Insights from the Early Cretaceous: The promise of Lycoptera aDNA sequencing

Abstract

We employed non-silica-based dipolar nanoparticle affinity bead technique to extract DNA from sedimentary rocks and successfully obtained aDNA from fossilized Lycoptera fishes from the Early Cretaceous in Beipiao, Liaoning Province, China. After library enrichment, high-throughput sequencing, nucleotide BLAST, and data filtering, 276 highly homologous ray-finned fish sequences were identified from 13,113 matched fragments. Molecular phylogenetic analysis showed that Lycoptera is closely related to Osteoglossiformes. At the same time, matching the 276 sequences to each Order of ray-finned fish showed that the fossil fish is closely related to Cypriniformes, but there are no genetic connections between fish groups geographically isolated from Eurasia. Gene exchange between these lineages has been blocked. In addition, analyzing the genetic connection between Lycoptera aDNA and modern genomes revealed unknown evolutionary relationships: The Cypriniformes genome has inherited many Lycoptera gene sequences. We propose the hypothesis that new transposase genes may arise through genome autonomous evolution mechanisms such as ‘progressive evolution’ and ‘overlapping coding region slippage replication recombination’. Evidence supporting this comes from observing the rapid expansion of gene families associated with transposons in aDNA.

Keywords: aDNA, Lycoptera, phylogenetic analysis, Cypriniformes, transposase

Introduction

The possibility of resurrecting ancient life has long been a fascination for science fiction enthusiasts. However, to achieve this goal, the genetic information of ancient species must first be obtained. Traditional silica-based DNA extraction techniques have been advanced to the point where they can extract the genomes of horse species[1], woolly mammoths[2], cave bears[3], and Neanderthals[4-7], as well as the partial genomes of mammoths that have been preserved in permafrost[8]. However, ancient DNA (aDNA) is usually fragmented and chemically altered during fossilization, making it difficult to recover intact sequences from most specimens[9]. Due to these challenges, previous studies have focused mainly on fossils from the Early Pleistocene onward. Some people believe that DNA cannot be preserved for over a million years under most conditions[9-11]. However, certain fossils in sedimentary rocks and the Jehol Biota have retained cellular structures, suggesting that some *in situ* DNA may persist even in older samples[12-16].

Lycoptera is an ancient freshwater fish that lived during the Early Cretaceous period in what is now northeastern China. It is one of the emblematic fossils of the Jehol Biota and an important biostratigraphic marker for the Early Cretaceous (J3–K1)[17-18]. Abundant Lycoptera fossils were discovered in Shangyuan Town, Beipiao City, Liaoning Province of China (Fig. 1). The fossils are located in the lower part of the Yixian Formation, and it is approximately 120-140 million years old. They were buried and preserved *in situ* with volcanic tuff generated under static water conditions immediately after rapid death, and the lithification process of these fossils is different from that of the fossils of prehistoric mammals and dinosaurs aforementioned.

Conventional DNA extraction methods are unsuitable for sedimentary rocks. In this study, we employed the non-silica-based dipolar nanoparticle affinity bead technique[19] to extract DNA from the fish fossil and established a DNA library for analysis using high-throughput sequencing (HTS). The fossil surface enables the absorption of external molecules. These molecules can precipitate, adsorb, accumulate, and degrade in a cyclical manner[20-22]. To distinguish endogenous aDNA from environmental DNA contamination (eDNA), we developed a rigorous data processing method. This approach allowed us to pinpoint endogenous aDNA from the fossil fish itself, along with that of contemporaries in its ancient food web parasites and prey alike, and filter out modern contaminants. We uncovered a window into this extinct world, preserved over vast stretches of time through the presence of its DNA.

Results and Discussion

Extraction and Analysis of DNA from Fossil Specimens

The Lycoptera fossil specimens were collected from the volcanic tuff formed approximately 130-140 million years ago in Beipiao, Liaoning Province, China (Fig. 1). The extraction and purification of fossil DNA were conducted using the Geobio® DNA Extraction Kit in accordance with the manufacturer's instructions. aDNAmix (ng/μL) represents a Maillard reaction product of proteins, polysaccharides, and DNA, whose concentration (ng/μL) can be estimated from the OD value at the 260 nm ultraviolet peak (Table 1). The aDNA in aDNAmix has a certain degree of structural damage, so its competitiveness for PCR Taq enzyme is poor. The core of high-throughput sequencing methods is PCR, while eDNA (especially modern eDNA) has a relatively intact nucleic acid structure and may be dominant in quantity, such as eDNA from local dominant species (such as mammals, human and livestock parasites and cultivated angiosperms). Since this study did not use the Target Enrichment method[23], the number of reads of eDNA in the sequencing results was far greater than that of aDNA reads.

Mapping aDNA Sequences onto Modern Genomes

The Fish 10,000 Genomes Project has nearly completed genome sequencing for representative species across fish Orders. To study how fossil DNA sequences compared to modern genomes, we used a DNA sequence similarity search tool (BLAST) to map

12,590,123 reads from the fossil fish to known fish genomes and counted the best matches. We found a total of 11,313 reads that matched ray-finned fish genomes (total matched reads, TMRs). To ensure accuracy, we only considered ‘qualified matched reads (QMRs)’ with an extremely high probability of similarity (E-value $< 1 \times 10^{-7}$), yielding 693 QMRs for ray-finned fishes. Another 10,620 reads likely came from more divergent regions or lower organisms lacking genome information (Table 2). The total number of reads from non-textured parts of the specimen was relatively low due to the dense, unfossilized nature of the specimen, as shown in Table 3.

Principles for Filtering *in situ* Reads

King and Jukes established a method to calculate how fast proteins evolve over time by estimating amino acid changes[24]. This can be converted to the rate of nucleotide changes per year in the genes coding those proteins. Most genes have now been sequenced, allowing their conservation to be estimated. For example, the myoglobin gene in Vertebrates changes at $\sim 0.39 \times 10^{-9}$ nucleotides per year. Based on this, we defined ‘highly similar reads (HSRs)’ and ‘highly identical reads (HIRs)’ as those with very low and extremely low divergence from modern counterparts, respectively. We found 59 HSRs and 34 HIRs matched to ray-finned fish genomes (Table 2). From this, we deduced several principles: 1) Long-term evolution has greatly increased the divergence between fossil and modern

DNA. This divergence depends on how conserved the DNA sequence is. The HSRs and HIRs for the fossil fish should be much lower than that of modern contaminants. 2) The HSRs and HIRs in the fossil fish should be in conserved genes. Those with modern DNA should be in less conserved genes. 3) If the HSRs and HIRs of the mapped fossil species are too high, especially in less conserved genes, modern DNA contamination is likely. 4) The qualified matched reads (QMRs) for the fossil fish should exceed those of contemporary species.

DNA Unveil Clues from Primeval and Contemporary Environments

Historically, Beipiao City was an agricultural and pastoral hub in northern China. As a result, DNA from humans, crops, livestock, as well as bacteria and fungi, would have dominated the eDNA that contaminated the fossils, as shown in Table 2. In contrast, ray-finned fishes had much lower percentages of HSRs (8.51%) and HIRs (4.91%), indicating that these reads likely came from the fossil fish itself rather than modern contaminants. In contrast, Decapoda， Diplostraca, Caligidae, aquatic mammals, Echinodermata, Hemichordata, Urochordata, Cephalochordata, Cyclostomata and Chondrichthyes had extremely low QMRs, HSRs and HIRs. This indicates the area likely lacked rivers, lakes, and marine influences after the fossil formed; the geographic evidence confirms this. The low HSRs and HIRs for cotton also indicate that the reads came from ancient angiosperm DNA *in situ* rather than modern contaminants, as this was historically not a cotton producing area. Bovidae now only inhabit southern Tibet and Xinjiang, but once ranged into Siberia and Alaska. Its high HSRs and HIRs suggest the Beipiao area was formerly part of its range. The fossil site has a colder climate and hilly terrain unsuitable for algae or amphibians. Reptiles were also rare. Therefore, their indicators can serve to establish background threshold ranges to rule out modern eDNA contamination.

Feature Differences between aDNA and Modern DNA

To find the difference between aDNA and later contaminated DNA, we analyzed the characteristics of QMRs matched to the genomes of Actinopterygii and *Homo sapiens*. We found that the lengths of the aDNA and contaminated DNA sequences were similar, but the GC content was slightly higher in the aDNA sequence. In addition, there were significant differences in the Per Identity (*Homo sapiens*: 98.32%; Actinopterygii: 68.65%), Query Cover (*Homo sapiens*: 98.82%; Actinopterygii: 77.02%) and average number of Gaps (*Homo sapiens*: 0.00; Actinopterygii: 2.25). The increased frequency of insertion/deletion polymorphisms compared to SNPs in the aDNA suggests that these genomic changes accumulate over time, implying the evolutionary process is irreversible, as evidenced by the differences seen between ancient and contaminated DNA sequences (Table 4).

Tracing Fragments to the Source

To investigate the phylogenetic relationship between fish and non-fish species, we selected 276 Actinopterygii sequences (Appendix 1, Table S1) using a specific method. First, we combined the 693 forward reads with their respective reverse reads. Next, we filtered the resulting aDNA sequences to ensure they had an E-value of less than $< 1 \times 10^{-7}$ and a length greater than 50 bp and were not identical to the other fragment.

To enhance the analysis of these sequences, we designed an innovative method called the ‘Three Stage Online BLASTN (3SONB)’. Employing this tool, we observed a reduction in consistency and coverage, coupled with an increase in E-values when aligning these sequences with the genomes across various Orders of Actinopterygii fishes, as well as between ray-finned and non-ray-finned fishes (Appendix 1, Table S1). Subsequently, further analysis revealed that some aDNA fragments displayed nearly identical or similar consistency and coverage metrics when matched to the genomes of both ray-finned and non-ray-finned fishes. In other words, this indicates that these fragments may originate from highly conserved genomic regions that span taxonomic phyla or even kingdoms. Nine fragments belong to ribosomal DNA and mitochondrial genes (Appendix 1, Table S2). Eight fragments belong to transposase genes (Appendix 1, Table S3). We analyzed forty-six other HIRs and HSRs in the same way (Appendix 1, Tables S4, S5), and used

NCBI Online Blastx in the protein standard database to find their relationship with protein coding genes/sequences. The results showed that most of the HIRs sequences were related to conserved regulatory genes. However, eight sequences were unable to be associated with conserved genes. Of these, four sequences were matched to non-local species (marine fishes) and four were matched to local fishes (1.1%). Eighteen HSRs could not be associated with conserved genes, of which five were matched to non-local species (marine fishes) and thirteen were matched to local fishes (4.7%). The above results can rule out the possibility that these aDNA come from eDNA contamination (Appendix 1, Tables S6, S7). These results suggest that aDNA fragments can provide valuable insights into ancient phylogenetic relationships, even across large evolutionary distances.

Phylogenetic Position of the Lycoptera

Greenwood[25] classified *Lycoptera* into Lycopteriformes of the Superorder of Osteoglossomorpha by analyzing their morphological related features to Hiodontidae. Zhang[26] conducted a systematic analysis of early Cretaceous fossil fish and contemporary fish based on the anatomical structures of 26 true bony fish and validated previous classification results. Our phylogenetic analysis combined ancient DNA and contemporary fish DNA sequences, focusing on four protein-coding sequences: *α-tubulin*, *NFkBp105*, *STK-PP1β*, and *SOGA3* (Appendix 1: ID192, ID259, ID196, and ID265). The findings reveal that Lycoptera is phylogenetically situated between

Osteoglossiformes and Clupeocephala, which suggests a close relationship between Lycoptera species and Osteoglossiformes, corroborating previous research findings (Fig. 2).

Beyond Anatomy How aDNA Tells Us About Evolutionary Relationships

Phylogeny reconstruction has traditionally relied on morphological characteristics. However, it is now well known that DNA or RNA has an important regulatory role in the development of morphology. Even individuals with the same genome can develop different adult skeletal morphologies when growing up in different environments[27-28]. Therefore, traditional methods of classification based on species morphological structures can benefit from DNA data for corroboration. In this study, we identified 276 QMRs using an ‘Order targeted’ search approach, and mapped these onto the genomes of various ray-finned fish Orders. Compared with other Orders, Cypriniformes has the highest number of mappings and Average Maximum Score, and there are significant differences between it and the closely related Orders of Ostariophysi or distant Perciformes ($P<0.01$, Fig. 3). Within the matching set of Cypriniformes, Xenocyprididae received the most matches, followed by Cyprinidae, and Danionidae received less (Fig. 4). In the system phylogeny, Lycoptera is closer to Acipenseriformes but farther from Polypteriformes. Although the two Orders accepted similar mapping numbers, the Polypteriformes had a greater Average Maximum Score than that of Acipenseriformes ($P<0.02$). We speculate that this may be because *Acipenser ruthenus* (the sequences were mainly mapped to *A. ruthenus* in Acipenseriformes) experienced 3R Whole-genome duplication (WGD) 58 million years ago[29-30], and its genome differentiation has accelerated; while Lycoptera may not have experienced 3R WGD 120 million years ago, its genome is more primitive and closer to the fishes of Polypteriformes.

Geographic Isolation and the Slow Fade of Genomic Similarity

Some ray-finned fish groups are strictly geographically isolated, and these tend to have very low Average Maximum Scores, such as the bowfins (Amiiformes) and trout-perches (Percopsiformes) in America, the southern smelts (Galaxiiformes) in Australia, and the argentines (Argentiniformes) in the deep sea (Fig. 3). Over long evolutionary timescales, the similarities between homologous DNA sequences in different fish genomes will decrease due to accumulated mutations. However, if there is gene exchange between these genomes, this can slow the decay of sequence identity through processes like gene conversion. A reasonable inference is that geographical isolation has blocked such gene exchange, resulting in very low Average Maximum Scores for these fishes.

Tracing the Roots of Cypriniformes

Early researchers proposed conflicting theories about the origins of Cypriniformes. Cockerell and Berg suggested that they evolved from the Jurassic fish Lycoptera[31-32]. Rosen argued that they were more closely related to Gonorynchiformes[33]. Tao analyzed 14 kb sequences of 1 mitochondrial and 15 nuclear genes using Bayesian and Maximum Likelihood methods. Tao found evidence that Cypriniformes originated around 193 million years ago, in the Early Jurassic[34]. The findings support Cockerell and Berg, suggesting Cypriniformes and Lycoptera shared a common ancestor. Nonetheless, the earliest Cypriniformes fossils emerged approximately 60 million years ago, a discrepancy of 133 million years from Tao's estimation. Presently, Cypriniformes inhabit freshwater environments across the globe, with the exception of Antarctica, South America, and Australia. Their extensive distribution implies that their origins predate the breakup of Pangea. We suggest that their evolution from ancestors with distinct morphologies involved the gradual dominance of recessive subgenomes. Over the course of 133 million years, these recessive subgenomes became prevalent, leading to the contemporary cyprinid morphology. The Cypriniformes genome contains more Lycoptera gene components, indicating Lycoptera likely was a direct ancestor.

The Origins of Transposase Diversity: A New Mechanism Unveiled?

Transposons, which constitute more than 40% of fish genomes, replicate, proliferate, and migrate within the genome, causing mutations, replication, inactivation, and alterations in tissue-specific expression of inserted genes, significantly impacting host evolution[35-43]. Transposase, the core component of transposons, is an essential factor in the transposition process. Transposases originate through two means: vertical descent within a lineage, leading to sporadic distribution even across distant relatives, and horizontal transfer between species, which is limited and rare. We observed that novel transposase genes evolve endogenously through two pathways: ‘progressive evolution’ and ‘overlapping coding region slippage replication recombination’ by aligning the transposase genes from the fossil aDNA and the genes from the modern species.

The first pathway, progressive evolution, involves ancestral DNA from species like Lycoptera evolving into new transposases that insert into host protein-coding regions. For instance, ancestral ID11 DNA (Appendix 1, Table S3) evolved into a Tcb2-like transposase that integrated into the growth arrest-specific protein 2 (GAS2) gene of *Anabarilius grahami*. The N-terminus shares 52 amino acids with GAS2 and Tcb2, while the C-terminus matches the conserved GAS2 sequence (Appendix 1, Fig. S1). This indicates the transposition occurred 120 million years ago in the Early Cretaceous, before the fossils formed. This result has a clock-like indicative effect. In addition, ID11 ancient DNA evolved into Tc1 transposase and inserted into the coding region of Arl13b of *Anabarilius graham*. As this sequence can simultaneously match the transposase sequence (ROL44406.1) and the coding region of Arl13b (ROL47080.1, Fig. S2), it is impossible to determine the fragment ownership, whether the transposition is complete, and the relationship between transposition and fossil formation time.

Previous research discovered overlapping gene expression in viruses and translational frameshifting in bacteria, where the reading frame changes to optimize the use of all genetic information[44-46]. In this study, we found an additional pathway by which ancient transposase genes can give rise to new transposases through a process involving overlapping, slippage, and recombinative innovation of transposases. This pathway is exemplified by the following scenarios: 1) Altering the reading frame to generate two distinct segments, which then recombine with other sequences to create novel transposases (Appendix 1: ID160; Fig. S3). 2) Producing two entirely different segments via 'slippage duplication' of DNA, followed by concatenation and recombination with other sequences to generate new members (Appendix 1: ID11; Fig. 5: ID80; Fig. S4: ID160; Fig. S5). 3) The introduction or removal of a single nucleotide can result in the formation of an entirely new transposase (Appendix 1: ID169; Fig. S6).

Further investigation is needed to pinpoint the mediators and intricate steps vital to the processes mentioned earlier, with the goal of thoroughly understanding the entire proposed mechanism. As expert Frith suggests, the numerous transposase sequences conserved throughout modern genomes can be considered 'protein fossils'[47]. Instances of homologous transposase sequences with high similarity are found concurrently in the genomes of Prokaryotes, Arthropods, and Mammals, consistently introducing new DNA into genomes. In contrast to gene replication, which can generate new functional genes[48], transposases persistently incorporate new DNA sequences into the genome. Through the pathways aforementioned, a multitude of new transposase sequences could be produced, contributing to the swift expansion of transposon families and genomes over time. This offers an explanation for the growth in genome content and the diversification of species through evolutionary processes. Incorporating these generative models into AI and analyzing transposons across modern genomes may reveal insights into ancient evolutionary events and critical turning points in evolutionary change.

Discovery of Ancient Symbiotic Relationships between Parasites and Prey in Fish Fossils

We have identified a 28S ribosomal DNA sequence (Appendix 1: ID282) associated with the genome of Mesomycetozoea, a clade of flagellated microorganisms belonging to the Opisthokonta group within Ichthyosporea[49]. This minute parasite occupies a distinct taxonomic position, situated between animals and fungi, and resides within fish hosts. Additionally, we identified seven aDNA sequences corresponding to two shrimp genera and a sea spider species, respectively (Appendix 1: IDs 284-289, ID283), which are likely related to ancient Decapoda species that served as prey within the food web. Using 3D X-ray computed tomography on the fossils, we found no other organisms, suggesting that these sequences represent symbiotic species of Lycoptera, including parasites and prey (Fig. S7).

Illuminating Evolution's Trajectory through aDNA

Since Higuchi pioneered ancient DNA research[50], modern biotechnologies have enabled increasingly sophisticated analyses of aDNA. Emerging techniques like AlphaFold, third-generation sequencing, and CRISPR enable conveniently studying ancient genomes to predict and reconstruct their functions. Species evolution occurs through gradual genomic changes. Utilizing both ancient and modern DNA enables us to elucidate intermediary stages, forecast future trends, and ascertain final outcomes. Although currently limited to a small amount of aDNA, advancements in techniques and increased sample amounts indicate the potential for obtaining greater information on Lycoptera genomes in the future. Thus, by assembling homologous fragments from aDNA and DNA databases, we aim to identify the genome of the earliest universal common ancestor and gain insights into the evolution of species‘ morphology.

Materials and Methods

Fossil DNA extraction, NGS library construction, Illumina NovaSeq Sequencing

Washing the stone by DI water with tap water pressure, the surface was clean. Dipping it into the Geobio® DNA cleaner solution for a few seconds and moving it out quickly, repeating the dipping step for 3 times. Washing the stone by DI water with tap water pressure until there is no additional cleaner solution. Cutting the stone in a 10×10×5.0 (cm^3) stock and slicing the multilayer stone into a single layer, about 0.8-1.2 mm. Scraping the texture portion of fishes, collecting the falling particles in a clean tube. Scraping the outline portion away at least 1.0 cm from the texture parts and collecting the falling particles in separate tubes. Using a set of pestles and mortar to grind the small particles into fine powders. Extracting the DNA from the fine powders by Geobio® DNA Extraction Kit; using NEB PreCR® mix (M0309S) for USER procedure (repairing aDNA chain).

Building the NGS DNA library by following SSDLSP Standard Procedure (Sangon Biotech, CORP). DNA sequencing was executed by following the Standard Programme of Double Index Sequencing on an Illumina NovaSeq 6000 (S-4 Reagent Kit); reads qualified by the quality control scores and quality distribution. Local NCBI Nucleotide BLAST with word size 11 and nucleotide sequence database：version 5. Control of the search results: the results qualified by a cut-off E-value $< 1 \times 10^{-7}$.

Molecular Phylogenetic Analysis of Lycoptera Position

In this study, four reads were obtained, specifically ID192, ID259, ID196, and ID265 (Appendix 1), which were identified as fragments of four protein-coding sequences: *α-tubulin*, *NFkB p105*, *STK-PP1β*, and *SOGA3*. These sequences were aligned with the corresponding teleostei sequences for phylogenetic analysis[51]. The concatenated protein sequences were analyzed using MrBayes 3.2.2 with default priors[52]. *Protopterus amnectens* and *Latimeria chalumnae* served as outgroup species.

3D X-ray Scans

The Xradia 620 Versa 3D X-ray microscope (Carl Zeiss X-ray Microscopy Inc., Pleasanton, CA, USA) were employed for the observation. The fossil samples were further decomposed and cut into small fragments (50.0 x 80.0 x 5.0 mm) containing complete fish (usually compressed into thin layers of 2.0 mm thickness by stone formation pressure) and surrounding species. We conducted high-resolution scanning imaging of the samples to observe the microscopic structure and changes of the fish and its surroundings. Using a voxel resolution of 45.7 μm, we successfully scanned and imaged the whole fish on the fossil and observed its structure. Using voxel resolutions of 6 μm, 6.5 μm, and 6.9 μm, respectively, we conducted high-resolution scanning imaging of the fish’s head, chest, and tail, and could clearly see the structures at different locations.

Data Availability

The data that support the findings of this study are available in the manuscript and its Supplementary Information, and from the corresponding authors on reasonable request.

Acknowledgements

We express our gratitude to numerous individuals for their invaluable assistance in various aspects of this project: Dr. H. J. Zheng of the Shanghai Institute for Biomedical and Pharmaceutical Technologies, for his insightful contributions to our manuscript discussions; Dr. J. J. Ma of the Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, for his expert guidance in conducting the phylogenetic analysis of MrBayes; Dr. A. M. Bailleul of the Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences，for her interesting discussions about the manuscript.

Conflict of interest

The authors have no conflicts of interest to declare.

Supplementary information

This file contains Figure S1-S7，Table S1-S7，and Appendix 1.

References

1. Orlando, L. et al. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* 499, 74-78 (2013).

2. Palkopoulou, E. et al. Complete genomes reveal signatures of demographic and genetic declines in the woolly mammoth. *Curr Biol* 25, 1395-1400 (2015).

3. Barlow, A. et al. Partial genomic survival of cave bears in living brown bears. *Nat Ecol Evol* 2, 1563-1570 (2018).

4. Green, R. E. et al. A draft sequence of the Neandertal genome. *Science* 328, 710-722 (2010).

5. Prüfer, K. et al. The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505, 43-49 (2014).

6. Prüfer, K. et al. A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science* 358, 655-658 (2017).

7. Mafessoni, F. et al. A high-coverage Neandertal genome from Chagyrskaya Cave. *Proc Natl Acad Sci USA* 117, 15132-15136 (2020).

8. Van der Valk, T. et al. Million-year-old DNA sheds light on the genomic history of mammoths. *Nature* 591, 265-269, (2021).

9. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* 362, 709-715, doi:10.1038/362709a0 (1993).

10. Allentoft, M. E. et al. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc Biol Sci* 279, 4724-4733 (2012).

11. Dabney, J., Meyer, M. & Pääbo, S. Ancient DNA damage. *Cold Spring Harb Perspect Biol* 5 (2013).

12. Pang, K. et al. The nature and origin of nucleus-like intracellular inclusions in Paleoproterozoic eukaryote microfossils. *Geobiology* 11, 499-510 (2013).

13. 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).

14. Yin, Z. et al. Nuclei and nucleoli in embryo-like fossils from the Ediacaran Wengan Biota. *Precambrian Research* 301, 145-151

(2017).

15. 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).
16. Zheng, X., Bailleul, A., M., Li, Z. Wang, X. & Zhou, Z. Nuclear preservation in the cartilage of the Jehol dinosaur audipteryx. *Commun Biol* 4, 1125 (2021).
17. Grabau, A. Stratigraphy of China. Part 2 Mesozoic. Beijing: *Geological Survey of China* 1-774 (1928).
18. Gu, Z. Jurassic and Cretaceous of China. Beijing: *Science Press* 1–84 (1962).
19. 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).
20. 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).
21. 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).
22. Zhou, Z. The Jehol Biota, an Early Cretaceous terrestrial Lagerstätte: new discoveries and implications. *National Science Review*

1, 543-559 (2014).

23. Fu, Q. et al. DNA analysis of an early modern human from Tianyuan Cave, China. *Proc Natl AcadSci USA* 110, 2223-7 (2013).

24. King, J. L. & Jukes, T. H. Non-Darwinian evolution. *Science* 164, 3788-3898 (1969).

25. Greenwood, PH. On the genus Lycoptera and its relationship with the family Hiodontidae (Pisces, Osteoglossomorpha). *Bulletin of the British Museum (Natural History) Zoology* 19, 257-285 (1970).

26. Zhang, J.Y. Phylogenetic relationships of arapaima superorder. *Vertebrata PalAsiatica* 44, 4-59 (2006).

27. Fraga, M. F. et al. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc Natl Acad Sci USA* 102, 10604-10609 (2005).

28. Hu, Y. & Albertson, R. C. Baby fish working out: an epigenetic source of adaptive variation in the cichlid jaw. *Proc Biol Sci* 284, 1471-2954 (2017).

29. Bi, X., Wang, K. & Yang, L. Tracing the genetic footprints of vertebrate landing in non-teleost ray-finned fishes. *Cell* 184, 1377-1391 (2021).

30. Du, K. et al. The sterlet sturgeon genome sequence and the mechanisms of segmental rediploidization. *Nat Ecol Evol* 4, 841–852 (2020).

31. Cockerell, T. D. A. The Affinities of the Fish Lycoptera Middendorffi. *Bull Amer Mus Nat Hist* 51, 313–318 (1925).

32. Berg, L. S. Classification of fishes both recent and fossil. *Trudy Zoologischeskogo Instituta* 5, 85–517 (1940).

33. 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).

34. Tao, W., Yang, L., Mayden, R. L. & He, S. Phylogenetic relationships of Cypriniformes and plasticity of pharyngeal teeth in the adaptive radiation of cyprinids. *Sci China Life Sci* 62, 553-565 (2019).

35. Warren, I. A. et al. Evolutionary impact of transposable elements on genomic diversity and lieage-specific innovation in vertebrates. *Chromosome Res* 23, 505-531 (2015).

36. Hayakawa, T., Satta, Y., Gagneux, P., Varki, A & Takahata, N. Alu-mediated inactivation of the human CMP-N-acetylneuraminic acid hydroxylase gene. Proc *Natl Acad Sci USA* 98, 11399-11404 (2001).

37. Etchegaray, E., Naville, M., Volff, J. N. & Haftek-Terreau, Z. Transposable element-derived sequences in vertebrate development. *Mob DNA* 12, 1 (2021).

38. Britten, R. J. & Kohne, D. E. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science* 161, 529-540 (1968).

39. Tafalla, C., Estepa, A. & Coll, J. M. Fish transposons and their potential use in aquaculture. *J Biotechnol* 123, 397-412 (2006).

40. Ting, C. N., Rosenberg, M. P., Snow, C. M., Samuelson, L. C. & Meisler, M. H. Endogenous retroviral sequences are required for tissue-specific expression of a human salivary amylase gene. *Genes Dev* 6, 1457-1465 (1992).

41. Yuan, Y. W. & Wessler, S. R. The catalytic domain of all eukaryotic cut-and-paste transposase superfamilies. *Proc Natl Acad Sci USA* 108, 7884-7889 (2011).

42. Fawcett, J. A. & Innan, H. High Similarity between Distantly Related Species of a Plant SINE Family Is Consistent with a Scenario of Vertical Transmission without Horizontal Transfers. *Molecular Biology and Evolution* 33, 2593–2604 (2016).

43. Gilbert, C. & Feschotte, C. Horizontal acquisition of transposable elements and viral sequences: patterns and consequences. *Curr Opin Genet Dev* 49, 15-24 (2018).

44. Barrell, B. G., Air, G. M. & Hutchison, C. A. 3rd. Overlapping genes in bacteriophage phiX174. *Nature* 264, 34-41 (1976).

45. Sanger, F. A. G. & Barrell, B. G. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* 265, 5687-5595 (1977).

46. Chandler, M. & Fayet, O. Translational frameshifting in the control of transposition in bacteria. *Mol Microbiol* 7, 497-503 (1993).

47. Frith, M. C. Paleozoic Protein Fossils Illuminate the Evolution of Vertebrate Genomes and Transposable Elements. *Mol Biol Evol* 39 (2022).

48. Ohno, S. Evolution by gene duplication. *Teratology* 9, 250-251 (1974).

49. 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).

50. Higuchi, R., Bowman, B., Freiberger, M., Ryder, O. A. & Wilson, A. C. DNA sequences from the quagga, an extinct member of the horse family. *Nature* 312, 282-284 (1984).

51. 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).

