAAATA CTGATTCTGCTCAGAGTGAATTTTTAAGCCCTGATGG
143	172	ATATGCCAGACCCAACAATGCAGAAGAGCTGAAGGCCACTATCAGAGCAACCTGGGCTCTTAT AACACCTGAGCAGTGCCACAGACTGATAGACTCCATGCCACGCCGCATTGCTGCAGTAATTCA GGCAAAAGGAGCCCCAACTAAGTATTGAGTTCTGTACATGCTCATA
144	250	CCAGGTGAGAAAAAGTAACACAACATAACATTTCCATAAAAAGTAACTATGTAACGCAGTTACTT TTTTAGGGAGTAACACAATGTTGTAATGTATTACTTTTAAAAGTAACTTTCCCCAAAACGGGTTAA AAATAAACTCATTTACAGCCACTTTTTCCAAATCCATTGTTTAGTTATACCGTTACTGTTAGTTGC TGTACCAGTTAGTATTGTTCTTGTCTTTGACTGTTAATTGCCAAGTTTATGTT
145	102	TATTGCAGGCAGATAATATTGGATATGTTCAGCTATTAGGATATAAAGCTTACAAAATGTTTATCA TAAACTAGCACTAAGACCCAGAATCTGCCTGCAATA
146	119	AAGCAGAATATGGCTGTTTCTAAAGCAGTGCCATCTGCTGTTAAAAACTAAGCTCAGAATCGGT TCAAGAGAGAATCACGATGCATTTGGAAAATCTCAGAATCGATCAGGAATCATTT
147	247	GGTGCATCACTAATAAAAAAAATATTTAAAAAAGCATTTTAAAAATCTAACAGTTTTTCACTGTAA AAACAAATGCAGTTTTAAGGAGCAGTGTACTAGTTTTCTTTTTTGTTTAGATGGTTAATATTAATG CAACTATCAGCGCACTTTTTTTTTTTTTACAATCCGGCCCTCAACCTAAAATGAGTTTGACACCC CTGATTTAAACTGTTCTGACCAGCAGGGGTCACCAGCGTGTGGTGTTCCG
148	210	CATACTGCCCTAGGTTAACAAGACCTTTAATGGAAGGGTTACCGTAATCTCCTCCTCCCTTCTTT TTATTTGTATATGAGTCTGGGAATAACTATGTCTTGCTATTGTCTAGTAGTGCCCTCACTGTCAT CTGATTCAGGGTCTCCAGAGAGCAGTTTCCTGTCTGCTGCATCAGCACATCGTGCTCACATTAC CTTCATAGTCAGGAGG
149	288	TTTACTGTTTAAGAAGAGTATCTTCTCACCTTGACCTTCCTTGTGTTGTAGGTAGTTTAGGCGAC GGCTGTATGTATGGCAAAGGGACCCGATTAACCTGTGGGTTGAAGAATCATGGGTAACCTCATT AGCAGACCTAGCTGTTTGGGACAAAAGTCCAAACAAGTTAAGTCAGATGAGAGCGTCCTGAAA GAGTGCTACCAGCAACGGAGAGAATGGCCGCTCCCGGAGCCTCCTAGAGAAGAGAAAAGAAA AGAGGATGTGGAAGATGTTCTTCTTAAACAGTAA
150	248	CCTGTTCGGGCGTAGGGGCCGTCACTTTGTTGGAGGCCTCAGTCTGCTGTGACTCTCTGCAAG AGCGCTTTTCAAACCTGGAGAGCCTCTGCGTGTTTTGTGAACCTCCCTTACTCAACACACCTGA TTCAGATCATCAGCTCTTTGGGAAAGAGCGCCAGTGAAGTCAGTTGAGTGCGCCTGATACGAG AAACGTACATACAAAATGTCCAGTGCTGTGGGTCCACGGGCGTAGAGACGTGGCAGCA
151	122	CGAATTACTTTTTTTGGCTATTATTTGGCTATTTAAGCTGCTTTTTTGCTACTCACTGCACAGAAC AGGCGCAGAGAGAGGTGATCCTGTGCTGGGGAGTCAGACCTCACGCTTGCGGGGCA
152	229	AGGGCAGACAGAAAATCCCAAATTCAGTATCTCAGAAAATTAGAATATTGTGAAAAGGTTCAATA TTGAAGACACCTGGTGCCACACTCTAATCAGCTAATTAACTCAAAACACCTGCAAAGGCCTTTA AATGGTCTCTCAGTCTAGTTCTGTAGGCTACACAATCATGGGGAAGACTGCTGACTTGACAGTT GCCCAAAAGACGACCACTGACACCTTGCACAAGGAG
153	77	TTTTAACAGACGACATACTGAACCGGTATGAAATTGGCAGAAAGCTTGGTGAGGGAGGATGCG GGTCCGTCTTTGAA
154	258	TCAACGCAGTTTTATTTTCATAAAAAACAAAACAAAAGTAAAACATAAGAACAAATAAAAAAAATT GCAAACAACAGCACAAATAAATGCAATAGAATAGAGAGCTATGAAAAAGTTTTTCAGGCTAAGTA GGTTTTTATTCAGGTAAGAAATACTACAGAAATTAAAGTAATCAAATGTAAAATAACACATCCAGT GTTCCCCATATGCGTCTTGTCTTATTATATTATTATTAATATTAATAAGACGAACGACACA
155	113	ATCCACTCAAAAAAAATATATTTTTGATATATTTACAGTAGGAAATTTACAAAATATCTTCATGGA ACATGATCTTTACTTAATATCCTAATGATTTTTGGCATGAAAGAAAA
156	249	AGTCCGTACAGCAAACAAATCCAAATCGTAATCCAACAAACAAGGTCAAGAACAGGCAAAGAAC ATACACGAGGCAATCAATGTGGCAATGAGAAAACGCTTGGTAGAGACAGGGGTTAGCTGGCAA TACTTCGCAATCTGTTGTACAACTGAGCTGGATTTAAACAGGAAGCTACAGGAAGTGGAGGAAG AAGAAGAGCACAGTCAGTAATGAGGAGATGGCGCCCTCTGGCGGCTTGATGAATGGAT
157	121	CGAATGTCCAATGTCAAGCCAACTTTATGCCTGTTTTTGTTTTGTTTTTTACTTTATTTTTCTTTTC TCTGTGCCGGATTGCTGATGTCTTTTACCCGGGCATAGATGTCCATCCATCAAT
158	93	CAGCTCACACTCCCAGATTCACTGTTTGATTGGTTACATGATCTGGACTGAAACACACCAGTTTC ACTTCTGCTGTAGTCAAGAGATAATTCA
159	192	GACCTCGGCCTCACTGCGCAGTGACTCGTGGCTCTGTTCCACCACATTATGGCCAATCACGGT TAATGACTAGTGGTCACGGCAAATCAGAGTTAAGTGGGAGAACTGGCCATTCTCTGATCGTGC CGCACAGCGCTTGGCTGTCTCTGATAGGTTAAATTGTTTGATTCTATTATCAGTGAAAACTCTGA G
160	130	TTATCCCTATTGCTTGTACACTTTTTTCTACCACATCTTTTCCTTCCCTTCGCCTCTCTATTAATGT GCTTGGACACAGAGCTCTGTGAACAGCCAGCCTCTTTTGCAATGACCTTTTGTGTCTTGCCCT

161	121	CTTAGATCCGGTAGTCTTACTGTTCCCAGAGCCAAATTAAAACAAAGAGGTGAATGTGCCTTTG CTGTTGCAGGCCCTAAACTGTTTACCAGTTCACAGGTCTGCACCTACATTACAGATT
162	164	GCTTTCTCTCTACCTATATCTCTTTATGTCTTTTTCCTTCCTTCACCCTCTCTCTCTCCTCCCC ATCACTCTCTCTCTCTTTGTGACACTATCTACCTCCTTCTCCCCTTCTCTCCCCCCCTCTCCGTG CTCTATCTCACTCAATCTAATTTTCTCTTTCCT
163	52	AGCGAAGGCTGACCCGTGTGGTTCGATCCAACAGAAGAGCTACTGTAGCTCA
164	141	TAAACATGTTTAAACATGTTTAAACATGTTTAGCGATGTTTAGCATGCTTAAACATGTTTAGCGAT GTTTAGCATGCTTAAACATGTTTAAACATGTTTAAACATGTTTAGCGATGTTTAGTGTGCTTAAGC ATGCTTAAA
165	133	GCTACAGGCGTTGGTGGAATGCTACAGGCGTTGGTGGAATGCTACAGGCGTTGGTGGAATGCT ACAGGCGTTGGTGGAATGCTACAGGCGTTGGTGGAATGCTACAGGCGTTGGTGGAATGCTACA GGCGTTG
166	165	CAACGCCTGTAGCATTCCCCCAACGCCTGTAGCATTCCACCAACGCCTGTAGCATTCCACCAA CGCCTGTAGCATTCCACCAACGCCTGTAGCATTCCACCAACGCCTGTAGCATTCCACCAACGC CTGTAGCATTCCACCACTAAAGGAATTGCGAATAATAAT
167	84	ACGTAAGTGCAGCAGAAATGTTTTTGTAACCTTGGCCAGATCTGTGCCTTGCCACAATTCTGTC TCTGAGCTCTTCAGGCAGTT
168	161	AGCAAGCCCAAAAACACCGCCAAACATCAGCAAGCCCACAAAAATCAGCAAGCCCAGAAACAC CGCCAAACATCAGCAAGCCCACAAACATCAGCAAGCCCACAAATATCAGCAAGACCACAAACA CCGCCCCCAAACAGCAAGACCACAAACACCGCCAA
169	238	TGATTTTTAATGACAGCACCAAGTCTTCTAAGCATGGAATGAACAAGTTGGCAACATTTTGCAAC ATCTATCTTTTTCCATTCTTCAAGAATTATCTCTTTTAGAGACTGGATGCTGGATGGAGAGTGAT GCTCAACTTGTCTCTTCAGAAATTCCCCATAGGTGTTCGATTGGGTTCAGATCAAAAGACATACT TGGCCACTGAATCAAGTTCACCCTGTTCTTCTTCAGAAAACCA
170	175	TGACTTGAACCCCATAGAGAATCTATGGGGTATTGTCAAGAGGAAGATGAGAGACACCAGACC CAACAATGCAGATGAGCTGAAGGCCGCTATCAAAGAAACCTGGGCTTCCATTACACCTGAGCA GTGCCACAGGCTGATCTTCTCCATGCCACGCTGCATTGATGCTGTAATT
171	176	GAATTACAGCATCAATGCGACGTGGCATGGAGGCAATGAGCCTGTGGCACTGCTCAGGTGTAA TGGAAGCCCAGGTTGCTTTGATAGCGGCCTTCAGGTCATCTGCATTGTTGGGTCTGGTGTCTCT CATCTTCCTCTTGACAATACCCCATAGATTCTCTATGGGGTTCAAGTCA
172	167	ACCCATGCATGCCATTCCTCTGCAGTGTACACCGTATTGTGTCACGGGAAATAGTCACCCCAGT TTGGCTTTCTACTTCTTTAGATAGCTGCAGTGAACTTGCTTGCCAATTTTCTTCAACCCTTCTCAT CAGAAGACGCTCCTGTCAAGGTGTTAACTTCTGTGGA
173	149	CAGGTAAGAGACACACACGCACACGGCCATCCACATGATGTAAAAGAAAACATGATTCATCAGA CCAGGCCATCTTCTTCCATTACTCCAGTTCTGATGCTCACGTGTCCACTGTTGGCTCTTTCGGC GGTGGACAGGGGTCAGCATGG
174	145	TCCCGTGGATGCGGATATCCATCCATCCATCCCGTGGAGGCAGATATTTATCCATCCATCCGTC CGTCCATCCATCCCTCCTGTGGAGGGGGATATTTATCCATCCATCCATCCCGTAGAGGGGGAT ATTTATCCATCCATCCAT
175	164	TGGGGAGGCCTGCGTTTCGGCATCCTCATCCTCTCAGTCTCAGATGCAACCATTATCTGAGCCA GGGACTCAGACCAGGAGACATCATGCCTGCATTTCACACACACACACGCACACACACACACGT ACGTACACACTGACACACACGGAAACATCCGTGTGTG
176	142	ATCTCCAGTGCAGACAGAGCCTCATGACACAGCCTTCAGGGGAGAGAGAGAAAGAAAAAGAGA GAGAGAGGGAGGGGAGAGAGGGGGAGAGAGACAGAGAGTGAGAGACAGAGAGGGAGAGAG AGGAAGAGAGAGAAAGAGA
177	147	GCTCGGAGGCAGAAAACTCAGGAGACTGGGGCTGCACGGAGGCAGAAAACTCAGGAGACTGG GGCTGCTCGGAGGCAGAAAACTCAGGAGACTGGGGCTGCACGGAGGCAGAAAACTCAGGAGA CTGGGGCTGCACGGAGGCAGGAG
178	152	TATCACTAGCCCAAACCTAAACCTAACCCTAACCCTAACCCTAGCTCTAACCCTAAACCAAGAA CTAACCCTAAACCTAGAACTAACTGTAACCATGAACTTAGCACTAACACTAACATTAGTCCTAAC CCTAACACAAATCCTAACCCTAA
179	178	AACCACCCCCCACCACCGCCTCCCCACTACTCCTCGCCTTGCAGAAAAGACAATGATTCTAATA CTGTCACGACCGCTGAGATATACTTCATGTAACACATGTTCCCTCGACCTTTACGGCACGTGCT AAAGAATTAATTCAGATATGTTTTGGGAATACATGTAAATGTGCCTGTCG
180	159	CACCTAGCCTCACCCCACCTCGCCCAGCATCACATGCCACGCCTCGCCTCGCCTTATCTCGCC CCATCTCGCTCCACCTCGTCCCGCCTCGCTTCCCCTCGCCCCACCTCGTTCCGCCTCGCCCCG CCTCACCCCGCCTCGCCTTACCCCGCCTAGCCT
181	92	TAACTCCGCTTCTCCGTGATCCTGCAGATCTTCCTCCACTTGGCGCTGCATCTGTGTGCAGTTG TTTTTGCCATCGCAGCCATTGTGCTGTA

182	178	GTGAGGACAGGAACATAGATTGACCAGTAAATTGAGAGCATTGTCTTTCGGCTCCCTCACCACA ACAGACTGATAGAATGACCACATCACTGTAGACACCGCACTGATTCACCTCTCAATCTCACACT CTCCCTCACTCCTGATCAAGACCCTGAGATATTGAGCACCTTTTTCCGGT
183	230	GTCCAACTAAATGCAAACTGGATGGGATGGCATGTCGCTGCAGGATGCTGTGGTAGCCATGCT GGTTCAGTGTGCCTTCAATTTTGAATGAATCCCCAACAGTGTCACCAGCAAAGCACCCCCACAC CATCACACCTCCTCCTCCATGCTTCACGGTGGGAACCATGCATGTAGAGACCATCCGTTCAACC CAAACCAAAGATCTCAAATTTGGACTCATCAGACCAAAG
184	110	TGGTATTACTTGCTTCCCCCGGGGAAATTAGGTCAAAAATTCCCCTTTCCCCGTTTCCCTGTATC CCTGCCTGCCTCTATCCCTGTCTATCTGTCTCTCTATCTGTCTCT
185	163	ATCAAATCCCAAGATTTCAGGAGAGAGCTTGTGAGTGAGGAAGTGTGTGTGTGTGTGTGCATGT AAATATGTGTGGTGTGTGTAGTATGTGGGGGGTGTGGTGTGTGAGTGTGCGTGTGGTGTGTGT AGTGTATGTGGTGTATGTATGTGTGGTGTGTGTGTG
186	259	CAGCGGATGGATGGATGGATGGACGGACGGACGGATGGATGGACGGACGAATTGCTAGATAT TGATGCATTGATAGATATTTGTTAATTATTGCCTAACACAAATAATAATAATTCCAGATTAATGGA ATGCTATTCGATATGGGATACTTTATTATATATGCTTGGACTGGGTAGGAGTGAACGAACGAAC GAACGAACGAACGATCGAATGAATGGATGGATGGATGGATGGATGAACAAACGGATGAA TGAA
188	249	GTCGATGGTGCATCATAAGAGGGGTGCCACGGGAATCTTGGAGTGGGGAGGGAATGGGGACT CTTCAAATGGAGAGGGAAACTCATGCAGGTCTACGGGTTAGGGTACAACATCTCTCCCTTATAT CACTCTCTCTCTCTCTCTCTCTCTCTCTCTAGGAGGTGTTAGCGTCGTTTGATAGGAGGAAGAT GGGATTGATGCCACAAGCCCTGCTTCTCGTGTTTTTTGAGTTCGGCTCTTGGGAGGTAT
189	159	GTCTCCTCAATTTATTTGGTCGTCTCTAGGCTGAGTCACTCACCCAGAGTTGACTCACTCACTG ATTTACTCACTCAGCCATCCAGCCACCCAGAGTTAGATCACTCATTCACTGATTTACTCGCCCG CTCACTCACTCACTCACTCACAGATCGGAAG
190	144	GGTAGGAAAGGTTCCACTGCAACTTTCACTTACATAACCTACTATACTTCTGTCAAAACATGTCT GCATGCACAACAGCACTGTGACTGGATTCACCCGGCGTAGACAGTGACAGAGGACTGCAGCT GCTGCTGACGTGGCGCT
191	117	GTAACCAAAACATAAGAGCCATGAAGATAGTTAACTTCACTTGAATGTACGGTTGTTAACATCAG CTAACATTGGCCACCAATATCCACTGGTCATACCAGACCACATCAGACTGCA
192	247	GTCAATGTTGCCATTGCCGCCATCAAGACCAAGCGCAGCATCCAGTTTGTGGACTGGTGCCCC ACTGGTTTCAAGGTTGGCATCAACTACCAGCCACCCACTGTGGTTCCTGGTGGAGACCTGGCC AAGGTCCAGAGGGCCGTGTGCATGCTGAGCAACACCACTGCTATCGCAGAGGCCTGGGCTCG GCTGGACCACAAGTTTGACCTGATGTACGCTAAGCGTGCCTTTGTTCACTGGTATGTGG
193	170	ATATTTGAAAAAAGGATCACAAGAAGGAAGGTATGGTTACTATAGTGGACGTTTAAAATATTTCT GCAGAATGCACTCCAACATCCATTATGATTTCTGTGTATTCACTGTTTTCTTTTGTTGACCTTTCT CTTAAAGCAAGTTGTGTGGCATTGAAGTGAAGTCGGAGG
194	206	ATTTTTCGGGCACACAGCTCATTTTTCAGGCACACAACTCACTTTTCGGGCACACAGCTCATTTT TCGGGCACACAGCTCATTCTTCAGGCACGCAGCTCATTTCAGGGCACACAGCCCACTTTTCGG GCACACAGCTCATTTTTCGGGCACACAGCTCATTCTTCGGGCACGCAGCTCATTTCAGGGCAC ACAACTCATTCTTCG
195	111	ATTTTGGTCAGAAATCAGCATTTTTAGTTACACCCACCCATGTTCTCCAGACAATTGCTAAGGAA TGGAAAGAGGTACAAACATTTTTTTTTTCTGATGATAAAGGAAAGT
196	118	CTTTTAGGTGACATCCACGGCCAGTACACGGACTTACTGAGGCTGTTCGAGTACGGGGGATTT CCGCCCGAGTCCAACTACCTGTTTCTGGGGGACTACGTGGATAGGGGCAAACAGT
197	131	CTGCTGTGATAGGGGTTCACACACACACACATACACAATCATAGACATGCACACAAACACATAC ACACTGATGCACAGGTGTGCATACACACTCACATACAGAAGCACAGACATGCAGACACATAGA CACA
198	173	AGGCCCTCCTCTCCTCTCTCGCTACCAACCTCTCAGTTGTCTCTCTTTCTCTATCCCTCTCTCTC TCAGCTCTCTCTCTCTCTCTCTATCCCTCCTCTCTAGCTGGAGGGAGATCCTCAACTCCATGAC GTCTTGGGGGAGGTCATCATATGAACGACGCCACAGCTCCTCCT
199	181	CACACACATTCCTAGCAGTATTGTACACACATTCGTAGCAGGTCCACACACATTGTAGCAGGTC CACACACATTGTAGCAGGTCAACAAACCTAGACGCGCGAAGCGCGGCTAGCTCAGCCGCGCA GCGCTAGACGCGCGAAGCGCGGCTTTAACTCATAAAATGACCGATTTCCGTTGTA
200	118	TCACTAAATGAGCCGATAAAGTGACGTTGAAATGATAAAGGTCAGAATTTAGTTCACAGATAAGA TCTTCCAGTGTAAATCATTGTTGTCAAATGGAGTGGAATGAAATGGCTGAGGA
201	248	ATCCATCCATCCATCCATCCATCCATCCATCTATCCATCCGTCCGTCCATCCATCAGACCGAAA ATGGATACTGTTTGAGGAATTGCTTGAGACTTGGAATGAGACACAGTCTGTCGACCTTCTGACC TCTGAGGTGACCTTGAATGGTCAGCCTGGCCAGGTGGCTGAGCCGTGGAACGCCGAGCAGGA GGGGACGATTCAGGACAACCGGGGTGTCGCACCGACCTGAACTTTCGACCAGCACGTT

202	155	TGGTAAGCAGAACTGGCGCTGCGGGATGAACCGAACGCCGGGTTAAGGCGCCCGATGCCGAC GCTCATCAGACCCCAGAAAAGGTGTTGGTTGATATAGACAGCAGGACGGTGGCCATGGAAGTC GGAATCCGCTAAGGAGTGTGTAACAACTCA
203	50	ATCCGCTAAGGAGTGTGTAACAACTCACCTGCCGAATCAACTAGCCCTGA
204	82	GGTTGATATAGACAGCAGGACGGTGGCCATGGAAGTCGGAATCCGCTAAGGAGTGTGTAACAA CTCACCTGCCGAATCAACT
205	222	GCTTCATACTTAAGATAATTAGCATGTATTAGCTTCATACTTAAGATAATTAGCATGTATTAGCTT TATACCCAGAATAATTAGCATGTATTAGCAAGTATTAGCAAGTATTAGTCTCATACCTAGAATAAT TAGTATGTGTTAGCTTAAATCTAGGATGATTAGCATGTATTAGCTTCATACCTAAGATAATTAGCA TGTATTAGCCTTATACCCAGAATA
206	109	GGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTCGATTGATT GATTCTGGATGATCTTCTGGGATTGACTGTTTATCACTGAAGTGG
207	98	ATTTCTACCAATTTCACCTGTGAAATTGCTGGTTCTCACATCAAAACATCCAATACTGTTAAAAAT TTAGGTGTAATCATTGACTCCACCCTTTCCTT
208	160	TGTTAACACAAATAGCTTATCAGCCAATCACATGGCAGCAACTCTGCCGAAGTTCAAACAGAGC ATCAGAATGGGGAAGAAAGGGGATTTTAAGTGACTTTGAAAGTGGCATGGTTGTTGGTGCCAG ACGGGCTGGTCTGAGTATTTCAGAAACTGCTGA
209	263	CGTTCTGTGGCAGTTTGCCTCCGTTTCTAAACACCTCTGAAACATTTATGTTAAAAAAAAAATATT GTCTGTGGGAGATGGGAGCAGCTGTGACATGTTTGTCAGCTGACCCACTTTTTCCTGATTATGA ATTAAAATTGAATGATGGCGTCTGAGATTGCCAATTAAAAATGCATGCAGTGCCACGGTCAGAG CTGAGAGGCCAAACGCAGTGTCGCCTCCGGCCTCCTACAGGAAACCATAATTCACTACTGGGC CTTTTA
210	171	GAGGCGTTCTAGCAACTGATGCTAGAGACTGCGGTCTTTAGCCTCCTTGTTAGAGGGCCCGAT TCCCATGCCAGCGGACCCGGGTTTGAGTCCCAGTTAGAACAGGACGTGTCACTTTGGTGCCAT GACCTGGATGGGAATGAGGTTTAGGGGGTAAGTGTAATGTACGTA
211	126	CTTCATCCTCGTCCTCCTCATCTTCATCTTCGTCATCAAGCATCGGCTCCTCATCATCAAAATCC TCCTCGTCCTCATCATCATCTTCCTCGAGGGTGTCATTCAGCGTTTTGAGAGTTTCGGTCT
212	252	GGCGACAGTCGCTCAAAGAAGCGTTCGATTTCAGTACTACTCAGTACTACGGTACTACTCTCTC CCACCACCCACCCATCCATCCATCATTTTGTTCTTCTTTCAATCTGGTCCTTTCGTCCATCTATC CATCCATCCGTCCGTCAGTCCGTCTGTCCATCACGGCCGCTAATTGTTCCGCTCGACCGATCA GGTGAGGAGTTAGTTTCGCACTCAGTCCCTTATCTCGTCTGAATTTTTGTTCTTCTTTCA
213	210	ATGCCAACACACACACATACACACAGACACAGACATGCACACACAGAACTTCACATGACATCAA TCTGTGTTGACTGTGTAGTGATCTCATTACAGACAACTTGCTGTTGATTAAAGCCTACTGTGTCT CTCTCATACACACACACACACACTTCTCTTTTCTCTTTTCCTCCTTTAGCCTGAGGCTCTCTCCA TTCTGCGGTTTCTCCT
214	127	AAAATAAATAAAGATCTAGACTTTAATAATAATGATAAATATACTAAGATATTAAGAAAATTAAAAT ATAGATTAATATCTTAGTATATTTATCATTATTATTAAAGTCTAGATCTTTATTTATTTT
215	122	CATTAAAATCATACCATTGACCAAGCACTTTTTGCCTCTGTGACTGTGGTTGACCAGTGGCATTC ATTTATGAGTTCACAACACATGAGAATCACCATGACTCTGTCGTACGTTTATGATGT
216	187	GCTCTTATTGGCCTACGTCCGTTACACTCACCCCCCTAAACCTCACTCCCATCCGGGTCATGGC ACCAATGTAACCCCTCCTCTTCGAACCTTGGTGGCCAGGTCAGGGGATCAGGTGAGTTAGCCT CCATTACACTGGCAACCACCCACCTCACCAAGCATTGCGGTGGCACGTTTTGCATGGGAA
217	264	ATGGAAAAGATACTCTTAAGATGAGACGGAAAAGTGTTTGGCCCAGTTTGTGACTACAGTTGGC TCTTAAATAATGAGTCTTGTCACAGGCATGTAGTTAGTACCCACAAGACCTGATCCCCTGTAATT TAGAATCAGTATTCCAAGGATAATGTCAGTTAGGATGTACACAGTAAACCTCAAAGGATATCCC CCGGTTATGCCTGAATATAACCTCTGCTATGCTTTGAATCTGCCTGTGATCAGCATGAGTATCTT TTCCAT
218	168	GATAGATAGACAGACTGATTGATTGATTGATTGATTGATTGACAGGTAGACAGACTGAATGCAT ATAGATTGTTAGATAGTTAGATAGATAGATAGATAGACAGACTGATTGATTGATTGATTGACAGG CAGGTAGACAGACAGAAGGCAGATAGATAGATAGACAGA
219	199	GTCAAGAAGGAGGTCAAGGAGGAGGTGAGCGTTTAAATTGCATTCTTCTCATTTCACCGCAGTG TATTGTTTTGTCTAAAATCTCAATCTCAAATCCTTAGTATATTGGCTCACTGACTAAATCTTATTTT TAATTTACTCTTCCTAGGATAAGGAACTGCGTGATGTTGGGGACTGGCGTAAGAATATTGAAGA CAAG
220	157	GACGTCTCAATGGTTGATTCGTTTGAACTTTGACCTACGTTGCATTGAACGGTTGAATGATTGAA AGCGAAACTGATTGATGGACAGGCAATAGGGTGAGGTAGAGGAGGAGGAGGAGGAGGAGGA GGGGAGGAACAGGGTGCTTTGGGGAGAGACT
221	276	GGCTATGGTTGCAGATATGAGATGATACATATCGTATCATATCACATCCATCAAGTCATTCGTTC ATTCGTTCATTCATTCATTCATTCATCCATCCATCCGTCCGTCCGTCCGTCCGTCCGTTCGTCCG AATATCAATACAGTCATCCTTCAATCCATCCATTCATCCGTCCAATCATTCATTTTTCTTAAAATC

		TCAAAGCAGTTTTACAGTCAGTTCAGAGTTTAAAATATGAAATTATTTCATTTGCGACGTATTCTT GATGTAATTTCAAG
222	188	CCTGGAGAGGGCCTCCGTCACCGGTGCCTCTGGTCGCCCGGTCGACCTCGAGGCGCTGCGC CCCGCCGCCGAGGACAGCGCTGCACCCGAGGAGAGCGCTGCACCCGAGGAGAGCGCTGCAC CCGAGGAGGGCGGCCGGGACGACGCCTGACGTCGTACGCCGTCGGCGAACCTCGTCGGCGC GTCGT
223	62	GTTCTCCTGCTGCACGGTATGGTTCTCCTGCTACACGGTATGGTTCTCCTGCTGCACGGTAT
225	105	TGTCGTGTGTTTGCAGAGGAGGCGGTGACTCGTGTGATCGATTTTAAAGTCTGATATTCGTCAA AATATTGTGTTTCAGTTCATATTTGTTGGCGTTTAATAACT
226	90	GTGGTTTACATGTTTTAGTCTACTTGAAAACAAATTTCTTATTGCAAACCTTTGTACAAGGAATGT AAACTTTCTCTTGCATATCAGTAA
228	129	CATGTGTGGAGGATCTGAGATTGCATCATCTATTTGTTGCCTGTAGATACGAAGCTCTGCGGTA GAAATGAGTTCCTCCTCAGGGATGCTGCTGAGATTGAAGACGAAGCGCAGCGGAGTCTCCTGT GA
229	266	ATCCTGAAAAAATTAATCACTGTTTCCACAAAAATATGAAGCAGCACAACTGTTTTCAACATTGA ATGTAGCCTTGGTGAGTATACAATACTTCTTTCAAAAACATTAAAAACATTTGCGGACTCCGAAC TATTGAATGGTAGTGTATATTAGGGTTGGTAACCGAGAACCGGTTCTCATCTGGAACCGGTAGT GTTTTTTGAAAAGAACCGGAACCGTGCAAGATTTCTAAGTTGCGGTTTCGAAAACGGTTCTGGT GCGATGGG
230	236	GAGCTGCCTGAACCGAAGAAACACATTGCATTACTTAAAAATACTTACATTTTCTTCATCTGCAG CTTTGCCTCGCTTTTGAAAACAAAATGGCAGCCCTGGGGGTGGAAATGTGCAGATTAAGGGGT GTTTAATATTATAATAAGATCCCCTTTTTAAATCACTAGGGGAGCGAAATCGGAACGGCTCGTTT TTTCACATGCTTGTAGAGAAAGGCTTGCCAAAACAAAGGTACT
231	225	GCCGTGGGCAGTGGCGGCGTCTTGCATATCAATAGTCGGGGACCCCTGGGCCGTTTCCCCCC CACGGTAGTGCCTGGGGGTGGTTAGGGAGTTACGTGGGTGGAATGTAAGCGGAGCGGTGCCG TACAGCTGTGCAGCAGTGTGGATGTTATGAAGTGAGGGAGCGCCCAGTCACGGCGCTGCGCT GTGTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGCG
232	202	AGATCCAGAGCACTCCGGAGCAGGTTTTCATTAAGGATGTCTCTTGACATTGCTTCATTCATCTT TCCCTCAATCCTGACTAGTCTCCCAGTTCCTGCCACTGAAAAACATTCCCACAGCATGATGCTG CGACCACCATGCTTCACTGTAGGGATGGTATTGGACAGGTGATGAGCAGTGCCTGGTTTTCTC CACACATATA
233	246	AACCTGCTACCAAGAGTCGGTCTGGCACCTGACGACGCATGATAACCTCTCAGCTCTCTGCCG ATACTCAGTGACAGACTGGAGGCTATCTGCGAGTGTGTTCGTGTGTGTGTGTGTGTGTGTGTGT GTGTGTGTGTGTGTATGTGTGTGTCAGCACGCTGACTGACAGATTAAAACAGGGAACAACAAAA TCAAACGGTGTCCCGCCCTACGGCTCTGATTTCCAATGACAAAGCAAGAAGTCTA
234	118	CGTCGAATTCTCGTTCTGCCTCGTAGAGAGGCTCGAGATTCCGTCAAAACCGACGTCTCCGGT GGGACTCGAACCCACAACCTCCGAATTGCTTTCGGATCAACCGCTAGAAGTCCAA
235	152	GATAGATAGACAGACTGATTGATTGATTGATTGATTGATTGACAGGTAGACAGACAGAAGGCAG ATAGATAGATAGATAGATAGACAGACTGATTGATTGATTGATTGACAGGCAGGTAGACAGACAG AAGGCAGATAGATAGATAGACAGA
236	193	ACGATAAAGAAGTTTCTAAGGGAGGAGAGGAATCACAGGACTGGAGACACACCTGTCTCTGGT CCAGGTAACCTTATTTCTGATGCCATAAATTACGCTTTTAATAAGAATAACTCCAGTATACTGCT GCAAGCCATTTGGACCATCCCCCTTTTTACTAACTAGAGATTTTCCACACAGCACAGCTCCACC A
237	235	ATGCTTAGCAAAAGAGTATTTTGACATTTTTGTGTGCTAGTTGACCGTTTAGGTGTGAATCCGAT GCCCATAGTATACTTTGCTTATGCACATTCTGCTCCGTTTCCGAGCTCACACACAAAAATCTTGT TTTCATAATCGTCGGAAGCCGAAATCAAAAATCAATTACGATTAATCGGACAGCCCTAATCAGCA CAACAATGATTGTGTGGTGTTCTTGTCCAATCAATTGCTG
238	235	CAGCAATTGATTGGACAAGAACACCACACAATCATTGTTGTGCTGATTAGGGCTGTCCGATTAA TCGTAATTGATTTTTGATTTCGGCTTCCGACGATTATGAAAACAAGATTTTTGTGTGTGAGCTCG GAAACGGAGCAGAATGTGCATAAGCAAAGTATACTATGGGCATCGGATTCACACCTAAACGGT CAACTAGCACACAAAAATGTCAAAATACTCTTTTGCTAAGCAT
239	136	ATCCGAACTCTAACTCTTACCCTGACCCTAACCCTAACCCTAAACATAAGCCTAAATTAAACCCT AAACATAACCCTACCCTAACCCTACCACAACACTAACCGTAACCGTAACCCTAAACCTAACCCC AACCCTA
241	129	AGCGTGACCTGCGAAGCAAAGAGACCCCAGAGAAGACGAACGTTGTGAAATCACCTGGACAG GCGGACCTGCATCCAGAGGGAGAAGACAGCAGAGAAACCGAAGAGCTTCCATACATGACAACA ACCG
242	116	CCTAAATAGCCAACCCAAGACTGACGTGGGTCATAGTCCACCTGTCGTATCTGTTTTTACACTA AATATCTGTAAACATATGGTATTTCAATCAGGACCACTTCCGTCAAGTATTA