52. Ronquist, F. et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539-42 (2012).

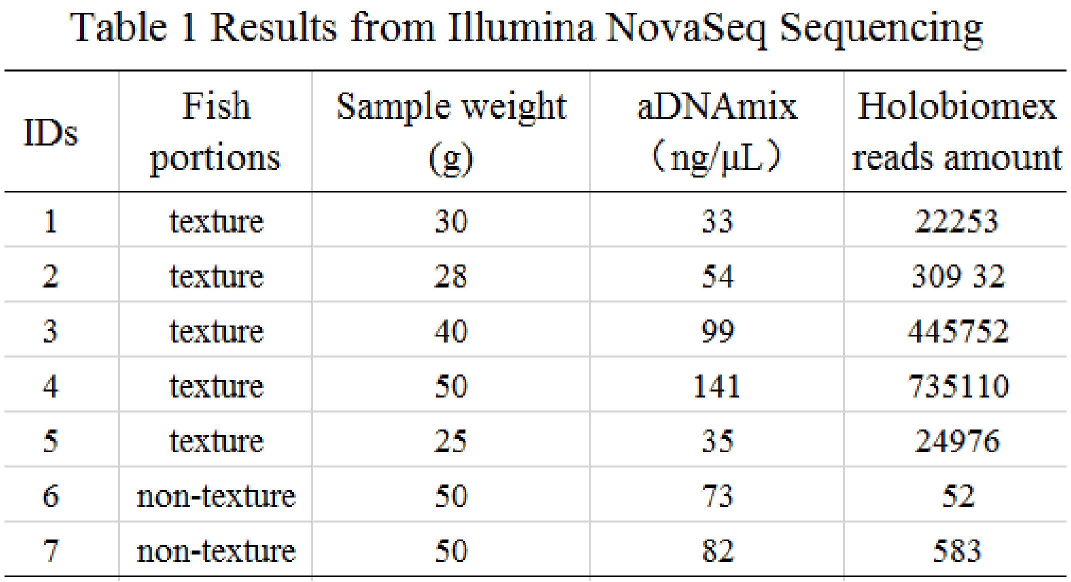
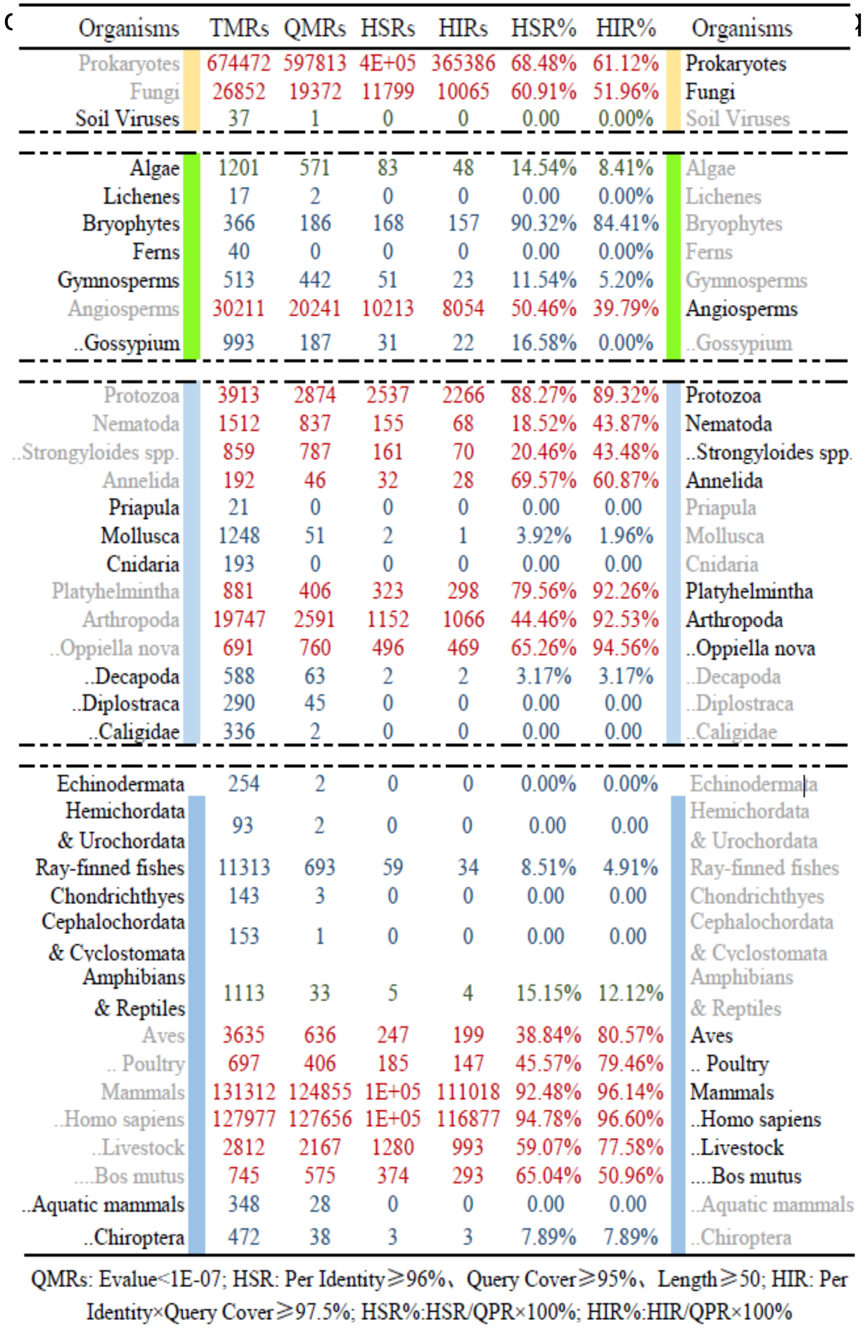


Table 1 Results from Illumina NovaSeq Sequencing

IDs	Fish portions	Sample weight (g)	aDNAmix（ng/μL）	Holobiomex reads amount
1	texture	30	33	22253
2	texture	28	54	309 32
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

Table 2 Matches c

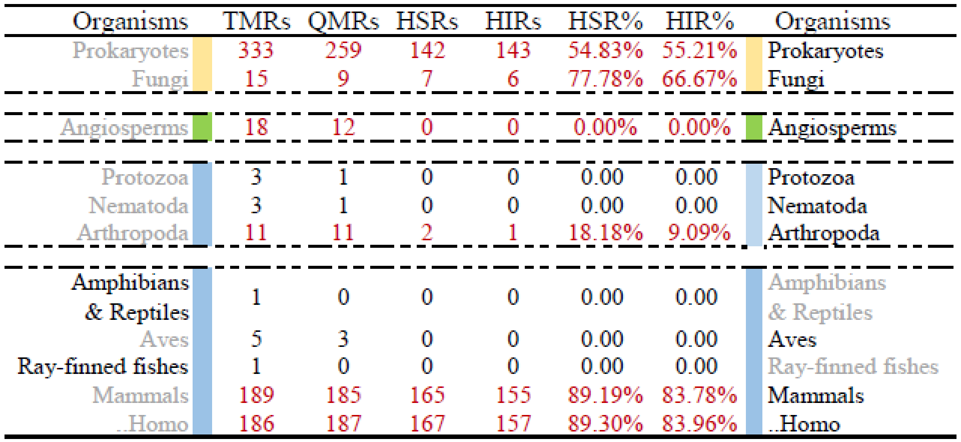
quenced genomes



Organisms	TMRs	QMRs	HSRs	HIRs	HSR%	HIR%	Organisms
Prokaryotes	674472	597813	4E+05	365386	68.48%	61.12%	Prokaryotes
Fungi	26852	19372	11799	10065	60.91%	51.96%	Fungi
Soil Viruses	37	1	0	0	0.00	0.00%	Soil Viruses
Algae	1201	571	83	48	14.54%	8.41%	Algae
Lichenes	17	2	0	0	0.00	0.00%	Lichenes
Bryophytes	366	186	168	157	90.32%	84.41%	Bryophytes
Ferns	40	0	0	0	0.00	0.00%	Ferns
Gymnosperms	513	442	51	23	11.54%	5.20%	Gymnosperms
Angiosperms	30211	20241	10213	8054	50.46%	39.79%	Angiosperms
..Gossypium	993	187	31	22	16.58%	0.00%	..Gossypium
Protozoa	3913	2874	2537	2266	88.27%	89.32%	Protozoa
Nematoda	1512	837	155	68	18.52%	43.87%	Nematoda
..Strongyloides spp.	859	787	161	70	20.46%	43.48%	..Strongyloides spp.
Annelida	192	46	32	28	69.57%	60.87%	Annelida
Priapula	21	0	0	0	0.00	0.00	Priapula
Mollusca	1248	51	2	1	3.92%	1.96%	Mollusca
Cnidaria	193	0	0	0	0.00	0.00	Cnidaria
Platyhelmintha	881	406	323	298	79.56%	92.26%	Platyhelmintha
Arthropoda	19747	2591	1152	1066	44.46%	92.53%	Arthropoda
..Oppiella nova	691	760	496	469	65.26%	94.56%	..Oppiella nova
..Decapoda	588	63	2	2	3.17%	3.17%	..Decapoda
..Diplostraca	290	45	0	0	0.00	0.00	..Diplostraca
..Caligidae	336	2	0	0	0.00	0.00	..Caligidae
Echinodermata	254	2	0	0	0.00%	0.00%	Echinodermata
Hemichordata & Urochordata	93	2	0	0	0.00	0.00	Hemichordata & Urochordata
Ray-finned fishes	11313	693	59	34	8.51%	4.91%	Ray-finned fishes
Chondrichthyes	143	3	0	0	0.00	0.00	Chondrichthyes
Cephalochordata & Cyclostomata	153	1	0	0	0.00	0.00	Cephalochordata & Cyclostomata
Amphibians & Reptiles	1113	33	5	4	15.15%	12.12%	Amphibians & Reptiles
Aves	3635	636	247	199	38.84%	80.57%	Aves
.. Poultry	697	406	185	147	45.57%	79.46%	.. Poultry
Mammals	131312	124855	1E+05	111018	92.48%	96.14%	Mammals
..Homo sapiens	127977	127656	1E+05	116877	94.78%	96.60%	..Homo sapiens
..Livestock	2812	2167	1280	993	59.07%	77.58%	..Livestock
....Bos mutus	745	575	374	293	65.04%	50.96%	....Bos mutus
..Aquatic mammals	348	28	0	0	0.00	0.00	..Aquatic mammals
..Chiroptera	472	38	3	3	7.89%	7.89%	..Chiroptera

QMRs: Evalue<1E-07; HSR: Per Identity≥96%、Query Cover≥95%、Length≥50; HIR: Per Identity×Query Cover≥97.5%; HSR%:HSR/QPR×100%; HIR%:HIR/QPR×100%

Table 3 Matches of Holobiomex reads from non-textured parts of fish fossils to sequenced genomes



Organisms	TMRs	QMRs	HSRs	HIRs	HSR%	HIR%	Organisms
Prokaryotes	333	259	142	143	54.83%	55.21%	Prokaryotes
Fungi	15	9	7	6	77.78%	66.67%	Fungi
Angiosperms	18	12	0	0	0.00%	0.00%	Angiosperms
Protozoa	3	1	0	0	0.00	0.00	Protozoa
Nematoda	3	1	0	0	0.00	0.00	Nematoda
Arthropoda	11	11	2	1	18.18%	9.09%	Arthropoda
Amphibians & Reptiles	1	0	0	0	0.00	0.00	Amphibians & Reptiles
Aves	5	3	0	0	0.00	0.00	Aves
Ray-finned fishes	1	0	0	0	0.00	0.00	Ray-finned fishes
Mammals	189	185	165	155	89.19%	83.78%	Mammals
..Homo	186	187	167	157	89.30%	83.96%	..Homo

QMRs: Evalue<1E-07; HSR: Per Identity≥96%、Query Cover≥95%、Length≥50; HIR: Per Identity×Query Cover≥97.5%; HSR%:HSR/QPR×100%; HIR%:HIR/QPR×100%

The major species contributing greater HSRs and HIRs include cultivated angiosperms such as wheat (*Triticum aestivum*), barley (*Hordeum vulgar*), rice (*Oryza sativa*), corn (*Zea mays*), peanut (*Arachis correntina*), rapeseed (*Brassica oleracea*), and soybean (*Glycine max*). In poultry, the major species include chicken (*Gallus gallus*), duck (*Anas platyrhynchos*), mallard (*Anas platyrhynchos*) and goose (*Anser Anser*). Livestock: pig (*Sus scrofa*), yak (*Bos mutus*), sheep, cattle (*Bos taurus*), and horse (*Equus caballus*). Fish: Zebrafish (*Danio aesculapii*), seabass (*Dicentrarchus labrax*), icefish (*Pseudochaenichthys georgianus*), whitefish (*Coregonus sp.*), blunt-snouted clingfish (*Gouania willdenowi*), carp (*Cyprinus carpio*), grouper (*Epinephelus fuscoguttatus*), silver carp

(*Hypophthalmichthys molitrix*) and mackerel (*Trachurus trachurus*). Protozo: *Plasmodium spp.*, *proteus*; Nematoda: *Paramorphic strongyloides*, *soybean cystic nematodes*; Arthropoda: *Oppiella nova*.

Table 4 Comparison of Actinoterpygii aDNA and Homo DNA

Markers	**Homo sapiens**		Lycoptera aDNA	
	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%
Gap/sequence	0.00	0.27	2.25	2.70
GC%	42.05%	9.73%	44.68%	11.00%
EHR*	99.70%		6.14%	

EHR: effective hit ratio; QMRs/TMRs×100%; QMRs: E-value<1E-07; TMRs: total matched reads; *Student's t-test, P<0.01

Figure Legends

Figure 1 A mental image of an early Cretaceous fossil site (left) and the Lycoptera fossil (right). The tranquility of the lake's surface was suddenly disrupted when a volcano erupted, releasing a cloud of black smoke and large patches of scorching lava flowing into the water. Lives in the lake was buried. Lycoptera davidi fossils found in Beipiao City, Chaoyang, Liaoning Province, with longitude:120.84 and latitude: 41.60.

Figure 2 Molecular Phylogenetic Analysis of Lycoptera Position. Upon examining four protein-coding sequences—ID192, ID259, ID196, and ID265—the phylogenetic placement of Lycoptera was deduced to reside between Osteoglossiformes and Clupeocephala, employing the analysis under the Whelan and Goldman (WAG) model for amino acid substitution.

Figure 3 Depicts matches of 276 aDNA sequences to genomes of species within Actinopterygii (Cutoff E -value < 1 x 10^{-7}). The search was constrained to a single Order. Each point represents the Order and Maximum Score for the best match to an aDNA sequence. The y-axis (n) shows how often each Order was the top match.

Figure 4 Depicts matches of 276 Lycoptera aDNA sequences to genomes of various species within Cypriniformes (size = 7, E-value < 1 x 10^{-7}). Each point corresponds to the genus and Maximum Score of the top match for an aDNA sequence. The y-axis (n) shows how often each genus was the best match.

Figure 5 ID11 aDNA produced two different fragments through slippage replication, which further recombined and fused into a new transposase coding region in the Cypriniformes genomes.

Figure 1

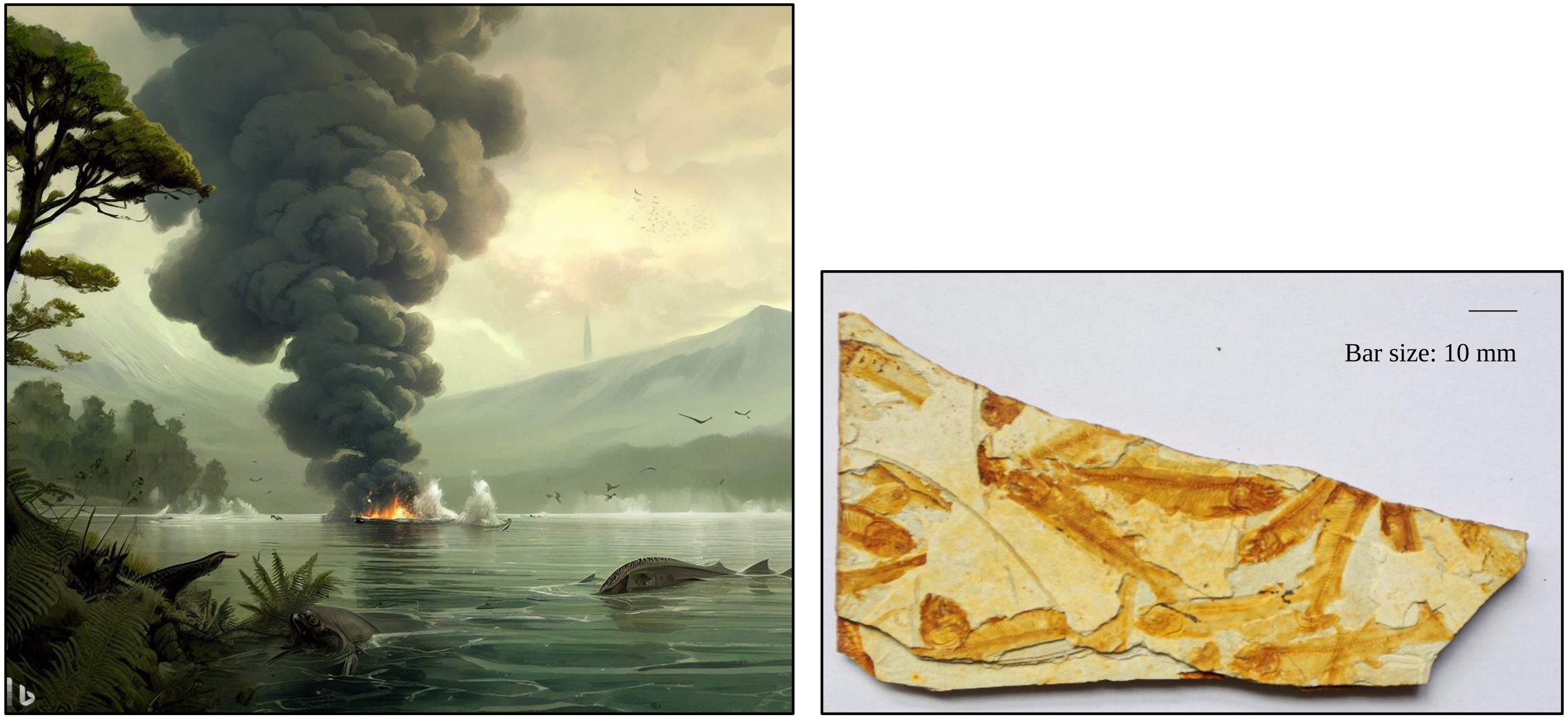

Bar size: 10 mm

Figure 2

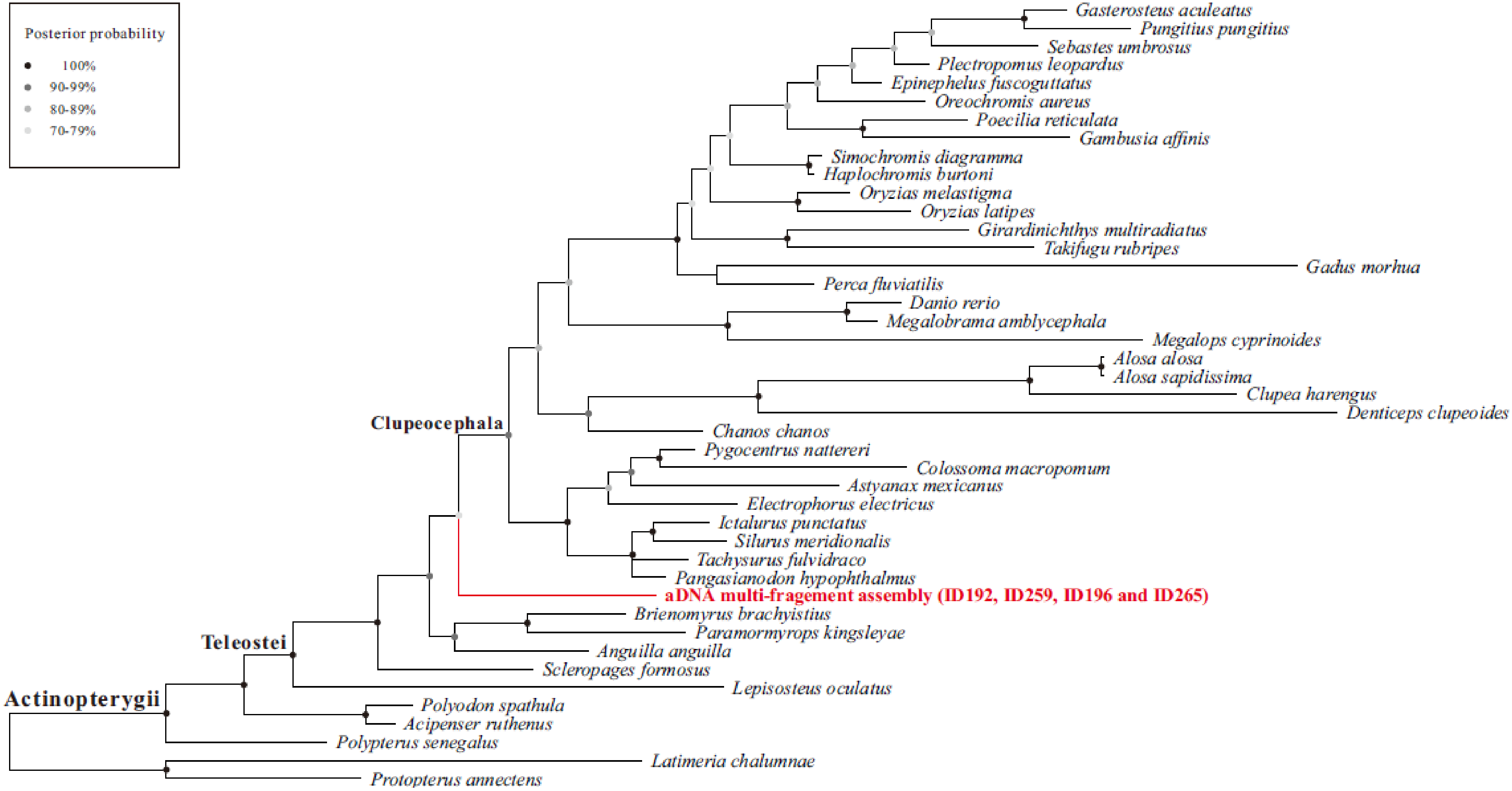

Posterior probability
● 100%
● 90-99%
● 80-89%
● 70-79%
Gasterosteus aculeatus
Pungitius pungitius
Sebastes umbrosus
Plectropomus leopardus
Epinephelus fuscoguttatus
Oreochromis aureus
Poecilia reticulata
Gambusia affinis
Simochromis diagramma
Haplochromis burtoni
Oryzias melastigma
Oryzias latipes
Girardinichthys multiradiatus
Takifugu rubripes
Gadus morhua
Perca fluviatilis
Danio rerio
Megalobrama amblycephala
Megalops cyprinoides
Alosa alosa
Alosa sapidissima
Clupea harengus
Denticeps clupeoides
Clupeocephala
Chanos chanos
Pygocentrus nattereri
Colossoma macropomum
Astyanax mexicanus
Electrophorus electricus
Ictalurus punctatus
Silurus meridionalis
Tachysurus fulvidraco
Pangasianodon hypophthalmus
aDNA multi-fragement assembly (ID192, ID259, ID196 and ID265)
Brienomyrus brachyistius
Paramormyrops kingsleyae
Anguilla anguilla
Teleostei
Scleropages formosus
Lepisosteus oculatus
Actinopterygii
Polyodon spathula
Acipenser ruthenus
Polypterus senegalus
Latimeria chalumnae
Protopterus annectens

Figure 3

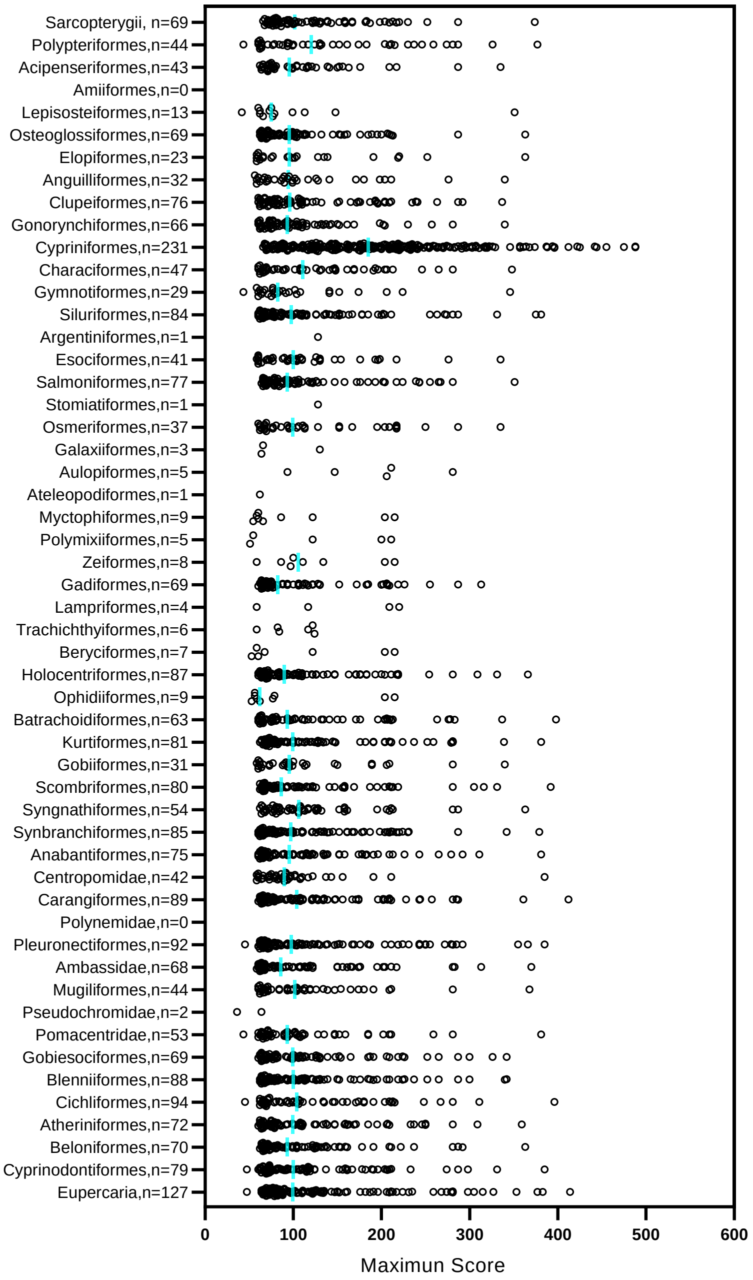
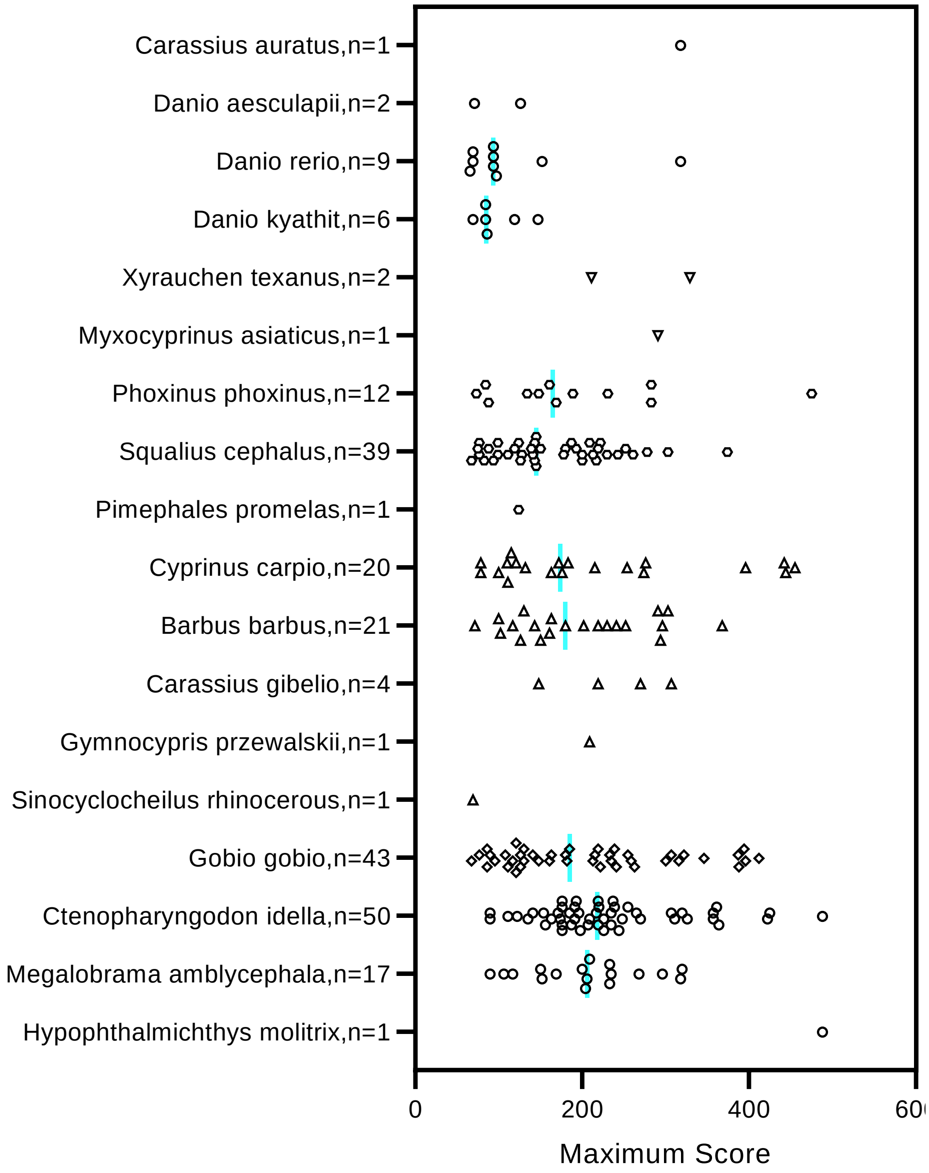
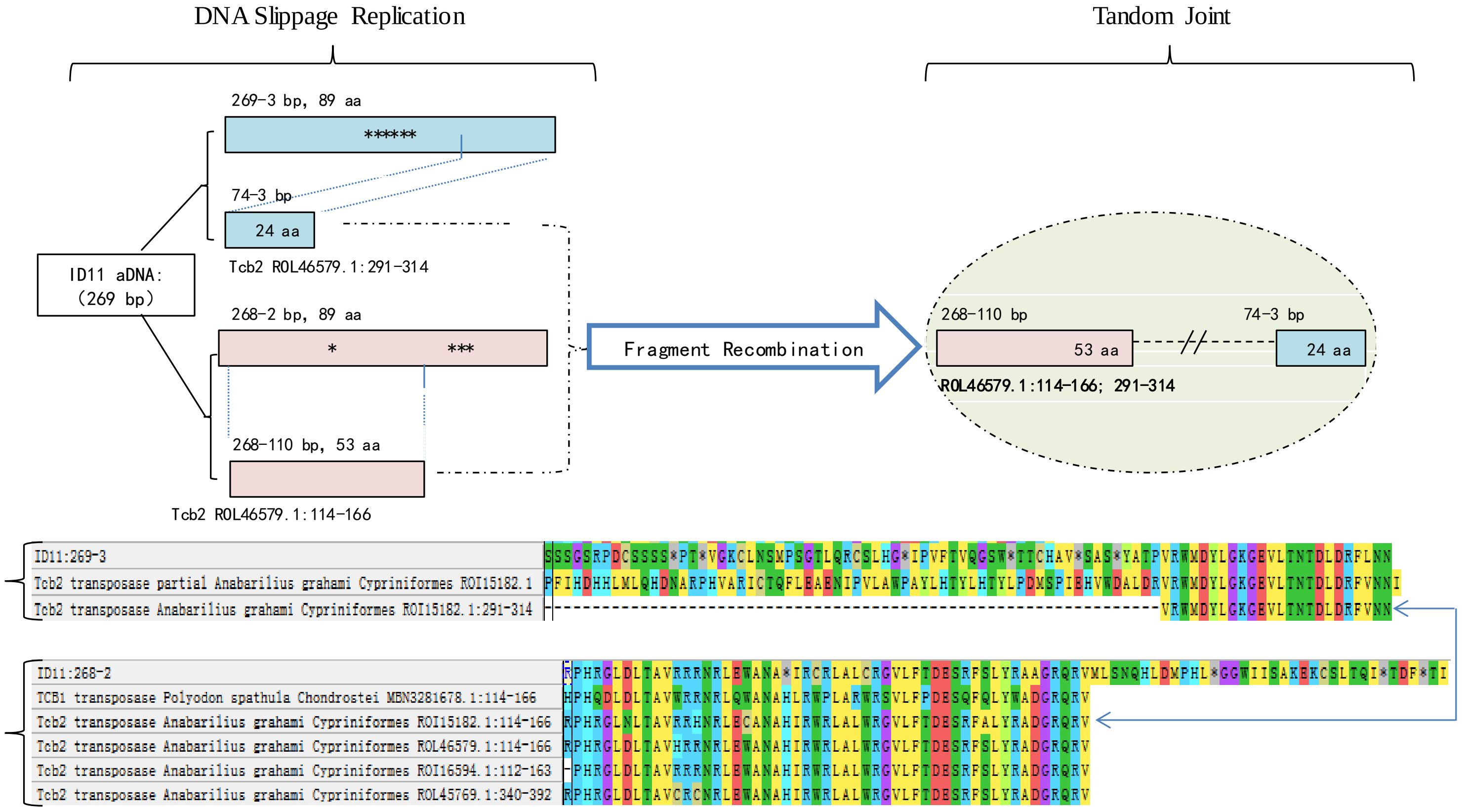

Sarcopterygii, n=69
Polypteriformes,n=44
Acipenseriformes,n=43
Amiiformes,n=0
Lepisosteiformes,n=13
Osteoglossiformes,n=69
Elopiformes,n=23
Anguilliformes,n=32
Clupeiformes,n=76
Gonorynchiformes,n=66
Cypriniformes,n=231
Characiformes,n=47
Gymnotiformes,n=29
Siluriformes,n=84
Argentiniformes,n=1
Esociformes,n=41
Salmoniformes,n=77
Stomiatiformes,n=1
Osmeriformes,n=37
Galaxiiformes,n=3
Aulopiformes,n=5
Ateleopodiformes,n=1
Myctophiformes,n=9
Polymixiiformes,n=5
Zeiformes,n=8
Gadiformes,n=69
Lampriformes,n=4
Trachichthyiformes,n=6
Beryciformes,n=7
Holocentriformes,n=87
Ophidiiformes,n=9
Batrachoidiformes,n=63
Kurtiformes,n=81
Gobiiformes,n=31
Scombriformes,n=80
Syngnathiformes,n=54
Synbranchiformes,n=85
Anabantiformes,n=75
Centropomidae,n=42
Carangiformes,n=89
Polynemidae,n=0
Pleuronectiformes,n=92
Ambassidae,n=68
Mugiliformes,n=44
Pseudochromidae,n=2
Pomacentridae,n=53
Gobiesociformes,n=69
Blenniiformes,n=88
Cichliformes,n=94
Atheriniformes,n=72
Beloniformes,n=70
Cyprinodontiformes,n=79
Eupercaria,n=127
0
100
200
300
400
500
600
Maximun Score

Figure 4



Carassius auratus,n=1
Danio aesculapii,n=2
Danio rerio,n=9
Danio kyathit,n=6
Xyrauchen texanus,n=2
Myxocyprinus asiaticus,n=1
Phoxinus phoxinus,n=12
Squalius cephalus,n=39
Pimephales promelas,n=1
Cyprinus carpio,n=20
Barbus barbus,n=21
Carassius gibelio,n=4
Gymnocypris przewalskii,n=1
Sinocyclocheilus rhinocerous,n=1
Gobio gobio,n=43
Ctenopharyngodon idella,n=50
Megalobrama amblycephala,n=17
Hypophthalmichthys molitrix,n=1
0
200
400
60
Maximum Score

Figure 5


DNA Slippage Replication
Tandom Joint
269-3 bp, 89 aa
******
74-3 bp
24 aa
Tcb2 ROL46579.1:291-314
ID11 aDNA:
（269 bp）
268-2 bp, 89 aa
*
***
268-110 bp, 53 aa
Tcb2 ROL46579.1:114-166
Fragment Recombination
268-110 bp
74-3 bp
53 aa
24 aa
ROL46579.1:114-166; 291-314
ID11:269-3
Tcb2 transposase partial Anabarilius grahami Cypriniformes ROI15182.1
Tcb2 transposase Anabarilius grahami Cypriniformes ROI15182.1:291-314
ID11:268-2
TCB1 transposase Polyodon spathula Chondrostei MBN3281678.1:114-166
Tcb2 transposase Anabarilius grahami Cypriniformes ROI15182.1:114-166
Tcb2 transposase Anabarilius grahami Cypriniformes ROL46579.1:114-166
Tcb2 transposase Anabarilius grahami Cypriniformes ROI16594.1:112-163
Tcb2 transposase Anabarilius grahami Cypriniformes ROL45769.1:340-392

```
ID11:269-3                                                            SSSGSRPDCSSSS*PT*VGKCLNSMPSGTLQRCSLHG*IPVFTVQGSW*TTCHAV*SAS*YATPVRWMDYLGKGEVLTNTDLDRFLNN
Tcb2 transposase partial Anabarilius grahami Cypriniformes ROI15182.1 PFIHDHHLMLQHDNARPHVARICTQFLEAENIPVLAWPAYLHTYLHTYLPDMSPIEHVWDALDRVRWMDYLGKGEVLTNTDLDRFVNNI
Tcb2 transposase Anabarilius grahami Cypriniformes ROI15182.1:291-314 ----------------------------------------------------------------VRWMDYLGKGEVLTNTDLDRFVNN

ID11:268-2                                                            RPHRGLDLTAVRRRNRLEWANA*IRCRLALCRGVLFTDESRFSLYRAAGRQRVMLSNQHLDMPHL*GGWIISAKEKCSLTQI*TDF*TI
TCB1 transposase Polyodon spathula Chondrostei MBN3281678.1:114-166   HPHQDLDLTAVWRRNRLQWANAHLRWPLARWRSVLFPDESQFQLYWADGRQRV
Tcb2 transposase Anabarilius grahami Cypriniformes ROI15182.1:114-166 RPHRGLNLTAVRRHNRLECANAHIRWRLALWRGVLFTDESRFALYRADGRQRV
Tcb2 transposase Anabarilius grahami Cypriniformes ROL46579.1:114-166 RPHRGLDLTAVHRRNRLEWANAHIRWRLALWRGVLFTDESRFSLYRADGRQRV
Tcb2 transposase Anabarilius grahami Cypriniformes ROI16594.1:112-163 -PHRGLDLTAVRRRNRLEWANAHIRWRLALWRGVLFTDESRFSLYRADGRQRV
Tcb2 transposase Anabarilius grahami Cypriniformes ROL45769.1:340-392 RPHRGLDLTAVCRCNRLEWANAHIRWRLALWRGVLFTDESRFSLYRADGRQRV
```

Supplemental Materials

Insights from the Early Cretaceous: The promise of Lycoptera aDNA sequencing

Wanqian Zhao[1]▯ , Zhanyong Guo[2] , Zengyuan Tian[1] , Tongfu Su[3] , Gangqiang Cao[1] , Zixin Qi[4] , Tiancang Qin5, Wei Zhou[6], Jinyu Yang[7], Mingjie Chen[8], Xinge Zhang[9], Chunyan Zhou[10], Chuanjia Zhu[11], Mengfei Tang[1], Di Wu[1], Meirong Song[3], Yuqi Guo[11]▯, Liyou Qiu[12]▯

[1]School of Agricultural Sciences, Zhengzhou University, Zhengzhou, China.