243	271	GCTCGCTCGGCTCCACCAGAGACCATCATCCTTATGGCTCCAACAGGCTCCCTTGTCCCTCCG GCTACGCCTTGGTCAGACGTCACTCTGCCTGTGCCATGGACTTCTGAGCTGTCTGCTGCACTC CGTCTCTCCACCCCTTCAGGTCTGTCTGGTTCCGCCTCGGATCTGACAGTCACTGCGGCTTCG CCTCAGCCTCCAGGACCTTCGGTGTTGACCAGTCGTATCTGCTGTCCGTCTGCACATAGGGCT CCACCTCCATCAGTTTCAT
244	209	AACTAACAATCCAACTACAAGTGTCAAACCATCTGAGACAACTAACAATCCAACTACCAGTTTCA AACCATCTGTGACAACTAACAATCTACTCACCAGTTTCAAACCAACTGTGACTATAAACAATCCA ACTACCAGTTGGATTGTTAGTTGAACTGGTAGTTGGATTGTTTATAGTCACAGTTGGTTTGAAAC TGGTGAGTAGATTG
245	126	CCTGAATCATAAGCATCGAATATTGAATATTGCATTCTCATATTTAACAGTTCAGTAGCGGTGAG ATGTGATACCTGTGGGGGTTTGTGGAGGTCTCTGAGCCTCTTGTCCTTTGCCTGCGTTAAA
246	206	CTCTGATGTTGACCCTGCCAACATAGCGTGCTTCAAAGCTGCATTCACAAAGGACCTAAACAGG CAAAAGGAGAACTCAAACCTATGGTGGTTAAAGGTGGCAACCGCTCTAGATCCTAAGTTCAAAA ACCTTAGGTGCTTGCCCAGGGCAGAGATGGGAGAGGTATGGCAAAAGCTGAGTGAGATGCTG AAGGATAGAGAACCTG
247	220	CGATAGGTAACATTGGTTTTGCTTTGTTTCACAGAATATATGCTGTTAAAACACACAGTAATACC AAAACACACTTCACTTATTGATGACATCCTACCCTCGGCCGCCTGGCACATCAGATCCTGCCAT AATCATTGCTGGGGTTGCGTACGTATGTGTGGGGTGGCGCTATCAAAATAGGGACGAGACCCT TTTGGGGTAGGGGCGTGTTTGTTTTGGT
248	148	TCCGCGAAAAGGCGCTGTGGTTGGAAGGCGCTGTGGTTGGAAGGCGCTGTGGTTGGAAGGCG CTGTGGTTGGAAGGCGCTGTGGTTGGAAGGCGCTGTGGCGGAGAGGCTCAGTTGCCCTGGCC GACGATGTACAGATATCCGGCGAG
249	85	CTGAGGCCGGGGAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATGGCGCC ACTGCACTCCAGCCTGGGCGACAG
250	92	ACAGTCACATAAAATCTAAAGTTAATAACTGAAGGGTCAAGATCCCAACTAAAGCAAACACTGAA CACTATAATGGTGATACACAAATTGTG
251	172	CTTATTCTTGTCTTGGATAACCAGGCGCTTTCGGGCAAAGTAATCAGTTTTTCCCTCTCTCCTTC TCCTGAACTTCACCTGGTACCTCTTGAAGTAGGCCTTGTTCTTCACTACTTTAACGAATCCCATT TTGTAGACTTTTTAGTCTCCTAAGCCGGAGCCTGGAGTCAAA
252	128	GGCTACGTAGCTACCGCCGGCTACGGCTACGCCATGCAGCAGCCGCTGGCCACCGCCGCTCC AGGGACCGCAGCCGCAGCCGCCGCCTTCGGTCAGTACCAGCCCCAGCAGCTGCAAGCCGAA CGGAT
253	103	GGAGGGTCAGCATACGGAGGGTCAGCATACGGAAGGTCAGCATACGGAAGGTCAGCATACGG AGGGTCAGCATACGGAGGGTCAGCATACGGAAGGTCAGCAT
254	144	CCCCCTACACCCAGATTTTTATTAGACATTAGTCTGATTCCTATTCCACAGGCGTTTCCCCTTCC ACTTTAACACTCAGCCACACCAGAGCCAATTTTAAAAAGGCTCTTCTGCCTAGCCTGGAAAAAA AAAAAAAACCCTTGG
255	138	CTAGCAACTAGCCATAATGCCACTCATAGCTAGCAACTAGCCATAATGCCACTCATAGCTAGCA ACTAGCCATAATTCCACTCATAGCTAGCAACTAGCCATAATGCCACTCATAGCTAGCAACTAGC CATAATGCCA
256	148	GTCGAGAATGCCGAGCTCGCTCTCTGCCTCGTCGACGCAGAAGACAAAGTAAAGGGTCGCGTA GTGGCGGTAGATGATTCGTAAAGGTTCTTGCGCGGTGTCGGGGTCCAGATCAAGTGCGCGAC CCTTGCCCTTGAGCTGCTGCTGG
257	93	GACAAGATGCACAGCATAGCCATATTCTAGGTAGACATGAGTTTTTTCTTTAACTATATATATATA TATATATATATAGTCGTAGTGGTCTTG
258	271	TGATGTTGATCTGAGAAACCCTGTTTTATTATAAAAAATAAAGCATTTTTATATATAGGGTGTGTT CACCCACCCACCACCATCCAACAAGCCATCTGATCACACATCCATTCCTTTTACTATTCATCCAT CCATCCATCCATCCATCCATCTCAAATACACCTGAACAAAGCCAACAATAAGAGAGTTCAACAA GTGTGCAGGCGTCAGTCTTTAGCTCCTTCTTCCTCTCACTGTATGCATGTTAAAGAGCGTCAGT ATATCATCTTAA
259	148	TTTGTCGCACAGCAGGTAGATTTCTTCACCACCGGTAACGCAACCAGCGGTACGGTCCATACG AACGATTTTCAGGTTAGAAGCGTTCGGAGCTTTAGAGTCGTAGATAGCGTCAGAAACAACCGGT TCCAGACGACGGGTGAAAGAA
260	82	GTGGTGATGGGGGTGGTGGTGGTGATGGTGGTGGAGGTGTGGGTGATGTGTGGGCAGAGGT GGAGACACGGGAGACAAAGGT
261	177	CTTGTAATTCTTGGCAAACATGAACCCACCCTCATTGATTCCTGGTGTGTTTGTGTTGTAGTTAT TCAACTTTTACACCACCACCCTGCAGAATTCTTTGTGTAGACATCGTTTCATCTGTGATTTGTCC TTAAAAACAATTCTACATTTTGAAAGCTGAATACAGATGTTAAAATA
262	86	TCTTAGCGTAGCCTGAACAAAACCCTGAAAATTAATTCTTTCCTTCACCCCCCATCATCCCGTCA CCTCCTCCCTCTTTATCTTCT

263	137	CATGGATTATCGGATTGTGCAAATTTAGCACCACAACCACCACAACCACCGCAACCTCCACAAC CACTACTACTTATAGAGCCTCCACAGCCTCCACAACCACCGCAACCTCCACATCCACCACAGCC ACCACTTGT
264	122	CATCTTCTTGCTTTTAGGCTAAAACAGAAGCACCCACAGCCTCCACAAACCTAGAGGAAGCAAT TTTTTAGGTTGAATGTTGACAGATGCATATTGAACTTCATCTTCCTGCAGCACATTCT
265	205	CCTGCTCAATCTTTCCGCCTCCATACGGATGTCAATCATCTGCTGGCGGTCCTTGAAGCTTCTC GATAAAAAACGTCCCTCTGGGGTGTAGAGGGAGCGAGTAGGAGTGAGGCACCGGATGTTGCG GACCTCCTCTGAGTCACTTTCGCCACCGACGGGACCCTCACGTTTGCGATGCGGCGGGAGCC GAGAGCAGTCATCGTCG
266	248	AATATGCAATCAGCTTGATGGTTCAGAACATAAAACACTCTAAAAGGAATGTAAAAGTGTTTATT CATTCCCCATTTCAAGATCGCTCATGCTTGTCTTTATCTTCCTGTAGGAGCTAGAAATGATGAGA AGGAGCGGCAAGTGGAAGGTATGGAGAGGAACAAGTCTCTCCAAGATTACAACGAGTATTACG ATTACTCAATGCAAATGCACCGCAAACTGTTCTCTGGCCAGAAAATTGGCAGAAT
267	185	TTAAATTAGAAAACAACCAAAAATTCTGTTTAAAATTAAGACTTGGTTGTTAAATGTTAAATGGAC CATGTCCGTAGTTAATGTACTACACCTGTTGTTAGGACACTTCTGTAAACTTGAGGGGAATCAA CTTCATACAGCTACTGTTACACCATTTGAATAAAAATAACAGCAAAAATGACTGA
268	134	CTGCGATCAATCAGAGAGCCGCGAGTCAGTGGAATTACTGGATCGATCACTGCGATCAGTGTT TATCAATCACCCACAACAGGCCATATGGACCGGCCGGTACGTCACAAAGTCCCCGGATTTGTT GCAGGAGA
269	148	CGTCTGCAGACGGGCTACAGAAGACGGGTAGTATAACCGGACCGAACCCAGGTCCCCAGCTC TCATTACACCCGGATCAACGCATCGCTTACAGTCTGTCGGTCGGAACGAGCTCGGTTATGGGA TATTAACTCAAAAGAGGATCGTG
270	180	GATGTGTTAAAAACGGACGCATGCAAGAAGACTCGTCTGGAGTTTTACATCATCAATCTGTGAC ACGTTGGGTGACTGTGAGCTGCATTTACATGTGGTGACTAGAAAACCAAATTTCTAGTGTTAAC AGGGTGTGAGAAAAAAAGTGAAGCGGGACTCGATTGAAATGTGAAAAAATAG
271	110	TAAACCCCCCCAAGGGGGTTTGACTGGGGGTTTGACTGGGGATTGACTGAGGGTTTGTCTGAG GGTTGACCGAGGGTCTGACTGGTGTTTGACTAAGGGTTGACTGAGGG
272	108	ATCACCCTCACTCTCCTGTCCTCTGTGCTGTCCCTGTCTTAGGTCACGTGCCTTCAGGACTTCT TTGGGGACGATGACATCTTCCTCGCATGTGGCTCTGAGAGGTTT
273	237	CAGCAAAATAAAGTATGGCATCGCAATAAACACAATCAAACCGGACAAAATATAACCATCTAGC CATTAAAAACACGAAAAATCAAGGAAAAAACGTCAGACAGGCAAATCTCGTGGTCTCGCGAATC CAGCAGCTTCGCTTCTGAAATGCTTCTGGTGTGAGTTGACACCGCACAGAGGACGGAACAAAA CCTTGTCGGAGCCGTAACACACCGCAGACGCATGCAGTCTAAACAA
274	184	GATTTGTTAACCGTCGTAATGCATTGGTTTTGTTATTCTCGTAGTCATCATCACCATCATCATCAT CATCATCATCATTACTGTAACACGTGTTCCAACTCTTCCAAATGTAAGGCACATAAACTCTTCCA AGGCCACGTGACTCCAGCAGCCTTGCCTTTATCATTATCATCATCATTATCGT
275	167	AGTTTTACGTTAAAGATGATAGATGTGTTATTAAAGTCTCAATTTTTTCCTTCAGGTCGCTAATCT GATCTTGCAAAGCTTCTGTCTGTGTTTACTGGATGCTTCCAGAATCAGCCCTCCTAACAGGTAC TGAAAACCTCTAAATGTAACTTTGGTTTCCTGAGATG
276	113	CACAGAACTTATTTTACTCTTAAATTCAAGGAACTCATTAAACCCCATTAAGAAAATCTTGAGGG AACCAGGGGTGGGGTTTTGACATCAGACTTCCTCCACTCTATTTCTTA
277	65	AGTACTAACCAGGCCCGACTCTGCTTAGCTTCCGAGATCAGACGAGATCGGGCGTGCTCAGAG TG
278	117	TAGATGGTTCGATTAGTCTTTCGCCCCTATACCCAGGTCGGACGACCGATTTGCACGTCAGGAC CGCTACGGACCTCCACCAGAGTTTCCTCTGGCTTCGCCCTGCCCAGGCATAGT
279	113	CTGCTGTCTATATCAACCAACACCTTTTCTGGGGTCTGATGAGCGTCGGCATCGGGCGCCTTAA CCCGGCGTTCGGTTCATCCCGCAGCGCCAGTTCTGCTTACCAAAAGTGG
280	100	CATGCGGACCAAGAAATAATCAAAATCTCTTCAGGTGAAAATGTTAACACAGAGTATGTATTGCT ACATGATCATGAGGGGGTCTCAATGTTTCCTCCAA
281	73	TGTAATAAAATACCTCACTCACCGGTCAAATAATAAAAACTCCATCATCACCGCATCACTTTTAC AACATTTT
282	149	ATTCGGCAAATTGGATCCGTAACTTCGGGATAAGGATTGGCTCTAAGGGTTGGGTACATCGGG CCCTGGTTGGAAGCCGCTGATGCTGGCTTGGACTGCTGCGTGGGAACATGCGGTGGACCGAG CCGGCGTCGGCGTGTGGACGGCCA
283	150	TTGTGCAACAGCGATAAAGTCGGTAGGAACAAGACGGGTGCCCTGATGCAGTAGCTGGTAGAA AGCGTGCTGACCATTAGTTCCAGGTTCACCCCAATAGATTTCTCCGGTGGAGGTAGTAACGGG GCTGCCATCGACACGAACAGACTT
284	121	ACGCAAGGCCTGAATTAGTGACACCATTTTGTTGTGCAGTCCAGGAACAGGTCGTCAAGACAA GGAAGGTTTGTCTGTGACGAAGCAGAAGACGGAGGACACACGATGACGTCACACCTTC

285	150	TTGTTGGTAAAGGCTCTTTTTTTAAATAAGTCCCAAATGGGTTTGTGATCGGTATGCACATTAAC CTTATGACCTAAAAGTATGTGTTTAAATTTCTTTAGACCCCCATATAAAGCTAAACACTCTCTTTC TGTTGTACTATAGTTTCTT
286	100	GTGAAAAAGCATTCAGTTTAGAGATATCAGCTACTTAGGATATGAATAGTCAAACAGAAGAACG GATTTCAGTACCTGGTATTTCTAATATTCATGACTT
287	150	ACGCAAGGCCTGAATTAGTGACACCATTTTGTTGTGCAGTCCAGGAACAGGTCGTCAAGACAA GGAAGGTTTGTCTGTGACGAAGCAGAAGACGGAGGACACACGATGACGTCACACCTACAAAAC GAAAAGCCCCCCGCAGGACCACAC
288	150	ATACAACACAATTTAGATATATTTTTTTATTTTTATGATCTTTATATGTATGTAAAAATAACTATGCA CAAACATTCTACATGCATGCATCTACTTATCAATATATATATATATATATATATATATATATATA TAATATTTTAGATAT
289	109	CAGCGGCTGCCCGGGCGGCGCAAGGAGGCCTGCCGACGCAGCGGCCTCGTCCAACTCCTTC CCGTTGTATCCCACGGCTCCTTCGACCTCGGACTCCGCCCCGGCGCCC

Table S3B. DNA fragments matching NHP's sequences

ID	Len	DNA Sequences
290	150	GCCCACTTCCGCGGCCACCGGTGCAGGGCATGCCCCCGGGCCGTGGAGGTGAGGGAAACAG GCGCGAGTCGCGGTTTAACAACCCCCCCCTCCCCCAAGGCCCCGCCGCGCAGCCAGGGCGG CAAACTTTAGCGGCCGGTAAGGTACAGA
291	150	ATCTTAGAGACCAATGACATATAGGATCTCAGCATTTTGTAATATTGAAAGGAGCAGGTAATTTT TCAACTAAATCAAAATACATAAAAGCTTGCTCATGAAATACCACCATCTTAGACAACACAGTTCT CTTGGGTGTCTTAAATATGA
292	150	AGTCGCTCCAGCTCTGAGACTTCATCTCCCTTCTGTAATGTGAAGGATCATGCTTGTCCTCACA GGATTACAAAGTCAAGTACAGACATACTGCTTTTTGTTTCCATAATGGTCTCCTGGAAACCACAT TTTGGGTCATATTAATGATTT
293	150	TAACCGAAGAATCACAAAAGAAGTGAAAATGGCCTGTTCCTGCCTTAACTGATGACATTACCTT GTGAAATTCCTTCTCCTGGCTCATCCTGGCTCAAAAGCTCCCCCCCTGAGCCCCTTGTGACCCC CGCCCCCCCCCCCCCCCGACCA
294	149	GAGCTGAGATCACACCATTGCACTCCAGCTTGGGCCACCGAGTGAGACTCCGTCTCAAAAAAA AAAAAAAAAAAAGAAAGGAAGGGAATTTGCCAAACCATTAATAATAATGCATTAATGCATTCTTT GTAGAATCAGGGCAGTGCACA
295	109	CCAGGAAGTTTAGAAAGTTGCTTTAACCAGGGCACATGCTCTGGAGACTGTGTGGTTAATAATG ACATTTAGTACTTTCTTATTCTTTACACCCTCCTCACGGGGCATT
296	150	AGAAAGCACACTGTACAACTTCAACAAGGACAGGTCAGTTTTCTTGTTTAAAACCACAGTTGGT GCCGGGCATGGTGGCTCACGCCTCTAATCCTAGCACTTTGGGAGGCCGAGATCGGAGGATCA CAAGGCAGAGAGCAAGTCACGACA
297	150	AATTATCCAATGTAATAAAAAAATACTAAAAACTGAAGCTCAACAAACTCCATCTTGGATAAAAAT AACAATAACAAGACACAATCAAAGCAAAGATTTTAAATTCAAATAAAAGAAAAGAATCTTCTTATT ATTTAATTATTTTTTTTA
298	74	GTCTCACCTGTCACCCAGGCTGGAGGGCAGTGGTGTGATCTCGGCTCACCCCAACTTCTGCCT CCCAGGCTCAA
299	119	ATAAATATTTGAGGTGATGAGAACATTTGAAATCCTCTCTTTTAGCTGTTTTGAAATATACATTAT TTTTAAGTATAATCATTATGCTGTGTCATTGCACAGCATGGTGACTATAGCAA
300	72	GGTTAGATCACTCCCACTGCACCAACGAGCACAGGGTAGGTCTCAGAGCAAGTCAGGGGAAGT GATCTAACC
301	150	GTCTCTAGCACGTCTTTATCTCCAACCCTCAGGGACTCTCCTGTGTGTCCCAGCTACTCTCCAA CCACGCCCACAATTCAGCGGGGGTCAGATCCAGGCACCAATCAAAAAACAACCAGATCCCACA ATACACACACTTACTCTCCAGCA
302	150	CGATTCCATTCCATTCCATTCCATTCCATTCCATTGCAATCGAGTTGATTCCATTCCATTCCATTC CATTCCGTTCCGGATGATTCCATTCCATTCCATTCCATTCCATTCCATTCCATTCCCCTG CAGTCGGGTTGATTCCATT
303	131	CCTGTAATGCCAGCTATTCAGGAGGCTGAAGCAGGAGAATCACTTGAACCCAGGAGGTGGAGG TTGCAGTGAGCCGAGACTGAGCAAGACTCTATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGAC ACGC
304	148	CTTTTTTTTTTCCTCTCCGCGCCCCCGCGGGCCCCTCCCTGCTCTGCGGGCGCAGAAATTCCA TCTCCAATTCCTGTTCCCAACATCCCCCAACCCCCCCACCCCCACGGTAGTACAAACCCCAAAC CGGTAGCAAAGATCGTGGGGT
305	87	ATTTGTATTCTTCCAAAATTCCTATGTTGAAACCCTAATCCCCAGTGGGATGCTGTTTGGAGATG GGGCCTTTGGAAGAATACAAAT

306	150	GGCAACAAGAGTGAAACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAACCAAAAAAAAGATA CCCAGTGGTTGGGGTGGGGTCTAAAGGGAGAAAGATAAACCAAATGGCAGATCCATAATGAAA GGCGTTGTGTTTTCCTCACATT
307	150	ACTCTCCAAATGAAGTTTCCACCTATCAATCCCTCTCCACTCAAGGCAGATGGTTCTTAGACCAA AAAAAAAAAAAAAAAAAAATGTCCCACCACCCACACAAGCCCCATAAATTATGTTGTGAATTAAA AAACAATTACAAGTAGGTTA
308	146	TCATAGTCAAACTGCTAAAATGCAATGATAATGAGAACATCTTGACAACCAGAGAAAAATAATCC ATTTCACACACAGGAACAATGATATGAATAACAGCTGCCTTCCCTCTCACCTCAGCCGCCCCGG AAGGTGGGTCGAGGGTC
309	149	GATATATACAAGAATATTCATAGCAGCTTATTTCACAACAGCTCCAAACTGGAAACAACTCAAAT GTCCATCCATATGGTAACAAGATCAACAAACTGTGGTATATTAACACTGGTATACTAACACAGTG GAATACCATACAGCAGTAA
310	150	AAAACTTTTCCTATTTGAGCCATAAACCCTTCCTTTGTTCTTCGGTGCACACACTTTCATTTGTGC TTCCCCAGTCTGCGGGAGTTTTTTTTGTCTTTTTTTTTTTTTTTTATGTTTTATATATAAAAAAACT CTTAAATTTTGATATTA
311	150	TTACCGGCGTGAGCCACTGCACCCGGCCTGAGTCTTTTTTGTTTTGTTTTTTTTTTTTTTTTTGAG TCAGAGTCTCTCTCAGTGGCCCGAGCGGGAGTGGGGTGGCTCAAACTAAGCTCACAGCAAAAT AAGCCACCGGGGGTCAAGAAA
312	150	GAATAATCAATCTTTTGTTTACGAAGCACTTCCACAATTATTACTTTTTTTTTTTTTTTAACAAATC CTCCTGTCCATTTCACTATGTAGGCAAGGCAAGTATTGAAATTATAATTATAATTTCATTGATGAG GAAATATCGTAAGTGCA
313	150	CACATTTTCTCTAAACTCATATAATATTTTACTTCTTTTTCTCTGTTCAAGCCAATTATTAAATACA TAGCATAGGCGATACCTGGATGCCCAAAGATGGATGCCACGCTCCCTGCCGTGGAGGGGGTC TCAGTCTATGCTATGTATTTA
314	149	CCAGGATGGTCTCGATATCCTGACCTCGTGATCCGCCTGCCTCGGCCTCCCAAAGTGCTGGGA TTACAGGCGTGAGCCACCGCGCCCTGCCCCTCATCCCGTTTCTTTTCCTCCCGCCCGCCCCCC ACAACCCCGGTCGCTCGCTGAGC
315	147	GGCTGGGCATGATGGCTCATGCCTGTAATACCAACACTTTAGAAGGCCAAGGCAAGCGGATCA CTTGAGGTCTGGAGTTTGAGACCAGCCTCATCAACATGGTGAAACCCCGTCTCTACTAAAAATA CAAAAAAGCATTAGCTGGGC
316	89	CGGTCCCTGTCCTCTGTTGCGATTTGGCAGGCAGCTCCAAGGATGTACGGAGGCACCCCTCCC TCCCCCCAAACCCCACCCCCGGAGG
317	94	AGCAAGTTTATAACAATAAACTTGACAACTTTGGTGAAATACATCAATTCCTTCAAAAACAAATTA CCAAATCCAACTAGTAGTTGGATTTGGT
318	150	AAACTCCGTCTCAGAAAAAAAAAAAAAAAAAAAAAAAAAAGAATGTATTTTGCGGGGAAGGGGTAA TTTGAACATAGAGTAACAAGGCAAAAATACAAGCAAATGCAAGGAAGCTGGAAAGGATACACAC CGGAGTTGCGGGAAATACATGA
319	150	ACTGTATTTTCCAGGGAAAAGAATGTCTAGCATCTCCAAAATCATCTCTCTGAGCTAAAATGGCA TCATGCAAAAGGAAACTTACTCATTCCCTGGGAAATCCAAGACAGAGACCTTTCTGAGAGCGAC CTGAAGTGAGGTTTGACAAGA
320	150	GGATGTGGGCATTTCAGTTTCCCAGGCCCCTCTATTTAGGAAGAAAACTCCAGGCTCCACCCA GATAAATACAGACTCCTCCTTCCTACCTAATCCAGGCTCAGGATGGGAGGCCCAGCCTCCACA TCCAGAACACCACACTACGTTCAC
321	112	TTTGGCAGTTTCTAATTTTCGAAGAATTGGTCCATTTCATCCAAGTTGTTGAATTTACCATATGTA TATACAAAGTATATCTTTAATATCCTTTCACTGGGAAATTTGAAAA
322	62	AAATGAAGAATGTTAGATGTCTCCACTGACCTACTGAGCCTACATCTCAACATTCTTCATTT
323	150	CTGGGATATTTTGTTTTTCACTATAGGCCTCAGTGAGCTCAAAAATGTCCATTTGCAGATTCTTC AAAAATAGAGTATCCACAAAGCGGAAACAGACGAAAGCTATAACTATGACAGATGAAACCAAAC ATAAAAAAGAAGTTTCACAGA
324	149	GAGACTCCATCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAGGATACAAAGAGAAAAAAGGACCAGCA AAAAAGAAAAAAAACCAATGATCTAAAAGATAACACAGGAATATTCAACAAACAACTTCAGTACA AAACAAAAAAGCAGGTAAAA
325	141	AATGTTTCTTTTTATTTATTTCTTTTTTTTTTTTTTTTTAGACAAGGTATCACAAATTCCCCCAGGCA GTAGTGAAGTTGCACAATAAAGGCTCAAAACAACCATTACCAACTGCCAAAGCCACCCAACTAG CTGTGACTAC
326	150	TAGGTTAAATTGTGTCACCACCACCCCCACCCCACACAAGCACAAAAAAAATTTAAATACAAAG CCCAAGTACAAGTGAAATTAAAATAAATTTTAAAATAGGCTTTTCAGATGTAATCAAAATTATCAG AGATAAGAATGGATAAGGGT
327	150	TAATGCATCACTGAATCAGGAAATGTTAGGAGCTTTGCTCACACATAGGTAAACCGTGTAACTT CTGGAGTCATTGAAAAATAATATTTCTTAATATCTGACAAGTCGTCAGCTATGCACTTAGGACAA GCCTCAAAAACACTTGTTTGA