[2]College of Agronomy, Henan Agricultural University, Zhengzhou, China.

[3]College of Sciences, Henan Agricultural University, Zhengzhou, China.

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

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

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

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

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

[9]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

✉e-mail: wqzhao@zzu.edu.cn; guoyuqi@zzu.edu.cn; qliyou@henau.edu.cn

Table S1 Evaluation of 213 Fish QMRs Using the 3SONB Method

ID	Len	NCBI Nucleotide BLAST in Actinopterygii		NCBI Nucleotide BLAST in Actinopterygii excluding the Order of fish species found in the first serach		NCBI Nucleotide BLAST excluding Actinopterygii	
1	101	Epinephelus fuscoguttatus Perciformes	115 115 100% 1e-22 85.15% 48981850 AP022677.1	Trachurus trachurus Carangiformes	115 115 100% 1e-22 85.15% 40754244 LR991628.1	NS	
2	111	Epinephelus fuscoguttatus Perciformes	77.9 77.9 45% 4e-11 94.00% 47245075 AP022679.1	Anabas testudineus Anabantiformes	75.2 150 63% 1e-10 83.10% 25912841	NS	
4	133	Megalobrama amblycephala Cypriniformes	161 161 96% 2e-36 87.50% 3498 XM_048160605.1	Plectropomus leopardus Perciformes	86.0 349 96% 8e-14 75.00% 42685566 AP022700.1	Sus scrofa Mammalia	57.2 57.2 23% 6e-04 100.00% 1159 AK350897.1
5	96	Argyrosomus regius Eupercaria	88.7 88.7 68% 2e-14 89.39% 26722334 OX336896.1	Nothobranchius furzeri Cyprinodontiformes	82.4 82.4 67% 7e-13 87.69% 4220 XM_015960765.1	Gari tellinella Protostomia	73.4 73.4 71% 4e-09 84.06% 98983066 OV277861.1
8	97	Scophthalmus maximus Pleuronectiformes	80.6 699 55% 3e-12 92.59% 29500699 CP026247.1	Phoxinus phoxinus Cypriniformes	78.8 140 54% 1e-11 92.45% 38982246 OX424416.1	Erebia ligea Arthropoda	77.9 198 56% 4e-10 90.91% 18887656 OU785232.1
9	181	Scophthalmus maximus Pleuronectiformes	281 13504 100% 3e-74 94.94% 25207873 CP026252.1	Myripristis murdjan Holocentriformes	100 2509 83% 7e-18 75.68% 36448408 LR597557.1	Hyalella azteca Arthropoda	108 535 87% 5e-19 78.48% 861 XM_018152871.1
12	125	Scophthalmus maximus Pleuronectiformes	150 150 93% 3e-33 89.83% 21783159 CP026257.1	Trachurus trachurus Carangiformes	59.9 537 44% 1e-05 83.64% 35709794 LR991637.1	NS	
15	164	Lateolabrax maculatus Pempheriformes	158 440 100% 4e-37 77.42% 21943731 CP027283.1	Siniperca chuatsi Centrarchiformes	155 155 98% 3e-34 77.96% 7950 XM_044166374.1	Vipera ursinii Amniota	72.5 202 74% 3e-08 74.77% 45961943 OX365972.1
16	163	Lateolabrax maculatus Pempheriformes	86.0 86.0 50% 1e-13 82.93% 23369250 CP032589.1	Salarias fasciatus Blenniiformes	79.7 79.7 49% 2e-11 81.48% 27169852 LR597439.1	NS	
17	231	Lateolabrax maculatus Pempheriformes	123 123 42% 8e-25 86.67% 21427312 CP032594.1	Trachurus trachurus Carangiformes	104 104 37% 8e-19 87.36% 25462409 LR991648.1	NS	

18	72	Ctenopharyngodon idella Cypriniformes	113 113 100% 3e-22 94.44% 3223 XR_007928914.1	Mastacembelus armatus Synbranchiformes	100 100 90% 1e-18 93.85% 28130073 LR535835.1	NS	
19	150	Ctenopharyngodon idella Cypriniformes	220 220 92% 8e-54 94.96% 10135 XM_051893824.1	Symphodus melops Labriformes	50.0 50.0 24% 0.007 89.19% 30971358 OX393528.1	Musotima nitidalis Arthropoda	51.8 51.8 34% 0.028 85.19% 14672933 OX421435.1
21	96	Ctenopharyngodon idella Cypriniformes	102 102 100% 9e-19 83.33% 6310 EU311637.1	Acanthopagrus latus Spariformes	53.6 53.6 30% 3e-04 100.00% 31400991 LR884464.1	Pammene fasciana Arthropoda	51.8 51.8 34% 0.015 93.94% 22451097 OU452278.1
22	282	Ctenopharyngodon idella Cypriniformes	374 374 100% 4e-100 89.47% 7936 XR_007926978.1	Ictalurus punctatus Siluriformes	83.3 83.3 24% 2e-12 86.11% 1887825 OL314503.1	NS	
23	128	Ctenopharyngodon idella Cypriniformes	155 285 100% 2e-34 83.92% 9852 XM_051877695.1	Silurus aristotelis Siluriformes	63.5 63.5 63% 8e-07 75.56% 30835389 OX327762.1	NS	
25	276	Ctenopharyngodon idella Cypriniformes	348 348 76% 5e-92 96.70% 1416 JQ033140.1	Ictalurus punctatus Siluriformes	226 226 72% 9e-56 85.00% 772 XR_001814710.3	Neospora caninum Phylum Apicomplexa	69.8 69.8 61% 2e-07 69.14% 3492925 LN714488.1
26	264	Phoxinus phoxinus Cypriniformes	275 5375 95% 2e-70 85.90% 38982246 OX424416.1	Pangasianodon hypophthalmus Siluriformes	129 129 59% 2e-26 77.07% 6005 XR_008304136.1	Gossypium hirsutum Spermatophyta	73.4 73.4 22% 2e-08 86.67% 255 GQ393980.1
27	240	Ctenopharyngodon idella Cypriniformes	256 256 92% 2e-64 85.40% 4544 XM_051865783.1	NS		NS	
28	103	Megalobrama amblycephala Cypriniformes	113 113 91% 5e-22 87.23% 1700 XR_007186334.1	Thalassophryne amazonica Batrachoidiformes	49.1 49.1 37% 0.014 87.18% 99266177 LR722979.1	Sciurus carolinensis Mammalia	59.9 59.9 40% 1e-04 92.86% 83138899 LR738605.1
29	133	Ctenopharyngodon idella Cypriniformes	200 200 100% 7e-48 93.23% 7719 XM_051861225.1	Pungitius pungitius Perciformes	127 127 100% 3e-26 79.17% 2605 XM_037479800.1	NS	
32	179	Megalobrama amblycephala Cypriniformes	89.7 89.7 30% 1e-14 96.30% 1519 XR_007186922.1	Silurus aristotelis Siluriformes	66.2 66.2 28% 1e-07 88.24% 17684804 OX327779.1	Allophyes oxyacanthae Arthropoda	53.6 53.6 32% 0.010 83.33% 15758794 OW028623.1
33	189	Ctenopharyngodon idella Cypriniformes	233 11094 100% 5e-58 88.17% 3581 XM_051875141.1	Salvelinus namaycush Salmoniformes	104 370 96% 5e-19 77.65% 3420 XM_039000777.1	Ypsolopha sequella Arthropoda	86.9 403 88% 2e-12 73.53% 25099185 OW204005.1

34	184	Ctenopharyngodon idella Cypriniformes	254 311 91% 5e-64 93.49% 6503 XM_051886010.1	Silurus aristotelis Siluriformes	95.1 95.1 57% 2e-16 79.09% 34926686 OX327758.1	Chrysoperla carnea Arthropoda	53.6 53.6 23% 0.010 86.36% 94407144 FR997756.1
35	146	Barbus barbus Cypriniformes	129 6410 99% 1e-26 87.27% 42984203 OW387171.1	Poecilia reticulata Cyprinodontiformes	67.1 552 74% 8e-08 75.45% 27716212 LR880662.1	Marchantia polymorpha Streptophyta	53.6 53.6 34% 0.008 84.31% 26794015 AP019870.1
36	169	Ctenopharyngodon idella Cypriniformes	361 361 76% 8e-96 97.64% 1754 XM_051867941.1	Ictalurus punctatus Siluriformes	226 226 72% 9e-56 85.00% 772 XR_001814710.3	Neospora caninum Phylum Apicomplexa	69.8 69.8 61% 2e-07 69.14% 3492925 LN714488.1
37	156	Ctenopharyngodon idella Cypriniformes	251 251 60% 3e-63 96.15% 5192 XR_007925807.1	Esox lucius Esociformes	105 105 60% 2e-19 73.08% 2396 XM_020048938.2	NA	
38	194	Ctenopharyngodon idella Cypriniformes	325 647 65% 2e-85 96.92% 3118 XM_051872334.1	Tachysurus fulvidraco Siluriformes	107 161 65% 6e-20 73.33% 1764 XR_007139352.1	Erynnis tages Arthropoda	53.6 53.6 25% 0.011 84.31% 12258325 LR990081.1
39	149	Carassius gibelio Cypriniformes	265 1929 59% 4e-67 99.33% 4172 XM_052561820.1	Ictalurus furcatus Siluriformes	59.9 59.9 16% 2e-05 92.86% 2603 XM_053642894.1	NS	
40	225	Ctenopharyngodon idella Cypriniformes	204 204 45% 1e-48 96.00% 1543 XR_007925360.1	Micropterus dolomieu Centrarchiformes	113 113 35% 1e-21 85.57% 2372 XM_046042353.1	NS	
41	203	Phoxinus phoxinus Cypriniformes	139 9705 77% 5e-29 87.61% 29721935 OX424430.1	Micropterus dolomieu Centrarchiformes	79.7 79.7 16% 3e-11 94.12% 2372 XM_046042353.1	NS	
42	225	Epinephelus fuscoguttatus Perciformes	115 115 100% 1e-22 85.15% 48981850 AP022677.2	Micropterus dolomieu Centrarchiformes	79.7 79.7 15% 3e-11 94.12% 2372 XM_046042353.1	Gossypium turneri Spermatophyta	53.6 53.6 18% 0.021 80.33% 63881757 CP032580.1
45	239	Megalobrama amblycephala Cypriniformes	308 308 93% 4e-80 90.58% 3341 XM_048185412.1	Ictalurus furcatus Siluriformes	117 117 62% 1e-22 78.00% 6837 XM_053632952.1	NS	
46	216	Ctenopharyngodon idella Cypriniformes	313 313 94% 8e-82 94.15% 9710 XM_051901286.1	Epinephelus fuscoguttatus Perciformes	51.8 51.8 24% 0.003 82.14% 52288159 AP022676.1	NS	
47	171	Gobio gobio Cypriniformes	162 11916 100% 2e-36 87.50% 69340224 OX442377.1	Plectropomus leopardus Perciformes	84.2 84.2 32% 4e-13 92.86% 28624681 AP022718.1	Shargacucullia verbasci Arthropoda	55.4 55.4 32% 0.003 83.93% 13081469 OX387299.1

48	91	Phoxinus phoxinus Cypriniformes	115 1514 97% 1e-22 89.77% 29721935 OX424430.1	NS		NS	
49	184	Gobio gobio Cypriniformes	142 4069 63% 2e-30 86.44% 65974243 OX442378.1	NS		NS	
50	186	Ctenopharyngodon idella Cypriniformes	235 362 77% 1e-58 95.86% 2751 XM_051911173.1	Pangasianodon hypophthalmus Siluriformes	160 160 65% 7e-36 89.26% 4077 XM_034309655.2	Gossypium herbaceum Spermatophyta	52.7 52.7 27% 0.038 85.19% 95846026 CP106702.1
51	131	Ctenopharyngodon idella Cypriniformes	180 180 100% 6e-42 90.84% 4757 XR_007927225.1	Platichthys flesus Pleuronectiformes	64.4 64.4 63% 3e-07 77.11% 28160870 OX438722.1	NS	
52	176	Megalobrama amblycephala Cypriniformes	216 216 93% 1e-52 88.89% 109382 KT695799.1	Paramormyrops kingsleyae Osteoglossiformes	56.3 56.3 33% 2e-04 81.36% 960 XM_023805191.1	Buxus sempervirens Mesangiospermae	52.7 52.7 23% 0.035 87.80% 36455880 OX387193.1
53	210	Gobio gobio Cypriniformes	306 5603 99% 1e-79 93.14% 49649979 OX442394.1	Brienomyrus brachyistius Osteoglossiformes	97.8 97.8 73% 8e-17 74.68% 4775 XR_007384855.1	Patella vulgata Protostomia	102 3474 81% 3e-17 74.69% 90894639 OW026350.2
55	149	Cyprinus carpio Cypriniformes	189 367 100% 1e-44 88.59% 11248262 LN590672.1	Coregonus sp. Salmoniformes	57.2 57.2 27% 4e-05 90.24% 93459789 LR778253.1	NS	
58	143	Squalius cephalus Cypriniformes	194 194 100% 3e-46 90.34% 37242670 OX439264.1	NS		NS	
60	154	Phoxinus phoxinus Cypriniformes	188 188 98% 2e-44 86.84% 38554997 OX424417.1	Sparus aurata Spariformes	48.2 48.2 16% 0.025 100.00% 4595 XM_030438549.1	NS	
63	166	Phoxinus phoxinus Cypriniformes	188 188 86% 2e-44 88.89% 33058111 OX424427.1	Lepisosteus oculatus Semionotiformes	68.9 68.9 67% 3e-08 74.36% 5709 XM_015352266.1	NS	
65	177	Barbus barbus Cypriniformes	166 166 92% 2e-37 82.63% 33697344 OW387177.1	Polyodon spathula Acipenseriformes	50.0 50.0 25% 0.008 86.96% 5456 XM_041264901.1	Macaca mulatta Mammalia	52.7 52.7 20% 0.035 92.11% 210829 AC204671.4
66	150	Barbus barbus Cypriniformes	166 1159 80% 2e-37 90.98% 26607564 OW387159.1	Myripristis murdjan Holocentriformes	55.4 55.4 30% 2e-04 89.13% 34950760 LR597559.1	Timema genevievae Arthropoda	58.1 58.1 29% 7e-04 88.64% 706695 OE839132.1

68	171	Megalobrama amblycephala Cypriniformes	213 213 94% 1e-51 89.51% 5273 XR_007186451.1	Pungitius pungitius Perciformes	137 256 98% 7e-29 98.72% 30300527 OX438543.1	NS	
69	237	Megalobrama amblycephala Cypriniformes	152 230 47% 2e-33 91.15% 5988 XM_048187844.1	Scophthalmus maximus Pleuronectiformes	69.8 122 43% 1e-08 80.52% 30934145 CP026246.1	NS	
70	167	Ctenopharyngodon idella Cypriniformes	214 214 100% 4e-52 88.76% 5262 XM_051878240.1	Epinephelus fuscoguttatus Perciformes	49.1 49.1 22% 0.027 89.47% 42713404 AP022691.1	Aspergillus flavus Fungi	57.2 57.2 41% 8e-04 79.17% 3968625 CP051040.1
71	181	Squalius cephalus Cypriniformes	89.7 202 66% 1e-14 82.76% 47495150 OX439254.1	NS		NS	
72	123	Barbus barbus Cypriniformes	138 138 100% 2e-29 85.37% 31637155 OW387187.1	NS		NS	
73	115	Squalius cephalus Cypriniformes	179 31970 100% 5e-42 94.74% 47495150 OX439254.1	Gadus morhua Gadiformes	99.6 99.6 96% 1e-17 80.53% 29710654 LR633950.1	Agriopis aurantiaria Arthropoda	56.3 56.3 40% 0.002 86.27% 22923517 OU611981.1
74	134	Phoxinus phoxinus Cypriniformes	178 178 96% 2e-41 90.30% 41353901 OX424415.1	NS		NS	
75	260	Squalius cephalus Cypriniformes	380 380 97% 8e-102 93.28% 70742303 OX439247.1	Hyperoplus immaculatus Uranoscopiformes	69.8 69.8 22% 1e-08 85.71% 28541892 OX442364.1	Bos mutus Mammalia	62.6 62.6 21% 3e-05 86.21% 59497618 CP027090.1
76	134	Ctenopharyngodon idella Cypriniformes	184 184 100% 5e-43 90.30% 5551 XM_051873604.1	Mastacembelus armatus Synbranchiformes	147 147 90% 3e-32 86.78% 24075063 LR535847.1	Spisula solida Mollusca	51.8 51.8 24% 0.024 93.94% 49312819 OX365953.1
77	274	Gobio gobio Cypriniformes	323 323 88% 2e-84 90.53% 59214519 OX442380.1	Chanos chanos Gonorynchiforme	82.4 82.4 23% 2e-12 87.69% 40347274 LR697115.1	NS	
78	218	Cyprinus carpio Cypriniformes	299 299 99% 2e-77 88.94% 15849037 LN590698.1	NS		NS	
79	217	Cyprinus carpio Cypriniformes	297 297 99% 6e-77 88.89% 15849037 LN590698.1	NS		NS	

81	163	Barbus barbus Cypriniformes	167 26004 100% 5e-38 83.75% 17274176 OW387166.1	Silurus aristotelis Siluriformes	101 1006 60% 5e-18 88.46% 44060567 OX327753.1	Chrysolina oricalcia Arthropoda	77.9 1427 58% 8e-10 80.00% 85697287 OX101808.2
82	144	Barbus barbus Cypriniformes	167 26004 100% 5e-38 83.75% 17274176 OW387166.1	Coregonus sp. Salmoniformes	66.2 66.2 32% 8e-08 91.67% 43663377 LR778289.1	NS	
83	218	Gobio gobio Cypriniformes	311 311 92% 3e-81 94.06% 80596314 OX442375.1	NS		NS	
85	192	Ctenopharyngodon idella Cypriniformes	227 227 87% 8e-56 89.60% 5967 XM_051864177.1	Hypomesus transpacificus Osmeriformes	106 106 84% 1e-19 73.78% 2065 XR_006957837.1	Montipora capitata Cnidaria	53.6 53.6 23% 0.011 86.96% 40290779 OX421454.1
86	109	Barbus barbus Cypriniformes	144 144 88% 3e-31 92.78% 31202023 OW387191.1	NS		NS	
87	156	Phoxinus phoxinus OX424409.1	Cypriniformes 211 211 100% 4e-51 90.38% 55893807	Chanos chanos Gonorynchiformes	65.3 65.3 43% 3e-07 80.88% 23837493 LR697118.1	NS	
88	141	Gobio gobio Cypriniformes	228 228 100% 2e-56 92.31% 87331821 OX442374.1	Chanos chanos Gonorynchiformes	65.3 65.3 43% 3e-07 80.88% 23837493 LR697118.1	NS	
89	276	Megalobrama amblycephala Cypriniformes	327 327 98% 5e-86 84.30% 901 XR_007187503.1	Chanos chanos Gonorynchiformes	55.4 55.4 21% 3e-04 80.00% 23837493 LR697118.1	NS	
90	205	Barbus barbus Cypriniformes	220 330 99% 1e-53 82.27% 31791306 OW387185.1	Mastacembelus armatus Synbranchiformes	167 1502 94% 5e-38 78.89% 24075063 LR535847.1	NS	
91	159	Gobio gobio Cypriniformes	158 354 92% 3e-35 83.66% 49649979 OX442394.1	NS		NS	
92	138	Gobio gobio Cypriniformes	158 354 92% 3e-35 83.66% 49649979 OX442394.1	NS		NS	
93	218	Ctenopharyngodon idella Cypriniformes	269 320 88% 3e-68 91.10% 3807 XM_051889826.1	Seriola dumerili Carangiformes	234 234 88% 5e-58 86.29% 4106 XM_022757098.1	Aplidium turbinatum Urochordata	77.9 745 35% 1e-09 83.12% 56992732 OU964919.1

94	178	Barbus barbus Cypriniformes	211 32285 98% 6e-51 91.84% 31187475 OW387192.1	Pangasianodon hypophthalmus Siluriformes	160 160 80% 6e-36 84.72% 4077 XM_034309655.2	Phalacrocorax aristotelis Aves	52.7 52.7 33% 0.036 81.67% 46833656 OX451230.1
95	108	Plectropomus leopardus Perciformes	92.4 92.4 72% 2e-15 85.90% 25625633 AP022720.1	Phoxinus phoxinus Cypriniformes	90.6 90.6 72% 6e-15 86.08% 27719707 OX424432.1	Drymonia ruficornis Arthropoda	93.3 93.3 70% 6e-15 86.84% 11016040 OX401925.1
96	141	Barbus barbus Cypriniformes	173 173 99% 1e-39 87.86% 42984203 OW387171.1	NS		NS	
97	173	Megalobrama amblycephala Cypriniformes	242 242 100% 9e-61 91.38% 561 XR_007187763.1	NS		NS	
98	287	Squalius cephalus Cypriniformes	313 313 99% 1e-81 82.32% 31211700 OX439271.1	Sphaeramia orbicularis Kurtiformes	127 127 38% 7e-26 85.71% 47041096 LR597477.1	Mesocricetus auratus Mammalia	132 132 36% 8e-26 87.62% 1620 XM_040751667.1
99	104	Cyprinus carpio Cypriniformes	100 100 96% 3e-18 82.69% 24322175 LN590718.1	NS		NS	
100	209	Squalius cephalus Cypriniformes	275 275 100% 2e-70 89.10% 31982984 OX439270.1	Sparus aurata Spariformes	50.0 50.0 22% 0.010 84.78% 35512760 LR537130.1	Ailanthus altissimus Spermatophyta	56.3 56.3 15% 0.004 96.97% 28675488 OX327701.1
102	146	Cyprinus carpio Cypriniformes	128 128 98% 4e-26 79.45% 1027199 LN590773.1	NS		NS	
103	213	Squalius cephalus Cypriniformes	216 4395 95% 2e-52 84.13% 55872136 OX439250.1	NS		NS	
105	216	Ctenopharyngodon idella Cypriniformes	244 244 93% 3e-61 86.60% 2602 XM_051912523.1	NS		NS	
106	165	Ctenopharyngodon idella Cypriniformes	244 244 93% 3e-61 86.60% 2602 XM_051912523.1	NS		NS	
109	136	Gobio gobio Cypriniformes	188 1037 98% 1e-44 91.04% 59214519 OX442380.1	NS		NS	

111	135	Ctenopharyngodon idella Cypriniformes	191 191 87% 4e-45 95.76% 5003 XM_051880559.1	Pholis gunnellus Perciformes	63.5 63.5 38% 9e-07 86.54% 21874052 OU342821.1	Sarcophaga rosellei Arthropoda	51.8 51.8 25% 0.024 94.29% 86778848 OV884021.1
113	97	Squalius cephalus Cypriniformes	133 4098 100% 5e-28 90.62% 47495150 OX439254.1	Cottoperca gobio Perciformes	50.0 50.0 44% 0.004 86.36% 25704503 LR131921.1	NS	
114	131	Gobio gobio Cypriniformes	185 185 96% 2e-43 92.13% 54584426 OX442387.1	Thunnus maccoyii Scombriformes	49.1 49.1 29% 0.020 87.18% 30220260 OU343208.1	NS	
115	133	Cyprinus carpio Cypriniformes	123 369 100% 2e-24 81.20% 23490330 LN590711.1	NS		NS	
116	133	Phoxinus phoxinus Cypriniformes	158 1737 100% 2e-35 87.60% 37204025 OX424418.1	Ictalurus furcatus Siluriformes	80.6 80.6 90% 3e-12 76.42% 1990 XM_053624557.1	Acinonyx jubatus Mammalia	52.7 52.7 23% 0.024 96.77% 5554 XR_003414647.2
117	193	Phoxinus phoxinus Cypriniformes	243 243 100% 1e-60 88.60% 38554997 OX424417.1	Salmo trutta Salmoniformes	48.2 48.2 19% 0.031 89.47% 73168960 LR584420.1	Mya arenaria Protostomia	52.7 105 25% 0.038 84.00% 70040107 CP111023.1
118	221	Squalius cephalus Cypriniformes	238 22716 100% 5e-59 85.38% 33754728 OX439268.1	Pangasianodon hypophthalmus Siluriformes	86.9 146 74% 2e-13 76.23% 897 XR_004578770.2	Gossypium hirsutum Spermatophyta	61.7 61.7 27% 9e-05 81.97% 255 GQ393936.1
119	189	Cyprinus carpio Cypriniformes	242 242 98% 4e-60 87.24% 442829 LN591170.1	Gadus morhua Gadiformes	52.7 52.7 21% 0.003 87.80% 23744039 LR633964.1	Oppiella nova Arthropoda	56.3 56.3 30% 0.003 82.46% 10122 OC918857.1
120	126	Phoxinus phoxinus Cypriniformes	192 192 100% 1e-45 93.65% 25938185 OX424433.1	NS		NS	
121	113	Phoxinus phoxinus Cypriniformes	192 192 100% 1e-45 93.65% 25938185 OX424433.1	NS		NS	
122	167	Barbus barbus Cypriniformes	244 1515 100% 3e-61 93.29% 30825632 OW387196.1	NS		NS	
123	142	Gobio gobio Cypriniformes	129 19609 89% 1e-26 80.92% 65974243 OX442378.1	Plectropomus leopardus Perciformes	80.6 2703 38% 4e-12 95.92% 42059005 AP022701.1	NA	

124	121	Squalius cephalus Cypriniforme	126 299 70% 1e-25 93.90% 38168245 OX439261.1	Erpetoichthys calabaricus Polypteriformes	97.8 3673 78% 4e-17 90.14% 89129353 LR536449.2	Bombina bombina Amphibia	78.8 78.8 68% 1e-10 80.72% 1069 XR_008401783.1
125	278	Phoxinus phoxinus Cypriniformes	232 870 82% 3e-57 83.33% 36753459 OX424419.1	Denticeps clupeoides Clupeiformes	78.8 78.8 25% 3e-11 83.78% 21430986 LR535827.1	Sus scrofa Mammalia	54.5 54.5 15% 0.017 88.10% 81313176 CP071562.1
126	171	Squalius cephalus Cypriniformes	236 15329 100% 1e-58 90.23% 45672686 OX439256.1	Silurus aristotelis Siluriformes	76.1 1877 64% 2e-10 75.23% 22864487 OX327772.1	Patella depressa Mollusca	54.5 107 31% 0.010 85.11% 90014739 OX419715.1
127	224	Ctenopharyngodon idella Cypriniformes	326 326 100% 1e-85 92.44% 3806 XM_051874321.1	Parambassis ranga Ovalentaria	49.1 49.1 12% 0.039 96.55% 32179582 LR131977.1	Homo sapiens Mammalia	56.3 56.3 23% 0.004 83.02% 102094 AL445437.4
128	229	Ctenopharyngodon idella Cypriniformes	170 170 45% 7e-39 96.15% 4009 XM_051862388.1	Ictalurus furcatus Siluriformes	124 124 41% 6e-25 89.90% 3380 XM_053652116.1	Pinctada fucata Mollusca	51.8 51.8 21% 0.048 84.62% 88058079 AP027118.1
129	167	Ctenopharyngodon idella Cypriniformes	200 3285 99% 9e-48 87.12% 4137 XM_051873850.1	Salarias fasciatus Ovalentaria	126 655 99% 1e-25 79.58% 34246461 LR597451.1	Agelas oroides Porifera	92.4 1795 100% 4e-14 72.39% 6990495 OX422204.1
130	140	Phoxinus phoxinus Cypriniformes	207 207 100% 5e-50 93.01% 32771178 OX424428.1	Poecilia reticulata Cyprinodontiformes	49.1 49.1 29% 0.022 87.80% 27016956 LR880652.1	NS	
131	101	Gobio gobio Cypriniformes	91.5 1007 82% 2e-15 84.88% 58069932 OX442384.1	Alosa sapidissima Clupeiformes	54.5 54.5 41% 3e-04 88.10% 2136 XM_042099293.1	NS	
132	141	Ctenopharyngodon idella Cypriniformes	178 227 75% 2e-41 96.33% 9375 XM_051881773.1	Silurus aristotelis Siluriformes	109 172 66% 8e-21 85.71% 38957035 OX327754.1	NS	
133	90	Cyprinus carpio Cypriniformes	113 113 86% 5e-22 92.41% 10116 LN593669.1	Ictalurus punctatus Siluriformes	74.3 74.3 58% 3e-10 90.57% 11790 XM_053674708.1	Gossypium hirsutum Spermatophyta	57.2 57.2 45% 3e-04 90.24% 255 GQ393936.1
134	105	Cyprinus carpio Cypriniformes	117 117 89% 5e-23 87.50% 185877 LN594174.1	NS		NS	
135	105	Cyprinus carpio Cypriniformes	122 122 89% 1e-24 88.54% 185877 LN594174.1	NS		NS	

136	112	Megalobrama amblycephala Cypriniformes	169 169 100% 9e-39 93.81% 6208 XM_048208505.1	NS		NS	
140	116	Gobio gobio Cypriniformes	167 6078 100% 3e-38 92.24% 49649979 OX442394.1	NS		NS	
141	249	Squalius cephalus Cypriniformes	164 5241 99% 1e-36 77.25% 41795634 OX439257.1	Epinephelus fuscoguttatus Perciformes	84.2 390 77% 6e-13 78.57% 36946246 AP022697.1	Alitta virens Annelida	58.1 58.1 24% 0.001 82.54% 36633593 OW028585.1
142	233	Barbus barbus Cypriniformes	145 1177 92% 3e-31 74.78% 30983155 OW387194.1	Silurus aristotelis Siluriformes	86.0 86.0 48% 2e-13 76.47% 22864487 OX327772.1	Bombylius discolor Arthropoda	54.5 54.5 23% 0.014 82.14% 42140605 OW584241.1
144	250	Barbus barbus Cypriniformes	150 7557 55% 2e-32 86.15% 31977477 OW387183.1	Silurus aristotelis Siluriformes	142 1290 55% 3e-30 88.03% 30835389 OX327762.1	NS	
145	102	Phoxinus phoxinus Cypriniformes	116 532 74% 4e-23 94.59% 35606265 OX424422.1	NS		NS	
146	119	Cyprinus carpio Cypriniformes	168 265 100% 3e-38 90.24% 136314 LN596051.1	Pholis gunnellus Perciformes	51.8 51.8 29% 0.001 94.29% 18512999 OU342825.1	NS	
147	247	Ctenopharyngodon idella Cypriniformes	200 267 80% 1e-47 85.71% 3203 XM_051887217.1	Sphaeramia orbicularis Kurtiformes	66.2 23623 33% 2e-07 95.35% 56493859 LR597468.1	Anopheles funestus◦ Arthropoda	66.2 637 25% 2e-06 85.71% 84636641 OX030924.1
148	210	Gobio gobio Cypriniformes	325 325 100% 1e-85 94.29% 80596314 OX442375.1	NS		NS	
149	288	Phoxinus phoxinus Cypriniformes	470 470 95% 7e-129 97.82% 37204025 OX424418.1	Colossoma macropomum Characiformes	178 178 59% 4e-41 83.04% 8810 XM_036588098.1	NS	
150	248	Barbus barbus Cypriniformes	113 1074 57% 2e-21 81.67% 29286462 OW387200.1	Silurus aristotelis Siluriformes	104 104 60% 7e-19 76.16% 12464635 OX327781.1	NS	
151	122	Barbus barbus Cypriniformes	107 834 72% 3e-20 88.37% 38967457 OW387174.1	Ictalurus furcatus Siluriformes	59.9 59.9 58% 1e-05 79.17% 2184 XM_053631430.1	NS	

153	77	Cyprinus carpio Cypriniformes	95.1 95.1 100% 1e-16 87.01% 2678220 LN598428.1	NS		NS	
154	258	Ctenopharyngodon idella Cypriniformes	279 279 77% 2e-71 90.14% 5078 XM_051905869.1	Ictalurus furcatus Siluriformes	61.7 61.7 19% 7e-06 85.19% 3356 XM_053625266.1	Vigna angularis Spermatophyta	65.3 65.3 21% 9e-06 87.27% 38115440 AP015044.1
158	93	Barbus barbus Cypriniformes	141 1804 89% 1e-30 97.59% 17274176 OW387166.1	NS		NS	
159	192	Barbus barbus Cypriniformes	297 297 100% 5e-77 94.27% 33248607 OW387179.1	Chanos chanos Gonorynchiformes	144 144 97% 5e-31 77.25% 43790413 LR697114.1	NS	
161	121	Squalius cephalus Cypriniformes	104 203 87% 4e-19 79.49% 41795634 OX439257.1	Mastacembelus armatus Synbranchiformes	66.2 66.2 38% 7e-08 91.30% 24075063 LR535847.1	NS	
162	164	Phoxinus phoxinus Cypriniformes	89.7 89.7 98% 1e-14 73.46% 38982246 OX424416.1	Gadus morhua Gadiformes	86.9 322 99% 1e-13 72.84% 23744039 LR633964.1	Vespa crabro Arthropoda	88.7 142 98% 4e-13 72.46% 12120248 OU342403.1
164	141	Argyrosomus regius Eupercaria	104 309 99% 4e-19 82.14% 24167187 OX336901.1	Neostethus bicornis Atheriniformes	82.4 672 87% 1e-12 76.47% 34385357 LR738543.1	Psylliodes chrysocephala Arthropoda	128 128 100% 4e-25 81.25% 38386302 OV651825.1
166	165	Gouania willdenowi Blenniiformes	149 7864 86% 1e-32 83.10% 72451976 LR132000.1	Acanthochromis polyacanthus Ovalentaria	100 366 83% 7e-18 79.17% 951 XM_051960172.1	Cryptosula pallasiana Protostomia	130 417 84% 1e-25 82.55% 51911826 OX183133.1
168	161	Sparus aurata Spariformes	109 7234 96% 1e-20 77.08% 38249111 LR537125.1	Pholis gunnellus Perciformes	94.2 1013 96% 9e-16 73.78% 20225815 OU342824.1	Eupithecia exiguata Arthropoda	87.8 237 95% 4e-13 74.05% 15327645 OX352229.1
173	149	Ctenopharyngodon idella Cypriniformes	221 221 100% 2e-54 91.03% 6215 XM_051916151.1	Cottoperca gobio Perciformes	203 203 100% 5e-49 88.46% 15932363 LR131931.1	Eptatretus burgeri Cyclostomata	144 144 100% 6e-30 80.13% 771 LC043689.1
174	145	Takifugu rubripes Tetraodontiformes	80.6 621 88% 5e-12 75.19% 16367601 LR584241.1	Danio rerio Cypriniformes	80.6 1618 90% 5e-12 79.17% 73995704 CP068741.1	Homo sapiens Mammalia	77.0 134 89% 7e-10 74.24% 180064 AC025265.21
175	164	Salmo trutta Salmoniformes	83.3 4665 54% 2e-12 83.75% 66900148 LR584415.1	Poecilia reticulata Cyprinodontiformes	76.1 76.1 43% 2e-10 82.67% 27716212 LR880662.1	Bos mutus Mammalia	85.1 85.1 89% 5e-12 74.15% 78623970 CP027084.1