328	150	GGAAAATGCAGTCTCTGGGGTAATTTTATGAGAAATACAGCTTCCTTGAGTTAACAATGAAGAG CTTTAAAATCAAATCTCATGTTGATTTCATCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGATATTG AGTATTGATATGGTTCC
329	150	CTTTGATCTTCTTCCCCTTCTTTTCTTCCCCTTCCCCTGCACCCAGAATAGGGCCTGGCAGAGG CTGGCAGGAGCCAGCGCTCTGTATCCCCTCTTACCTTGGTAATGATGTCCTTGAACTGACCCTC CTTCCAGGCCTCAGTGGGAGAT
330	150	ACAGGAAGGGGAACATCACACACTGGGGACTGTTGTGGGGTGGGGAGAGGGGGGAGGGATA GCATTAGAAGATATACCTAATGCTAAAGGAGAGTTAATTGGTTCAGCAAAACAAAATTGCATAAG TAAAAATAAGAAACGAATAGCAAA
331	150	GATTCAGAGGAGATGCTTCCCAGAGGGCTCAGTCCTGGAGTCAGAGATGCTTCCTGGAGGGCT CAGTCCTGGAGTCAGAGATGCTTCCTGGAGGGCTCAGTCCTGGAGTCAGAGAAGATGCTTGGC GGGCGGCGAGGTCGTGTAGGCGGA
332	79	ACACTATGGGGACTCCAACAGAGCCATACCTTCCTGTCTACGGCGGTTGGACCTCCTGGCTCC GGGCACCAGCTTCTCC
333	145	GTGTCTAGGGCTTCACCAATTGTATGTTCATTCATTCAATCTTGATGGCTTGCTTTCCTCCCTTC CTTCTTTCTTTCGTTCCTTCTAGCCTGCCTTCCTGCTGTCCTTCTGTCCTGCCTACGGACCCCG CGTCGCCCGACCCCGC
334	148	ACTTAACTATGGAAAAGATATTTGTTGGTGGCATTAACGAAGACACTGAAGAATATCACCTAAGT GATGATGTTGTACAGTATGGAAAAAATGAAGTGATTGAAATCGTGACTGACCAACAAATATCTTG TCGATAGTTAAGTAGAGC
335	149	GAATGGAATGGAATGGAATCAAACCGAGTGGAATGGAATGGAATGGAATGGAATGGATTGGAA TGGAATGGAATCAAACCGAGTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGCAATGG CATCGGATGGACACAACACGGCA
336	149	GTTACGTGTAAGGGTGATCAAAGATTTAAACTGTCATGTCGACAGTAAGCCGAGGAACAGGCT GTGCAGGGGCCTTAATAAATCTGCCCTTGACTTATGGCCGAACCTCATTCAAGATCGGAAGGA GCACACGTCTGAACTCCAGTCAC
337	150	AGGGAGTAATTAAACCTTTTTTCCCACATGATAACAAGTATGATCGTTCACACTTGATATCATGT GGGAAAAAAGGTTTAATTACTCCCTAATTGTCTGGGAAACAAAAGATTTCAAAATACGGAGCTC CTGGGGGAAGGAGAGTTAATA
338	150	TAACAGATTCAGGATTGAATAAAAAAGAAGTTATGGGCATAAGTAGTGCAACATGAAAAGGTAC ATTACAAACAGTGTCACGGGTGGAAAAGGGGCAATGCATGGTAGCTCCTAGTTTCCAAACGGG TTTCAGAAAGTATGTCACACTTT
339	150	CAAAGCATAAAAAAAATAAAAAAGGTAAAGAGTACTAAGATCCAGAAAATTAACAAGTAGTAAAA TTTTTGTTTCAGTCTGTCAAGATTATCATTCCATTAAGCATCCAAAAATTTTACTACTTGTTAATTT TCTGGAGCTTAGTAATCT
340	96	TAGTCTTGATAGTTTGTTTTTTTATGAAAATTTGTGCATTTCATTTAGGCTAATTTATTGGCATGTA ATTGTTTATATAATTCCCTTATGTATACC
341	125	TTATTCAATTCAAAATTGATGACCTTCTACTCTGTAGTCCCTCCTTTGAGTCTTCTCAAGCAGATA CCCTCCTGCTCCTTCAACATTTATACACTGGAGAATAAAAGTTGAAGGAGCAGGAGGGT
342	84	AACTTCAAAGCAAAATGAAGCTCTTTTGGTTGCTTTTCACCATTGGGTTCTGCTGGGCTCTAAGA TCTCAGGCCAGATTCTCCT
343	143	ATAACTGTTTTGTCTTTCCCTCCCCCGCTTCGTTCCCAGGTTTCGGTCTAAGTACCACCCAGAT GAGGTGGGGAAGCGTCGGCAGGAGGCCCGGGGGGCCCTCGAGGCCCAGGGGGACCAGGAG GTCCTGGAGGGCCTGGGAT
344	150	TATCAAGGAGTGACTGATAGTAATATTCTTTTTTTTTTTTTTTTTTTTTTGGAGATAGCATCGCACTA TGTTGCTCTGGCTGGAGTGCAGTGGCTTTATCTCGGGTCACAGCAAGAAAGGAAGAGCACACC TAGGAAAAACAGTAAAGACA
345	150	CTTGCTTGGTTTCTTTCAGAAATTGTTGCCAAGGTCAGAGAAGTCTCCCAGCCCGACTGGACGC CCCACCAGAAGTCACGCTTGTGTTGACCAAAGAGAACTTTGACATCATTATGAATCTGTTTTACT CTTCCCAGTGCCTTGAAAAAA
346	105	ATACATGGGGCCTTCAGAGTTCATGGAAAAAAGATTATTATGAAAAAAGTATGCATGGATTTCAA TTTTTTTGCACCAAAACAAATATATCTCATTTCCGTTTTC
347	148	TGCATCTTTTATGGCAGGGAGTCGTCGCATGGTGGATGTCATGGACGTGAACACACAGAAAGG CATTGAAATGACCATGGCTCAGTGGACACGCTACTATGAGAGGGGGTGGCCTGGGCGGCTCC TGCAGCCTGGAACATGGCTCGTT
348	150	GACAGAGCGAGACTCCATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGATATAAAAGCAAAATTAA GGGGGCAGAGGGGTATTAAAAGGTGAAATATGGGGAAGAAGATGGACCAAAATGCGTTGGTTT GATATACAATAACCACAAAACA
349	118	AAGTGTCTGAAATGCCGTGGCAGTGGAGACACTTGATTTGAGATCTCCAGTAGCCTCCAAATAT TTGGTGGATGAGGGTGGTAGAAGCAGCAGAAGGGGATGGCCTTGGCAAGGTGGT

350	150	AGAAAGCACACTGTACAACTTCAACAAGGACAGGTCAGTTTTCTTGTTTAAAACCACAGTTGGA GCCCGGCATGGTGGCAAAACCCTCTAAACCCACCACATTGGGAGGCCCAGGAGCGAAGAACA CCACGTCTGAAGTAAGTGACCAAA
351	150	CCCCATTCTTCCCTCTCTCTGTCCTCAGAACACTGCCTCATATCCTTCCCTGGTCCCTGGCTCT CTGAGTCCCGCGGTTTTTTTTTTTTTTTTTTTTTTGTTTAGAGAGATAGGAAGAGCACACCGATG AACACCACTCACGACCAGAGA
352	150	CTCACACCTATAATCCCAGCACTTTGGGGGGATCGCAAGCACTTGCTTCCAAACTTGGATCAAT GAATGAGCACTGCTTTTTGCAGACATAGAGCAGGGCTCTAGATAGAGTCTACTCAGTCCTACTC TGGAGTAGCCCATAGTCCTATA
353	108	TTTTTAGTGACAAAACATAAAGGTACTGAGAAAAGAGAATCCCCTTCACCAGCACCGAAGCCTA GAAAGTAGAGACTAATGTTGGAATCTTCAAATTCTAATTCCTTT
354	149	TCTTAAAAACATTTTTTTTTTTTTTTTTTTTGGCCAGGCATGGTGGCAAACCCCAGTAAAACCAG CAATTTGGGAGGCCGAGGCAGGCAGAAAAGCTAAGGTAAGGAGAATAAAGGAAGAGCACACG GATGAACACCAGGCACGACAA
355	148	TCCAATCTGCTATCTGCTTGTGCCTGAGTTTGTGCAATATCCTCCTCATTGAGCTTCCTGCCTCT TCTCTTGCCTACCCCCACGACCCGCCCCCATGTTTTTTATTCCGCACGGGGGGGGTCGTGTGT TTTATGGGGGGGAAGGTGTA
356	92	AGGAAGTGCTGGTGGATTATTACATCGACCCGGCCGATGCAAGCCCTGACCAAGAGATCAGCA AGATTAGCCTGCAAGCCAAGAGCCAGCTC
357	100	CTTAGTTTTCTATATACTCATTGGCCAATTAAGCAACCACACAAATCAATGCTTCACTACAATATC TTAGGCCTAGACTTGAAGTTAGGATTTCAATTTT
358	150	CAATCCTTCACTCGGCTTTCTTTTTCCTTCATGGTACTTAAAGCTGATATTAGACGGTCATTACTT GTCTGTAAGTTCCCGCAGTGTCTTCACCTACGAGATCGGAAAGAGCACACGTCTGAACTCCAG TCACGACAAGAGATCTCGTAT
359	149	GAATCTCTTCATATAAAATCAAGACAGAAGCATTCTCGGAAACATCTCTGTGATGTTTGCATTCA AGTCACAGTATTGAACAAAAAAATTCATAGGGCAGGTTTGGAAAAAACGTTCGTTCGATCGGAG GAGGACGCGAGTTCACTCCA
360	114	TCTCATAAACTGATGAGGCTGTGATATCAGGCTCACTTGACAGTCGTCCTCAGCTGAGAAGGAA GCCAGCACCACTTCTGTGCGTCCGTGAATCTCACATTGTGGGTGCTGGCT
361	149	AGAGTGAAGAGGCTTGGAGGATCTGATCATCACATCACAAGGGAATTCATTTTCTGCCAGTTAC GTGGTCCCCACAGGATGCCGAGTTGCAAAGCCCACAGATTAATCCGGAACAAATAGCGCCACA AAAAATAATTTCACCAGGTCAT
362	150	TTCCTTCTAACAGACAGGACTCTCAGATGCAGGTATGTAGGAGTTTTCTAGAGGTCAAATCAAG AACCTGTTTGCCTGGGTATCAGCAGCGGTGTCTGAAGAAAATAGGAGATCGGAAGATCACACG AATTAACTCAATTCACGACACGA
363	150	AGGCCTTTTAAATATACCTATAGTTTGGATATCCTCCATCTTTAATTAAGCTGACTTTTAACCGTC ACATTTCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGAAAGATTATTTCTTTGTTGCCAAGGCTG GATTGCAGTGGCTTGA
364	118	AGGCTGGAGTGTAGTGGTACAATCTCGGCTCACTGAAATCTCTGCCACCCGGGTTCAAGGGAT TATCCTGCCGCAGGAGGCTGAGGCAGGAGAATACCGTGAACCCGGGGGGCAGAGA
365	150	AGACCAGCCTGGGCAACATGGCAAAACCCTACCTCTACAAAAAATACAAAAATTAGCTGAGGGT GATGGCACATGCCATTCCCCAGCTGGGTGACCTTGGGTTAGGGCCCTCTCCTGTCTCAAAGAG CAGACACGCCACATGCTGGGTCT
366	150	CGGCGGGGCCAAGGGGAGCATGAACCGGCACGTGGCGGCCATCGGGCCCCGCTTCAAGTGA GGGCCCTCTTCCTGGGGAGCACAGGGCCCCGTTCTTGCCGATGCCCATGTTCTGGGACACAG CGACGATGCAGTTTAGCGAACCAACCA
367	150	AGAATGTAGCGCCTGCTTCCCCCAACCCCCAACAGATTTCCTGTTTTTTTTTTTGTTTCTTTCTTT ATTTATTTTTTTTATAAAAGCATACAAAATAGAATATAGGAAGAGCACACGTCTGAACAACACTCA CGACAAGAGATAGCGTAA
368	86	TGTCCCTTATTTGCCCTGGCCCTTCCCTGGCTCTGCCATACCCCTTCTCTGGATTTGGATGTGT CCTGTCCCTTATTTGCCCTGGC
369	150	TCCACTTGAGGTGATTCCATTCCATTCCATTCCATTCCATTCCATTCCATTCCATTCCATTCCACT CTGTTTGATTCAGTTCCATTCCATTCCATTCCACACGATTTGATTCCATTCCATTCTATTCCGTTT CATTCCATTCCACGCCAT
370	150	GGGATTACGGGCGCCCCCCACCTCGCCTGGCTAATTTTTTGTGTGTATTTTTACTAGAGACAGA GTTTCACCATCTTGGCCAGGCTGGGGTGGGGGTGGAGCGCGCCCAGGGGGGCCACACCACA CCCACCCACAACGGAGGGGGGACCA
371	150	TGACGTTTCTGTGTCCAAGGCCTGAACTGAGGTGTCCGCGCCTTGTAAGCGTGCTGGCTCTGT TTCCAGTTGGAAGATATAGCAGACAAGATCGGGAAGAGCACACGTCTGAACTCCAGTCACGAC AAGAGATCTCGTATGCCGTCTTCT

372	125	CCTCATTCTTTTCTAAGACATTACCCTGCCACAGGGAAAATTAAGAAAGAAATCATCCTAACTCA GTGAGGATGATTTCTTTCTTAATTTTCCCTGTGGCAGGGTAATGTCTTAGAAAAGAATGA
373	150	GGAGAGATGGACCTGGGTTGGATTTTATCCATTATCAGTCTTTCGATAGTTTGACTTTTTTTAAC CATGTGCACATTTTGATTTGAGGTTTTTATATGTTTATCAAACCAATATGTGCACATGGTTAAAAA AAGTCAAACTATCGAAAGA
374	150	GGATCTCATTATATTGCTCAGGCTGGTCAAGAACTCCTGGCCTCAAATGATCCTCCCACCTTGG CTTCCCAGAGTGCTAGGAGTTCTTGACCAGCCTGAGCAATATAATGAGATCCCATCTCTACAAA AAATTTTTTTTAAAGTAGCTGA
375	150	ACAACCTCATCTCAAGACCCTTAACTAATTACATCTGCATTCACCCTATTTCTTTTTCTTTTTTTTT TTTTTTTTTTTTTGTTTTGGTTTTTGTTAGGTAGAGGACAAATGTGAACATACAGAACGCAAAGAG ATCGGGGTGTGCGGAGA
376	150	TTGTGCCAGCCACAGCGCTAGACATGGGCATACGGCAGCTTTACTGCTGCTCAGCTCCTGTGT GCCCATCCTGGAGCAATTGTTCTTTTTTTTTTCTTATTATAGATATGTTTTTAAAAATTTCAATAGC TTTTGGAGTACAAGTGGTTT
377	150	AAAATCTATAGAGACAAAAAAATTAGTGATTGCCTAGGGCTGGGAATTTTTTTTTTTTTTTTTTTTT TTTTTTTTTTTTTAAGGAGTAAAGGTGTGTAGCCCAAGGATGAAAGAAAATGTGAAATAAACGCA CAAGACAACAAACGCCCC
378	148	GATCGGAATAGCACACGACTGAACACCATACAAGACAGTACATCTCCAATGACAGCGTCTGGT GTGAAGGGTGCGGCGGCAGAGGGCGGCGGGCTGGCGAGTGGAGTCATGGAGAGAACGCGG GACTCGTCGCCTGGGGCCGGGCGGC
379	78	AGGAGCTGCTAAGAGAAAAACAGAAGGAGCAGCAGCAAATGATGGAGGCTCAAGAGAGAAGC TTCCAGGACCTGAAAC
380	116	TGGAGCTCAGCGAGGGGAAGGTGCTTTCCTTACCTCTAAACAGCAGTGCAGTCGTGAACTGCT CCGTGCACGGCCTGCCCACCCCTGCCCTACGCTGGCCCAAGGGGTCCAGATCT
381	132	CAGGACTTTATATTGCCTAAGGTTAGGAGAACGGAAGCTCTCTGGTTCCAAGCTGGGCTCTGA GCGGTAGCTGGGGTGACGGCTGCTCTCTGCAATTGAGGTTGGCTCCTTCAAGCTGTCCCCACG ACTCAA
382	150	CTTTCTTTCTTTTTTTTTTTTTTTTGTCTTTTTATTATATATTGGGTTTAACAGTGTTTGCCAGGCT GGTCTCGAACTCATGACCTTATTTGATCCACCCACCTTGGCCACACCGAGAGCTGGGATTAATC CTTGAGCCACCGCACTCAA
383	150	TGAAAAACTGGTCCAGCGGGCGGCGGACAAAAGTTTCCGAGCGCGCTCCACAGCTGGGCGGC GGAGCGGGCCCCGGCGGCCGGCCCCGCCTCGGATGCTCCCGGCCTCCTTCCTGTGCTCGTG ACGGCGCCGAGGGCGGCCCCCCCCCCC
384	131	AACTTAATATAACAAACAAATTGAAATAGCTTATTAAAAGTGTCCGTATATAAAAATGGCCTTGTG ATGTACAGTAAAATGCAATAAAAATTATCCTCCTTTGTTCCGGCAGCCCACCCCCCGGACCGAC A
385	150	ATTACAGGCGTGTGCCACCACGCCCCGCCAAATCAGCAGTTTATGTAAAAATTCTTTAACCAGG GAGGTGGATGTTGCAGTGATCTGGGATCGCACCACTGCACTCCAGCCAGGGTGAGTAGGGCA AAACTCATGAGAAACTGCTCCAGC
386	150	GACCGCCAGCAGCCCTGTACTCTCAGCCAGACCCCCACGTACCAGGCGGGAGATGAAGATTC CTAGAACCCGCTCCAGGCCCGGCTCAGCCACCCGCACGCACCTGCAATCACGCATGAACAAG CCGAGGTGGCGGTCTCAAGCATGCTT
387	79	GCACCAACGTCCGGAGAAACCTGGCCTTCCACACACTCAGCCAAGAAGTCCTGCTCAAGGAGT TCTCCACCACCTTCTC
388	150	AGGAGCTGGACAGCCTTCTCCACCTCCTCATCTTCTCCTGGAACTCCTTTGTCTCCTGGAAACC CATTTTCCCCTGGAACCCCCATTTCCCCAGGGCCCATAGTCCCCAACACATCACCCCCAAGGC AACCCCACCCCCCCCCCGACACC
389	74	GTCTCACCTGTCACCCAGGCTGGAGGGCAGTGGTGTGATCTCGGCTCACACCAACTTCTGCCT CCCAGGCTCAA
390	95	TCTTCAACTGATGCAGCTACATCTGAGTCCAAGGAGACCCTTGGCACTCTGCAATCCTCACAAC AGCAACCAACACTCCCAACACGTCTAGACCT
391	120	GGCTGCGGCCCGGGCGGCGGCGATGCTGTGGCCGCGGCTGGCGGCGGCCGAGTGGGCGGC GCTGGCCTGGGAGCTGCTGGGCGCCTCGGTGCTGCTGATCTGCTGGTACTCGCAGTCCTT
392	150	CTGAGACCAGGTCACACAGTCCTCCCATCTTCACTTCTTTGATCACCACCACAGAGACCACCTC ACACAGTACTCCCATATTCACTTCTTCGATCACCACCACCGAGACCACCTCACACAGTACTCCA AGCTTCACTTAATCGATCACCA
393	89	GTGTACGGCACCCTTGAAAACAAGCAGGCGCTAAGTTTGAATCTTGAAACTTCCTAGCTTTCAG CCACTTCCTCAGCTCTATAAATTCA
394	44	AAGTTGCCGAGTGATGGGGTGGGGGGACATTCAGACGGCAACTT
395	97	ACCTTTAAAGGTTTGGCAGTTGCAGAGGAAGGAGCCGAACACGTTCACGCAGGAGCCTCCGTT CTCACACTCCTCTCGTTCGCACTCATTGATGTCC