176	142	Danio kyathit Cypriniformes	124 124 82% 5e-25 83.90% 89469171 LR812525.1	Hyperoplus immaculatus Uranoscopiformes	120 120 86% 4e-24 82.40% 26757407 OX442366.1	Ecdyonurus torrentis Arthropoda	119 119 85% 2e-22 82.93% 50407882 OX439131.1
177	147	Ctenopharyngodon idella Cypriniformes	238 6325 97% 3e-59 96.53% 4536 XM_051885310.1	Myripristis murdjan Holocentriformes	66.2 181 82% 9e-08 76.19% 42201344 LR597550.1	Candidatus Eisenbacteria Bacteria	66.2 125 99% 1e-06 82.50% 2690560 CP070702.1
178	152	Syngnathus acus Syngnathiformes	145 5545 100% 2e-31 82.31% 28444102 LR594595.1	Takifugu rubripes Tetraodontiformes	141 26699 100% 7e-30 82.14% 20050447 LR584233.2	Haliaeetus albicilla Aves	145 20314 100% 2e-30 82.31% 51297233 OX381641.1
179	178	Salarias fasciatus Blenniiformes	204 204 94% 8e-49 86.98% 41296897 LR597441.1	Thunnus albacares Scombriformes	79.7 79.7 34% 2e-11 88.52% 38830372 OU607594.1	NS	
180	159	Girardinichthys multiradiatus Cyprinodontiformes	110 250 92% 1e-20 77.30% 753 XM_047367909.1	Myripristis murdjan Holocentriformes	94.2 464 100% 9e-16 73.03% 47119472 LR597552.1	Gallus gallus Aves	243 2.173e+05 100% 9e-60 94.27% 3700872 CP100592.2
181	92	Salarias fasciatus Blenniiformes	92.4 184 79% 2e-15 87.67% 36057378 LR597446.1	Astatotilapia calliptera Cichliformes	59.0 59.0 51% 8e-06 87.23% 1347 XM_026184664.1	NS	
182	178	Sphaeramia orbicularis Kurtiformes	157 6315 95% 1e-34 79.89% 63265450 LR597460.	Sparus aurata Spariformes	141 1069 94% 8e-30 77.46% 31047742 LR537139.1	NS	
184	110	Gadus morhua Gadiformes	75.2 1333 57% 1e-10 86.89% 28829685 R633955.1	Coregonus sp.Salmoniformes	67.1 2095 62% 7e-08 81.16% 55641933 LR778267.1	Acomys russatus Mammalia	68.0 545 64% 2e-07 81.69% 69034057 LR877218.1
185	163	Pseudochaenichthys georgianus Perciformes	124 237 93% 5e-25 81.29% 48837966 LR792554.1	Trachurus trachurus Carangiformes	118 271 82% 2e-23 82.98% 37400736 LR991632.1	Ophion luteus Arthropoda	119 735 87% 3e-22 81.88% 67162817 OX101784.1
186	259	Pseudochaenichthys georgianus Perciformes	123 234 35% 6e-24 81.17% 48837966 LR792554.1	Trachurus trachurus Carangiformes	116 726 51% 2e-22 82.86% 37400736 LR991632.1	Ophion luteus Arthropoda	117 902 52% 3e-21 81.76% 67162817 OX101784.1
188	249	Chanos chanos Gonorynchiformes	75.2 196 20% 4e-10 92.16% 53909382 LR697109.1	Epinephelus fuscoguttatus Perciformes	71.6 71.6 19% 5e-09 91.84% 47245075 AP022679.1	Pipistrellus pipistrellus Mammalia	75.2 75.2 20% 5e-09 92.16% 53281546 LR862369.1
189	159	Phoxinus phoxinus Cypriniformes	90.6 18950 59% 2e-14 77.50% 44815072 OX424413.1	Silurus aristotelis Siluriformes	89.7 50805 64% 1e-14 78.45% 30578072 OX327764.1	Tridacna crocea Mollusca	89.7 711 58% 2e-13 78.45% 48376938 OX031069.1

190	144	Argyrosomus regius Eupercaria	233 362 97% 1e-57 96.45% 33736778 OX336883.1	Toxotes jaculatrix Percomorphaceae	75.2 75.2 49% 2e-10 83.10% 4212 XM_041057369.1	NS	
191	117	Argyrosomus regius Eupercaria	176 461 100% 6e-41 93.16% 30788448 OX336885.1	Thunnus albacares Scombriformes	69.8 645 45% 6e-09 88.68% 30700134 OU607607.1	NS	
193	170	Nibea albiflora Eupercaria	113 113 85% 1e-21 77.78% 26229440 LR699038.1	Pleuronectes platessa Pleuronectiformes	99.6 99.6 92% 2e-17 73.78% 27621514 OX374657.1	Tyria jacobaeae Arthropoda	53.6 53.6 17% 0.009 100.00% 23541150 OX387214.1
194	206	Thalassophryne amazonica Batrachoidiformes	146 12897 99% 2e-31 77.83% 92122960 LR722981.1	Coregonus sp. Salmoniformes	88.7 1456 92% 5e-14 72.59% 54335267 LR778272.1	Erithacus rubecula Aves	150 4724 95% 2e-31 78.28% 6516792 LR812129.1
195	111	Thalassophryne amazonica Batrachoidiformes	159 1055 99% 5e-36 91.96% 74844089 LR722983.1	Sparus aurata Spariformes	75.2 75.2 99% 1e-10 74.56% 31425404 LR537142.1	NS	
196	118	Coregonus sp. Salmoniformes	141 141 95% 1e-30 87.61% 68138733 LR778260.1	Esox lucius Esociformes	140 140 94% 4e-30 87.50% 1507 XM_020043446.3	NS	
197	131	Danio kyathit Cypriniformes	100 1724 87% 5e-18 80.87% 67457037 LR812526.1	Neostethus bicornis Atheriniformes	95.1 7333 89% 2e-16 78.07% 31437157 LR738547.1	Pan troglodytes Mammalia	98.7 152 87% 2e-16 80.00% 218680 BS000171.1
198	173	Coregonus sp. Salmoniformes	86.9 86.9 42% 1e-13 87.84% 40727438 LR778286.1	Gadus morhua Gadiformes	80.6 80.6 63% 6e-12 78.38% 21723002 LR633959.1	Anopheles ziemanni Arthropoda	84.2 84.2 59% 6e-12 80.95% 94852749 OX030919.2
199	181	Pseudochaenichthys georgianus Perciformes	72.5 2249 38% 4e-09 88.14% 28991432 LR792564.1	Silurus aristotelis Siluriformes	51.8 51.8 50% 0.003 75.51% 17684804 OX327779.1	Apteryx mantelli Aves	63.5 63.5 44% 2e-05 79.01% 405270 LK065603.1
200	118	Ctenopharyngodon idella Cypriniformes	112 170 87% 2e-21 95.65% 7080 XM_051897731.1	NS		NS	
201	248	Danio rerio Cypriniformes	102 102 31% 3e-18 89.74% 59354185 LR812068.1	Silurus aristotelis Siluriformes	93.3 93.3 26% 2e-15 90.91% 38141395 OX327755.1	NS	
208	180	Danio kyathit Cypriniformes	224 67335 100% 2e-55 90.30% 73064863 LR812520.1	Thalassophryne amazonica Batrachoidiformes	218 17128 100% 2e-53 90.62% 77050466 LR722982.1	Bombina bombina Amphibia	170 228 100% 5e-38 93.10% 987 XR_008401981.1

209	110	Gobio gobio Cypriniformes	395 395 99% 4e-106 93.51% 49156386 OX442395.1	Myripristis murdjan Holocentriformes	91.5 91.5 38% 4e-15 80.39% 39034921 LR597563.1	NS	
210	108	Ctenopharyngodon idella Cypriniformes	177 273 95% 1e-40 81.46% 10940 XM_051871826.1	Chanos chanos Gonorynchiformes	99.6 214 95% 2e-17 75.16% 24698569 LR697117.1	NS	
211	237	Danio kyathit Cypriniformes	90.6 1091 71% 9e-15 83.52% 74256860 LR812524.1	Coregonus sp. Salmoniformes	82.4 485 71% 9e-13 82.56% 92224161 LR778256.1	Apteryx rowi Aves	88.7 88.7 70% 3e-13 80.90% 2847 XM_026074609.1
213	184	Barbus barbus Cypriniformes	258 307 89% 5e-65 90.10% 28318050 OW387153.1	Denticeps clupeoides Clupeiformes	109 109 57% 1e-20 81.30% 22537163 LR535824.1	Dolichovespula media Arthropoda	66.2 66.2 21% 2e-06 91.30% 10456037 OU426856.1
214	167	Danio kyathi Cypriniformes	70.7 1079 99% 8e-09 75.19% 57421999 R812530.1	Acipenser ruthenus Acipenseriformes	60.8 108 98% 3e-06 74.53% 14920740 OV754640.1	Macrobrachium nipponense Arthropoda	68.9 284 68% 3e-07 79.76% 83001933 CP062038.1
215	113	Squalius cephalus Cypriniformes	190 190 95% 3e-45 96.58% 35998302 OX439267.1	Taurulus bubalis Perciformes	51.8 51.8 31% 0.001 89.47% 50132026 OU342711.1	NS	
216	73	Ctenopharyngodon idella Cypriniforme	142 142 48% 3e-30 94.51% 10940 XM_051871826.1	Thalassophryne amazonica Batrachoidiformes	95.1 1657 34% 2e-16 93.55% 70111961 LR722985.1	NS	
217	160	Gobio gobio Cypriniformes	391 391 95% 5e-105 94.47% 57298966 OX442385.1	NS		NS	
220	263	Acanthochromis polyacanthus Ovalentaria	71.6 71.6 35% 3e-09 89.29% 3710 XM_051950156.1	Argyrosomus regius Eupercaria	66.2 66.2 26% 1e-07 95.35% 24167187 OX336901.1	Mus musculus Mammalia	67.1 67.1 28% 1e-06 93.48% 97401718 OX390159.1
221	171	Oryzias latipes Beloniformes	133 11724 55% 2e-27 81.29% 29208093 CP020784.1	Gasterosteus aculeatus Perciformes	133 273 50% 2e-27 81.29% 127186 GU130437.1	Yponomeuta sedellus Arthropoda	132 737 51% 8e-26 80.71% 26282666 OW203724.1
222	126	Salarias fasciatus Blenniiformes	73.4 1028 37% 1e-09 82.86% 34145682 LR597455.1	Acanthopagrus latus Spariformes	59.0 320 30% 2e-05 83.87% 26679254 LR884471.1	Harmothoe impar Annelida	70.7 141 32% 1e-07 85.25% 87888585 OX381707.1
223	210	Acanthopagrus latus Spariformes	91.5 1692 96% 8e-16 93.33% 26170885 LR884481.1	Anguilla anguilla Anguilliformes	64.4 124 96% 1e-07 83.33% 2408 XR_004765272.1	Pinctada fucata Mollusca	56.3 56.3 61% 5e-04 92.11% 57001201 AP027113.1

226	127	Larimichthys crocea Eupercaria	140 200 100% 3e-30 90.82% 7081103 LT972186.1	NS		NS	
229	122	Ctenopharyngodon idella Cypriniformes	187 187 43% 1e-43 95.69% 6325 XM_051860994.1	Mastacembelus armatus Synbranchiformes	79.7 79.7 24% 3e-11 89.23% 28732325 LR535853.1	NS	
230	187	Ctenopharyngodon idella Cypriniformes	222 222 79% 4e-54 86.77% 5624 XM_051880066.1	Oryzias latipes Beloniformes	49.1 49.1 14% 0.041 91.18% 30564142 CP020635.1	NS	
231	264	Myripristis murdjan Holocentriformes	76.1 76.1 24% 4e-10 91.07% 28934276 LR597568.1	Taurulus bubalis Perciformes	76.1 132 25% 4e-10 91.38% 50132026 OU342711.1	Nigrospora oryzae Fungi	393 393 100% 1e-104 98.67% 4222472 CP096792.1
232	157	Silurus aristotelis Siluriformes	301 41365 99% 5e-78 93.47% 30835389 OX327762.1	Myxocyprinus asiaticus Cypriniformes	297 297 98% 6e-77 92.96% 5798 XM_051659601.1	Rana temporaria Amphibia	247 247 98% 1e-60 87.44% 3196 XM_040322511.1
233	276	Pholis gunnellus Perciformes	360 360 100% 7e-96 92.68% 27911860 OU342810.1	Thunnus maccoyii Scombriformes	345 420 100% 1e-91 90.87% 34909826 OU343200.1	Lumbricus rubellus Annelida	102 102 42% 3e-17 82.24% 49323427 OX243815.1
234	188	Pholis gunnellus Perciformes	76.1 76.1 55% 2e-10 87.69% 16414073 OU342826.1	Scophthalmus maximus Pleuronectiformes	71.6 71.6 55% 2e-09 84.62% 23836697 CP026244.1	Lumbricus rubellus Annelida	77.0 134 59% 5e-10 84.29% 32701251 OX243827.1
236	62	Sparus aurata Spariformes	95.1 95.1 78% 3e-16 75.50% 35791275 LR537133.1	Thunnus albacares Scombriformes	88.7 88.7 41% 5e-14 85.00% 32838394 OU607604.1	NS	
237	90	Phoxinus phoxinus Cypriniformes	98.7 216 36% 3e-17 93.75% 27719707 OX424432.1	Thunnus albacares Scombriformes	82.4 278 28% 2e-12 87.69% 30510322 OU607609.1	NS	
238	266	Phoxinus phoxinus Cypriniformes	98.7 216 36% 3e-17 93.75% 27719707 OX424432.1	Thunnus albacares Scombriformes	82.4 278 28% 2e-12 87.69% 30510322 OU607609.1	NS	
239	236	Phoxinus phoxinus Cypriniformes	147 1756 98% 4e-32 84.73% 38982246 OX424416.1	Acipenser ruthenus Acipenseriformes	144 1523 98% 4e-31 84.56% 29433752 OV737706.1	Gallus gallus Aves	226 1614 100% 5e-55 97.08% 15361164 CP100574.1
242	202	Sinocyclocheilus rhinocerous Cypriniformes	73.4 73.4 47% 6e-10 89.09% 1838 XM_016522593.1	NS		NS	

243	246	Megalobrama amblycephala Cypriniformes	268 346 93% 4e-68 96.32% 4719 XM_048153829.1	Paralichthys olivaceus Pleuronectiformes	49.1 49.1 14% 0.049 87.18% 820 XM_020090886.1	NS	
244	118	Ctenopharyngodon idella Cypriniformes	199 19830 100% 1e-47 91.43% 6477 XM_051868490.1	Hyperoplus immaculatus Uranoscopiformes	48.2 48.2 17% 0.036 88.89% 35090712 OX442354.1	Cryptosula pallasiana Protostomia	58.1 111 50% 0.001 73.87% 46264604 OX183135.1
245	193	Squalius cephalus Cypriniformes	87.8 87.8 55% 3e-14 87.50% 55872136 OX439250.1	NS		NS	
246	235	Ictalurus furcatus Siluriformes	291 291 100% 3e-75 91.26% 1789 XM_053623338.1	Barbus barbus Cypriniformes	247 485 100% 9e-62 86.89% 42984203 OW387171.1	NS	
248	235	Salvelinus alpinus Salmoniformes	96.0 1082 72% 2e-16 83.17% 1497 XM_024141615.1	Taurulus bubalis Perciformes	94.2 2589 62% 8e-16 83.15% 28944496 OU342720.1	Branchellion lobata Annelida	95.1 1174 62% 3e-15 85.86% 10374467 OX387253.1
253	129	Cottoperca gobio Perciformes	111 483 98% 2e-21 84.16% 1276 XM_029438179.1	Coregonus sp. Salmoniformes	76.1 410 96% 1e-10 76.77% 44662967 LR778285.1	Branchellion lobata Annelida	108 876 100% 2e-19 88.89% 8106404 OX387259.1
254	116	Sparus aurata Spariformes	212 212 88% 1e-51 96.85% 9291 XM_030409705.1	Thunnus maccoyii Scombriformes	205 205 93% 2e-49 92.25% 20451074 OU343214.1	NS	
255	271	Oncorhynchus kisutch Salmoniformes	72.5 213 98% 3e-09 73.57% 1548 XM_031800097.1	Chanos chanos Gonorynchiformes	62.6 751 93% 1e-06 70.54% 50300218 LR697111.1	Ostrea edulis Mollusca	94.2 8682 96% 8e-15 77.69% 48989042 OX387713.1
256	209	Anarrhichthys ocellatus Perciformes	95.1 95.1 52% 2e-16 87.01% 1559 XM_031848587.1	Xiphias gladius Istiophoriformes	91.5 91.5 50% 3e-15 86.67% 1585 XM_040124043.1	Epithele typhae Fungi	86.9 86.9 64% 1e-12 81.25% 606 XM_048026079.1
257	126	Acipenser ruthenus Acipenseriformes	69.8 69.8 40% 5e-09 100.00% 6266 XM_034902013.1	Parambassis ranga Ovalentaria	67.1 117 41% 6e-08 97.44% 26339914 LR131975.1	Macrobrachium nipponense Arthropoda	72.5 280 58% 2e-08 90.74% 58878174 CP062055.1
258	206	Pangasianodon hypophthalmus Siluriformes	310 310 100% 1e-80 87.07% 14800 XM_026929691.3	Colossoma macropomum Characiformes	146 146 82% 3e-31 70.63% 10849 XM_036566310.1	Piscicola geometra Annelida	80.6 80.6 34% 1e-10 80.68% 7659117 OX030966.1
263	148	Salvelinus namaycush Salmoniformes	112 3729 76% 3e-21 83.81% 1035 XM_038961910.1	Gouania willdenowi Blenniiformes	107 1834 76% 3e-20 83.02% 36801689 LR131983.1	Piscicola geometra Annelida	123 27090 81% 2e-23 85.71% 9301023 OX030961.1

264	92	Megalobrama amblycephala Cypriniformes	115 193 86% 2e-22 98.48% 1313 XM_048153573.1	NS		NS	
266	103	Squalius cephalus Cypriniformes	288 288 99% 4e-74 87.50% 49162881 OX439252.1	Cottoperca gobio Perciformes	48.2 48.2 21% 0.041 82.69% 22902521 LR131919.1	NS	
267	144	Megalobrama amblycephala Cypriniformes	232 232 91% 2e-57 90.00% 5181 XM_048161894.1	Parambassis ranga Ovalentaria	48.2 48.2 20% 0.031 91.89% 24575694 LR131963.1	NS	
270	138	Squalius cephalus Cypriniformes	141 1323 70% 2e-30 93.55% 37306196 OX439263.1	Mastacembelus armatus Synbranchiformes	86.0 86.0 37% 1e-13 88.06% 19675934 LR535850.1	NS	
271	148	Cheilinus undulatus Labriformes	88.7 1315 86% 2e-14 82.83% 996 XM_041785140.1	Sphaeramia orbicularis Kurtiformes	86.9 665 87% 8e-14 81.52% 62138941 LR597462.1	Chondrosia reniformis Porifera	83.3 580 88% 1e-11 81.19% 6903244 OX359201.1
272	93	Alosa alosa Clupeiformes	108 108 62% 3e-20 95.52% 5476 XM_048234029.1	Scleropages formosus Osteoglossiformes	92.4 92.4 58% 2e-15 92.06% 4578 XM_029251347.1	Calidris pugnax Aves	74.3 74.3 58% 5e-09 85.71% 2667 XR_001412922.1
273	271	Megalobrama amblycephala Cypriniformes	326 326 99% 1e-85 91.14% 1305 XR_007179981.1	Thunnus albacares Scombriformes	76.1 187 37% 3e-10 80.68% 38830372 OU607594.1	NS	
274	137	Danio kyathit Cypriniformes	78.8 312 31% 2e-11 89.66% 48699847 LR812522.1	Chanos chanos Gonorynchiformes	77.0 136 32% 6e-11 89.47% 59736555 LR697107.1	Branchellion lobata Annelida	79.7 79.7 25% 3e-10 97.83% 6982612 OX387261.1
275	122	Ctenopharyngodon idella Cypriniformes	122 122 43% 2e-24 97.22% 4918 XM_051883090.1	NS		Calidris pugnax Aves	74.3 74.3 58% 5e-09 85.71% 2667 XR_001412922.1
276	248	Gobio gobio Cypriniformes	128 341 99% 3e-26 93.98% 65974243 OX442378.1	Denticeps clupeoides Clupeiformes	86.9 86.9 96% 6e-14 78.76% 1711 XR_003742578.1	NS	
281	185	Gobio gobio Cypriniformes	91.5 844 97% 1e-15 88.57% 87331821 OX442374.1	Notothenia rossii Perciformes	49.1 49.1 47% 0.008 91.67% 93745879 OX449115.1	NS	
3	149	Gadus morhua Gadiformes	86.0 86.0 91% 1e-13 75.91% 30713045 LR633953.1	Danio rerio Cypriniformes	73.4 293 92% 8e-10 78.18% 38183376 LR812087.1	Morus notabilis Spermatophyta	75.2 75.2 57% 2e-09 80.46% 1413 XM_024175174.1

212	252	Sphaeramia orbicularis Kurtiformes	88.7 13038 44% 7e-14 79.61% 62138941 LR597462.1	Danio kyathit Cypriniformes	88.7 23694 53% 7e-14 80.61% 63723151 LR812531.1	Mus musculus Mammalia	89.7 211 37% 2e-13 81.91% 204907 AL772249.6
218	168	Barbatula barbatula Cypriniformes	165 19559 100% 2e-37 82.25% 22832458 OX344785.1	Chanos chanos Gonorynchiformes	160 47440 100% 6e-36 81.76% 49187548 LR697112.1	Biomphalaria glabrata Protostomia	162 55150 100% 3e-35 81.55% 55047444 OX365777.1
219	199	Phoxinus phoxinus Cypriniformes	241 413 100% 4e-60 87.25% 25938185 OX424433.1	Silurus aristotelis Siluriformes	91.5 364 56% 3e-15 89.19% 32150870 OX327759.1	Scyliorhinus canicula Chondrichthyes	78.8 78.8 26% 3e-10 92.45% 1828 XM_038807777.1
235	152	Danio kyathit Cypriniformes	159 42731 100% 7e-36 84.21% 74256860 LR812524.1	Thunnus maccoyii Scombriformes	159 27456 100% 2e-35 83.66% 38771177 OU343192.1	Biomphalaria glabrata Mollusca	159 24732 100% 3e-34 83.77% 47236035 OX365780.1
205	222	Danio rerio Cypriniformes	95.1 2059 89% 4e-16 76.24% 55327790 LR812070.1	Notothenia rossii Perciformes	57.2 106 85% 7e-05 69.54% 85894744 OX449117.1	Diadumene lineata Cnidaria	90.6 844 99% 2e-13 71.54% 42014270 OU974069.1
206	109	Danio aesculapii Cypriniformes	123 14676 73% 3e-25 100.00% 55069896 LR812495.1	Oryzias latipes Beloniformes	120 4875 71% 3e-24 100.00% 28246893 CP020790.1	Mya arenaria Protostomia	123 3832 70% 3e-24 100.00% 71350902 CP111022.1
183	230	Dicentrarchus labrax Eupercaria	333 637 100% 9e-88 90.46% 14803407 FQ310507.3	Myripristis murdjan Holocentriformes	329 49257 100% 9e-87 91.42% 47119472 LR597552.1	Bufo bufo Amphibia	253 253 91% 8e-63 80.97% 459 XR_005773886.1
269	148	Megalobrama amblycephala Cypriniformes	236 236 100% 1e-58 95.27% 3225 XM_048208442.1	NS		NS	
165	133	Gouania willdenowi Blenniiformes	146 7736 100% 1e-31 84.21% 72451976 LR132000.1	Acanthochromis polyacanthus Ovalentaria	95.1 345 96% 2e-16 79.10% 951 XM_051960172.1	Cryptosula pallasiana Protostomia	122 463 100% 2e-23 82.39% 51911826 OX183133.1
61	153	Gobio gobio Cypriniformes	241 46427 100% 3e-60 94.77% 50665609 OX442393.1	Acanthopagrus latus Spariformes	223 932 100% 5e-55 92.16% 22160155 LR884473.1	Eptatretus burgeri Cyclostomata	158 158 99% 3e-34 82.89% 771 LC043689.1
259	148	Nematolebias whitei Cyprinodontiformes	125 125 97% 1e-25 79.17% 3188 XM_037675489.1	Echeneis naucrates Carangiformes	118 118 97% 2e-23 77.93% 3033 XM_029513065.1	Tupaia chinensis Mammalia	121 121 97% 6e-23 78.47% 3860 XM_027772931.1

Table S2 Identifying 9 HIRs through the 3SONNB Method with Emphasis on Ribosomal and Mitochondrial Sequences

ID	Len	NCBI Nucleotide BLAST in Actinopterygii		NCBI Nucleotide BLAST in Actinopterygii excluding the Order of fish species found in the first serach		NCBI Nucleotide BLAST excluding Actinopterygii	
202	155	Gobio gobio Cypriniformes, Ribosomal RNA	280 5335 100% 3e-72 100.00% 59061776 OX442381.1	Urophycis chuss Gadiformes	280 280 100% 2e-72 100.00% 6818 OP151440.1	Mus musculus Mammalia	280 679 100% 3e-71 100.00% 89877872 OX439032.1
203	50	Clarias gariepinus Siluriformes, 28S ribosomal RNA	91.5 91.5 100% 4e-16 100.00% 4151 XR_008356643.1	Pleuronectes platessa Pleuronectiformes	91.5 91.5 100% 4e-16 100.00% 4017 XR_008337966.1	Malaclemys terrapin Amniota	91.5 91.5 100% 4e-15 100.00% 3856 XR_008441993.1
204	82	Gobio gobio Cyprinirmes, 28S Ribosomal RNA	149 2834 100% 5e-33 100.00% 59061776 OX442381.1	Hyperoplus immaculatus Uranoscopiformes	149 745 100% 4e-33 100.00% 37741853 OX442351.1	Mus musculus Mammalia	149 386 100% 5e-32 100.00% 89877872 OX439032.1
278	117	Oncorhynchus keta Salmoniformes, 28S ribosomal RNA	212 212 100% 9e-52 100.00% 3926 XR_008108006.1	Acipenser ruthenus Acipenseriformes	212 212 100% 9e-52 100.00% 3821 XR_004662302.1	Mus musculus Mammalia	212 535 100% 1e-50 100.00% 89877872 OX439032.1
249	85	Salvelinus alpinus Salmoniformes	154 154 100% 1e-34 100.00% 220 XM_024146018.1	Takifugu rubripes Tetraodontiformes	70.7 70.7 60% 4e-09 90.20% 22784 AH005559.3	Homo sapiens Mammalia, 5S ribosomal RNA	158 158 100% 1e-34 100.00% 2231 X12811.1
279	113	Carassius gibelio Cypriniformes, 28S ribosomal RNA	209 209 100% 5e-51 100.00% 4078 XR_008154644.1	Acanthochromis polyacanthus Ovalentaria	209 209 100% 4e-51 100.00% 3942 XR_007939936.1	Mus musculus Mammalia	209 209 100% 7e-52 100.00% 4731 MN537140.1
251	172	Carassius auratus Cypriniformes	311 311 100% 2e-81 100.00% 1098 XM_026264368.1	Chanos chanos Gonorynchiformes	261 261 100% 2e-66 93.60% 1048 XM_030791247.1	Pristis pectinata Chondrichthyes,60S ribosomal protein L5	160 160 76% 9e-35 87.02% 1102 XM_052011409.1
277	65	Culter alburnus Cypriniformes,	121 220 100% 1e-23	Betta splendens	113 15114 100%	Gopherus	109 109 100% 3e-20

		5S ribosomal RNA	100.00% 376 ON156356.1	Anabantiformes	2e-22 98.46% 21851952 LR132024.4	flavomarginatus Amniota	96.92% 119 XR_007772947.1
247	220	Cyprinus carpio Cypriniformes, pyruvate dehydrogenase [acetyl-transferring]-phosphatase, mito-like	389 389 100% 1e-104 99.09% 4026 XM_019091877.2	Tachysurus fulvidraco Siluriformes	120 120 66% 7e-24 78.08% 3678 XM_047819351.1	NS	

Table S4 Results from the 3SONNB Method for the 20 HIRs

Table S3 Outcomes of the 3SONNB Method on the 8 Ultra Conserved Ancient Transposases Sequences

ID	Len	NCBI Nucleotide BLAST in Actinopterygii		NCBI Nucleotide BLAST in Actinopterygii excluding the Order of fish species found in the first serach		NCBI Nucleotide BLAST excluding Actinopterygii	
11	269	Barbus barbus Cypriniformes	381 9971 100% 6e-103 91.54% 27814229 OW387156.1	Thalassophryne amazonica Perciformes	306 89795 100% 1e-79 85.92% 70111961 LR722985.1	Bufo gargarizans Amphibia	188 188 59% 3e-43 86.16% 2002 XM_044287322.1
54	157	Gobio gobio Cypriniformes	260 59452 100% 1e-65 96.84% 45801292 OX442396.1	Pangasianodon hypophthalmus Siluriformes	255 255 100% 9e-65 96.20% 4734 XM_053239047.1	Nanorana parkeri Amphibia	208 208 100% 2e-49 89.94% 737 KC818538.1
80	249	Cyprinus carpio Cypriniformes	436 17180 100% 1e-118 98.80% 22665327 LN590700.1	Silurus aristotelis Siluriformes	367 6098 100% 4e-98 93.17% 44060567 OX327753.1	Vipera ursinii Amniota	244 1.129e+05 97% 4e-60 81.38% 45961943 OX365972.1
160	130	Gobio gobio Cypriniformes	235 11659 100% 9e-59 100.00% 43006573 OX442398.1	Epinephelus fuscoguttatus Perciformes	231 231 100% 3e-57 99.23% 43073509 AP022689.1	Lithobates catesbeianus Amphibia	145 145 100% 1e-30 84.62% 1177 BT081731.1
169	238	Sparus aurata Spariformes	377 377 100% 9e-101 95.38% 31859785 LR537135.1	Epinephelus fuscoguttatus Perciformes	372 372 100% 1e-99 94.96% 52788389 AP022675.1	Lithobates catesbeianus Amphibia	233 233 100% 2e-56 89.20% 1177 BT081731.1
171	176	Barbus barbus Cypriniformes	300 32087 100% 4e-78 97.73% 30825632 OW387196.1	Sparus aurata Spariformes	282 2532 100% 8e-73 95.45% 37948182 LR537137.1	Lithobates catesbeianus Amphibia	233 233 100% 2e-56 89.20% 1177 BT081731.1
172	167	Ctenopharyngodon idella Cypriniformes	270 270 100% 6e-69 95.81% 4826 XM_051913678.1	Sparus aurata Spariformes	261 514 100% 2e-66 94.61% 31425404 LR537142.1	Bombina bombina Amphibia	150 150 96% 4e-32 80.12% 2095 XR_008399815.1
167	84	Denticeps clupeoides Clupeiformes	148 5722 100% 2e-32 98.81% 35364899 LR535816.1	Squalius cephalus Cypriniformes	147 412 96% 2e-32 100.00% 37306196 OX439263.1	Vipera ursinii Amniota	103 954 94% 7e-18 88.61% 77666749 OX365971.1

ID	Len	NCBI Nucleotide BLAST in Actinopterygii		NCBI Nucleotide BLAST in Actinopterygii excluding the Order of fish species found in the first serach		NCBI Nucleotide BLAST excluding Actinopterygii	
20	264	Ctenopharyngodon idella Cypriniformes	477 477 100% 4e-131 100.00% 7224 XM_051898784.1	Megalops cyprinoides Elopiformes	291 291 100% 2e-75 83.33% 3290 XM_036547297.1	Vipera ursinii Amniota	230 1.050e+05 100% 1e-55 78.68% 45961943 OX365972.1
260	82	Micropterus salmoides Centrarchiformes	149 149 100% 5e-33 100.00% 9809 XM_038703445.1	Morone saxatilis Eupercaria	125 125 100% 6e-26 92.94% 7772 XM_035658193.1	NS	
10	125	Scophthalmus maximus Pleuronectiformes	226 226 100% 2e-57 100.00% 25101323 CP026253.1	Thunnus albacares Scombriformes	57.2 57.2 44% 5e-05 80.33% 36935597 OU607596.1	Chamaenerion angustifolium Spermatophyta	51.8 51.8 36% 0.022 86.67% 36737576 OX328272.1
64	78	Squalius cephalus Cypriniformes	141 141 100% 8e-31 100.00% 46137648 OX439255.1	NS		NS	
107	239	Cyprinus carpio Cypriniformes	432 432 100% 1e-117 100.00% 796870 LN590943.1	Chanos chanos Gonorynchiformes	210 210 69% 6e-51 88.48% 63121898 LR697106.1	NS	
62	93	Cyprinus carpio Cypriniformes	168 168 100% 7e-39 100.00% 16759857 LN590681.1	NS		NS	
7	74	Scophthalmus maximus Pleuronectiformes	134 134 100% 1e-28 100.00% 24130524 CP026245.1	Mugil cephalus Mugiliformes	47.3 47.3 50% 0.038 91.89% 2772 XM_047594801.1	NS	
261	177	Micropterus salmoides Centrarchiformes	320 320 100% 4e-84 100.00% 6493 XM_038706206.1	Cottoperca gobio Perciformes	196 196 93% 1e-46 87.27% 27680345 LR131936.1	NS	
262	86	Micropterus salmoides Centrarchiformes	156 156 100% 4e-35 100.00% 3792 XM_038738965.1	Taurulus bubalis Perciformes	93.3 93.3 96% 4e-16 85.54% 19286841 OU342730.1	Erithacus rubecula Aves	52.7 52.7 41% 0.012 91.67% 2330025 LR812135.1
30	276	Hypophthalmichthys molitrix Cypriniformes	480 480 100% 3e-132 98.55% 2025 KC189212.1	Salmo trutta Salmoniformes	55.4 55.4 19% 3e-04 81.82% 48699937 LR584439.1	NS	
252	128	Ctenopharyngodon idella Cypriniformes	223 223 100% 6e-55 98.44% 3122 XM_051873391.1	Salarias fasciatus Blenniiformes	163 163 100% 3e-37 87.79% 1933 XM_030120230.1	Ornithorhynchus anatinus Mammalia	119 119 100% 2e-22 80.47% 2607 XM_007666431.4

265	205	Ctenopharyngodon idella Cypriniformes	352 352 100% 8e-94 98.05% 2841 XM_051866111.1	Electrophorus electricus Gymnotiformes	204 204 100% 7e-49 81.95% 2607 XM_027002262.2	Elephantulus edwardii Mammalia	118 118 86% 3e-22 75.14% 2829 XM_006878971.1
228	129	Megalobrama amblycephala Cypriniformes	229 229 100% 1e-56 99.22% 1726 MF175178.1	Clarias gariepinus Siluriformes	90.6 90.6 71% 6e-15 81.52% 1941 XM_053509956.1	Ursus arctos Mammalia	59.0 59.0 59% 2e-04 76.62% 1625 XM_026480362.3
280	100	Ctenopharyngodon idella Cypriniformes	185 185 100% 7e-44 100.00% 4304 XR_007926179.1	Echeneis naucrates Carangiformes	48.2 48.2 26% 0.014 100.00% 20416991 LR584053.1	NS	
225	105	Larimichthys crocea Eupercaria	174 305 100% 2e-40 98.02% 14415159 LT972178.1	Cyprinus carpio Cypriniformes	52.7 52.7 34% 0.001 91.67% 42439 LN599672.1	Autographa pulchrina Arthropoda	51.8 51.8 26% 0.017 100.00% 14022394 FR997769.1
157	121	Carassius gibelio Cypriniformes	214 214 100% 3e-52 99.17% 899 XR_008153833.1	Ictalurus furcatus Siluriformes	61.7 61.7 50% 3e-06 81.97% 2011 XM_053632581.1	Erebia aethiops Arthropoda	55.4 55.4 34% 0.002 90.48% 30006606 OV281084.1
143	172	Ctenopharyngodon idella Cypriniformes	302 302 100% 1e-78 98.84% 2863 XR_007925961.1	Gouania willdenowi Blenniiformes	295 7346 98% 1e-76 98.81% 43149712 LR131997.1	Vipera ursinii Amniota	161 4183 95% 3e-35 84.34% 45961943 OX365972.1
14	68	Scophthalmus maximus Pleuronectiformes	122 384 98% 5e-25 100.00% 21452505 CP026258.1	Myripristis murdjan Holocentriformes	56.3 56.3 51% 6e-05 94.59% 34950760 LR597559.1	NS	
67	165	Phoxinus phoxinus Cypriniformes	279 279 100% 1e-71 97.59% 48540618 OX424411.1	Silurus aristotelis Siluriformes	91.5 91.5 58% 3e-15 81.82% 20108747 OX327775.1	Anas platyrhynchos Aves	58.1 58.1 41% 7e-04 79.41% 74035964 LS423640.1
156	249	Cyprinus carpio Cypriniformes	445 1134 100% 2e-121 99.60% 154453 LN598968.1	Tachysurus fulvidraco Siluriformes	187 187 91% 7e-44 79.06% 763 XR_003444404.2	NS	

Table S5 Results from the 3SONNB method for the 26 HSRs

ID	Len	NCBI Nucleotide BLAST in Actinopterygii		NCBI Nucleotide BLAST in Actinopterygii excluding the Order of fish species found in the first serach		NCBI Nucleotide BLAST excluding Actinopterygii	
192	247	Argyrosomus regius Eupercaria	411 3098 100% 4e-111 97.14% 28515118 OX336891.1	Echeneis naucrates Carangiformes	410 472 100% 4e-111 96.76% 1540 XM_029505790.1	Rhinopithecus roxellana Mammalia	379 379 100% 3e-100 93.93% 1610 XM_030938818.1
207	98	Ctenopharyngodon idella Cypriniformes	173 173 100% 6e-40 98.98% 9543 XR_007925706.1	Epinephelus fuscoguttatus Perciformes	47.3 47.3 33% 0.046 90.91% 47245075 AP022679.1	Chilo suppressalis Arthropoda	64.4 64.4 45% 2e-06 91.11% 18582568 OU963916.1
138	111	Megalobrama amblycephala Cypriniformes	196 196 100% 6e-47 99.10% 13643 XR_007181673.1	Siniperca chuatsi Centrarchiformes	96.9 96.9 79% 3e-17 84.09% 2977 XM_044213697.1	Lachnospira eligens Bacteria	51.8 51.8 35% 0.019 90.00% 2144199 CP085938.1
112	148	Cyprinus carpio Cypriniformes	248 417 100% 2e-62 97.97% 470726 LN591007.1	Neostethus bicornis Atheriniformes	56.3 56.3 25% 2e-04 92.11% 36095363 LR738542.1	Dolichovespula sylvestris Arthropoda	54.5 54.5 42% 0.008 76.81% 5461102 OU964980.1
110	251	Ctenopharyngodon idella Cypriniformes	421 421 100% 3e-114 96.85% 1656 XR_007925717.1	Mastacembelus armatus Synbranchiformes	68.9 2650 33% 5e-08 82.43% 28130073 LR535835.1	Galaxea fascicularis Cnidaria	56.3 56.3 19% 0.004 85.42% 33853504 OX419340.1
104	76	Barbus barbus Cypriniformes	129 129 100% 4e-27 97.37% 31682544 OW387186.1	Pygocentrus nattereri Characiformes	46.4 46.4 86% 0.031 76.47% 6109 XM_037535378.1	NS	
101	113	Ctenopharyngodon idella Cypriniformes	191 191 100% 3e-45 97.35% 2266 XM_051881989.1	Ictalurus punctatus Siluriformes	147 147 89% 2e-32 92.08% 4447 XM_017470597.3	Rhinatrema bivittatum Amphibia	58.1 58.1 56% 4e-04 79.69% 80992113 LR584401.1
44	207	Ctenopharyngodon idella Cypriniformes	356 356 100% 7e-95 98.07% 9303 XM_051871373.1	Gouania willdenowi Blenniiformes	55.4 55.4 38% 2e-04 75.00% 45933286 LR131998.1	NS	
268	134	Ctenopharyngodon idella Cypriniformes	233 233 100% 3e-58 98.51% 2991 XM_051910782.1	NS		NS	
152	229	Gobio gobio Cypriniformes	387 17132 96% 5e-104 99.09% 43225543 OX442397.1	Epinephelus fuscoguttatus Perciformes	378 511 95% 2e-101 98.17% 39642397 AP022695.1	Vipera ursinii Amniota	231 1231 95% 9e-56 82.02% 19983490 OX365978.1

250	136	Carassius gibelio Cypriniformes	149 149 100% 7e-33 95.65% 1827 XM_052589436.1	Silurus aristotelis Siluriformes	71.6 293 85% 1e-09 82.19% 30334140 OX327765.1	NS	
155	113	Squalius cephalus Cypriniformes	178 14811 100% 2e-41 95.58% 39398959 OX439259.1	NS		NS	
139	152	Squalius cephalus Cypriniformes	243 30915 100% 8e-61 95.39% 47495150 OX439254.1	NS		NS	
241	225	Megalobrama amblycephala Cypriniformes	206 206 100% 5e-50 95.35% 2965 XM_048170279.1	Clarias gariepinus Siluriformes	80.6 80.6 94% 3e-12 75.61% 2793 XM_053512090.1	NS	
108	121	Barbus barbus Cypriniformes	193 193 100% 9e-46 95.12% 32261525 OW387182.1	NS		NS	
31	100	Ctenopharyngodon idella Cypriniformes	163 163 100% 3e-37 96.00% 9916 XM_051915796.1	Ictalurus punctatus Siluriformes	48.2 48.2 62% 0.013 77.78% 2043 XR_008393465.1	NS	
170	175	Barbus barbus Cypriniformes	289 34074 100% 7e-75 96.57% 28985711 OW387201.1	Sparus aurata Spariformes	271 2505 100% 1e-69 94.29% 37948182 LR537137.1	Lithobates catesbeianus Amphibia	226 226 100% 8e-55 88.57% 1177 BT081731.1
163	52	Cottoperca gobio Perciformes	86.0 232 100% 2e-14 96.15% 24104793 LR131929.1	Gobio gobio Cypriniformes	86.0 12516 100% 2e-14 96.15% 63192444 OX442379.1	Rana temporaria Amphibia	55.4 55.4 96% 4e-04 84.00% 5093 XM_040342415.1
137	178	Barbus barbus Cypriniformes	295 11289 100% 2e-76 96.63% 32261525 OW387182.1	Silurus aristotelis Siluriformes	290 10190 100% 5e-75 96.07% 36636424 OX327756.1	Xenopus tropicalis Amphibia	187 187 100% 7e-43 83.71% 39225 AC175588.2
43	269	Ctenopharyngodon idella Cypriniformes	427 427 100% 7e-116 95.17% 3918 XM_051900741.1	Echeneis naucrates Carangiformes	68.0 942 41% 5e-08 74.55% 24574459 LR584045.1	NS	
24	149	Ctenopharyngodon idella Cypriniformes	242 380 99% 2e-60 96.62% 7122 XM_051895884.1	Coregonus sp. Salmoniformes	143 436 100% 4e-31 80.65% 44616205 LR778284.1	Bombina bombina Amphibia	102 102 51% 2e-17 89.47% 14018 XM_053711925.1
56	250	Gobio gobio Cypriniformes	412 412 100% 1e-111 96.44% 58069932 OX442384.1	NS		NS	

13	104	Scophthalmus maximus Pleuronectiformes	166 28831 100% 7e-38 96.97% 21783159 CP026257.1	Sphaeramia orbicularis Kurtiformes	69.8 788 87% 6e-09 81.11% 47724062 LR597475.1	Tyria jacobaeae Arthropoda	65.3 177 95% 3e-06 76.77% 26178051 OX387208.1
57	140	Gobio gobio Cypriniformes	230 230 99% 4e-57 97.12% 49649979 OX442394.1	Scleropages formosus Osteoglossiformes	117 117 56% 5e-23 92.59% 43577283 LR584067.1	Canis lupus Mammalia	98.7 98.7 43% 2e-16 96.72% 47610717 CP050588.1
84	240	Gobio gobio Cypriniformes	394 394 100% 4e-105 96.28% 54078157 OX442388.1	Salmo trutta Salmoniformes	205 205 75% 2e-49 85.79% 35205783 LR584417.1	Osmia bicornis Arthropoda	60.8 60.8 20% 1e-04 88.00% 14187857 OU015505.1
59	133	Squalius cephalus Cypriniformes	213 51008 100% 1e-51 96.15% 37306196 OX439263.1	Mastacembelus armatus Synbranchiformes	142 207 97% 1e-30 84.62% 28732325 LR535853.1	Gossypium hirsutum Spermatophyta	60.8 60.8 32% 5e-05 90.70% 255 GQ393936.1

Table S6 Online NCBI Blastx Identifing the 20 HIRs

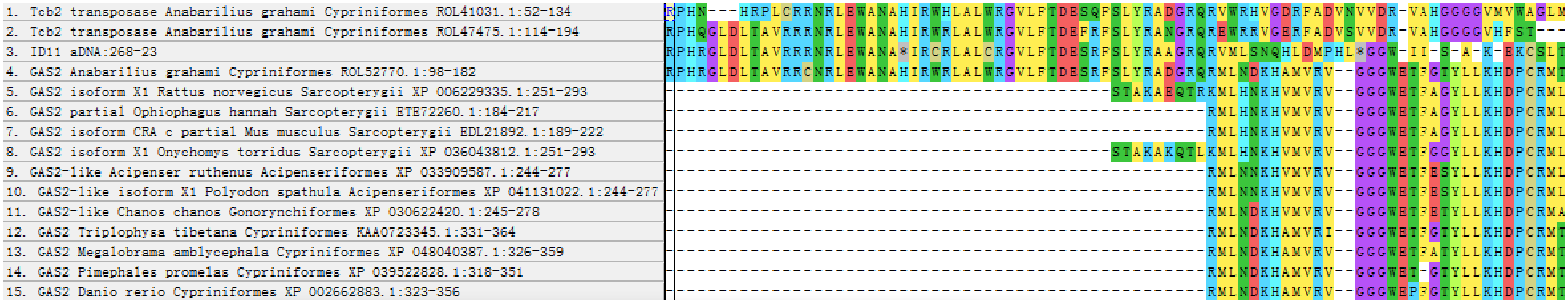
ID	Len	NCBI Nucleotide BLAST in Actinopterygii		Online NCBI Blastx	
20	264	Ctenopharyngodon idella Fc receptor-like Cypriniformes	477 477 100% 4e-131 100.00% 7224 XM_051898784.1	Labeo rohita **Tcb1** transposase Cypriniformes	79.3 141 93% 4e-24 71.70% 331 KAI2667082.1
260	82	Micropterus salmoides nsd3 Centrarchiformes	149 149 100% 6e-32 100.00% 9809 XM_038703445.1	NS	
10	125	Scophthalmus maximus Pleuronectiformes	226 226 100% 2e-57 100.00% 25101323 CP026253.1	NS	
64	78	Pimephales promelas C2orf69 Cypriniformes	122 122 96% 5e-24 96.00% 1293 XM_039649191.1	Pimephales promelas mito protein Cypriniformes	48.1 48.1 92% 7e-05 87.50% 338 XP_039504995.1
107	239	Cyprinus carpio Cypriniformes	432 432 100% 1e-117 100.00% 796870 LN590943.1	NS	
62	93	Cyprinus carpio Cypriniformes	168 168 100% 7e-39 100.00% 16759857 LN590681.1	NS	
7	74	Scophthalmus maximus Pleuronectiformes	134 134 100% 1e-28 100.00% 24130524 CP026245.1	NS	
261	177	Micropterus salmoides znf 319b Centrarchiformes	327 327 100% 3e-85 100.00% 6493 XM_038706206.1	NS	
262	86	Micropterus salmoides Wnt-5b Centrarchiformes	159 159 100% 4e-35 100.00% 3792 XM_038738965.1	NS	
30	276	Hypophthalmichthys molitrix microsatellite Hysd511-1 Cypriniformes	480 480 100% 4e-131 98.55% 2025 KC189212.1;	Anabarilius grahami Cypriniformes, Hypothetical protein	80.9 80.9 47% 9e-17 84.09% 148 ROI69421.1
252	128	Ctenopharyngodon idella rbm24a Cypriniformes	226 226 100% 7e-55 98.44% 3122 XM_051873391.1	NS	
265	205	Ctenopharyngodon idella SOGA3 Cypriniformes	357 357 100% 4e-94 98.05% 2841 XM_051866111.1	Megalobrama amblycephala SOGA3 Cypriniformes	145 145 99% 3e-38 100.00% 774 XP_048008368.1
228	129	Megalobrama amblycephala bmp Cypriniformes	233 233 100% 4e-57 99.22% 1726 MF175178.1	Gobiocypris rarus bmp4 Cypriniformes	85.5 85.5 97% 2e-20 97.62% 76 AXZ78342.1
280	100	Ctenopharyngodon idella ncRNA	185 185 100% 8e-43 100.00%	Anabarilius graham	62.8 62.8 99% 4e-10 81.82% 239

		Cypriniformes	4304 XR_007926179.1	hypothetical protein Cypriniformes	ROL52981.1
225	105	Larimichthys crocea Eupercaria	174 305 100% 2e-40 98.02% 14415159 LT972178.1	NS	
157	121	Carassius gibelio Cypriniformes, ncRNA	219 219 100% 1e-52 99.17% 899 XR_008153833.1	NS	
143	172	Ctenopharyngodon idella ncRNA Cypriniformes	307 307 100% 4e-79 98.84% 2863 XR_007925961.1	Labeo rohita **Tc1** Cypriniformes	103 103 87% 9E-26 98.00% 200 KAI2661099.1
67	165	Phoxinus phoxinus Cypriniformes	279 279 100% 1e-71 97.59% 48540618 OX424411.1	NS	
14	68	Scophthalmus maximus Pleuronectiformes	122 384 98% 5e-25 100.00% 21452505 CP026258.1	NS	
156	249	Cyprinus carpio Cypriniformes	445 1134 100% 2e-121 99.60% 154453 LN598968.1	NS	

Table S7 Online NCBI Blastx Identifing the 26 HSRs

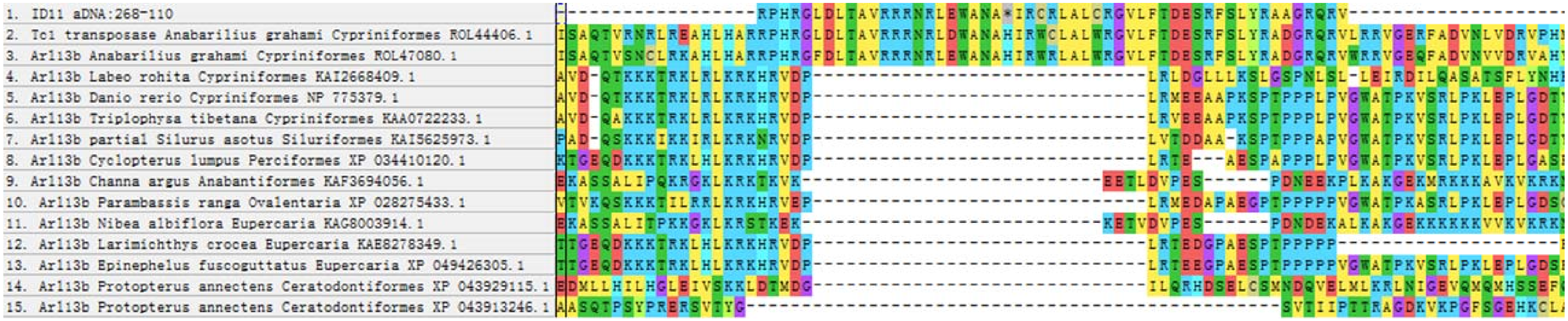
ID	Len	NCBI online blastn in Actinopterygii		Online NCBI Blastx	
192	247	Argyrosomus regius Eupercaria	411 3098 100% 4e-111 97.14% 28515118 OX336891.1	Rousettus aegyptiacus Eupercaria,tubulin alpha 8	182 182 99% 4e-55 98.78% 277 KAF6498422.1
207	98	Ctenopharyngodon idella arnt Cypriniformes	173 173 100% 6e-40 98.98% 9543 XR_007925706.1	Labeo rohita putative RT from transposon Cypriniformes	58.5 58.5 97% 1e-08 87.50% 245 KAI2646354.1
138	111	Megalobrama amblycephala Cypriniformes, ncRNA	196 196 100% 6e-47 99.10% 13643 XR_007181673.1	NS	
112	148	Cyprinus carpio Cypriniformes	248 417 100% 2e-62 97.97% 470726 LN591007.1	NS	
110	251	Ctenopharyngodon idella zbp Cypriniformes	421 421 100% 3e-114 96.85% 1656 XR_007925717.1	Acinetobacter baumanni hypothetical protein Bacteria	137 137 83% 1e-35 95.77% 555 OIC78210.1
104	76	Barbus barbus Cypriniformes	129 129 100% 4e-27 97.37% 31682544 OW387186.1	Anabarilius grahami Cypriniformes,Y+L amino acid transporter 2	44.7 44.7 94% 0.001 70.83% 603 ROK35778.1
101	113	Ctenopharyngodon idella ncRNA Cypriniformes	191 191 100% 3e-45 97.35% 2266 XM_051881989.1	Carassius auratusapoptosis regulator Bcl-2-like Cypriniformes	68.9 68.9 95% 2e-12 83.33% 234 XP_026057171.1
44	207	Ctenopharyngodon idella Cypriniformes	356 356 100% 7e-95 98.07% 9303 XM_051871373.1	Labeo rohita Ubiquitination network signaling protein acrB Cypriniformes	135 135 98% 1e-38 92.65% 147 KAI2668389.1
268	134	Ctenopharyngodon idella Cypriniformes	233 233 100% 3e-58 98.51% 2991 XM_051910782.1	Megalobrama amblycephala uncharacterized protein Cypriniformes	94.0 94.0 98% 1e-21 97.73% 270 XP_048025757.1
152	229	Gobio gobio Cypriniformes	387 17132 96% 5e-104 99.09% 43225543 OX442397.1	Labeo rohita 30S ribosomal protein S16 Cypriniformes	63.5 63.5 43% 2e-10 93.94% 132 KAI2645783.1
250	136	Carassius gibelio Cypriniformes	149 149 100% 7e-33 95.65% 1827 XM_052589436.1	NS	
155	113	Squalius cephalus Cypriniformes	178 14811 100% 2e-41 95.58% 39398959 OX439259.1	NS	
139	152	Squalius cephalus	243 30915 100% 8e-61	Pimephales promelas hypothetical	57.8 57.8 96% 8e-08

		Cypriniformes	95.39% 47495150 OX439254.1	protein Cypriniformes	64.00% 297 KAG1949499.1
241	225	Megalobrama amblycephala msl1b Cypriniformes	206 206 100% 5e-50 95.35% 2965 XM_048170279.1	Megalobrama amblycephala msl1 Cypriniformes	79.7 79.7 95% 1e-15 95.12% 509 XP_048026236.1
108	121	Cyprinus carpio Cypriniformes,microsatellite	108 108 55% 2e-19 95.52% 1500 JN754197.1	NS	
31	100	Ctenopharyngodon idella collagen alpha-6(VI) Cypriniformes	163 163 100% 3e-37 96.00% 9916 XM_051915796.1	Anabarilius grahami hypothetical protein Cypriniformes	64.3 64.3 99% 1e-10 90.91% 283 ROL45339.1
170	175	Barbus barbus Cypriniformes	289 34074 100% 7e-75 96.57% 28985711 OW387201.1	Ctenopharyngodon idella uncharacterized protein Cypriniformes	115 115 99% 2e-28 91.38% 410 XP_051722091.1
163	52	Cottoperca gobio Perciformes	86.0 232 100% 2e-14 96.15% 24104793 LR131929.1	NS	
137	178	Barbus barbus Cypriniformes	295 11289 100% 2e-76 96.63% 32261525 OW387182.1	Sinocyclocheilus anshuiensis uncharacterized Protein Cypriniformes	105 105 96% 1e-25 87.72% 321 XP_016306211.1
43	269	Ctenopharyngodon idella Cypriniformes	427 427 100% 7e-116 95.17% 3918 XM_051900741.1	Ctenopharyngodon idella uncharacterized protein Cypriniformes,	177 177 99% 2e-52 94.38% 321 XP_051756719.1
24	149	Ctenopharyngodon idella gpr84 Cypriniformes	242 380 99% 2e-60 96.62% 7122 XM_051895884.1	NS	
56	250	Gobio gobio Cypriniformes	412 412 100% 1e-111 96.44% 58069932 OX442384.1	NS	
13	104	Scophthalmus maximus Pleuronectiformes	166 28831 98%) 7e-38 96.97% 21783159 CP026257.1	NS	
57	140	Gobio gobio Cypriniformes	230 230 99% 4e-57 97.12% 49649979 OX442394.1	NS	
84	240	Gobio gobio Cypriniformes	394 394 100% 4e-105 96.28% 54078157 OX442388.1	NS	
59	133	Squalius cephalus Cypriniformes	213 51008 100% 1e-51 96.15% 37306196 OX439263.1	NS	



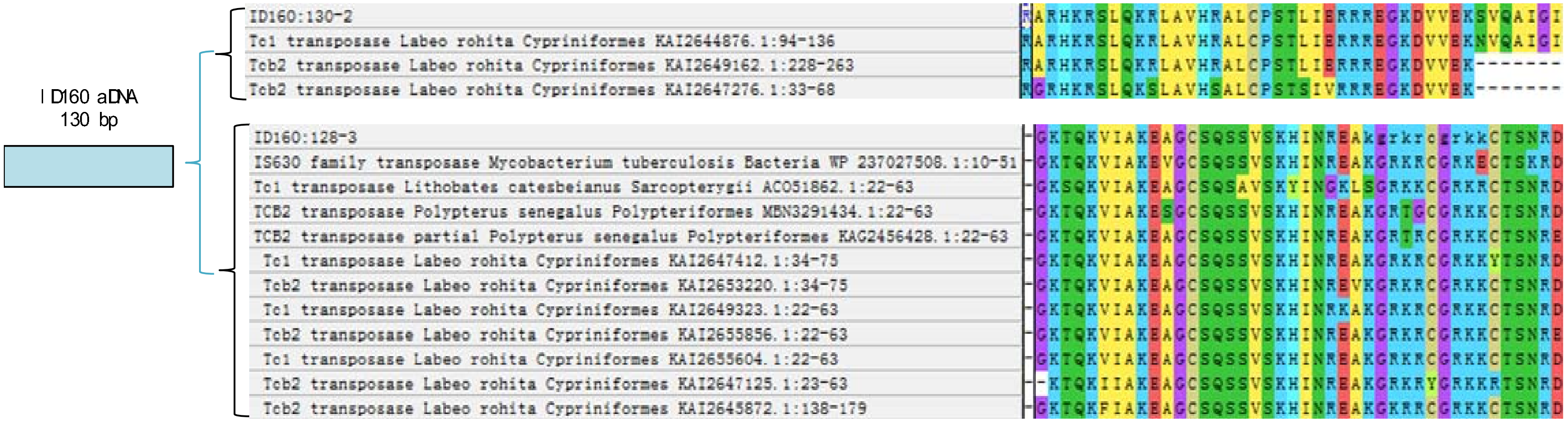
```
1. Tcb2 transposase Anabarilius grahami Cypriniformes ROL41031.1:52-134        RPHN---HRPLCRRNRLEWANAHIRWHLALWRGVLFTDESQFSLYRADGRQRVWRHVGDRFADVNVVDR-VAHGGGGVMVWAGLM
2. Tcb2 transposase Anabarilius grahami Cypriniformes ROL47475.1:114-194       RPHQGLDLTAVRRRNRLEWANAHIRWRLALWRGVLFTDEFRFSLYRANGRQREWRRVGERFADVSVVDR-VAHGGGGVHFST---
3. ID11 aDNA:268-23                                                             RPHRGLDLTAVRRRNRLEWANA*IRCRLALCRGVLFTDESRFSLYRAAGRQRVMLSNQHLDMPHL*GGW-II-S-A-K-EKCSLT
4. GAS2 Anabarilius grahami Cypriniformes ROL52770.1:98-182                     RPHRGLDLTAVRRCNRLEWANAHIRWRLALWRGVLFTDESRFSLYRADGRQRMLNDKHAMVRV--GGGWETFGTYLLKHDPCRMT
5. GAS2 isoform X1 Rattus norvegicus Sarcopterygii XP 006229335.1:251-293      ------------------------------------------STAKAEQTRKMLHNKHVMVRV--GGGWETFAGYLLKHDPCRML
6. GAS2 partial Ophiophagus hannah Sarcopterygii ETE72260.1:184-217            ---------------------------------------------------RMLHNKHVMVRV--GGGWETFAGYLLKHDPCRML
7. GAS2 isoform CRA c partial Mus musculus Sarcopterygii EDL21892.1:189-222    ---------------------------------------------------RMLHNKHVMVRV--GGGWETFAGYLLKHDPCRML
8. GAS2 isoform X1 Onychomys torridus Sarcopterygii XP 036043812.1:251-293     ------------------------------------------STAKAKQTLKMLHNKHVMVRV--GGGWETFGGYLLKHDPCRML
9. GAS2-like Acipenser ruthenus Acipenseriformes XP 033909587.1:244-277        ---------------------------------------------------RMLNNKHVMVRV--GGGWETFESYLLKHDPCRML
10. GAS2-like isoform X1 Polyodon spathula Acipenseriformes XP 041131022.1:244-277 ---------------------------------------------------RMLNNKHVMVRV--GGGWETFESYLLKHDPCRML
11. GAS2-like Chanos chanos Gonorynchiformes XP 030622420.1:245-278           ---------------------------------------------------RMLNDKHVMVRV--GGGWETFETYLLKHDPCRMA
12. GAS2 Triplophysa tibetana Cypriniformes KAA0723345.1:331-364              ---------------------------------------------------RMLNDKHAMVRI--GGGWETFGTYLLKHDPCRMT
13. GAS2 Megalobrama amblycephala Cypriniformes XP 048040387.1:326-359        ---------------------------------------------------RMLNDKHAMVRV--GGGWETFATYLLKHDPCRMT
14. GAS2 Pimephales promelas Cypriniformes XP 039522828.1:318-351            ---------------------------------------------------RMLNDKHAMVRV--GGGWET-GTYLLKHDPCRMT
15. GAS2 Danio rerio Cypriniformes XP 002662883.1:323-356                     ---------------------------------------------------RMLNDKHAMVRV--GGGWEPFGTYLLKHDPCRMT
```

Fig. S1 Tcb2-like transposase fused in the GAS2 coding region



```
1. ID11 aDNA:268-110                                                  ------------------RPHRGLDLTAVRRRNRLEWANA*IRCRLALCRGVLFTDESRFSLYRAAGRQRV-------------------
2. Tc1 transposase Anabarilius grahami Cypriniformes ROL44406.1       ISAQTVRNRLREAHLHARRPHRGLDLTAVRRRNRLDWANAHIRWCLALWRGVLFTDESRFSLYRADGRQRVLRRVGERFADVNLVDRVPH
3. Arl13b Anabarilius grahami Cypriniformes ROL47080.1                ISAQTVSNCLRKAHLHARRPHRGFDLTAVRRRNRLEWANAHIRWRLALWRGVLFTDESRFSLYRADGRQRVWRRVGEQFADVNVVDRVAH
4. Arl13b Labeo rohita Cypriniformes KAI2668409.1                     AVD-QTKKKTRKLRLKRKHRVDP------------------------------LRLDGLLLKSLGSPNLSL-LEIRDILQASATSFLYNH
5. Arl13b Danio rerio Cypriniformes NP 775379.1                       AVD-QTKKKTRKLRLKRKHRVDP------------------------------LRMEEAAPKSPTPPPLPVGWATPKVSRLPKLEPLGDT
6. Arl13b Triplophysa tibetana Cypriniformes KAA0722233.1             AVD-QAKKKTRKLRLKRKHRVDP------------------------------LRVEEAAPKSPTPPPLPVGWATPKVSRLPKLEPLGDT
7. Arl13b partial Silurus asotus Siluriformes KAI5625973.1            PAD-QSKKKIKKIRLKRKNRVDP------------------------------LVTDDAA-KSPTPPPAPVGWATPKVSRLPKLEPLGDT
8. Arl13b Cyclopterus lumpus Perciformes XP 034410120.1               KTGEQDKKKTRKLHLKRKHRVDP------------------------------LRTE---AESPAPPPLPVGWATPKVSRLPKLEPLGAS
9. Arl13b Channa argus Anabantiformes KAF3694056.1                    EKASSALIPQKRGKLKRKTKVK---------------------------EETLDVPES------PDNEEKPLKAKGEKMRKKKAVKVKRK
10. Arl13b Parambassis ranga Ovalentaria XP 028275433.1              VTVKQSKKKTILRRLKRKHRVEP------------------------------LRMEDAPAEGPTPPPPPVGWATPKASRLPKLEPLGDS
11. Arl13b Nibea albiflora Eupercaria KAG8003914.1                   EKASSALITPKKGKLKRSTKEK---------------------------KETVDVPES------PDNDEKALKAKGEKKKKKKVVKVKRK
12. Arl13b Larimichthys crocea Eupercaria KAE8278349.1               TTGEQDKKKTRKLHLKRKHRVDP------------------------------LRTEDGPAESPTPPPPP--------------------
13. Arl13b Epinephelus fuscoguttatus Eupercaria XP 049426305.1       TTGEQDKKKTRKLHLKRKHRVDP------------------------------LRTEEGPAESPTPPPPPVGWATPKVSRLPKLEPLGDS
14. Arl13b Protopterus annectens Ceratodontiformes XP 043929115.1   EDMLLHILHGLEIVSKKLDTMDG------------------------------ILQRHDSELCSMNDQVELMLKRLNIGEVQMQMHSSEF
15. Arl13b Protopterus annectens Ceratodontiformes XP 043913246.1   AASQTPSYPRERSVTYG------------------------------------------------SVTIIPTTRAGDKVKPGFSGEHKCL
```