396	80	GCGGTGTATGCCTACCGCCACCAGATTCATCGCCGGAGCCATCAGCATATGTCTCCTCTTGCT GCCTCATAATCCTTTAT
397	89	CCTCAAGTTCCGGAATACTCACCTAGGCAAGAAAGGATCCGAGATCTAAGTGGCAATCTTTGG GAGCGTTCCAGGAGGCCCAAGGAGGT
398	150	ATGCGGATGGCCATTGAGTCGGCAGGTGAGACTCCAGTTCCTCCAGTTAAATGCTGCGTTCGC CCCGTGGGTCAGGGTCTTCACCGGTGGGTTCACCAACTCCCATTTTCCTTCCTCTTTATTTTCTT CACAGAAGCACATGCCCAGAGA
399	149	GTGGGCGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAGGATAATGGCGTGAACCCGGGA GGCGGATCTTACGGTCACACATCTCGCCACACTTCACTACAGCAGAGGCGACAGAGGGAGACT CTGTCGCAAAAAAGAGGGGGGTTA
400	90	GTCAGACCTCCAGGACTAAGATGAAGTCAAAGAGCGACCACAGCTCTCGGAGGTTATTTTGCA TCGGTGAGCCAGACAGAATGATCCGAT
401	150	ATGCGGATGGCCATTGAGTCGGCAGGTGAGACTCCAGTTCCTCCAGTTAAATGCTGCGTTCGC CCCGTGGGTCAGGGTCTTCACCGGTGGGTTCACCAACTCCCATGTTCCTTCCTCTTTATTTTCT TCACAGAAGCACATGCCCAGAGA
402	86	ATTAAAGGATGTTCTTGAAGAATTCCATGGGGCTATTTTTTGATCTACTGAGGGAATCTGATGTT CTTGGAGAACTAACTGCTTCT
403	126	CCTACTGATGAGTATTCTTTCTTCCTTCCCAGGATATTCTCCAGGGAAGATGCCATCCAAATGCA GGACTCCAAGATTGTATGACGCATTGGGTTCCCCATTTCTTCTGCTTTTAACCAGTACTTT
404	81	TGGACGTGGATAAATCACTTTGTGATTTGGGGTTCTTTAGCCTTCTATGTATTTTTCTCATTCCTC CCCAGAAGAATGAGA
405	103	CCATTCCTAATTCTCCAGTCCACGATATTGAGTTCAACAGCAGCAAACCACTTCCACAGCCAGT GCCACCTAAAGGGCCCGTCTGGGCTGGGTCTCCAGATGT
406	116	CGGAATCAAACACATTATGCTTTGCAGGCATCACTGAAGCTACTTGATTTTTATGAAAAGTACTT TGATATCTACTATCCACTCTCCAAAACTTTTCCTCCCTTAGAAAGTGGAAA
407	149	GATCTTTTATTAAAGTTATGGAAACTGAACTAAAAAGGATTTATAAATTTTACAAGATGGATTTGT TTTCTCATGGTGCCCCTATCTTTTCAGAGCAGATATCTTCCCTCCAAAAGGCAAAGAAACCAACT TGCAAATCTAGTCCTCAA
408	102	AGATGGCTCCTGTGGGCTTGTGTACGAGCATGCTGCAGCGGAGGGGCCCCCTATTGTCACCCT TCTGGACTATGTCATTAGATTATGCTTGGCTTTCACTGA
409	76	TCCTGCAGGGTGATTCTGGGGGTCCGCTGGTATGTCAGTTCAAGAAGAATGGTCAGCAAAAAG CTGACACAGATGA
410	150	GAATGCACTTTTGGTTTTTGGTCATGTTCGGTTGGTCAAAGATAAAAACTAAGTTTGAGAGATGA ATGCAAAGGAAAAAAATATTTTCCAAAGTCCATGTGAAATTGTCTCCCAGATCGGAAGACACAC GACTGAACTCCAGTCACGACA
411	83	ATTCCCCAACTGGCATTGACTTTTCTGTATTAAGCTTGAACTTATTCATGCAATCTTCTGCTAAGT TAAGATGGACAACTTGC
412	132	AATGCTGATCCAGCCCCTATTTATGTTTGGAGTCCTACTGGGATCGGGGTACTTAGGAGTTTGT GAGGAAGGAGGAAGGCTCTCACTTTCATCTGAATCTGAGGATGAGGTATCTGTTTCTATGATTT CCCT
413	118	GGAATGGACTTTGTTGCCCAACAGAAAATGAGAACTCAAACAGAGGAGCTACACTATAAATACA CTGTATGTGATAAAAGCTTCCACCCTCTCCTCTTCCACCCTCAGTGGATGATAA
414	149	CTCTGCCCAGGCCCGGGCACATGCCCCCGCCCAGCCGGCGGCGACAGCAGCAGCCAGACCA GTCCCTGCGACCCCACCTCGGAGGCAACTCCATTTCGGACCTGGGACTTCTGACGCGGATCCT CAGCTACTTCTCACCTCGGGGCTGT
415	78	GAAGTTTTGCTATTACACAACAATGACCAGAACGCAGAGGAAGATCAAGATCATTTCCACGTGC ATGGAATCGTGGCC
416	109	GGAGCTGAAAAGATTGCAGGATTAAGCCAGATTTACAAAATGGGAAGCTTGCCTGAAGCTGTTG ATGCTGCCAGGCCGAAGTCAAAGCGGCATTGCTCGTTCATGGAGT
417	137	TACGAGGGCCTTTGATGCTTAGTGGAATGTGTGTCTAACTTGACTCTCTTCCTCTTTGGGCCCG AAGAATTCATAGGTTCATTTTTAATTTCTTTCTCATAGACATAGTCAGAGGGTCTGAGTAGTTGA CTATGAAA
418	148	CATTCCATAGTATTGATGGTAATTACCACTGGTTCTGAACCCAGTTTTGTCCATGGTACATGAAT CCTCAATTCATGAATATGTCCACTTAAAAAAGTGAATGGTAATTACCATCAATACTATGGAATGC ATTTTGAAACTTAAGGAT
419	150	GCATAGAAACATGATCGAAGTGATCATTTACTTTATCATGTCATAAAACACAAGTGCTTATTGTA GAACAGTGATATAATTTAAATGGCAAGGAAATTTGTTATGAATCACAAATTTCCTTGCCATTTAAA TTATATCACTGTTCTACAA
420	111	ACCAAGCTCAAACCTGTGGGTGACAAGTCTGGTTTCTCTGCAACACAACTGAACCCGTGACCA CGTCTCCTGTATGGACAGGGACTGGGTCGTCCATCATGAACAGCGTCT

421	85	CCATCTGCAGGTACCTGGTGGACGTGCGGCAGGAGCAGTGGCTGCTGCCCGGAGCCGTACC GAGGCAGTGGCATGGTCAGCTGCT
422	140	CACCACGCAAAGTGATGTGTAAGTGTGGGTGTTGCTCTCTTGGGGTGAGTGTGGGGGAGAGA GGGTCATCCTCACGAGCACAGGCATCTCGTCGTCCGACATCGAGCTGGCTGGCTTGTCTCTGG CGGAGGGGGCGGCAA
423	150	ATATGATGGAATTGACAGCAACTTTGAACCTGAGGTTTGTAAAGAGCCATCACTGACCATGTGA CAAATTTTGCCAAACGCATCATCAACTAAGCGTATTTCTTCATTAGAAGAAGGAATTGGGACAAT GCTAAATTAAAAGAAAAGAAA
424	150	GAAATAATCTGGTTACCATCTACTCCTTCATGAGTCTGTTCTAAAGTCACAGGGATTCTCTGGCA GGGGCGTGTTGGGAATCTGACCGCCCATGTGCATCCGAATATGTTGCTGCAGCATCACGGCAT TAGTGAACTTCTTCTGGCAGAT
425	134	CAGGAACAGGACTCTGCTGCCCAACACAACCTTGACCTATGACACACTTTCCATAGGCCTTCCT CACCACATCACAGAACTCAGAGAAGAGGCTGTCTTCTTGTTTGATACACAGCTCGTCTCTTAGG AACACC
426	150	AATTTCTTCTAGGTAATTGATCTGGGTGGTGAGCCAATTAAAAGCAGTGACTACTTTGGTAGGA GGGATGTCTGTAGAGAAAAGGTCTATTTCAGTATCAGCCAGCAAAACAGCCTTCATTCCCCTAG AGCTGAGCACTTGTTTACAGAA
427	150	GCTGCCAATTCTGCAACAACTCAGGTTCTGATTGGGAACAACATTCGATTAACTGTACCTTCTG ACTCATACATCCTCTTCTGCACATCTTTTAATGCATCTTCTAATTGCTTTATCTGTTCATCTGAAT CCTCCTTGACCCTTTCACAT
428	150	CCCAGGAGCGTCTGGCAACAGCTTTGCAGAAGCTGGAGGAAGCTGAGAAGGCAGATGTAAAA GGAAGCAGGTTATGACAGGCCATCCTATATTCTTCATATTGGGAACTACAGTTCCCAAAGCTGC CATCAAGCAGTGCTTCAGAAACCT
429	93	TGGTGGCCCATCCTGTTTTTTAAAGTTTCAGCTGTAAAACATATCCATGGAAAATACTTCACTCC AAAGATATCCATAATATTGGCCATCGTA
430	76	AGAAGTCAGCTTCTAAGTTGAGTATGTTGGTGAGACCAATTTGGGGAAGTATGAAGGTTTAAAA AAGGTATGCAAT
431	85	GCTCCCAGCATGTCCTCGCTCAACTCCTTGGCCTCTTCTTGCCTTTCAGTCTCTGGCGAGCAAA GTCCTCAAATACAATGTCGTC
432	111	CCTGCCCAACTGGTCTCCACCTACCACCACCTGGAGTCTGTCATCAACACAGCCTGTTTCACCC TCTGGACCCAGCCTGACCCGAGGCTCCTCCACCACCTCCCACAGCTC
433	150	GGACATATTTGACAGTGATTGGTACACTTCTCGAAATCTAATTGGGGGCGCTGAATCATTGTGA TCAAATACAACGTTAATGACAAGTTTTCATTCCATGAAGTCATCATAGAGACAGCGTTGACGCTT GGAGAGCCTGCAGCGGATAAC
434	104	ATGTGCAGGCCATGACTGCTCCCTGCGTCTCTGGAGCCTGGACATCCACAAACTCCTCAAAGA TCCTATGGGCCTTAGAGACCAGTTTTGCAGTTGACCTGGTC
435	68	AGTCACCCCCAGCCTGCAGGAACCAGGTGGAGGCTGATCAGTGGCTTGTCCATCTCTTCTGGC ATTTT
436	137	CTTGGTCCTGAGCAGCCAACACACCAGCCCAGACAGCTGCAAGTCACCATGGATCTAGCTCAG GTCATTCCTCTACCCATGGGCAACATGGTTCTACATCAGGACAGTCATCGAGCTGTGGCCAACA TGGAGCTAGT
437	70	GAGGCTCAAGAAGCAGAATGTCAGGCGTTGAGCAGCTCCATTTCCTGCTGGAGCTGTAGCCTC AAGGCCT
438	88	ATCAAAGTATGAGGGGAAATGGCAGAAAACAGTATCAAGATTCACCTAATCAAAAGAAAAATATA CAACAAAGGGAACAAGAAAGGTA
439	113	ATCTCCTCGGCATAAAAAAAGTGATTCTTCAGGTCAGGAGATAGAGTAGTCTTTGTAGGAGATG GGGACGCAGTACCAGGTGAGTCAGTTCCTCCATCGGAAGTACTTGGTGA
440	116	ATTTTGTTTTAGCTTCCTCCCTCCCAAGGCTGACTCCCTGGATAGTGGAGTGGTGAGCACCATT GATTTTAGTGTTACTGTGCAAAGGTGATAATGAGAAAGAATTAGTTCTGTGT
441	94	AAAAGTTTATCCAGGAAAACATGTGAGTACTTCTTTGTATCTTGTCTGTCAAGAGCAGCTTCAAC AGCATCAAAGAACTCTTCTTCATTAATCA
442	141	TGGGCCTGGCATGTGTCTTGTCCCTTTGGAACACAGGTGGGTCTAGCTGGAGTTCTCCTTGAC CTTGTGTGACATCCCTACAGAACCGGGATCCATTTCTTGTCCTCGTTGTAGAGGAAGGACTGCT TTCCTGGAAGCTCT
443	150	TGAAACCTGTAGCAAGCCAGGCTGGACTTGGCTTTATTGATCAGCTGTCCTGCATGGCTTGCTA CAGGTTTCACCAGTCAGTGCTTGCAAAACCTTTGTCATTTTCTAAAAATAAAAAAGCTAAAAGTT ATTGTAGAAGAAGGGGCTTTT
444	150	AAAAGAGAATGTAATCTTTTTTCAGCAAGGAATGCTCCCCGCCATGAGTTTTGATGGGAAAATTA TTTTGGAAGAGAAGAACAAAGTTTCTATGGCTCCAGGCATCAGTGGAGAGCTGATGAAGGAAG TGCAGTAGCTGGAATCAAGCTA