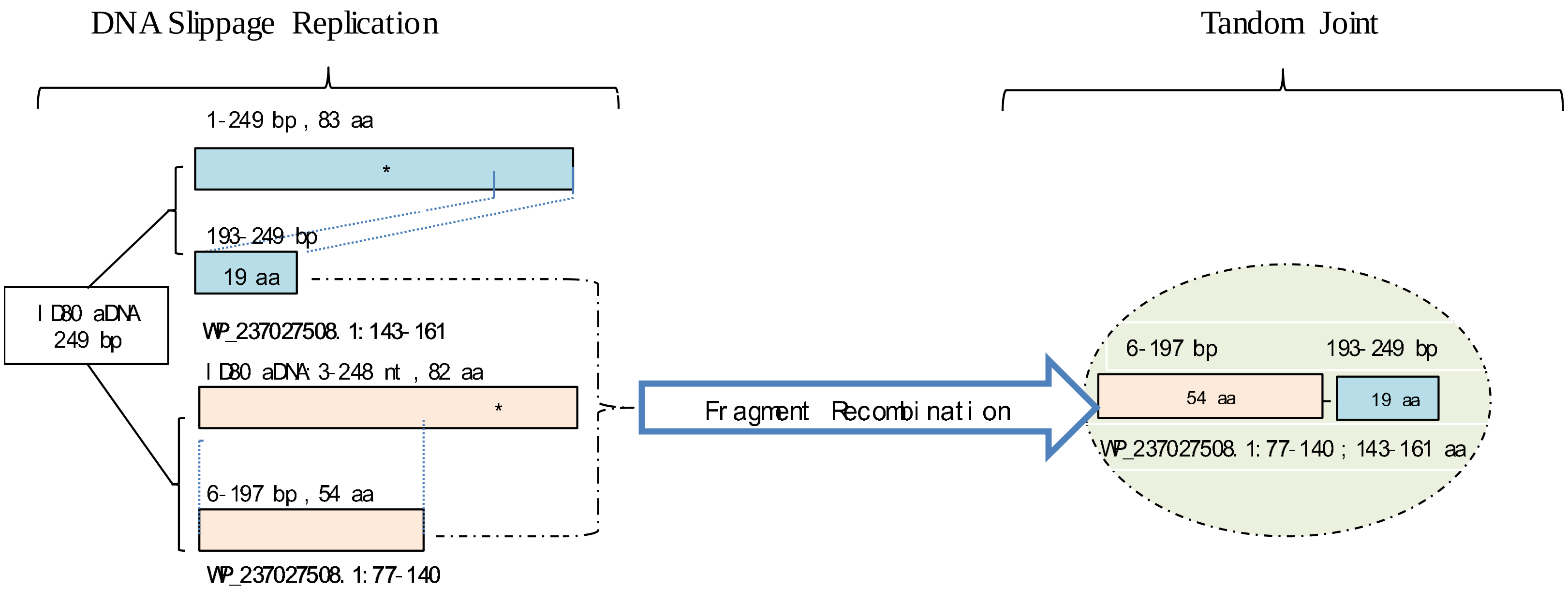
Fig. S2 Tc1-like transposase aligned with Arl13b and Tc1 coding region



```
ID160:130-2                                                      RARHKRSLQKRLAVHRALCPSTLIERRREGKDVVEKSVQAIGI
Tc1 transposase Labeo rohita Cypriniformes KAI2644876.1:94-136   RARHKRSLQKRLAVHRALCPSTLIERRREGKDVVEKNVQAIGI
Tcb2 transposase Labeo rohita Cypriniformes KAI2649162.1:228-263 RARHKRSLQKRLAVHRALCPSTLIERRREGKDVVEK-------
Tcb2 transposase Labeo rohita Cypriniformes KAI2647276.1:33-68   NGRHKRSLQKSLAVHSALCPSTSIVRRREGKDVVEK-------

ID160:128-3                                                                          -GKTQKVIAKEAGCSQSSVSKHINREAkgrkrcgrkkCTSNRD
IS630 family transposase Mycobacterium tuberculosis Bacteria WP 237027508.1:10-51    -GKTQKVIAKEVGCSQSSVSKHINREAKGRKRCGRKECTSKRD
Tc1 transposase Lithobates catesbeianus Sarcopterygii ACO51862.1:22-63              -GKSQKVIAKEAGCSQSAVSKYINGKLSGRKKCGRKRCTSNRD
TCB2 transposase Polypterus senegalus Polypteriformes MBN3291434.1:22-63             -GKTQKVIAKESGCSQSSVSKHINREAKGRTGCGRKKCTSNRD
TCB2 transposase partial Polypterus senegalus Polypteriformes KAG2456428.1:22-63     -GKTQKVIAKEAGCSQSSVSKHINREAKGRTRCGRKKCTSNRE
 Tc1 transposase Labeo rohita Cypriniformes KAI2647412.1:34-75                       -GKTQKVIAKEAGCSQSSVSKHINREAKGRKRCGRKKYTSNRD
 Tcb2 transposase Labeo rohita Cypriniformes KAI2653220.1:34-75                      -GKTQKVIAKEAGCSQSSVSKHINREVKGRKRCGRKKCTSNRD
 Tc1 transposase Labeo rohita Cypriniformes KAI2649323.1:22-63                       -GKTQKVIAKEAGCSQSSVSKHINRKAKGRKRCGRKKCTSNRD
 Tcb2 transposase Labeo rohita Cypriniformes KAI2655856.1:22-63                      -GKTQKVIAKEAGCSQSSVSKHINREAKGRKRCGRKKCTSNRE
 Tc1 transposase Labeo rohita Cypriniformes KAI2655604.1:22-63                       -GKTQKVIAKEAGCSQSSVSKHINREAKGRKRCGRKKCTSNRD
 Tcb2 transposase Labeo rohita Cypriniformes KAI2647125.1:23-63                      --KTQKIIAKEAGCSQSSVSKHINREAKGRKRYGRKKRTSNRD
 Tcb2 transposase Labeo rohita Cypriniformes KAI2645872.1:138-179                    -GKTQKFIAKEAGCSQSSVSKHINREAKGRKRCGRKKCTSNRD
```

Fig. S3 ID160 aDNA can produce two different fragments through slippage replication with further recombination of the coding regions of another two types of transposases


DNA Slippage Replication
Tandom Joint
1-249 bp, 83 aa
*
193-249 bp
19 aa
ID80 aDNA
249 bp
WP_237027508.1: 143-161
ID80 aDNA 3-248 nt, 82 aa
*
6-197 bp, 54 aa
WP_237027508.1: 77-140
Fragment Recombination
6-197 bp
193-249 bp
54 aa
19 aa
WP_237027508.1: 77-140; 143-161 aa

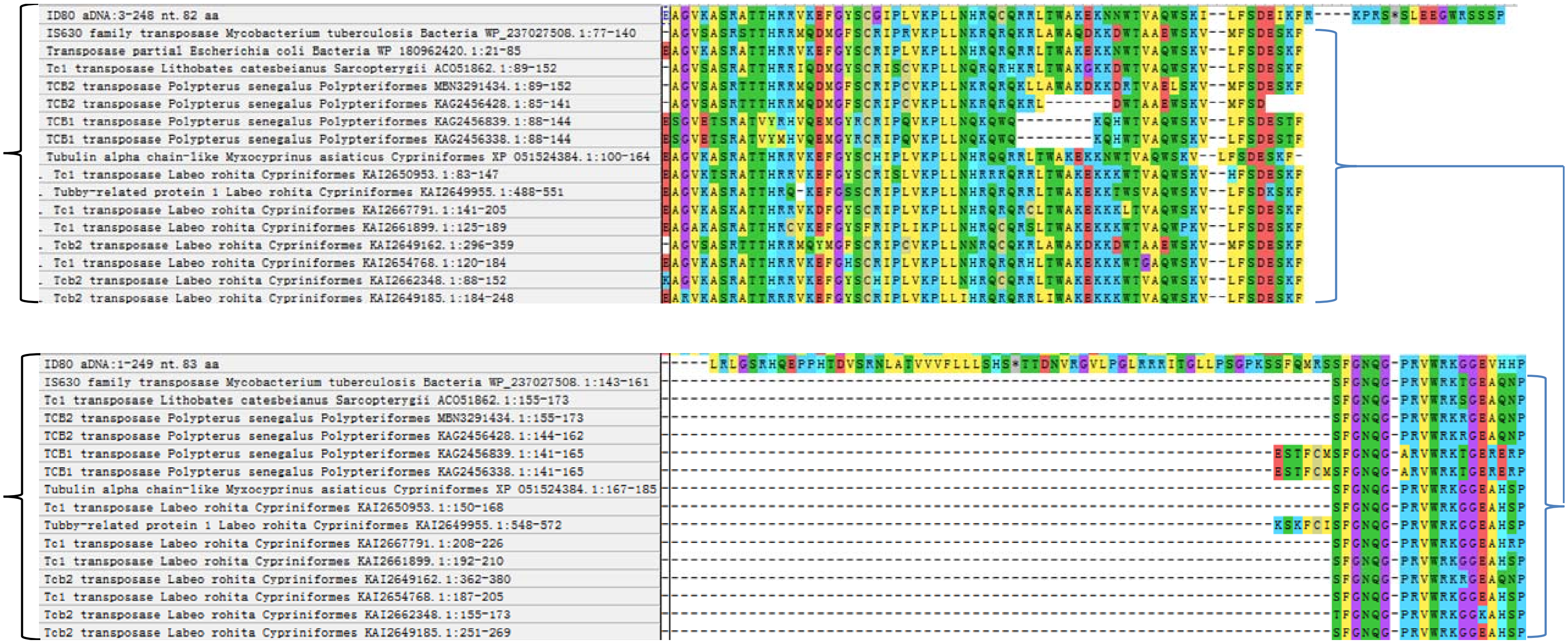

ID80 aDNA:3-248 nt.82 aa
IS630 family transposase Mycobacterium tuberculosis Bacteria WP_237027508.1:77-140
Transposase partial Escherichia coli Bacteria WP 180962420.1:21-85
Tc1 transposase Lithobates catesbeianus Sarcopterygii AC051862.1:89-152
TCB2 transposase Polypterus senegalus Polypteriformes MBN3291434.1:89-152
TCB2 transposase Polypterus senegalus Polypteriformes KAG2456428.1:85-141
TCB1 transposase Polypterus senegalus Polypteriformes KAG2456839.1:88-144
TCB1 transposase Polypterus senegalus Polypteriformes KAG2456338.1:88-144
Tubulin alpha chain-like Myxocyprinus asiaticus Cypriniformes XP 051524384.1:100-164
Tc1 transposase Labeo rohita Cypriniformes KAI2650953.1:83-147
Tubby-related protein 1 Labeo rohita Cypriniformes KAI2649955.1:488-551
Tc1 transposase Labeo rohita Cypriniformes KAI2667791.1:141-205
Tc1 transposase Labeo rohita Cypriniformes KAI2661899.1:125-189
Tcb2 transposase Labeo rohita Cypriniformes KAI2649162.1:296-359
Tc1 transposase Labeo rohita Cypriniformes KAI2654768.1:120-184
Tcb2 transposase Labeo rohita Cypriniformes KAI2662348.1:88-152
Tcb2 transposase Labeo rohita Cypriniformes KAI2649185.1:184-248
ID80 aDNA:1-249 nt.83 aa
IS630 family transposase Mycobacterium tuberculosis Bacteria WP_237027508.1:143-161
Tc1 transposase Lithobates catesbeianus Sarcopterygii AC051862.1:155-173
TCB2 transposase Polypterus senegalus Polypteriformes MBN3291434.1:155-173
TCB2 transposase Polypterus senegalus Polypteriformes KAG2456428.1:144-162
TCB1 transposase Polypterus senegalus Polypteriformes KAG2456839.1:141-165
TCB1 transposase Polypterus senegalus Polypteriformes KAG2456338.1:141-165
Tubulin alpha chain-like Myxocyprinus asiaticus Cypriniformes XP 051524384.1:167-185
Tc1 transposase Labeo rohita Cypriniformes KAI2650953.1:150-168
Tubby-related protein 1 Labeo rohita Cypriniformes KAI2649955.1:548-572
Tc1 transposase Labeo rohita Cypriniformes KAI2667791.1:208-226
Tc1 transposase Labeo rohita Cypriniformes KAI2661899.1:192-210
Tcb2 transposase Labeo rohita Cypriniformes KAI2649162.1:362-380
Tc1 transposase Labeo rohita Cypriniformes KAI2654768.1:187-205
Tcb2 transposase Labeo rohita Cypriniformes KAI2662348.1:155-173
Tcb2 transposase Labeo rohita Cypriniformes KAI2649185.1:251-269

Fig. S4 ID80 aDNA produced two types of transposase fragments through slippage replication, which further recombined and fused into a new type of transposase coding region.

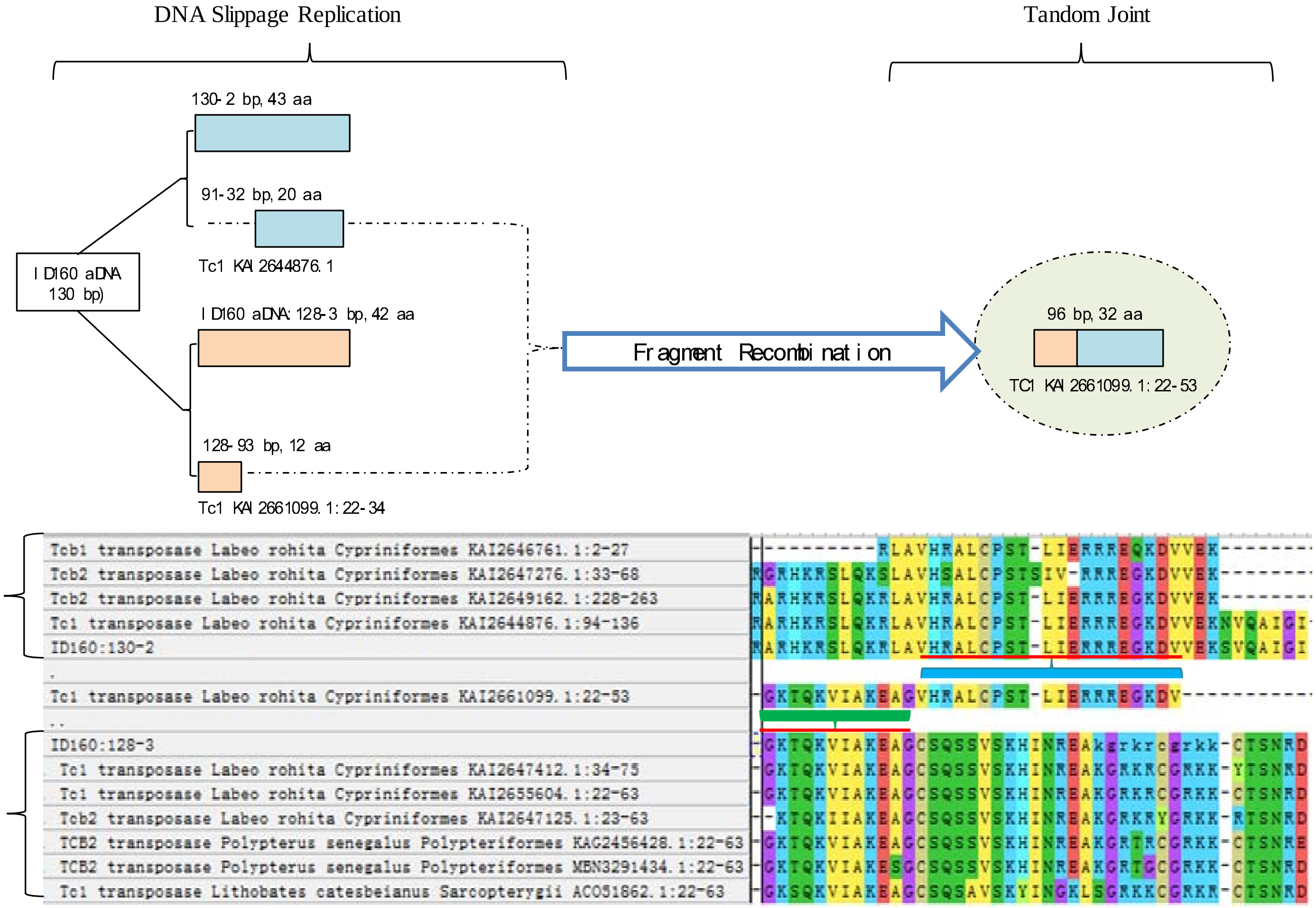

DNA Slippage Replication
Tandom Joint
130-2 bp, 43 aa
91-32 bp, 20 aa
Tc1 KAI 2644876.1
ID160 aDNA 130 bp)
ID160 aDNA: 128-3 bp, 42 aa
128-93 bp, 12 aa
Tc1 KAI 2661099.1: 22-34
Fragment Recombination
96 bp, 32 aa
TC1 KAI 2661099.1: 22-53
Tcb1 transposase Labeo rohita Cypriniformes KAI2646761.1:2-27 ----------RLAVHRALCPST-LIERRRREQKDVVEK--------
Tcb2 transposase Labeo rohita Cypriniformes KAI2647276.1:33-68 RGRHKRSLQKSLAVHSALCPSTSIV-RRREGKDVVEK--------
Tcb2 transposase Labeo rohita Cypriniformes KAI2649162.1:228-263 RARHKRSLQKRLAVHRALCPST-LIERRRREGKDVVEK--------
Tc1 transposase Labeo rohita Cypriniformes KAI2644876.1:94-136 RARHKRSLQKRLAVHRALCPST-LIERRRREGKDVVEKNVQAIGI-
ID160:130-2 MARHKRSLQKRLAVHRALCPST-LIERRRREGKDVVEKSVQAIGI
.
Tc1 transposase Labeo rohita Cypriniformes KAI2661099.1:22-53 -GKTQKVIAKEAGVHRALCPST-LIERRRREGKDV----------
..
ID160:128-3 -GKTQKVIAKEAGCSQSSVSKHINREAkgrkrcgrkk-CTSNRD
Tc1 transposase Labeo rohita Cypriniformes KAI2647412.1:34-75 -GKTQKVIAKEAGCSQSSVSKHINREAKGRKRCGRKK-YTSNRD
Tc1 transposase Labeo rohita Cypriniformes KAI2655604.1:22-63 -GKTQKVIAKEAGCSQSSVSKHINREAKGRKRCGRKK-CTSNRD
Tcb2 transposase Labeo rohita Cypriniformes KAI2647125.1:23-63 --KTQKIIAKEAGCSQSSVSKHINREAKGRKRYGRKK-RTSNRD
TCB2 transposase Polypterus senegalus Polypteriformes KAG2456428.1:22-63 -GKTQKVIAKEAGCSQSSVSKHINREAKGRTRCGRKK-CTSNRE
TCB2 transposase Polypterus senegalus Polypteriformes MBN3291434.1:22-63 -GKTQKVIAKESGCSQSSVSKHINREAKGRTGCGRKK-CTSNRD
Tc1 transposase Lithobates catesbeianus Sarcopterygii ACO51862.1:22-63 -GKSQKVIAKEAGCSQSAVSKYINGKLSGRKKCGRKR-CTSNRD

Fig. S5 ID160 aDNA produced two different fragments through slippage replication, which recombined and fused into a new transposase coding region in Labeo rohita.

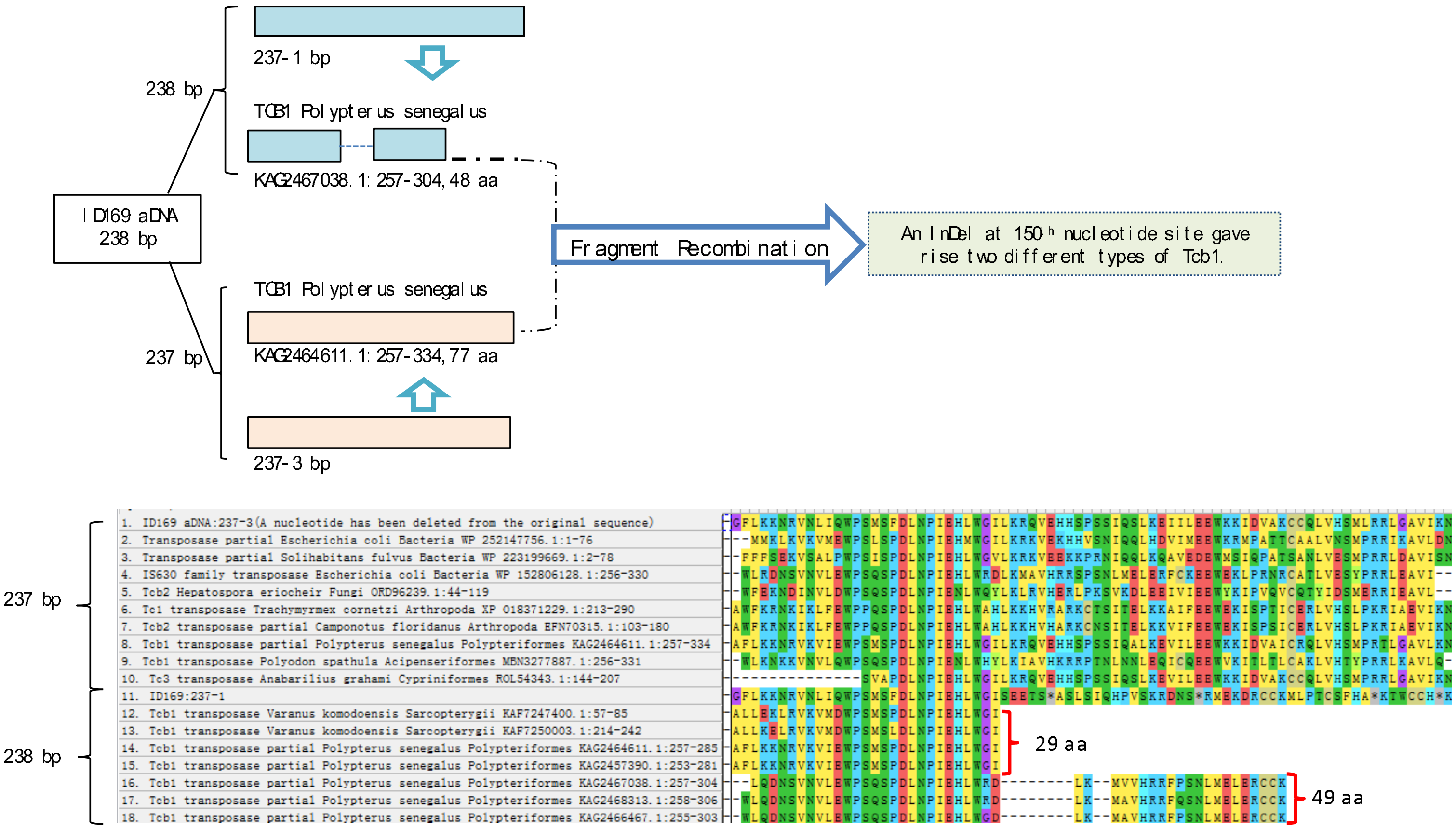
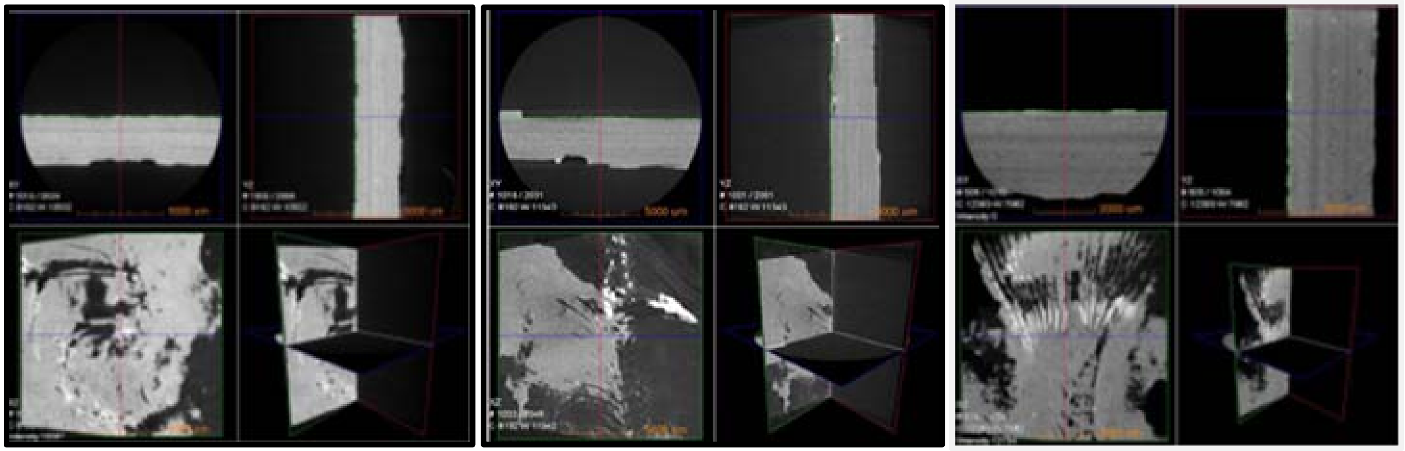

238 bp
237-1 bp
TCB1 Polypterus senegalus
KAG2467038.1: 257-304, 48 aa
ID169 aDNA
238 bp
Fragment Recombination
An InDel at 150th nucleotide site gave rise two different types of Tcb1.
TCB1 Polypterus senegalus
237 bp
KAG2464611.1: 257-334, 77 aa
237-3 bp
237 bp
238 bp
1. ID169 aDNA:237-3(A nucleotide has been deleted from the original sequence)
2. Transposase partial Escherichia coli Bacteria WP 252147756.1:1-76
3. Transposase partial Solihabitans fulvus Bacteria WP 223199669.1:2-78
4. IS630 family transposase Escherichia coli Bacteria WP 152806128.1:256-330
5. Tcb2 Hepatospora eriocheir Fungi ORD96239.1:44-119
6. Tc1 transposase Trachymyrmex cornetzi Arthropoda XP 018371229.1:213-290
7. Tcb2 transposase partial Camponotus floridanus Arthropoda EFN70315.1:103-180
8. Tcb1 transposase partial Polypterus senegalus Polypteriformes KAG2464611.1:257-334
9. Tcb1 transposase Polyodon spathula Acipenseriformes MBN3277887.1:256-331
10. Tc3 transposase Anabarilius grahami Cypriniformes ROL54343.1:144-207
11. ID169:237-1
12. Tcb1 transposase Varanus komodoensis Sarcopterygii KAF7247400.1:57-85
13. Tcb1 transposase Varanus komodoensis Sarcopterygii KAF7250003.1:214-242
14. Tcb1 transposase partial Polypterus senegalus Polypteriformes KAG2464611.1:257-285
15. Tcb1 transposase partial Polypterus senegalus Polypteriformes KAG2457390.1:253-281
16. Tcb1 transposase partial Polypterus senegalus Polypteriformes KAG2467038.1:257-304
17. Tcb1 transposase partial Polypterus senegalus Polypteriformes KAG2468313.1:258-306
18. Tcb1 transposase partial Polypterus senegalus Polypteriformes KAG2466467.1:255-303
29 aa
49 aa

Fig. S6 Several sequences nearly identical to ID169 aDNA (238 bp) were found with high consistency spanning immense distances across Prokaryotes, Fungi, Arthropods, Polypteriformes, Acipenseriformes, and Cypriniformes. Compared logically with the ID169 aDNA sequence, the key 150th nucleotide (Site 150) lacks an ‘A’ base. This indicates an ‘A’ base insertion likely occurred in the ancient aDNA sequence originally.

The sequence following the inferred insertion site would have undergone a reading frame shift, generating incomplete peptide chains. ID169 (238 bp) itself plausibly encodes some short sequences of Polypteriform Tcb1, started from the 5' end. These DNA chains may then have integrated with other transposase types, producing new transposase variants.



Cranial Structure **Thoracic Structure** **Caudal Structure**

Fig. S7 3D Scanning an Ancient Fish

3D X-ray scans of the entire fish fossil were performed using a ZEISS Xradia 620 Versa scanner in order to rule out the possibility of co-deposition of other, larger species. The fossil was plausibly scanned at 45 μm resolution initially to visualize overall structure. The head, thorax and tail regions may then have been scanned at high 6 μm resolution for greater detail. ZEISS 3D Viewer software was likely used to observe and reconstruct 2D orthogonal sections from the scans, with each section representing a different virtual slice through the specimen. No definitive outlines of other species were discerned in the scans and subsequent reconstructions obtained.

Appendix 1 aDNA Sequences of Symbiotic Parasites, Prey, and Lycoptera in Fossilized Fishes

ID	Len	aDNA Sequences
Actinopterygii Sequences		
1	101	TGCATATTTCATTAAAGTCACCAGTGGAGGTGGGAAAGGAGTAGGTACAGCTCTCAGGGGAAGTAAAGGTTGTTTAGCTCAAGTGAATGTTACAGGAGCTG
2	111	GGGATCATTACAAGATTGTTCACTTAGCATAAGTTTTGGAGCTCATCTGAGAAGATGCCTCACTTTACCAAAGGTGTGTGACGGTAATAAAGGAAGGTGGCAAATGTAGCA
3	149	GACCAAGCAGAAAGAACCGACCAAAAACTGAACAGAACCGACCAAAACTGACCAAAACCGACAAAAAACGACCAAAACCGACCAAGAACCGACTAAGAACCGACTAAGAACTGACCGTATACAGAACCGACCAAAATCAACCAGAACCG
4	133	TTTTATTCCTCAAAAATGAGTCTTTGTAAAAAAAAAAAAAAAACTTTACCAAGTATGATTTAGAAGCTTGTCCTTTGGCATGCAGAGAGGGATCTCCACAGAGCACATGCAAAATTAGTTTCATTGTTTCTAG
5	96	ATTGTGAGTCCTTCAGCAGTTCACAGATATCACTTCTCGTTCCTTCTCCATTGGGCAACCTGGCAGCACTGTCATGGACTGTTTCGGAATCAGAAG
7	74	GATTTTTTGGGGTTCATCCGCAGGTTGGACATCGGGATTTTTCGTTTCTTCCAAAGCGTGCGTAAATTGACCGG
8	97	GCCGTGCGCGATCTGCTCGAATTCAAGCCGCTCACGTCCGTTCCTGTGAGCCCTGTGACCCCTGTAACCCCTGTGACCTCTGTGACCCCAGTGAGCA
9	181	GTCAACACTGAAACACAAATCAGCTCGTCAGTTACTACAGCTCGTGAGCTACTACAGCTCATCAGTTACTACATCTCTTCAGTTACTACAGCTCGTTTGTTACTACAGCTCGTCAGTTACTACAGCTCGTCAGTCACTACATCTCGTCAGTCACTACATCTCGTCTGTTACTACAGCTGGT
10	125	CAAGGTTCCATCTGAGCTAAACACAAAACATGTATTCAAAGAAGAAAAAAAGGAATTTCCCTGAGCTATTTTTGCACATTCAACAAAAAGATGATCACACGTGTCATGCAAAAGAAGCACAATAT
11	269	ATATTGTTCAAAAATCTGTCTAAATCTGTGTTAGTGAGCACTTCTCCTTTGCCGAGATAATCCATCCACCTCACAGGTGTGGCATATCAAGATGCTGATTAGACAGCATGACACGTTGTCTACCAGCTGCCCTGTACAGTGAAAACCGGGATTCATCCGTGAAGAGAACACCTCTGCAAAGTGCCAGACGGCATCGAATTTAAGCATTTGCCCACTCAAGTCGGTTACGACGACGAACTGCAGTCAGGTCGAGACCCCGATGAGGACGA
12	125	TGAACTAGTGACCTCATGAAAACATAGAAATGCTTGGTGGAATGACTACAAGAACATTGGCTTGTTGGGAACATGTTCTCTTTAACAGCAACACCTATAAAAGCCTCCTCCTTCAAGCCATTTTG
13	104	ATCGGTGACTGTATATGAAGACGATCACTGACTGTATATAAAGACGATCACTGATTGTATATAAAGACGATCGGTGACTGTATATGAAGACGATCACTGACTGT

14	68	TGTTGAATAAAAGTTGCCATACTACCAAAATGCCTGGACGCAATTAGATAGACCTACATGCAGTGAGA
15	164	AAAAGATGGTCCAACGAGTCAAAGCTGGATTTACAGGCCTGCATTGAGTGTTTTTGAAGTTGCAGCCACTGACCTGGACTGACACTGTGACTTCTTACATCAGTTTTTGTGAGGACGTGTGTGCTGACTAAAACCTTACGCACATACAACAACCATAAACCCTG
16	163	CCTTCATCATTCCTCCTCTTCCTCTTCCCACACACGGTTTCAGCTGACAGTGGGGTGCTCTGGACTCCACTTACAGCCTCTGTGCGGCGGGTAAATACGAGAGTTGTTAATAATAATTCTGACCTTATGTTTTAATTTATTCCTCTATCACAGTGATCCCCAA
17	231	GTGTGTGAGTGTGTGTGCCCTCATGCTTCACTCTCACCTCGTCTTCAGGGTGGTGGTTGGTGAGAATCAGATGCAGTGATGCTGACTCGAGTATCACGTGATATACTTGCAGGTAGTTCGCAGCGGGATTCTCTAACAAGTGAGTAAAAAATTAGTAAAGCTTTAATTTGCTTTTCTGCAAATGGCTGAAAAGTGCAGCTCCCGACTCTCAATCAGTCCAAAGATCTGCTG
18	72	ATACAGCATGACTGGGCGGTTCTTCTGATTTCCAGTGTATAGCTATACATTTCTTTGCTGCTATCAGAGCCA
19	150	AAAGACCACACATGTGCAGTATATGTTGTCCTTTTTCATACTATTACAGCAGTATGTAAAGGGACAAAAGAAAATAATCACGAGGATAGAAAGCAGCGACTGAAAAGAGCTTTTGCCGTAGATATTCTTGTTATAACATGTGTGGTCTTT
20	264	GCCTTCAGATCATCTGTATATTGTTGGATCGACTGTTTCTCATCTTTCTCTTGAAAATATCCCATAGATTCAGGTCAGGCATATTGGCTGGCCAATAAAGCACAGTAATATCATGGTCAGCAAACCACTTGGAAGTGGTTTTTGCACTGTGGGCAGGTGCTAAAGTCCTGCTGGAAAAGGAAATCAGCATCTCCATAAAGCTTGTCAGCAGATGGAAGCATAAAGTGCTTCAAAATCTCCTGGAAGATGGCTGCATTGACTTTG
21	96	GTGTGTGTGTGTGTGTGCTTCAGATATACCCTACATTATTGGGACAAAATATCCCTACTAATATCTGAAATTGTTCACCTTGTGGGAACATTTTTT
22	282	GTTAACTCCACATCAGCTACATAAATTCATCAACTAACCATTCAGAAACATCCTGTTGCAATCTACATGTTGTCACTTCTTCTTAAATCTCTCCATCATCGTCCGATTCCGGTTTGAACATAAAAGGCTGAACAGTTTCTGACATTTTCAGTGAGCCTGTGTGAGGTAATCGGAGGTGCTAAAATGAGCTCTTGAAACTCTGTCCTCTTCTTGGGAGCAGCAGCTCATTTGCATTTAAAGGGACACACAAAATCTGAGTGTTTTTGCTCACCCTCAAAAAGT
23	128	CTACTAAAGACATATTTAAAACAGTTCATGTGACTGTTATGAAGTGACAAGAATACTTTTTGTGCGAAAAAAAAAATGGCTTTATTCAACAATATCTAGTAATGGGCGATTTCAAAACACTGCTTCAT
24	149	AAAGGAAAAGAATTGAAAGTGGGGGGAAAAATCACATTATGAAATAAATGTTTTTCTCCAAAACACGTTGGCCACAATTATTGGCACCCTTTTATTCAATACTTTTTGCAACCTCCTTTTGCCAAGTTAACAGATCTGAGTCTTCTCCT
25	276	GCAATTATTTAAGACAATTCTCCTAACCTTTTTGGCATAAGGAGGACCTATTATGCCCCCCATCATATAAGTCTCTGATGTCCCCAGAATGTGTCTGTGAAGTTTCAGCTCAAAATACCCCACAGATCATTTATTATAGCTTGTCAAATTTGCCCCTATTTGGATGTGAGCAAAAACACGGCGTTTTTGTGTGTGTCCCTTTAAATGCAAATGAGCTGCTGCTCCCGGCCCCCTTTCCAGAAGAGGGCGGAGCTTTAACAGCTCAACAACAACAAA
26	264	TGTTTGTGTACACATACACACATATATAAAAATTCCTGTTTGGATTTAAATAGGCATATAATATACTTAAGAGTCTAAGACTGGATCATTATTGCAGTGATTATTATGTTTCTAGTATGTTATATGGTTGACAACAGTTATTCTAACCCTGACTGATGGAGTGTGTAGCTTTTCATTTCTTTAACAACCATGTAGGAAGTCAAATTGTTGCCATATTCCAGGATGACAATGTCAAGATTCCACAAATAGTGAAAGAATTGTTGG
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	TAAATGTTCATTTATTAAACATATTAATGTTAATTTCCAACATATTTGGGTTCATTTTAAGCAACAATACAGTAATTTTTAAACAATATTTGGGTTTTAAAAAAAACTACCCAGCACATTGGGGCAAATATTTAACCCAACCGCTGGGTTTGTCCATTTTCAACCAAACTTTTAGAGTGTACAAGTTCGATTAAACCATTGAAATATGAGGAAAATATTTTTATA
43	269	CTAGCCCGCAAGGTTATGATTATCGACACAATCCTCGTCAGAAAGGCAAAGGGGGAGGTGTTGCTGTAATTTATAGTAATATTTACAGTATTAGTCAAAAGTCTTCCAAATATAATTCCTTCGAAGTGATGGTGCTTTACGTAACATTATGTAAGTTGACATTTGTGCTGGCTACTGTATACAGGCCACCAGGACACCATACAGACTTCATCAAAGAATTTGCTGCTTTTCTTTCAGAGTTAGTACTAACTGCAGATAAAGTCCTTGTT
44	207	CCTTTCCCTTCTGCTGGCCTCCGTAACAGACAAAGAGCTGATCTGAGGTCCTGAAGCTTCGGGTTCTGTCCACGTACAGGTGCAGAGCGCGGACGGGACAGAGCAAAGCTAGGGCTGGGTCTGCCTCCTCCGAAGGCAGCGCTTGCAGGTTCACTACCTGATCTCGAAAGGGAGTAGTAGGAACCTTGGGCACATATCCAGGCCGGG
45	239	TTCTCAGGCACTCTTTACTTAAAGCTGCAGTCCGGAACTTTTTTTGATTAAAAATTATCCAAAATCAATTTTTGAGCAAGTACATAACCAGCCAGTGTTCAAAACTATCGAATTATCTTAGCTCGATTCACAACGGTAAGCTTGTAATAACGTTTTATAATAAGAGCGACCTGGTGGATTTCCGTGGGAAATTCAAGCATGCAGCAGTTCGTCTTTGTGTCATTACATCACGTCTGTTT
46	216	ATACTAATAAGCAAAAATGACTTTAGATTAGATTTAGAAAGAAGAGTGTTGAACTTCCTTCTACTGTATCCTAATCTTCTTTAATCCAGAATGGCAGCACAGCTGAGAGGTTTGTTTGAGCGGCGCCCTCTACTGAACAGGTGTAAATATGCATTTCCTTCAGCCTGAGGATTATTCATTTCACTTTTGGTGTGAAAGGGCCTTTAAAGCAATTAA
47	171	GGTTTGAGTTTTATGAACTTATTTCTTTTAATTTAGACAAACTTAAGTTTAACTGAAACTGGGTTGGGTTTTATATTTCCCAGCATGCTTTGCTCATGACACTCGAAAGGCCCATGCAACTGATTGCCTCAATTTTTTTGAGTTTTGCCAACTTATTCAGGTTTACAGTGT
48	91	ATAATCATGGACATGATTCGAAAAGGTTAAATGGGCCTGAAAACAAGTTACACTTGTGCTCCTCCCTGTCAAAAATGGCACATTGGAAATG
49	184	GAAAAGGGGGTTTCGGGCCAGAAAAATTGTTTCAATATATCTCCAAAAAGGACAGATGAGAGAACCAGGAGTTTAAGCCCCGGGTATACTTCGGCCGTCTGCGTTTGTGAGTAATTTCGGGACCGCGCGGATGGTGCGCGTGCAACAGCGCATGCGCAAGCAGGACAGAAGAGTACACTAGGTG

50	186	TGGAGGCCTCACAGAGCCATCGGATTTCATCAAAAATATCTTAATTTGTGTTCTGAAGATGAACGAAGGTCTTATGGGTGTAGAACGACATGAGGGTGAGTAATAAATGACAGAATTTTCATTTTTGGGTGAACTAACTCTTTAAAATTTTACTTGTCCAATGTTAAGAATGCATGCAGTTTTACT
51	131	CTGGAGTTGTGTTTACCTAACGTTAAAAATAAGCGGGGCAGCCAGAAATCTGCCCCAATGGCTCTCTTTGAGAGCGCGCTGCAGTTCCATTGTACACCACAGATTTACATAGGAAAGTTGTGGGGCGCTGT
52	176	ATGGCGTAATAATGTAACTTCTGTTGCTGTTCACGGTCAGGGACTAATTTTTTTCCGGCGGTAGGAATGTATTAGTGAAAATGTACTTCATGAAAGTTGCATTGATATATTTTTGGCTTTAATATTTGTATTGTGTGGTAACCGCTTCGCATCGTGCCTAACGACAGGGAACGACA
53	210	CGAGCAGAGAAAAACAGCATTTTTCATGGACAAAGGAAAGGTACTCCTCTCAGAAAACATACAAATAAAGATACTTTTAGATGTTTAATTGCTATTTGGAGCATCTGGATCTCTAGATGAGGGCTTAATGAACTAATTTTTGTGAAAACCTCAAATTCGACCAAAACTGGCATTTCTCACTTGTTTTGCCTGTAGCTCATAATTAGCTGT
54	157	CTCGTCAACATCACTGACTGACCTCACAAATGCGCTTCTAGAAGAATGGTCAAAAATTCCCATAAATACACTCCTAAACCTTGTGGAAAGCCTTCCCAGAAGAGTTGAAGCTGTTATAGCTGCAAAGGGTGGGTCAACTCCATATTAAACCCTACGG
55	149	GGAGAAGAGTAAAATTTAGCTTCAAAATCAATTGAATGTTTTGTCGTGTGCGTGTGAGAGCACTGCTAGCATTCATTTCACTGTAAGACAGGAAAAGAGGAGCTGTCTGTGAGGATGTTGACTTTTTCCTTCGCTGATTTAATGACCTG
56	250	ATAGCAATGCAACTTATATATTTCAGGACATAATTAAAGCATCAAAACATTTAAATGCATTGAAAATCAGAGGGAACTCATATGTTAGAAAGTTTATGTGAATATGAGTAACTCAAGAGCTTAAGCTTTGAGGCTTGGACGGTACTCATTGCTGTCCACAACAGCCTTCTGTGGTGTTAAAGAGTAAAAATACCACTGAACTCAAATCCCTCTCAGAATTTCTGTATCCGGCATGTGTTTGATTTAACCA
57	140	CATATTTAATGTCCTGATGAACGCTTCCACCTGCTGACACAGGAAAAGTGACAGAACAATACACCTCCCAACCCCCTCCATCCCTCCGTCCAGCAGCTTCAAAGCCCGTCTGCCCGGAACATCGCCATCATCAATCCAGA
58	143	TAATGACTTCACCCTTATCAGAGTTCAGTGAGAGTTCAGGGAATGCAGAGTTATTCCCTAACATTCATCCACAACATAAAGCAGCATTCACAGCTTAATCATTCACTTTTGGTGCAGAGGAATCTTAAAATGCAATTTACAAT
59	133	TATTTCAGTGTCACATGATCCTTCAGAAATCATTCTAATATGCTGATTTGATACTCAGTTATTATCAGTTTTGGAAACAGTTTTGCTGCTTAATATTTTTTTGGAACCTGTGATATTTTTTTTCAGGATTCTTT
60	154	CATGCTGATACGGATGAGGTGACTAAACACATACAGTACTTATGTAGAAACACAAATGATGCCCGTATAATCTCCAAACACCAGTCTTATTCAGTTCAGGGTGTATTATCAAAGAAAATACCACCACTTTTTCATATTTCATGAACTTTAGCAT
61	153	GGATTGAGATCTGGAGAATTTGGAGGCCAAGTCAACATCTCAAACTTGTTGTTATGCTCCTTAAACCATTCCTGAACCATTTTTTCTTTGTGGCAGGGCGCATTATCCTGCTGAAAGAGACCACAGCCACCAAGGAATACCATTTCCATGAAA
62	93	ATTTAAATCAGAGTAGGTCACAAGGCCAAAGCCTCACCATCAGTTTGAAGAGCCCCTCTATTGATCACAGATAAAATCAAGGTAAAACCATTG
63	166	GCAGGCCCATGACAGTGAACTGGGGCACAAAGATGCACTGAAGATCCTCCCCAACCTGATCATTTGTTCCAGAAGTTTCCGTCTGGCAGATGTTTCAAGCAATCCAAACCAAGACCAACAGACAACCGCACTGCCAACACTGTACACATTCACATACATATATTCA
64	78	AAACAGAGAACCTGCAGAATGGAAACGATCACGTAGTCTTCTTTCCTGGAGACATTCAGGTTGGTGCATGTCATCAAT
65	177	TGTAATTCTTCTTGCTCTTTTCTCTGCATCTCAAGCACATTATTTCATTCAGTCTGTGCTCCATGTGCTCAGCTTTTGCCTTTTCAGCTTTCTGGCCACAAAGTCTCTGCTGTTGCTTCTCCGTTCTTTCTTTATTTCACAAATTCTGCCATCTCATTCTCTTTGTCACTGAACCAA
66	150	CGACGCAAAATGCACTGCGCACCTAAAAATAAATATGTATAAAACGATGTTGATTTTTTTTTTTTTAAATAACGTAAATCTTTCATGGCTTTCTACTACATATTTCTGTAAGAGACATAATTTACGTTATATTAAGCCATACAAGTCTCA
67	165	ACACAACAGGAAATAGCTCCATGTTTTTTCCTGGGCGAGAGGGGCCAATAACTGTCACTTTAGTTTAGATGGCTCGAGCGCAGTGGTCCAATGTCTGTTTAACAGCTTGGGTTGGAAACTTCTGCAATTAAAGAGTCCATGTGCACACGGAGATGCACTTCCCAG
68	171	AACCTATGTTTATAGACCCTTTTACAGCCCACGTAATCAAATAGTCACGCGTGCATTTTGGCAACGGAAGTGGTGTTGTTCCCCAGTGGTTCTAACGTATTCATAAACATCCAGTGATTTATATGCTCGTAATTTCTCTCGGGTGTACACGCTCGGTTTCTCAACTAAATA

69	237	GATGACCGAAAATCACCTGGTGAAGGACTTTTATTGACAACTTGCGAGCAGCCAGCATTCCATGGGATGAGGCTGAAGTCGCCTCTGTGGATCGAAACTGATGGAGAACACCTGTCCTCCCATGTGTAACAGGCGTATGCTAGTCAGAGCTGTGCGAGTAAACCTCACTCCCCGATCTCAAGAGATGCTCTAGCGACTGACGCTAGAGGTTGCAGCCTTTAGCCTCTGTGTTAGAGT
70	167	AGTCATTTAGTCCACAGCATGTTAGTTCATTAGGGAAATGAACCTAAGAGGGGAGCTAAATTTCACTAAATATATGCACTATATTAGCTTATATATTCATTATATTTACTTCACCTGTTAGTAATGTCTTAATTATTTAATATATGTTTAAATAAATTTGTCATAAA
71	181	TGTCATGATCAGACTATCACATTGTTAGCTTAGCAGCAAGCTAACACCGAATTGGTGCAATTTTTTTCCTAAAAAAGGCAAGCACAAATCCCACCTGCAGCTTATTTGTTTGGTCAACATGACCACTCAGTCACCTGTCAAGAAAAGTTCAGACCTCATGGATTTCAGAAATATAATGGAT
72	123	ATAAAGTAGTTTTCATTTGCGCATTTCGGCAAATTTAGCAGTTCAGTCCGTACATACAGAAAACTTGCCTTCTTCATTACTACTCTTTTTGTAAATCAGTAGGTGGTATGGTATTCTGAACAT
73	115	TGAAGAGTTCAGATGCAAAAGCCTCTAAGTGCCGTCTGAAATTTTCTTCTAAAATGTACATTTTTATCAAGCTTGTATGTTTATGTTAAGTTATTTAACTTTAATGGCAATTAAT
74	134	GACTATGTCTGCGCTTTCAAAAGTATGCAGTCTCAAAAATGAGGAGACCCCTCCTCGTTGCACTTTGGTGTGTGGACAGGAAGAGAGACAGCACAGCTGTCTTTACAAATATGTTCAGCTGAGTCAACACAGGA
75	260	ACTGTCCCTTAGAGCACACACACACACACACAACACACACAAAACCATATGGCCCCAGCATCAAGTGTCATGTTTACAAGCGTAGCTCATTGGCACTGCACTCGTCCTTCTTACCATTATCTTTCCATCTGAACTCATTCAAAGTAAGTCCAGCGTAGAAACATTAATCCTCAAGTTATTCCAAAACAGTTTCTTGCCATCCTGCTCATCATTTTAGCGCTTTGAAATGTGGGCTAGTTACACAAAGCCCCATCAAGCCT
76	134	TCATGTTGTTCCACACCCATAAAACTTTCGTTCTTCTTCGGAAGACAAATTAAGATATTTTTGATGACAGAATGCTGTCAGATTTTCTTCCATTGACTGCCTTTGCAACTACCACTTTGACGCTTCAAGAACTT
77	274	GGCGTTTAACAGGGGCATGGGCAGCTCGGGAGCTCAGGGCTTCGACTACGGCATCTCCAACAACAAAGGTGAGAGTAAAGATCTGGGAGAGATAGAAATAAAATGTGATTTATTATTGAGTGACGCTTACATCTCTGTTTATATGGTTACAGTCCACCTGCCAAATGAACATTTTTTTTCATGCCATGCTCCTGCTAAATATCTGTCATATTTAATCAGATATATCTTGGTAGTTTATTGCCATATACATTCAATTGATTTCAATCCAATGGTG
78	218	GTTGCCTTTCACTGAAGGGTCAGTTGAAGTCAAATCCTGTGTTCTGCTTAAGGTCACACTGATAGAAGAGACTCATTGTGCTCAAATAGAGAAATACTGAGAAGAACTGTCACTCACTCGACTAATAAAAAGCTCTCCCTTAAATAAAAACAAGTGCTTCACAAAACCTGTTTATTTCCTCACGCTTTAAGCAGACTGGTCAGGATGCTGTGGCTCAT
79	217	ATGAGCCACAGCATCCTGACCAGTCTGCTTAAAGCGTGAGGAAATAAACAGGTTTTGTGAAGCACTTGTTTTTATTTAAGGGAGAGCTTTTTATTAGTCGAGTGAGTGACAGTTCTTCTCAGTATTTCTCTATTTGAGCACAATGAGTCTCTTCTATCAGTGTGACCTTAAGCAGAACACAGGATTTGACTTCAACTGACCCTTCAGTGAAAGGCAA
80	249	CTGAGGCTGGGGTCAAGGCATCAAGAGCCACCACACACAGACGTGTCAAGGAATTTGGCTACAGTTGTGGTATTCCTCTTGTTAAGCCACTCCTGAACCACAGACAATGTCAGAGGCGTCTTACCTGGGCTAAGGAGAAGAATAACTGGACTGTTGCCCAGTGGTCCAAAATCCTCTTTTCAGATGAGATCAAGTTTCGGAAACCAAGGTCCTAGAGTCTGGAGGAAGGGTGGAGAAGTTCATCACCCA
81	163	TGAGTTATAAAGTCAGAATTCAAAGAAAAAACAGTCTTTTTTTTCCTCACAACTGGACATAACACGCAATTGCAAGTTTATATCTCACAATTCTGACTTTATAACTCGCAATTGTGATTTTATATCTTGAAATTTTGAGAAAAAAAGTCAGAATTGTGAGATA
82	144	GAAGCCCACTAACTCTCATTGATTTTGCTTCAGTGTCTCTTTATTGAAACTTTATCAGGTGGATCTGAGCGCATGGAGAGGGAATCCGGAACGAGTGATGAACTTTTCTGTCTCTCCTCCTCATGCGTACACCCCGCTACATCG
83	218	TGGAATCGAGATGAGGGCATTTCAGGGTACATGAAAGAAAAGCTTTTCTCTTGCAGGGCTTTCTAAGGCAAAACAGCACTTCATAAATTAAAATTACTTGTCCACAAGCAACACTGGATCTTGAGTTCAGTTCCTGAGGCAATGTTTTAGACATTAATATAACAAAACCATTTGGGGATAAAGTGTAGTTCAGCATGTGTGCGCGCATCTCGATTCCA
84	240	GAGTAATTGATTTAGAGAATGTTAAAATCAATTGGCAGACAGCTGGGGTCTGGGAGATTAGAAGTGAAGTATAGGTGGAGTGGAGAGCAGCAGACGGCCCCAGAGCAAACAGTAATCTCAGCATTATAGACCAGGGGGAGATGTACCTGTTTTCTCTTATTTCACTTTCACTCTCTCTCTTGCTCCTTTTCTCCTGTTGGGGACGGTCCATAGTCTCTTACACCAACCTTGTCTCCATCT
85	192	TTAATCATTTTCATTTATGTCACTGCTATTGCTAGACGGCATAGCGGCAAAATGTCCGAAGGAAATTATAGTGGAGCACAAAAAAACGGAAAATACAAAGGTTAAAAGAACAACGTGAACAATATATACTGAAAAAATACCCAAATTGACTTGATACTTCAAGGCATTGCATACATCAAGACAGAAGTGGCG

86	109	CCTATGCCAACATGTCCAGGTCACAGTCCAACGGGCACCATACCAACATTCCCTCACTGACACATTAACACAATGAAACAGGTTGCAACCATTGTTAACAGCCTGCAAA
87	156	GGAATTTGTCTAAAATTAGCCCAAGAAGAAAGAAAGAGCTAATTATTGTCAACTTCACATGTCTTATCCCACGCAAAGATGGTGAAACACTCACCCAACGTAAAGAAATTCTGATACGTCTACCACCAATAACGCTTCCCTCTATGTGTTTCACAA
88	141	CTTAAATAACTTTTTAATTTTTTTTTTTTTTTGAGTGATTCCAATATTCAGCAATAATTTGCCAAATGTCTTCTTTCAAAATAGACCTTAGATCTGTATTCCAAAGCATTCATGAATTACAACCATTTATTTCGCATGTCGT
89	276	AGCACTTACCGAATACAACAGACGAGTCAGTGCCATTGTGAGGGTACCGATAGTGACAGTAGTGGCAGCTGAAAATCTTGTCCAATGTATAAAATGAAACTCGGTGGACTGGACGAGATCATTTGTAGTCAGAGATGTTGCCTGTATTTAGCAAGAGCTGGTCTGAGGCATGTGGTTTACAAGCACCCAATTTAAGGGTGTCCATGCCCCCTTTGCTATCCTATGGTCGTTGACACTCTACTTTTCTGTGAAGCCCATCAAGTTTTGATTGCCGTG
90	205	AACGATGGCAGGGGAAAAACGAACCGTTTTGAAAGAGTAATTACGTTGTAGCCTACTAGAATTGATGCCGACCATGTTTATCATAAGTGCAACAAGTCATTTGCATGTGAAAGTAGAAACATGCAATTCAGACGAAACTTCTAAAACGTGCATATTAACTTTCCTCATTAAATGTGTCTGTGTGTCCGACACAGCCTCCACGACG
91	159	GATCGGAAGAGATGAAGAATATACGTCATAGATTGCTTTGGAAGTCAGGCATTTTGGTACATTATCATGTGCCAGTTTGGTACATCCATACTACAACCAGAGGTCAGGAAAGATGCCACTATTGGCCAAAAATGATGACTAAACAGGTCAAAATTATGG
92	138	TAGAGTAAAGAGACGCAATTAAGCAGGTGGTTTTCCCCTCAATCTGGCGCCCCTGTGCGGTGGGCTGTAGGGGAGGGTTGATGTCGGAGCACTAGCCCCTCTCATCCTCAAGCACATGGTCCACATTAGTGTGGAATG
93	218	CAAGCCAAAAACTGTTTTTAAAGTAAACTGGTGCCTTCCACTCCTCCCATGTGCTTCTCCTTTTCACAGTCTGGCAAAAATAATGGGTAGTTCCAAAATATGTGGTCACATGGCAATAAAAGGGTACGGTTGCTAAGATACAGGGGGTGGGTCAACATACCCACTGGCTCATCTTACCCCACTCTCCCCTACACATAGAAAACTCATAGGCCCGGTTT
94	178	GTGGACTATTTTAACAATGTCTTAGAGTCAGACAGCTCTCGGATTTCATCAAAAATATTTTAATTAATTTGTGTTCCGAAGATGAACGAAGGTCTTTTAGGTGTGGAACGACATGAGGGTGAGTAATTAATGACAGAATTTTCATTTTTGGGTGAACTAACCCTTTACGTTGCTTTGT
95	108	AGTGAATTGAATTTCACTATGTGTCTGTCTATGTGTCTCTGTGTGTGTGTGTGTGTCTGTCTATGTGTCTGTGTGTGTGTAGAGAAGTATTGGATATAGTGGTGGAAT
96	141	GCCACAAGCTGTTTTGTTCCTTGCCTTCAGGCAAGCGTTATTGCACTGTCAAAGCAAAGACTAGCAGACTAAATATTCCAACCCACTGCAGTTCATAGACTAAATAACACTTTTATAGGACATGTTCAAGGTTTTTTGCTG
97	173	GCGCTTATCCGGATTGTTCAAACCGACTTAAAACTAAAAGATTATGTTCTCTTGCGCGAACTCATTCAGGAGTGGTGACTTTCCACGTATTCCCTATTAACTTCTGAGCTGCTCAGCTGAAGTTATGGTTGATTGATCTTCGGCTGGATATCAACACACCCATTAATGTACTA
98	287	GAAAAAGCCGAAGGGAAGCTGGAGGGGGAAAAAAAGTTCAGATTAGTTAAATAATGATTAAAATTATTCAGAAAAATATATACTGATTATGTCTTACTTTCCCAACTGCTTTTGGAGGTCAACCATGTACTGGTAACACTGGGCGTAATACGTCATCTGAGCCTCCACAAAATCATTTAGACAGCGGAGATGATGGGCCTGGGAGAGAAATGATATGGTTTGTTATAGGTCAGACGTTGAGTAAATCGAGGGGGGTTGTGAGATGTGGATTGGCAGGCGGCCTAGTGG
99	104	GACAAGTTGAAGAAAAAATGGAAGAGAAAAGATTATTATAGATATCATTGGGTTTTTTCAACAGACATATATACAACATCAGTCCTTTAATCTGACTGGTGATA
100	209	CTTCACATTTGAGAGTCACAGATGTGCTGGAGGTCCATGAGACCACAATTTCTGTTGCAATTGGTCTGTCTTAAAATGAAATGTGTTAATGAGCAAACAGATAAAATAGATGAATTGTATATGAATAAAAATGACTCTGATATGTCTTGACTTACATTGAATAATTAATGTAAAGTTTTGGGATTTTATAGTTCCATTTTTTCTAACTA
101	113	TCTCAAAATCTTACAGTCACTGCTGGAAAGGGTTCAAATATGCAGAAGATGCTGGAAAACTGAAGAATCTGCAGGACCTGGAGGATTTTTCTGAAGAACAGAGCTCAATTTAA
102	146	CATTGTTATGATTCCCTAGTCTTTACAGTAAAAGGACCTGGAGCCCTAGGGGTCAGTAGATGTTTTACAGGACTTCCTGTGTTGTGCGAAAGGGGGCTCTGGGGTCAGTACACATTATGCAACAAGGGGCCCGCAGGCCTAATTCT
103	213	ACTAGCGTTTCTGCTGAAGTTTTGCTTGTAAAAAAAGTTCATTGTCTATTGTCAATAGTGTAGTGATGTATGAAGCACAGCTCACACAGAAGCCACTTTCAAGCAGCTTTTTGTTGGTCAATGCAATGAATTGTGAAATCTTTCCCTTGTCTGAAATTCCCAATTGTGTGGATATCTATTGGAATTACTTGATTTTGCTTGTGAAAATTCATT
104	76	GCGCCAGATCGTGCCCACAGGACGTCATCGACCGGCATCAGCCCAAACTCTTCCTCTTGTCCCCCAAACATGAGCT

105	216	CTTATCCAAACGCTGCAGTGACTTGACTGAACTGTTTTTCTTTCATGTAACTGTCGCGTCCTTACAAAACTTGCAGAAACCGATGTTAATATATCACCACCTTCAGTATTTTTCAGAATTTACTTGGAATGCACGTTCATCAGCGACAGCAGACAGTGTGTCTTTGTGTGGGACGTCCTGCAAAAGAACCCCACAAACTCCGTAATAAGAAACGCG
106	165	TGAGAGCCATTTCCCCAAACACATTTACATCAAATTAAACATATATGCCAACATTTAAATCACGTACTGTGAATTTCTCTAATGGAAATGTTATTTGACATGGCTAATTTATTTTATTAAAAAAATTAGACAAAAGGGTAACTTTTTAAACTTAGAAAACCCAAA
107	239	GCTGTGTGCCTGGAGCTGAGGGGCTGAGGTGAAGAGCATCCAGTGGCGTCTCACAGTTCCGCTAGGCAGCACACTAATTGAGAAGTCGCTCAGGGGTTGAGCGCAGGGGGCTCGCTGCACTGTGCTCCCTATCCAAATCAAGCCCCTGTAAGTGGGGCTGACATGGACCAATCACCATCTGGAGGTTCCTCCACTAAAACTCCACCATTCCTGACTACCATTAAAGGATCTGCTGCTAT
108	121	CACACGGACAGGTACTATTTTATGCACTGTGGCAAAATACAGTGTATTAAAAGGAATGAGTCACTGCTGACCTTTCAGCAAGTTGTACTGCCATGTTTCCTTAAACATAGATTTGCTAAGA
109	136	TTTGGGTGGTTTCTTGGTGGGGGCTTATTGGCCCAAATGTTGTGAAATTTACAGGACCTTTATACAGTGCATGGCATTGTAAAGACCTTAAGTCTATCCTTCACAGCCTCTCTTCATCTGTGTACAATTCAGTATG
110	251	AAAAATTGGATGACCAGTAGTTTCCTACTTCTAATTTCAGAAAAAAACAGAGATTCTAATTTTTGGACCAAAAACTTCACGTAATAGCCTAGAATACTGTCTAACACTTGATGGCTGCTCTGTTAAGTCTTCGTCGTCAGTTAGGAACCTGGGTGTGCTCTTTGATACCAATCTTTCATTTGAAGGCCATGTTACTAGCATCTGTAAAACCGCATTCTTCCATCTTAAAAATATATCTAAACTACGACATA
111	135	GCATGGTTATACTTTTCTCTCTAATATTGTCAATCTTGCCAGTAAAGAAGTTCATAAAGTCATTACTGTTGTGCTGTTTGGAAACATCAGAAGTTGAAGCTTTATTTCTTGTTAATTTTGTTGTTTAGATTTGTG
112	148	TTACCCTAACCAAAAGGACAAACTGTTGAGAAGGACAAAAAGGAGACTTTTTATCCAAAAAAGTGACCAGTTTCTGTTTTTTTTTTCTTTTCCAAAGAGAATGACCCTCATTGGACTGAATGTAAGAAATACTGTCTGAATCGGGCAG
113	97	TGTCTCTTTTGTCCACTTTCAACCACTTTTGTCCTGATTTCTTCGAGGGGATGGTAAATATATTGGCTTTTTCTGTTCTTTTGATCTAATGGACGAA
114	131	CACAGCGATCAGGTTATTGCCATGTCCACCTGCAAGAAAAACATAACAGTTGCATGAAAATCCAGTGTTGGGAACAGAGCAATCTAGCAGAAAGGTTAATTATTTATTATTGTAAGCAATCCCTGACATAT
115	133	CACTTTCTCTAATAAAAAACATCAGTATGGTGAGGATGACGTTGCTGTTTTATCCACCGTTTGTGGCAGAATTGTGGTCAGGGGCTTGAAGGCTGACAGTGAGCCGTTAGTGTATATGTGCACTGAAACACGG
116	133	ACAACAAAATGGAAATATTTGGATACTTTCTAACATTTACAGATGGAGAATATACCTGTGAAATGTAAGTTATGTACTAAGGAAAATGCCATATTTAAAATACATTAAACAGCGCAAGTATCTGTAAGAGCAT
117	193	AGGAAATGATATAGATAGAAAATAAAAACATGGAATATAAATCTAGTATGGATGATGCTAGGATAATAAAAAACGTGCTGTGGATCCACTATTGCTCAAACAAGACTTCAGATGAAACTTTAGCAAGTTTTCTAGCAATCCATTATTAGCTTATACACCAATTCTGCAATGTTTCAGCACCATCTCACAATCA
118	221	GTTACATATGATTTCTATTTCAAATAAATGCTGTTCTTTTGAACTTTCTGTTCAGTAAAGTTTCCTGAAAAAAAAAAAAAGTCACTGTTTCCACAAAGATTTTAAGCAGCACGACTGTTTTCGACATTGATAATAATAACCATTATTAATAGGTGAGTACCAATATTAATTGATGATAATTGAGCACCAAATCAGCATATTAAAATGATTTCTGAAGGATCA
119	189	CATATAGGTCCTATACCTATAAAACACACAAAAAAACATTTGAATGAAAACTTGAAATGTAATAGTTTGGACACTGTTGCAATAAATCTTGCCAAGAATTTCCTAACTTTCAACGTTTCGTTTGAATTGTGAAATACATGGGACTGAGACAATGAGACAATCAATCAATACAATCACTACACAAAACTA
120	126	GCCAATTGGGATATACTGTAGATAAATGAATGGGGAGGATAGGCTTATATAAGACTGTATAATTTGGTGTTTCGGTGCCAACTTGAGCCATGATGGTTTCAAGCCTTTGGACAAGTAGTAATTTAT
121	113	CATGGGTACCATGCAACGACATTCTGGACCCTAATCTACTGGACACCTTCCATTCTGCTCATCCCAACAGACCTGCCCCTCGTGGAAGAGGACAACCACCACGATGTCGGGGT
122	167	GTGTTGTGCCGAATTCTGACCCCTGTCAGATTATTAACCCCCTAGTATTGGTAGGTTTAGGATTAGATGTGGGGGAGAGGGGTTAGGATTAGGGGTTGGGTTAGGATTAGGCAATCAGGTAGCGACTTCAACGAGGGGGGTAATAATTTGGCAGCGAGTCGAAATTG
123	142	CTTTTTTAAAACCAGTTGGCAACATAATGACAGAATTTTCATTTTTGGGTGAACTGTCCCTTTAAATAAATGTGGTGTAATGTTCTTAACTATGATTCAAGTGGTCTGTTG