445	150	ACCAGGTGACTGATATTTTCCAGCATCACATCTCTGTACAAGGCCCTCTCGGCAGGCACCATGATGGCCCATTCCTTCGTAGTGAAGTACACAGCCACGTCCTCAAAAGCCACTGGGATCGGAAAGAGCACACGTCTGAACTCCAGTCAC
446	143	GACACGTGCTGCTGAACTTTCACTTGCTTGTTAATGATGCTGTGTTACAGGCTCTTCTCACAGCCTTCCTGCTGAGCACACACACACAGGCCGTATCTATCCGGATAAACCGCCAGGCAGCCTGCTTGCCATCCATGGTCAGC
447	73	AATGGATCTCTTTCTCCACCCATCCCGAGACAGGGCTTGGAGATGCTCATCTGTTAATGGCTACTCAAAGTGT
448	149	GGTTACCACTACAGATGGCACAGGGTATTCTGCACAAAAGGTAAGAGCTAGGGCAGTTGGGCTAGCTGCTCCAGTTCCTGCTCCATATGCTCCTGCAGCTCACTGGGCCTCAGCAGGACCTGGCAGATGGGGCAAATTGGGGCCTGGCT
449	57	ATCTACGAATTCTGAATGAACTTCCTACCTTCACGGGCACCGTCCGCCCATCATGCA
450	149	ATACTTTGGTTGTTCCTTTGATCTGACCAAAGTGAAGGATTCCAGTTTTGAACAACACAGTGTCCAAATAATGGTCAAGGATAATGCAGGAAAAATTAGATCGGAAAGAGCACACGACTGAACTCCAGTCACGACAAGAGATCTCGTAT
451	73	GAAAGAGGAATGTGGTACCTTGAAGGCAGTAGATCCACTGCAATCCAAAGACGTCAGTCCTGCCTGGAAACCA
452	96	AGGACCAACCATGGGCCTGGGGAAAACGTTCAGCATTGGTAAAATTGCTGTAGCTGTATTAGAAGAAACCACAAGAGAGAGGACAGATGTTATTGA
453	58	GAGCCATAATGTCGACCTGCCGGAGTCGAGGATTCAGGATAAGAGAGGAGAGTGTCTT
454	127	CAGAGGCAACCGTTTCTAAGTTCCAGCCACTCTTCATAGAGCTCTCCACCTTGGCTGAATCCTGCGTGGGGATGTCAGAGAGGTTTAGTGTCTTAAGAAACTCATCCTTCATGCTCTGGAGCAGTGT
455	150	TACACTGTCCGTGTCACCTGCCTCTACCCTGGGGGTGGCTCCTCTACGCTGACTGGCCGGAAGTCCTCATTGATGTAGGTGAAGCTGGTGACAGGGACAGCCTTCCTAGGGGGCACAGCTGGGCACAGCGGCTCCTGTGCCTCATAGCCA
456	150	GCCCAGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTGGTCTCGATCTCGTGACCTCGTGATCCACCCGCCTCGGCCTCACGAACTGAAGAGAAAACAGCAATAATCCACACACCCACCCAGCTCGCTTT
457	120	ATGTGTTTCCATTGATTGTTTTCTCATTTCTTTGAAGACATGTGGTCCTCTGTGGGCTGTCCAGCAGGAGGTGCCTGAGGCTGCTGGGGATTGCCTCCTGCTGGAGGTGGGGGACCTTGA
458	150	ATTATTTTCTTTAGTATGATTCTGACTGAAATGGAAACTCATGAGGATGCATGGCCTTTTCTACATTCATTCTGTCTTTAGTGAGTAATAAACTGATGTTCTCATGCTGTAATGAGCTGGCATGAGTATTTGTGCCACATGGCTCCACAT
459	88	CCCCAAGGAGAGGAGCAGTCCCATCATCATCCAGCTGCTGTCAGACAGGTGTCTGCTCCTTGCTCAGGCCTGGGGGCAGTGCCTTGTT
460	106	CTGCCCTTTTGTTAACAGGAAGGCCTAGGGAGTCAGGTAAAAAACGGGGATCTACATGGGTGTTTTCTGCTTCGACTTTCAGAGTATAAAGTCTTCGGCTCTCATA
461	122	CAGACCCCGTGGAGGCAGGGGAGGAAGTTAAGGTGAGAATGGACCTGCTGCCGCTCCACATGGGCCTCCACAAGCTGTGCAAGGGCTCCTGGGCTTGGCCCAGTGCCAACCCCAGGGATTCC
462	81	TATTCGGATGATAGTCAGAGGGACTCTTGGCAGCTGTAGCAGCAGCAGAGAAAGGAAAGCATTCATTCAGCGTTCATTCCA
463	139	CCCAATGACCACCTCAGTCTCCTCCCTAAGTGCCTCCCAGCCACCAGAACCTTCCCTTCCCCTAGAACACCCCTCACCCTCACCCACCTGGAAGGCCAGGGCCAGGATGCTCCTTGCTGCAGGCACATCCCCTGCCAGC
464	64	CGATGTGTGCTGAAATCTTGCATTTGTTGCCTTTGTCCTGGACTGCCTCAAACACCAGGTCATT
465	150	GGAGAAGGCGGTGCTGGACGAGCTGGGCCGACGCACGGGGACCCGGCTGCAGCCCCTGACCCGGGGCCTCTTCGGAGGGAGCTGAGGGCCGCGCCCAGGGCGGCAAGCGCTGGGGGAAGCCTACTCCGCATCCTTGGCCTCCGAGCTAGA
466	118	CAAGCATGTCCCCGGGCCAGGCAAGCCGGCTCTGTCTAAAAAGAACCCCGGCGTGGTGGCCTACCAAGGAGGCTACTGGAGATCTCAAATCAAGTGTCTCCACTGCCACGGCATTTCA
467	149	CTTCCAAGAGCCTTTCCTCCCCATATCCACTTTGATGGCTGGAGTCATGGCTATGATTTCTGGTACTCCACTCTTGCCAAGAACCCTCTGCACATGGCCTGGGTTCGGGTAATCAGCTGGGCCAGCTTCTCATCTCGCATCTCCTCTAA
468	106	GGACCAAAGTGGTTCCCTGCCCTGCCTGCTCCTGGCCAAGCACTCTCAACCCCTCAGTGACCCACTTCGAAGCACTTCTGTTCACCTTGTTGATGATACAGTTTTT
469	72	ACAATTGAAAAGGTTTCCTGTGAAATGAGATACAAGGCTACAACTTACGGGTGCAGGGCATAGAGGCAGCAT
470	125	ACGCTTCCAGAACGCCTGCCGCGACGGCCGCTCGGAAATCGTAAGTCGGCTGGCCCGGACAGGAAATCCCGCCCACAGGGCAGGAAGTCCCGCCCGCAAACGCTCCAGTCGCCGGCGTAGCGGTT

471	89	ATGGAGATGCCCAGAGTGCTACCCTCAACTTTCTACAGGCCGCGGCTGCCCACTTTGTGACCT GGGAGAAAAGAGGGTGAGAATGTTTA
472	150	CCCACAGTCCAGCCACAACTTTTAGGTCCCGCCCACAGCCGCCGGCCCAGGCTCTGGCCACG CCCCCCCCCCGCCCCCCCCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCACAC AAAAACAAACACACACAAAAAAACACA
473	150	TCTCTTTGTCGAAGCGGAGGTTGTGGAGTCTCAGCAGGCAATGGAAAAGGAGTTAGGATTGAA TTTGATGATGAAAATGATATAAATGTTGATGATCTTAAAGCAAGTTCTAAGATTTTCTGAACTTAG TTCTGCCAAAACAATAAATGT
474	150	TTCTCACCCTGTCTGCTTTTACTTTTTCTCTTTCTTCCTCCCTTTCTCCTTCTCCCCTCCTTCTTTT TCATTATTCTGTTCTGTATTCCTTTCCCCTTCATTCTCACCCTGTCTGCTTTTACTTTTTCTCTTTC TTCCTCCCTATCTCCA
475	137	CATAATTCTGCCTTCAAACAAGCCATTCAAGCCTCTCACAAAAAGGTATCAAAAATGGCTTGTTT GAAGGCAGAATTATGGTTGATATTTTCACTTCTCTTTGAGGAATAAAAAATTGATACACTGAAAG AGCTCCC
476	91	TTTCACCTCCTGGCCAGACCACGGTGTTCCCGACACCACTGACCTGCTCATCAACACCTACCC CAAGTCTGGTAAGTGAGGAGGGCCACCC
477	121	CCGGCGGCTAGGCGCGGGAGAAGTGCGGAGGAGCCATGGGCGCCGGGAGCTCCCTGCGTC CCCCTCAGCCATGTGGCCCCTGGACCTGCCCTGCAGCCGCCGCCTGCCAGGTCCGGAGCGC
478	150	GTCTCAAATTCCTAGCCTCAAGTGATCCTCCCATCTTGGCCTCCTAAAGTGCTGGGATTACAGG CATGAGCCACTGTGCCCCGCCAAGAGTATAATATTTAATTATACTAGATTTTATTTATATATCTTG TGACACTGAGGCAGTGATGA
479	120	TGAAAAGGCATTATGTGAAGGCGAAATAATTGATAAGCGAAGAGAGAGGGAAACTCTAGAGAA GGCTGAAAAAAACGCCTAGAGGCACGTGGAGTAGCCCCAGACAATTGAGTGCAGCAA
480	149	GCGGTTTCCTGCTTGCTGATTGCGTACGGAGCAAGTAAACCAAACGGTGAGTGTCCTCTCCCT CCATCTTCTGTCAGGTTCAGCTTCCCATTGGCGAACGCCCGTCTGGTGAACTCGCCTGCCTCC GCCGGTCGAAGCCCTGGCACGCT
481	149	AGAGATCGCGCCACTGCACACCAGCCTGGGCGACAGAGTGAGACTCTGTCTCAAAAAAAAAAA AAAAAAAAAAAAAAACAAAAAAAAGAAAGCAAAAGACCACGACTGAACACCACTAACGACACGA GAGGAAGGATTGCGACTGCTGA
482	107	CTCACTGTTGGCTCTCTTAACTGTGTCCACTCCTTCATGGTGTCAGAGCACTGAAGCATCTCCT GATTAATTTTTAGTTCTTCTCCTGGTCCACTTTCCTTATGAAT
483	150	GGCTCCGGCTGGCCCTGAGTAGCCGAGAGGATGAGCTGGTCCGCACGCAGGCCTCCCTGGA GGCCATCCGAGCTGAGAAGGAGACCTGGAACGAAAGGGCACTTTGGACTGGAAACAGCCACA GAGCAGGAGCCACAGGGCGCGCTGGAT
484	150	AAGAACTCAACATTCACGTATTTCACCTCTGACCATGGAGGACATTTAGAGGCAAGAGATGGAC ACAGCCAGTTAGGGGGATGGAACACGTAGGAAACATACAGAGGGACTCGAAGGGTCACATAG GACTGCACTAGGTTTAGGACCTAA
485	150	ACTTCGGGCACGGATGGCCAGCCTGCGTCAGGGCTGCGGGGACCTCCGAGGTTTGGTCAGCA CCTTTACCCAGAGCTGTCAGGGTTCTTACAGGTTTCTATCCACCAGCGGTGAGGTGTCTCCTCT GTGTCCTCTCCCACCTGAAAAATA
486	102	CAAACTCAAGTATTTTTTTGCTGTGGACACAGCCTACGTGGCCAAGAAGCTAGGGCTGCAGGG TACAAAAAGCCAATTTCAGAAGTAAGTGGACAAGGCAGA
487	150	GAGCCTGTCGGTGATGCTGGTGCGAAACCGCGACGAGGTGCAGGCGCTGGCCTTCGACGAG CAGCGGCGGCCACTCCCGCGCCGCGCAGCGAGGGAGACGCGGACGAATGCCCGGGGCAGC GTACAAAAACTAAAGAGGACGAAACAGCA
488	113	CATCAGATGCCGAGCATCAGTGTCAGTCCCCCCAACACGAGCCCCTTCAGCTCCCTGCACCTG TTCGGGTGCTCCGTGCTGTTCCGCCATCTCCCCAGGAACACCTCTGGCTT
489	125	TGCCTCCAGCACCCTGAAGATGCTGTGTGATATGTCCTCCTTACCAAGTTGGTGATGGCTGCCT CAGTGGTACTTAATGCTGATATGCTCACAATTTCTGTATAAAATAGAAGTTGAGAAAGGGA
490	150	AGTTGTGTGGCCGGGACCCTACTCTCACTGATGAGCTGCTGAATATTCTCACAGAGCTAACTCA ACTCAGTAAGACCACCAATGAAACAAAATTTTAGCACACACATACCTGTGGAACCATTGCAACC CTCTGGTTTCTTTTAGGTGACC
491	122	AAACATGGCAGAAATGGAGAAAGAAGGGAGACCTCCCGAAAATAAACGGAGCAGGAAGCCGG CTCACCGCGATCAAGTTCAACACCAACAGCACTTTGACTCTTCTCAGTGAAGCCATCAAA
492	149	GCGTCCAGTTGGTACGCAAGTTTGCCCAGTCCATCTTGCAATCTTTGGATGCCCTCCACAAAAA TAAGATTATTCACTGCGATAGATCGGGAAGAGCACACGTCTGAACTCCAGTCACGACACGAGAT CGCGTATGCCGTCTTCTGCGT
493	150	CCACTGCTCTTTCCTCCTCCCCAGCTGAACCTCCACAGTCCTCTACACTCTTCCAGGAGCAGCA GAAAATGAACATAAATGTCCCCGCAAAACAACAAAAAAACATCCAATCAGGGTTTGTCCCCAGC TGCACTGCAATACAAGTTCCAA

494	114	TTATAAATGTAGCAAATGTGAGAAGAGCTTTTGGCATCACTTAGCGCTTTCAGGACATCAGAGC AGGCGGGACAGCGGGTGGCTGTCCAGGGCTCCAATCAGCTCCGAGCAGGC
495	150	CAACTGAGTTTGAATTCAATCTCTATTACTTAAAGTTCTTGGCTTTCCACTATGTGTCTAATCGCT TTAAAACATTTCTCCTTGCGCACCTGCTGCATGCAAAATCTTCAAAAGCGGCAGTTTTGGTCGTA CCTGATATAAACATAAAAT
496	150	TGCATTGGATAAAACAGAAGAGCTCGAGGTGAGCAACGGCCACTTAGTGAAGCGTCTGGAAAA AATGAAAGCAAATCAATAACCTCTCTCTTATCTACCACAGGACTTGAGTCAGCATTATCCATCAA AGGAATATAAGTAGTTGCATGA
497	100	GAGAAATAGTCAGCCTAGCAGTTCTACCATGATCAGCGTGCTTCGAGCGGGTGGCGCCGCTCG GCCAGGCCCCCGCGGCCCAGCACTTCGCCGCTGAACT
498	150	GAGAGGAAGCCGAAGCTAAAGAGGAAAAGAAAGTGGAGGAAAAGAGTGAGGAAGTGGCTACC AAGGAGGAGCTGGTGGCAATGGAGGGTCTGCAATAATCTGAATAGCCTCTTCCTCTGAAATTG GCTGCATTGGAGACAATGGAGGCAA
499	150	GTCAATAATTTTGACCAAGACTTTACCCGGGAAGAGCCGGTACTCACCCTTGTGGACGAAGCAA TTGTAAAGCAGGGTCAGAGCAGTCACTGGGTTTACCCGGAGGTACTTGCAGTTGGGCTGACCA TAGTCAGCAAGGTAGGTTGACCA
500	150	GCTATATATGATTTCTTCCAAGTGGCCTCCAGCAAGGAGCGGTATAAGCTGACAGTTGGGAAAT ACAGAGGCACGGCAGGAAAGGCGGCCCGAACCCGCCGCCGCTGCCGCTAACGCCAACAACA CCAGCAACGCTGCTCGCCCCGGGGC
501	122	CCAGAAGCCGGGCCTGGAGCTGGCCCCGGCCGAGCCCGCGTACCCGCCGGCGGCGCCGGA GGAGCCGTTCGATGATGTTCAGCAGAATGTCCCACGCCACCACCTGGAGCTCCTTCCTATAC
502	150	GCTGAGACAGGGTCTCACTCTGTCTCAGCAAAAAAAAAAAAAAAAAAAAAAAAGAAAAAACAAATA CAAAAGTGGGAAAAAAAAAAATGTCCTAAATGAAAAAAAAAAATTTTCTAAAACGGAAGAGCACA CGTATTAACACCAGAACCCA
503	150	GAGTATGTGAATTCTGTCAAGCAGTTTTCCCACCATCCATTACATCCAGGGGGGATTTCCTTCG GCATCTTAATTCACTAATAAAAGTCTCGAAGTAACCCCAATCCTGATTGGAATTTGCAAGCAAAG CTTCTTCTGTATTGCACTTGT
504	150	GATCCAAACAGCACACCAATGAACTCAACTCACGACACGAGATCTCCTGTGCCGTCGTCAGCA TCCTAGACCAGCCGGTGGCCAGGGGGGTGCGCGGTGTACGCGGGAGCGGGGTAGGGATAGG GGTGAGGCGGCGGGTGCCGCTGCTCC
505	150	CTGGACTTGAGCAAGCTGTCACCATCTTCTTCTTCTTCCTCATCCTCATCCAGCTCCAGCTCCC GGCGTGGGCTGTCCGGGGGGCCGCCCCCCAGGGGCGGCGGGGAGAGCACACGTCGGAACG ACAGTAACGACCACAGAGAAGGTTAT
506	104	AGAATTTTCTATTTTTCTACACTACTCTGTGCTATATGTCATCCCATTTAAAAAAAATCTGCACAA TTATGATTCAGTTGAAAACAAAGATAGTACCACTATAA
507	150	GAGCACTCTCAAAAGTCTTACTTTAAAAAAAACAAAAACAAAAACAAAAAATGGCCGGGCACGG TGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGAGGGGGGATCATTAGGTCAGGA GTTAGAGAGCACCATCTCAATAAT
508	66	CATAGCAAATTACAGAAACGGAGTAAGAACGTAAGAAATCGCAAAGCACGTGACGTAATTTGCT AT
509	150	GTCTCTTCAGTGGGGATGGTGGGTTTTTCCGTGGGGATGGTGGGTTTTTCCGTGGAGATGGTG GGTTTTTCTGTGGGAATGGTGGGCTTTTCTTTGGGGACTGTGGGCTTTTCTGTTGTGACACTGG GTTTTTCTATGGAAACTGTAGAT
510	150	GACCTCGTGATCCGCCCGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGC GCCCGGTCAATAAAGGACTCTTAAATCGTCTCAAAGTGTGGCGTTCTCTCTAACACCCCTGGGC ACAACAAGAGGTTGTGGTGAGCGA
511	140	GTCTTAAGACAGTTGCAGCTTTTTAACTGTTACTAGTTTTAAGAATAAAAATTTACGTTTGACCTT TGTAACAGGAGTACAGTAATCAGAGGATTGTATTTTCCCAGATGGATGTAGCATTTTCCCATCTA TTGTAAACA
512	150	GGGGCAACAACTGAGTGCAACTGAGATTCTGTTGGAAGGAAGACGAGAACTTCATCCTGTGCT TTCTCCATCACCCTGCCCTCTTGGTCTCCATCATGGCACTCCAGTCAGGAGACAATCACAAATT GGAATGCAAGTGGGTGTGTAGAC
513	123	AAGAACAATCTGCCAGTTATCAAGACAGGAGGCAAATGAAGGGACTTCGACTGAAGTCTTTCCT GGTGGCTGAGATCACAATTTCCCCCTTTTTTAGGGCTAGAGCCAAGGTGAATTCTTTCC
514	149	CTTATAGCAGCAGCCAGTCTAGACTGATAAGATTCAATGTCAGCTTCCAGTCTTTTCTTGCTTTC TTTTTCCTTCAACAGTTCGGCATTGAGCCTTGTATTCTCAGCCTAGATCGGAAGGAGCACACGT CTGAACTCCAGTCACGACAA
515	123	CAAGTGATGGCACATTCCTCCAGTGAGTCCTCAGGGACTTTGCTTCCTCCTTCGTTGGAGACTT GGCCTTTTCCGGGGACTTGACCTCAGCTGGAGATTTTGCTTCTTCCTTCACTGGGGACT

516	139	CCACAAGGGAATTTGAAGAGATTCTGTGAGAGAGATATGAAATATTTTATAGTTAAAACACCAAT ATATTCAACTTTACACTTCTTGTGTTTTTCATAAACATTCCTGTTCTTTTCCTTTCAGTAATAAAAT TTAAAAA
517	87	CATCCCCCAGAGTGCCGAAGCTGCAGCCGAGGCCACCAAGAATATGGAAGCTGGAGCCGGAA GAGCTTGGCCACGCGCCACACACCC
518	118	CCCACACTGGGTTAGTTGATCCTCTGGGGTCCTCTGCCTGGCCATCTTCCTCTCAAGATTTGCC CTAAGGGTCACGGCCTCTATGAAGCTTTCCCTGACCCTCCCACACTGGGTTAGT
519	110	ATTATTTTTTTAGTAAAATTGTTTGAAGTTATTGAAACAAGAGAGTCAGACATTTGCAAAATATTA CAATATACAGAATCCACCATCTTTTGGAGGTTTTTTCCAGTATA
520	113	CCACACACCAACAATTATAGCAGGTCCACAGAACATAACAACATCTCTTCATCAGACTGTAGTTT TGGAATGCATGGCCACAGCCTTTCCCCTCAGAAGCAGCATTCTTTGAA
521	65	ACGCCCTCTGGCCTTTCCTCGGTTCCTTCCTCCCCCATGTCCAGGAAGGGCCTGGGGGGCGG CTC
522	150	TCCAATGAAACAACAAGAATCATATTTCTTGGCCACCATTTTAAATCATGTTAATCCATTTCTCAG AATACTGCATCTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGTTAAAAAAAAAAAAATGTTTT AAAACCAAAAAAAAACCC
523	103	AAGTACTGGAAACCGAAACCCTGGGAGCTCTGGGGTCTGGCTGCCTCACACTTGGGGTGGAG AATGTGGCATCCCTCGCTGACAGGATGGAGAGTCCCGGCCC
524	104	TGACCGGGTTTCTTTGATGCGACCTCAGGAAGGAGATGAGATTTGTGGCTCTTGATTCCTCTTG AAGTCTCAGTTTTGATCAAATCTTCTTTAGAGTGTGATTT
525	150	CCAGGAACAACGGCCACCCAGACGACAGGCCCACGTCCAACCCCAGCAAGCACCACAGGCCC AACCACCCCACAGCCAGGACAACCCACGAGGCCCACGGAAACGCTCTGTGGAAGACACAAAG CCTTTGGCCTGCTGGGCTCTCCGGGA
526	137	GATCTGGGGCTAGTGTAGAAGGTGGTTGATTTCTCACTGAGGCTTGGTGTTGTGGGACTGCCG GGGCAAGAGAAGACTGCCTAGAATTAAGTGTGGAGGGTCTTATAGGATACTCTTCGTTTTGTAC ATTTCCAGCA
527	87	GAAGGTCAGGATTTTAAGAAGAAGAAATGGAGCAGAATGCATCCTAGATCTGGGCTACCATATG AGACAGCCCAGATCTAGGATGCA
528	150	CCTCTCAAAGTGCTAGGATTATAGGCGTGAGCCACTGAGGCCTGCCTGCCTGCAGTTTTCATTG AGGTGTTTTTTTTTTTTTTTTTTTTTTTTTTATACGGAGGAGTAAAAAGTACCCCAAGCAGGAGGA CAGTGGCAGTAAAATTGCTC
529	150	ACTTCTTTGTGATGTGTGTATTCAACTCACAGAGTTTAACCTAACTTTACAAAGAGCAGATTTGA AACACACTGTTTGTAAAGTATGCAAGTGGAGATATCAACCGAAGAGAACACATCTGAACACCAG GCACTAACATAAAAAAAAAAC
530	150	ACTGAGATCCTACCATTGCACCCCGGCTTGGGAGACAGAGAGAGACTCTATCTCCAAAAAAAAA AAAAAAAAAAAAAAAGGGGAAAGTAATTGTTGGAAGGGGAGGGAAGGGCAACACTGGGAAAAAC CTGACACCACAGGAGAGGGGGTT
531	149	TCAGTGGGATGTACATGGGGGAGCTGGTGAGGCTTATCCTGGTGAAGATGGCCAAGGAGGAG CTGCTGAAGGAGCCGTTCGATGATGTTCAGCAGAATGTCCCACGCCACCACCTGGAGCTCCTT CCTATACTTCTTGATGAGCCTGGT
532	102	GTTTTATCCTGGGTGTCATTTTGGATCAAAAAAGATGCTACGCCAGCAAGAACAGCATTATAAG GCCAGACGTGACAATAGGAGAGATCCCTAGCTCCATCA
533	150	TTTCTCTGGAAGCTCACCCACTGCAGCCTTCAGCCGAGTCAGCTCCTGGTAGATCTCCTGCATC TTAGATTTCTCTGGAAGCTCTCTGGGATGTCGCAAACATCCAGCTGTGACCACTGGCTGAATTC AAAGGTGAAATTGTTATTAACT
534	149	TGGGGTTGTTTGTTTTTTTCCTGTAAATTAGTCTGAGTTCATTGTAGATTCCTGATATTAGCCCTT TGTCAGATGGATAGACTGAGAAGTTTTCTCCCATTCTGTAGGATGTCTGTTCACTCTGATTAAAG TTTCTTTTGATGTGCGAA
535	150	AGAGACGGGGTTTCACCGTGTTAGCCAGGATGGTGTCGATCTCCTGACCTCGTGATCCACCTG CCTCAGCCTCCCAAAGTGCTGAGATGACAGGCGTGAGGCACCGCACCAGTCCTTTTTTTTTTTA TTATTTGAGAATGAGTTTCACTA
536	149	GTGTAAGGAAGGACTCACATTCTTTTTTTTTTTTTTTTTCAACTGGGGCCTTGTTATGTTGCCCA TACAGGAGTGCAATTGTTTTATCATGGCTTACATCAGCCATTGTCTTATGGGCCCAATCAAACAT CATGCCTCACCCTCCAGA
537	73	CCACTCACTGCAACCTCTGCCTCCCAGCTTCAAGCAATTCTCCTGCCTCAGCTTCTGGAGTGGC TGGGAGCAC
538	150	ATCCTATGCCAGAATAACAGAGATAGATAGATACATAGATAGATAGATAGATAGATAGATAGATA GATAGATAGATAGATAGACAGATGATAGATAGATAGATAGATAGATAGATAGATAGATATTGATTGATA GATGAAAAACTGTAACCACA

539	150	TAATATCTCTTAATAAATCCCTTTCTGTGACCTGCACGTACACATCCAGATGGCCGGTTCCTGCC TTAACTGATGACATTCCACCACAAAAGAAGTGAAAATGGCCTGTTCCTGCCTTAACTGATGATAT TACCTTGTGAAATTCCTTCT
540	150	TCTTGCTCAAATTTTCCATCTTGGACCATTGGACTCTCTTTCGGTTGGCTCCTGTGTCCTTTGTC CTTTTTTTTTTTTTTTTTTTTTTTTTTAAAGTTTAAAAAAATTTTGCCCAGGCGGGCGGACCGTTTT GTTATAACCAAAAAAAGC
541	149	GTGTAAGGAAGGACTCACATTCTTTTTTTTTTTTTTTTTTAAAAAGGGGCCATGTTATGTTGCCCA GACAGGAGTTAAAAGGTTTTATAAATGCATACAGCAGCCATTGTCTTCTGGGCCCAAGCAAACA TAAAGCCACAGCCACCAGA
542	150	CCCTCCATATATGTATATATACATGCACACATCCTCTCATCTACCAATCTATCATCCATCCATCCA TCCTCCATCCTTTTATTCATCCATCCATCCTCCCATCCATATACATATATATACATACATACATCC TCTCATCTGCCAATCTAT
543	150	CATTCCACTCCATTCCATTCCATCCCATTCCACTCCATTCCATTCCATCCCAATCGGGTTGATTC CATTCCATTCCATTCCATTCCATTCCATTCCTTTCCTTTCCATTCCGTTCCATTCCATTAGGGTTG ATTCGATCGCAATAGAAAA
544	149	AAACCTTTTTTTCATTCAGCAGTTTGGAATCACTGATTTGTTAGGATCAGAGACGAGAAAGGTAG TAGTGAATAAATGAGCCTTAAGGTGAAAAAGCAAATATTTCAGATAAGAGCTAGAACGAAGCTG TCTGAGAACCTGCTGGGTGA
545	150	CTCAGTGCAGATCACTGGTTTTATCAGAAGGTCAGTGGTTGGAGACATCCATTTTCTGAAATCA GGTAAAGTGTTGAATGCTGTATATAGCACCCGAGAGGTGAGGCAGATATTTATCAATTTAGGTA CACAGGTGTCAGTTTGATTATT
546	150	CTTTTTCATCTCTTATGCCAGTATATAGTTGGTTTTTGAAGTCAGGTGTGTAGGCAGGCTTATTT CCTAGGCATATCATGACAGGAAAATGAAGAACGCTTTTTTTTTTTTTTTTTTTTAAGGGTTTTTTC CCATGTCGCCCCAGCAGGA
547	149	CAGGTTATGAAACCAGTTAGTTTTTGTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT GACGTGTTTTAACCGTTTTGCCAAGGCTTGGTCCGAGGGAACCACCGGCCCGCGGCCACAAAA CGGAAGAGCACCCGACA
548	150	GGCGGAGGTTGCAGTGAGCTGAGATCGCGGCCACTCCACTCCAGCCTGGGCGACAGAGCGA GACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGAAAACAACACGGGAACGAAAGAC AAAACGAATAAAAACAAGGAACAAC
549	150	CCCCATTCTTCCCTCTCTCTGTCCTCAGAACACTGCCTCATATCCTTCCCTGGTCCCTGGCTCT CTGAGTCCCTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTAGATAGATAGGAAAAGCAAACCTCTGAA AAACACTCACAACCACAGA
550	150	GGCAGAGTTAGACTGTCGGAAAAAAAAAAAAAAAAAAAGAAATGTATTGAAGTAAAAAATAGACA AAAAAATTTTAAAATTGGAGAGTAATTTTGATAAATTCTGAGGGCAGTTGGAGAAGAGAACGGAA AAGCAAAATTAAGAACTCCA
551	149	CAACTCAATAAGGTGAAAAGTCCAAACCTGACAGGTCGAGAATAAATAAATACAGCATCTAGAT CATTATTGTGCCAGGGGAAAAACTGGAAAAACTGACTTAGACATTGGCATTAAACTTTAATATCT AACGGGGGCTTTGGTTTTGA
552	148	ACTCGATTGTTATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATTAACCC GAATACAATGGAATTGAATGGACTGGAACGGAATGGAGTGGAATGGAATGGAATCAACCCGAG TGCAGGGGAATGGAATGGAAT
553	147	TCTTCCCAAGGGAGGAGGAGCTCAAGTGTCGGGAACTGTCTAACTTCAGGTTGTGTGAGTGCG TTAAAAAAAAAAAAAAAAAGAAGCCAAAAACAACATTTTTAAAACGAAATACCAACCGAATAAAAA CCAGACACAACAAGAGAA
554	149	ACTTCAGTACCACAGGTTAAGGGAGTGGTAGTTGAGTGACAGCTATGCTTCTGCTGTGTAACCT TTAGCAAACCAAATAATTTTCTAATGCAAAACTTTTTTCTGATCAAATCTACAGTACTATTATTCT AAATTCCTTTTCAATATA
555	150	CCAGTCCTCCATCTCTGGAGTGACATGGTCAATGCTGCATCTTCACAAAGCAGGGACACTAATC CCATTTAAGAGGGCTCCATCTTCAGACCTAATTATCTTGCAAAGGCACCACCTTTAACATATGAA TCAGGAGGGTGGGAATACACT

Table S4. Geographical classification of DNA sequences from the subset of ray-finned fishes

Group	Sequence ID	Number	Average affinity	E-value	Blast Mode
Non-native	#001 #002 #003 #007 #009 #010 #012 #013 #014 #015 #016 #164 #165 #166 #168 #169 #017 #049 #176 #179 #180 #181 #182 #183 #190 #191 #192 #193 #194 #195 #196 #199 #205 #208 #223 #225 #226 #233 #236 #240 #246 #253 #254 #258 #259 #260 #261 #262 #272	49	80.34%	≤ 4.00E-9	the MS mode
Native	#004 #011 #018 #019 #020 #021 #022 #023 #024 #025 #026 #207 #028 #029 #030 #031 #032 #033 #034 #035 #036 #037 #038 #039 #040 #041 #042 #043 #044 #045 #046 #047 #048 #050 #051 #052 #053 #055 #056 #057 #058 #059 #060 #061 #062 #063 #064 #065 #066 #067 #068 #069 #070 #071 #072 #073 #074 #075 #076 #077 #078 #079 #080 #081 #082 #083 #084 #085 #086 #087 #088 #089 #090 #091 #092 #093 #094 #096 #097 #098 #099 #100 #101 #102 #103 #104 #105 #106 #107 #108 #109 #110 #111 #112 #113 #114 #115 #116 #117 #118 #119 #120 #121 #122 #123 #124 #125 #126 #127 #128 #129 #130 #131 #132 #133 #134 #135 #136 #138 #139 #140 #141 #142 #143 #144 #145 #146 #147 #148 #149 #151 #152 #153 #154 #155 #156 #157 #158 #159 #161 #170 #171 #172 #173 #177 #200 #207 #209 #210 #213 #215 #216 #217 #219 #228 #229 #241 #242 #230 #237 #238 #243 #244 #245 **#247** #250 **#251** #252 #264 #265 #266 #267 #268 #269 #270 #273 #275 #276 #280 #281	180	83.85%	≤ 3.00E-14	
Other ray-finned	#008 #054 #137 #150 #160 #162 #163 #167 #174 #211 #218 #227 #232 **#277**	14	77.51%	≤ 5.00E-12	
Non-specific	#175 #178 #184 #197 #198 **#202** **#203** **#204** #206 #212 #220 #221 #222 #235 #248 **#278**	16	65.45%	≤ 3.00E-9	the lowest mode

The transposase sequences are underlined in the table while the mitogenomic and ribosomal sequences are presented in bold. All sequences are listed in Table S3.

Table S5. Results from DNA Extraction and Sequencing

ID	Fossil portions	Sample weight (g)	aDNAmix (ng/mL)	Sequences amount
1	texture	30	33	22253
2	texture	28	54	30932
3	texture	40	99	445752
4	texture	50	141	735110
5	texture	25	35	24976
6	non-texture	50	73	52
7	non-texture	50	82	583