		TGAATTGGGTCTTTGATTTGGATCTTTTATA
124	121	AAAGTATCATTCTGACTGGATTATAAGAGCAGACTGATCTCTTAGAATTCTGCTCAGAAACACTGCCATGAATAAAGAATGGTATAAAAACTTCCTGCAAGAGCAGCTTCTCCCAACGATA
125	278	ACTAGTGTGCGTTACTGTACAAGTTACGGGTTTACTACAGAAACGGAGAAGGTGAATTCAGGGCAGGATAAGGAAAAAAATAAATAAAAAAAAATTGTGGTTTGGGAATTTGGATACAGGAGAGATGAAGTATCACAGAGCCAGGTAATTTGCAAGACCTGTCAAACTGTCGTTGCGGCAAAACAGAGTAGGGGCCATTCACACAGAATGTGTTTTTCTGGTTAAAAACATGAGACGCAGGTCTAGAGACTTGTTTTTTTAAAAATGTAACTTATTTTA
126	171	ATCTCTTTCAGTAGCTTAGTCGGTATAGGGTCTAACATACATGTTGTTGATTTTGATGATTTAACAAGTTTAGACAATTCTTCCTCTCCTATAGCAGAAAAGGAATGGAATTGTTCCTCAGGAGATCTGTAGCACGACATCTGTGATACTGTAGTAAATGGTTGCATGGTT
127	224	CCCCATGATCCTTTTCACAGGGAAGTTGTTTGGCATCACAGAATGGGTACAGGAGGCTCAGAGTTCGATTTTACCTGTAAATGTATGTATAGTATTTTCTATACTACTGTAATATTTATACAAGCTAGATTTGGTACATTATTTTGGTAAAAACGACATTGTACTTCAACAAAAATACTTTACATTTACCTTTATCAGTCCATTCAAATGTTTCAAATACATAG
128	229	TTTTTCTTTTGTTGCTTGATTAACATTAATGACAGACAGCAGCAGGTTTATTAGGCTGCTGTCACTTTAAGGACAAATGCACAGACCGAATATACTTATGGTTATTCAAAACTTTTTAACTTATTTAAAACTGTGTTCACGAGGTCACTGGCCAAAAAAAGCATTTTGACATAATTTTGTGTTTTTGTCCATTAAAGTGCGACCAGACAGGCGCTAAACAGACAGTGTG
129	167	CACAATATTCTATTTCGTTCCCTCGATTTGCTAAATCGCGTACACAATTTACTATTTCGTTCCCTCGATTAGCTAAATTGCACGCATTATTTACTATTTCGTTCCCTCGATTTGCTAAATCTTGACGACGATTTACTATTTCGCTCCCTCGATTTGCTAAATCTTGA
130	140	TTGAGCTTCAAGAAAAGACGTGCTGGAATTATAGTAAACAAAATAATAACATGAAAACGCTTTGCTGAAGTTCTGCTTGTCAGTGATGAAGAAATAAAAAATACTGCTCGCAGCCTCTTTCTTGTTCTGCCCCACATTTT
131	101	ACTGTACTGCTGGTTGCTATTAGGCATTTCACATCAAGCACGATAACTATAACAATAAAGTTTTAATAAACATTGTTTCTTTGAGAATAGGGAAGTCCACA
132	141	CCCTCAACTGTATCGTAAAAATGTTCATTACACTAATGCTGTGTTCGAAATCGCCCCCTATACCCTCATTCACTATTCCCTACATTACTCCACTAATATAGTCCACTTGGAGTGAATGAAACGAGTGAGTGAATTCGGACA
133	90	CTATCCTGACAAAAAAAAAAAATTATCCTAGTTTCCACAACTGTTTTCAACACTGATAATAATCAGAAATGTTTCTTGAGCAGCAAATCAT
134	105	ATATAAGACTGGTTTTGTGCTTCAGGGTCACATATAACATTAATTTACTTTGTCACTTGTGTTTTAGTCTCGCCCACAGGTGTTTGCTGGCATGTCTGATGGTCT
135	105	ATATAAGACTGGTTTTGTGCTCCAGGGTCACATATAACATTAATTTACTTTGTCACTTGTGTTTTAGTCTCGCCCACAGGTGTTTGCTGGCATGTCTGATGGTCT
136	112	TGTGCGATTTGGTTGATATCCACAATTAAAACTTATTTAATACAGACTTCAGCACATCGTTTGACATGTTTTTGCATTAGGATTTGAATGAATCGATTACCACTTGGCTGTT
137	178	TTGACTTCTGAGAGACACTGCCACTCTGAGAGGCCCCTTTTATACCCAACCATGTTGCCAATTGACCTAATAAGTTGCAAATTGGTCCTCCAGCTGTACCTTATATGTACATTTAACTTTTCCGGCCTCTTATTGCTACCTGTCCCAACTTGTTTGGGATGTGTAGCTATCATGAAAT
138	111	TTTTTTTCTTGTTCAAGTTAAGCATAAAAGGCAGTGCTAAATCTTTTGTCCAGTGCAGATGTTAAATTTAGGTAGGAATATGATGGGTAATTTAACACATTATTGCAGGAA
139	152	TAATTTGGACTTCTCTGTGCATGCTCTTTGAGGTGGTGACCATAGACCTGGATTATATGAGTCACAGACAAGAACAGTTTGACGTAAAAATATAAAAATGGAAACTACTGCGGAAACAAAGACATCTACATCTTGGATGCCCTTGAGGTAAG
140	116	TAAAATGGTGTAAGATTTAGGAATTTGATAATGTACTTATCCATTTTGTTGCGAATGCAGTGCACTTGCAGAGAGAGACCGCCCACACACTCTTCTGATTGGCTGTGATTTTTGAT
141	249	TTTGCGTTCTCTCACAAAACTTTTGTGTTCCGCTAAGAAACTTTGTGTTCCCTCGCAAATCTTTTGCATTTCCCCCGACAAACTTGATTTGCAGGCAACTAAAAAGTTTTGCAAGCGAACGCAGTTTCTCAGGGGAACGCAAACATTTTGTGAGTGAACACAAAGTTTCCTGGGGAATGCAAAAAAAAAATTGCGAAAGAATGCAAAAGTATTAACATATAATCTTTGCGTTCTCTCGCAAATTTTTTTG
142	233	AAAATCTTGTAATAAAAGATGATGTTGTATTGGCTATGACTAGACAGCAAATACAGACTCTCCATGTTTTTTGAGTATTATTAAAGTACAGGCAAAAAAAAAAGTAAAAAAGTTATTAACCAGTATTTTTTCTAATCAAATCGGTTTCAGATTGGTAATAATACAGAAGGAACACATGAACAGAAACGTTAAAATACTGATTCTGCTCAGAGTGAATTTTTA

		AGCCCTGATGG
143	172	ATATGCCAGACCCAACAATGCAGAAGAGCTGAAGGCCACTATCAGAGCAACCTGGGCTCTTATAACACCTGAGCAGTGCCACAGACTGATAGACTCCATGCCACGCCGCATTGCTGCAGTAATTCAGGCAAAAGGAGCCCCAACTAAGTATTGAGTTCTGTACATGCTCATA
144	250	CCAGGTGAGAAAAAGTAACACAACATAACATTTCCATAAAAAGTAACTATGTAACGCAGTTACTTTTTTAGGGAGTAACACAATGTTGTAATGTATTACTTTTAAAAGTAACTTTCCCCAAAACGGGTTAAAAATAAACTCATTTACAGCCACTTTTTCCAAATCCATTGTTTAGTTATACCGTTACTGTTAGTTGCTGTACCAGTTAGTATTGTTCTTGTCTTTGACTGTTAATTGCCAAGTTTATGTT
145	102	TATTGCAGGCAGATAATATTGGATATGTTCAGCTATTAGGATATAAAGCTTACAAAATGTTTATCATAAACTAGCACTAAGACCCAGAATCTGCCTGCAATA
146	119	AAGCAGAATATGGCTGTTTCTAAAGCAGTGCCATCTGCTGTTAAAAACTAAGCTCAGAATCGGTTCAAGAGAGAATCACGATGCATTTGGAAAATCTCAGAATCGATCAGGAATCATTT
147	247	GGTGCATCACTAATAAAAAAAATATTTAAAAAAGCATTTTAAAAATCTAACAGTTTTTCACTGTAAAAACAAATGCAGTTTTAAGGAGCAGTGTACTAGTTTTCTTTTTTGTTTAGATGGTTAATATTAATGCAACTATCAGCGCACTTTTTTTTTTTTTACAATCCGGCCCTCAACCTAAAATGAGTTTGACACCCCTGATTTAAACTGTTCTGACCAGCAGGGGTCACCAGCGTGTGGTGTTCCG
148	210	CATACTGCCCTAGGTTAACAAGACCTTTAATGGAAGGGTTACCGTAATCTCCTCCTCCCTTCTTTTTATTTGTATATGAGTCTGGGAATAACTATGTCTTGCTATTGTCTAGTAGTGCCCTCACTGTCATCTGATTCAGGGTCTCCAGAGAGCAGTTTCCTGTCTGCTGCATCAGCACATCGTGCTCACATTACCTTCATAGTCAGGAGG
149	288	TTTACTGTTTAAGAAGAGTATCTTCTCACCTTGACCTTCCTTGTGTTGTAGGTAGTTTAGGCGACGGCTGTATGTATGGCAAAGGGACCCGATTAACCTGTGGGTTGAAGAATCATGGGTAACCTCATTAGCAGACCTAGCTGTTTGGGACAAAAGTCCAAACAAGTTAAGTCAGATGAGAGCGTCCTGAAAGAGTGCTACCAGCAACGGAGAGAATGGCCGCTCCCGGAGCCTCCTAGAGAAGAGAAAAGAAAAGAGGATGTGGAAGATGTTCTTCTTAAACAGTAA
150	248	CCTGTTCGGGCGTAGGGGCCGTCACTTTGTTGGAGGCCTCAGTCTGCTGTGACTCTCTGCAAGAGCGCTTTTCAAACCTGGAGAGCCTCTGCGTGTTTTGTGAACCTCCCTTACTCAACACACCTGATTCAGATCATCAGCTCTTTGGGAAAGAGCGCCAGTGAAGTCAGTTGAGTGCGCCTGATACGAGAAACGTACATACAAAATGTCCAGTGCTGTGGGTCCACGGGCGTAGAGACGTGGCAGCA
151	122	CGAATTACTTTTTTTGGCTATTATTTGGCTATTTAAGCTGCTTTTTTGCTACTCACTGCACAGAACAGGCGCAGAGAGAGGTGATCCTGTGCTGGGGAGTCAGACCTCACGCTTGCGGGGCA
152	229	AGGGCAGACAGAAAATCCCAAATTCAGTATCTCAGAAAATTAGAATATTGTGAAAAGGTTCAATATTGAAGACACCTGGTGCCACACTCTAATCAGCTAATTAACTCAAAACACCTGCAAAGGCCTTTAAATGGTCTCTCAGTCTAGTTCTGTAGGCTACACAATCATGGGGAAGACTGCTGACTTGACAGTTGCCCAAAAGACGACCACTGACACCTTGCACAAGGAG
153	77	TTTTAACAGACGACATACTGAACCGGTATGAAATTGGCAGAAAGCTTGGTGAGGGAGGATGCGGGTCCGTCTTTGAA
154	258	TCAACGCAGTTTTATTTTCATAAAAAACAAAACAAAAGTAAAACATAAGAACAAATAAAAAAAATTGCAAACAACAGCACAAATAAATGCAATAGAATAGAGAGCTATGAAAAAGTTTTTCAGGCTAAGTAGGTTTTTATTCAGGTAAGAAATACTACAGAAATTAAAGTAATCAAATGTAAAATAACACATCCAGTGTTCCCCATATGCGTCTTGTCTTATTATATTATTATTAATATTAATAAGACGAACGACACA
155	113	ATCCACTCAAAAAAAATATATTTTTGATATATTTACAGTAGGAAATTTACAAAATATCTTCATGGAACATGATCTTTACTTAATATCCTAATGATTTTTGGCATGAAAGAAAA
156	249	AGTCCGTACAGCAAACAAATCCAAATCGTAATCCAACAAACAAGGTCAAGAACAGGCAAAGAACATACACGAGGCAATCAATGTGGCAATGAGAAAACGCTTGGTAGAGACAGGGGTTAGCTGGCAATACTTCGCAATCTGTTGTACAACTGAGCTGGATTTAAACAGGAAGCTACAGGAAGTGGAGGAAGAAGAAGAGCACAGTCAGTAATGAGGAGATGGCGCCCTCTGGCGGCTTGATGAATGGAT
157	121	CGAATGTCCAATGTCAAGCCAACTTTATGCCTGTTTTTGTTTTGTTTTTTACTTTATTTTTCTTTTCTCTGTGCCGGATTGCTGATGTCTTTTACCCGGGCATAGATGTCCATCCATCAAT
158	93	CAGCTCACACTCCCAGATTCACTGTTTGATTGGTTACATGATCTGGACTGAAACACACCAGTTTCACTTCTGCTGTAGTCAAGAGATAATTCA

159	192	GACCTCGGCCTCACTGCGCAGTGACTCGTGGCTCTGTTCCACCACATTATGGCCAATCACGGTTAATGACTAGTGGTCACGGCAAATCAGAGTTAAGTGGGAGAACTGGCCATTCTCTGATCGTGCCGCACAGCGCTTGGCTGTCTCTGATAGGTTAAATTGTTTGATTCTATTATCAGTGAAAACTCTGAG
160	130	TTATCCCTATTGCTTGTACACTTTTTTCTACCACATCTTTTCCTTCCCTTCGCCTCTCTATTAATGTGCTTGGACACAGAGCTCTGTGAACAGCCAGCCTCTTTTGCAATGACCTTTTGTGTCTTGCCCT
161	121	CTTAGATCCGGTAGTCTTACTGTTCCCAGAGCCAAATTAAAACAAAGAGGTGAATGTGCCTTTGCTGTTGCAGGCCCTAAACTGTTTACCAGTTCACAGGTCTGCACCTACATTACAGATT
162	164	GCTTTCTCTCTACCTATATCTCTTTATGTCTTTTTCCTTCCTTCACCCTCTCTCTCTCTCCTCCCCATCACTCTCTCTCTCTTTGTGACACTATCTACCTCCTTCTCCCCTTCTCTCCCCCCCTCTCCGTGCTCTATCTCACTCAATCTAATTTTCTCTTTCCT
163	52	AGCGAAGGCTGACCCGTGTGGTTCGATCCAACAGAAGAGCTACTGTAGCTCA
164	141	TAAACATGTTTAAACATGTTTAAACATGTTTAGCGATGTTTAGCATGCTTAAACATGTTTAGCGATGTTTAGCATGCTTAAACATGTTTAAACATGTTTAAACATGTTTAGCGATGTTTAGTGTGCTTAAGCATGCTTAAA
165	133	GCTACAGGCGTTGGTGGAATGCTACAGGCGTTGGTGGAATGCTACAGGCGTTGGTGGAATGCTACAGGCGTTGGTGGAATGCTACAGGCGTTGGTGGAATGCTACAGGCGTTGGTGGAATGCTACAGGCGTTG
166	165	CAACGCCTGTAGCATTCCCCCAACGCCTGTAGCATTCCACCAACGCCTGTAGCATTCCACCAACGCCTGTAGCATTCCACCAACGCCTGTAGCATTCCACCAACGCCTGTAGCATTCCACCAACGCCTGTAGCATTCCACCACTAAAGGAATTGCGAATAATAAT
167	84	ACGTAAGTGCAGCAGAAATGTTTTTGTAACCTTGGCCAGATCTGTGCCTTGCCACAATTCTGTCTCTGAGCTCTTCAGGCAGTT
168	161	AGCAAGCCCAAAAACACCGCCAAACATCAGCAAGCCCACAAAAATCAGCAAGCCCAGAAACACCGCCAAACATCAGCAAGCCCACAAACATCAGCAAGCCCACAAATATCAGCAAGACCACAAACACCGCCCCCAAACAGCAAGACCACAAACACCGCCAA
169	238	TGATTTTTAATGACAGCACCAAGTCTTCTAAGCATGGAATGAACAAGTTGGCAACATTTTGCAACATCTATCTTTTTCCATTCTTCAAGAATTATCTCTTTTAGAGACTGGATGCTGGATGGAGAGTGATGCTCAACTTGTCTCTTCAGAAATTCCCCATAGGTGTTCGATTGGGTTCAGATCAAAAGACATACTTGGCCACTGAATCAAGTTCACCCTGTTCTTCTTCAGAAAACCA
170	175	TGACTTGAACCCCATAGAGAATCTATGGGGTATTGTCAAGAGGAAGATGAGAGACACCAGACCCAACAATGCAGATGAGCTGAAGGCCGCTATCAAAGAAACCTGGGCTTCCATTACACCTGAGCAGTGCCACAGGCTGATCTTCTCCATGCCACGCTGCATTGATGCTGTAATT
171	176	GAATTACAGCATCAATGCGACGTGGCATGGAGGCAATGAGCCTGTGGCACTGCTCAGGTGTAATGGAAGCCCAGGTTGCTTTGATAGCGGCCTTCAGGTCATCTGCATTGTTGGGTCTGGTGTCTCTCATCTTCCTCTTGACAATACCCCATAGATTCTCTATGGGGTTCAAGTCA
172	167	ACCCATGCATGCCATTCCTCTGCAGTGTACACCGTATTGTGTCACGGGAAATAGTCACCCCAGTTTGGCTTTCTACTTCTTTAGATAGCTGCAGTGAACTTGCTTGCCAATTTTCTTCAACCCTTCTCATCAGAAGACGCTCCTGTCAAGGTGTTAACTTCTGTGGA
173	149	CAGGTAAGAGACACACACGCACACGGCCATCCACATGATGTAAAAGAAAACATGATTCATCAGACCAGGCCATCTTCTTCCATTACTCCAGTTCTGATGCTCACGTGTCCACTGTTGGCTCTTTCGGCGGTGGACAGGGGTCAGCATGG
174	145	TCCCGTGGATGCGGATATCCATCCATCCATCCCGTGGAGGCAGATATTTATCCATCCATCCGTCCGTCCATCCATCCCTCCTGTGGAGGGGGATATTTATCCATCCATCCATCCCGTAGAGGGGGATATTTATCCATCCATCCAT
175	164	TGGGGAGGCCTGCGTTTCGGCATCCTCATCCTCTCAGTCTCAGATGCAACCATTATCTGAGCCAGGGACTCAGACCAGGAGACATCATGCCTGCATTTCACACACACACACGCACACACACACACGTACGTACACACTGACACACACGGAAACATCCGTGTGTG
176	142	ATCTCCAGTGCAGACAGAGCCTCATGACACAGCCTTCAGGGGAGAGAGAGAAAGAAAAAGAGAGAGAGAGGGAGGGGAGAGAGGGGGAGAGAGACAGAGAGTGAGAGACAGAGAGGGAGAGAGAGGAAGAGAGAGAAAGAGA
177	147	GCTCGGAGGCAGAAAACTCAGGAGACTGGGGCTGCACGGAGGCAGAAAACTCAGGAGACTGGGGCTGCTCGGAGGCAGAAAACTCAGGAGACTGGGGCTGCACGGAGGCAGAAAACTCAGGAGACTGGGGCTGCACGGAGGCAGGAG

178	152	TATCACTAGCCCAAACCTAAACCTAACCCTAACCCTAACCCTAGCTCTAACCCTAAACCAAGAACTAACCCTAAACCTAGAACTAACTGTAACCATGAACTTAGCACTAACACTAACATTAGTCCTAACCCTAACACAAATCCTAACCCTAA
179	178	AACCACCCCCCACCACCGCCTCCCCACTACTCCTCGCCTTGCAGAAAAGACAATGATTCTAATACTGTCACGACCGCTGAGATATACTTCATGTAACACATGTTCCCTCGACCTTTACGGCACGTGCTAAAGAATTAATTCAGATATGTTTTGGGAATACATGTAAATGTGCCTGTCG
180	159	CACCTAGCCTCACCCCACCTCGCCCAGCATCACATGCCACGCCTCGCCTCGCCTTATCTCGCCCCATCTCGCTCCACCTCGTCCCGCCTCGCTTCCCCTCGCCCCACCTCGTTCCGCCTCGCCCCGCCTCACCCCGCCTCGCCTTACCCCGCCTAGCCT
181	92	TAACTCCGCTTCTCCGTGATCCTGCAGATCTTCCTCCACTTGGCGCTGCATCTGTGTGCAGTTGTTTTTGCCATCGCAGCCATTGTGCTGTA
182	178	GTGAGGACAGGAACATAGATTGACCAGTAAATTGAGAGCATTGTCTTTCGGCTCCCTCACCACAACAGACTGATAGAATGACCACATCACTGTAGACACCGCACTGATTCACCTCTCAATCTCACACTCTCCCTCACTCCTGATCAAGACCCTGAGATATTGAGCACCTTTTTCCGGT
183	230	GTCCAACTAAATGCAAACTGGATGGGATGGCATGTCGCTGCAGGATGCTGTGGTAGCCATGCTGGTTCAGTGTGCCTTCAATTTTGAATGAATCCCCAACAGTGTCACCAGCAAAGCACCCCCACACCATCACACCTCCTCCTCCATGCTTCACGGTGGGAACCATGCATGTAGAGACCATCCGTTCAACCCAAACCAAAGATCTCAAATTTGGACTCATCAGACCAAAG
184	110	TGGTATTACTTGCTTCCCCCGGGGAAATTAGGTCAAAAATTCCCCTTTCCCCGTTTCCCTGTATCCCTGCCTGCCTCTATCCCTGTCTATCTGTCTCTCTATCTGTCTCT
185	163	ATCAAATCCCAAGATTTCAGGAGAGAGCTTGTGAGTGAGGAAGTGTGTGTGTGTGTGTGCATGTAAATATGTGTGGTGTGTGTAGTATGTGGGGGGTGTGGTGTGTGAGTGTGCGTGTGGTGTGTGTAGTGTATGTGGTGTATGTATGTGTGGTGTGTGTGTG
186	259	CAGCGGATGGATGGATGGATGGACGGACGGACGGATGGATGGACGGACGAATTGCTAGATATTGATGCATTGATAGATATTTGTTAATTATTGCCTAACACAAATAATAATAATTCCAGATTAATGGAATGCTATTCGATATGGGATACTTTATTATATATGCTTGGACTGGGTAGGAGTGAACGAACGAACGAACGAACGAACGATCGAATGAATGGATGGATGGATGGATGGATGGATGAACAAACGGATGAATGAA
188	249	GTCGATGGTGCATCATAAGAGGGGTGCCACGGGAATCTTGGAGTGGGGAGGGAATGGGGACTCTTCAAATGGAGAGGGAAACTCATGCAGGTCTACGGGTTAGGGTACAACATCTCTCCCTTATATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTAGGAGGTGTTAGCGTCGTTTGATAGGAGGAAGATGGGATTGATGCCACAAGCCCTGCTTCTCGTGTTTTTTGAGTTCGGCTCTTGGGAGGTAT
189	159	GTCTCCTCAATTTATTTGGTCGTCTCTAGGCTGAGTCACTCACCCAGAGTTGACTCACTCACTGATTTACTCACTCAGCCATCCAGCCACCCAGAGTTAGATCACTCATTCACTGATTTACTCGCCCGCTCACTCACTCACTCACTCACAGATCGGAAG
190	144	GGTAGGAAAGGTTCCACTGCAACTTTCACTTACATAACCTACTATACTTCTGTCAAAACATGTCTGCATGCACAACAGCACTGTGACTGGATTCACCCGGCGTAGACAGTGACAGAGGACTGCAGCTGCTGCTGACGTGGCGCT
191	117	GTAACCAAAACATAAGAGCCATGAAGATAGTTAACTTCACTTGAATGTACGGTTGTTAACATCAGCTAACATTGGCCACCAATATCCACTGGTCATACCAGACCACATCAGACTGCA
192	247	GTCAATGTTGCCATTGCCGCCATCAAGACCAAGCGCAGCATCCAGTTTGTGGACTGGTGCCCCACTGGTTTCAAGGTTGGCATCAACTACCAGCCACCCACTGTGGTTCCTGGTGGAGACCTGGCCAAGGTCCAGAGGGCCGTGTGCATGCTGAGCAACACCACTGCTATCGCAGAGGCCTGGGCTCGGCTGGACCACAAGTTTGACCTGATGTACGCTAAGCGTGCCTTTGTTCACTGGTATGTGG
193	170	ATATTTGAAAAAAGGATCACAAGAAGGAAGGTATGGTTACTATAGTGGACGTTTAAAATATTTCTGCAGAATGCACTCCAACATCCATTATGATTTCTGTGTATTCACTGTTTTCTTTTGTTGACCTTTCTCTTAAAGCAAGTTGTGTGGCATTGAAGTGAAGTCGGAGG
194	206	ATTTTTCGGGCACACAGCTCATTTTTCAGGCACACAACTCACTTTTCGGGCACACAGCTCATTTTTCGGGCACACAGCTCATTCTTCAGGCACGCAGCTCATTTCAGGGCACACAGCCCACTTTTCGGGCACACAGCTCATTTTTCGGGCACACAGCTCATTCTTCGGGCACGCAGCTCATTTCAGGGCACACAACTCATTCTTCG
195	111	ATTTTGGTCAGAAATCAGCATTTTTAGTTACACCCACCCATGTTCTCCAGACAATTGCTAAGGAATGGAAAGAGGTACAAACATTTTTTTTTTCTGATGATAAAGGAAAGT
196	118	CTTTTAGGTGACATCCACGGCCAGTACACGGACTTACTGAGGCTGTTCGAGTACGGGGGATTTCCGCCCGAGTCCAACTACCTGTTTCTGGGGGACTACGTGGATAGGGGCAAACAGT

197	131	CTGCTGTGATAGGGGTTCACACACACACACATACACAATCATAGACATGCACACAAACACATACACACTGATGCACAGGTGTGCATACACACTCACATACAGAAGCACAGACATGCAGACACATAGACACA
198	173	AGGCCCTCCTCTCCTCTCTCGCTACCAACCTCTCAGTTGTCTCTCTTTCTCTATCCCTCTCTCTCTCAGCTCTCTCTCTCTCTCTCTATCCCTCCTCTCTAGCTGGAGGGAGATCCTCAACTCCATGACGTCTTGGGGGAGGTCATCATATGAACGACGCCACAGCTCCTCCT
199	181	CACACACATTCCTAGCAGTATTGTACACACATTCGTAGCAGGTCCACACACATTGTAGCAGGTCCACACACATTGTAGCAGGTCAACAAACCTAGACGCGCGAAGCGCGGCTAGCTCAGCCGCGCAGCGCTAGACGCGCGAAGCGCGGCTTTAACTCATAAAATGACCGATTTCCGTTGTA
200	118	TCACTAAATGAGCCGATAAAGTGACGTTGAAATGATAAAGGTCAGAATTTAGTTCACAGATAAGATCTTCCAGTGTAAATCATTGTTGTCAAATGGAGTGGAATGAAATGGCTGAGGA
201	248	ATCCATCCATCCATCCATCCATCCATCCATCTATCCATCCGTCCGTCCATCCATCAGACCGAAAATGGATACTGTTTGAGGAATTGCTTGAGACTTGGAATGAGACACAGTCTGTCGACCTTCTGACCTCTGAGGTGACCTTGAATGGTCAGCCTGGCCAGGTGGCTGAGCCGTGGAACGCCGAGCAGGAGGGGACGATTCAGGACAACCGGGGTGTCGCACCGACCTGAACTTTCGACCAGCACGTT
202	155	TGGTAAGCAGAACTGGCGCTGCGGGATGAACCGAACGCCGGGTTAAGGCGCCCGATGCCGACGCTCATCAGACCCCAGAAAAGGTGTTGGTTGATATAGACAGCAGGACGGTGGCCATGGAAGTCGGAATCCGCTAAGGAGTGTGTAACAACTCA
203	50	ATCCGCTAAGGAGTGTGTAACAACTCACCTGCCGAATCAACTAGCCCTGA
204	82	GGTTGATATAGACAGCAGGACGGTGGCCATGGAAGTCGGAATCCGCTAAGGAGTGTGTAACAACTCACCTGCCGAATCAACT
205	222	GCTTCATACTTAAGATAATTAGCATGTATTAGCTTCATACTTAAGATAATTAGCATGTATTAGCTTTATACCCAGAATAATTAGCATGTATTAGCAAGTATTAGCAAGTATTAGTCTCATACCTAGAATAATTAGTATGTGTTAGCTTAAATCTAGGATGATTAGCATGTATTAGCTTCATACCTAAGATAATTAGCATGTATTAGCCTTATACCCAGAATA
206	109	GGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTTGGTCGATTGATTGATTCTGGATGATCTTCTGGGATTGACTGTTTATCACTGAAGTGG
207	98	ATTTCTACCAATTTCACCTGTGAAATTGCTGGTTCTCACATCAAAACATCCAATACTGTTAAAAATTTAGGTGTAATCATTGACTCCACCCTTTCCTT
208	160	TGTTAACACAAATAGCTTATCAGCCAATCACATGGCAGCAACTCTGCCGAAGTTCAAACAGAGCATCAGAATGGGGAAGAAAGGGGATTTTAAGTGACTTTGAAAGTGGCATGGTTGTTGGTGCCAGACGGGCTGGTCTGAGTATTTCAGAAACTGCTGA
209	263	CGTTCTGTGGCAGTTTGCCTCCGTTTCTAAACACCTCTGAAACATTTATGTTAAAAAAAAAATATTGTCTGTGGGAGATGGGAGCAGCTGTGACATGTTTGTCAGCTGACCCACTTTTTCCTGATTATGAATTAAAATTGAATGATGGCGTCTGAGATTGCCAATTAAAAATGCATGCAGTGCCACGGTCAGAGCTGAGAGGCCAAACGCAGTGTCGCCTCCGGCCTCCTACAGGAAACCATAATTCACTACTGGGCCTTTTA
210	171	GAGGCGTTCTAGCAACTGATGCTAGAGACTGCGGTCTTTAGCCTCCTTGTTAGAGGGCCCGATTCCCATGCCAGCGGACCCGGGTTTGAGTCCCAGTTAGAACAGGACGTGTCACTTTGGTGCCATGACCTGGATGGGAATGAGGTTTAGGGGGTAAGTGTAATGTACGTA
211	126	CTTCATCCTCGTCCTCCTCATCTTCATCTTCGTCATCAAGCATCGGCTCCTCATCATCAAAATCCTCCTCGTCCTCATCATCATCTTCCTCGAGGGTGTCATTCAGCGTTTTGAGAGTTTCGGTCT
212	252	GGCGACAGTCGCTCAAAGAAGCGTTCGATTTCAGTACTACTCAGTACTACGGTACTACTCTCTCCCACCACCCACCCATCCATCCATCATTTTGTTCTTCTTTCAATCTGGTCCTTTCGTCCATCTATCCATCCATCCGTCCGTCAGTCCGTCTGTCCATCACGGCCGCTAATTGTTCCGCTCGACCGATCAGGTGAGGAGTTAGTTTCGCACTCAGTCCCTTATCTCGTCTGAATTTTTGTTCTTCTTTCA
213	210	ATGCCAACACACACACATACACACAGACACAGACATGCACACACAGAACTTCACATGACATCAATCTGTGTTGACTGTGTAGTGATCTCATTACAGACAACTTGCTGTTGATTAAAGCCTACTGTGTCTCTCTCATACACACACACACACACTTCTCTTTTCTCTTTTCCTCCTTTAGCCTGAGGCTCTCTCCATTCTGCGGTTTCTCCT
214	127	AAAATAAATAAAGATCTAGACTTTAATAATAATGATAAATATACTAAGATATTAAGAAAATTAAAATATAGATTAATATCTTAGTATATTTATCATTATTATTAAAGTCTAGATCTTTATTTATTTT
215	122	CATTAAAATCATACCATTGACCAAGCACTTTTTGCCTCTGTGACTGTGGTTGACCAGTGGCATTCATTTATGAGTTCACAACACATGAGAATCACCATGACTCTGTCGTACGTTTATGATGT

216	187	GCTCTTATTGGCCTACGTCCGTTACACTCACCCCCCTAAACCTCACTCCCATCCGGGTCATGGCACCAATGTAACCCCTCCTCTTCGAACCTTGGTGGCCAGGTCAGGGGATCAGGTGAGTTAGCCTCCATTACACTGGCAACCACCCACCTCACCAAGCATTGCGGTGGCACGTTTTGCATGGGAA
217	264	ATGGAAAAGATACTCTTAAGATGAGACGGAAAAGTGTTTGGCCCAGTTTGTGACTACAGTTGGCTCTTAAATAATGAGTCTTGTCACAGGCATGTAGTTAGTACCCACAAGACCTGATCCCCTGTAATTTAGAATCAGTATTCCAAGGATAATGTCAGTTAGGATGTACACAGTAAACCTCAAAGGATATCCCCCGGTTATGCCTGAATATAACCTCTGCTATGCTTTGAATCTGCCTGTGATCAGCATGAGTATCTTTTCCAT
218	168	GATAGATAGACAGACTGATTGATTGATTGATTGATTGATTGACAGGTAGACAGACTGAATGCATATAGATTGTTAGATAGTTAGATAGATAGATAGATAGACAGACTGATTGATTGATTGATTGACAGGCAGGTAGACAGACAGAAGGCAGATAGATAGATAGACAGA
219	199	GTCAAGAAGGAGGTCAAGGAGGAGGTGAGCGTTTAAATTGCATTCTTCTCATTTCACCGCAGTGTATTGTTTTGTCTAAAATCTCAATCTCAAATCCTTAGTATATTGGCTCACTGACTAAATCTTATTTTTAATTTACTCTTCCTAGGATAAGGAACTGCGTGATGTTGGGGACTGGCGTAAGAATATTGAAGACAAG
220	157	GACGTCTCAATGGTTGATTCGTTTGAACTTTGACCTACGTTGCATTGAACGGTTGAATGATTGAAAGCGAAACTGATTGATGGACAGGCAATAGGGTGAGGTAGAGGAGGAGGAGGAGGAGGAGGAGGGGAGGAACAGGGTGCTTTGGGGAGAGACT
221	276	GGCTATGGTTGCAGATATGAGATGATACATATCGTATCATATCACATCCATCAAGTCATTCGTTCATTCGTTCATTCATTCATTCATTCATCCATCCATCCGTCCGTCCGTCCGTCCGTCCGTTCGTCCGAATATCAATACAGTCATCCTTCAATCCATCCATTCATCCGTCCAATCATTCATTTTTCTTAAAATCTCAAAGCAGTTTTACAGTCAGTTCAGAGTTTAAAATATGAAATTATTTCATTTGCGACGTATTCTTGATGTAATTTCAAG
222	188	CCTGGAGAGGGCCTCCGTCACCGGTGCCTCTGGTCGCCCGGTCGACCTCGAGGCGCTGCGCCCCGCCGCCGAGGACAGCGCTGCACCCGAGGAGAGCGCTGCACCCGAGGAGAGCGCTGCACCCGAGGAGGGCGGCCGGGACGACGCCTGACGTCGTACGCCGTCGGCGAACCTCGTCGGCGCGTCGT
223	62	GTTCTCCTGCTGCACGGTATGGTTCTCCTGCTACACGGTATGGTTCTCCTGCTGCACGGTAT
225	105	TGTCGTGTGTTTGCAGAGGAGGCGGTGACTCGTGTGATCGATTTTAAAGTCTGATATTCGTCAAAATATTGTGTTTCAGTTCATATTTGTTGGCGTTTAATAACT
226	90	GTGGTTTACATGTTTTAGTCTACTTGAAAACAAATTTCTTATTGCAAACCTTTGTACAAGGAATGTAAACTTTCTCTTGCATATCAGTAA
228	129	CATGTGTGGAGGATCTGAGATTGCATCATCTATTTGTTGCCTGTAGATACGAAGCTCTGCGGTAGAAATGAGTTCCTCCTCAGGGATGCTGCTGAGATTGAAGACGAAGCGCAGCGGAGTCTCCTGTGA
229	266	ATCCTGAAAAAATTAATCACTGTTTCCACAAAAATATGAAGCAGCACAACTGTTTTCAACATTGAATGTAGCCTTGGTGAGTATACAATACTTCTTTCAAAAACATTAAAAACATTTGCGGACTCCGAACTATTGAATGGTAGTGTATATTAGGGTTGGTAACCGAGAACCGGTTCTCATCTGGAACCGGTAGTGTTTTTTGAAAAGAACCGGAACCGTGCAAGATTTCTAAGTTGCGGTTTCGAAAACGGTTCTGGTGCGATGGG
230	236	GAGCTGCCTGAACCGAAGAAACACATTGCATTACTTAAAAATACTTACATTTTCTTCATCTGCAGCTTTGCCTCGCTTTTGAAAACAAAATGGCAGCCCTGGGGGTGGAAATGTGCAGATTAAGGGGTGTTTAATATTATAATAAGATCCCCTTTTTAAATCACTAGGGGAGCGAAATCGGAACGGCTCGTTTTTTCACATGCTTGTAGAGAAAGGCTTGCCAAAACAAAGGTACT
231	225	GCCGTGGGCAGTGGCGGCGTCTTGCATATCAATAGTCGGGGACCCCTGGGCCGTTTCCCCCCCACGGTAGTGCCTGGGGGTGGTTAGGGAGTTACGTGGGTGGAATGTAAGCGGAGCGGTGCCGTACAGCTGTGCAGCAGTGTGGATGTTATGAAGTGAGGGAGCGCCCAGTCACGGCGCTGCGCTGTGTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGCG
232	202	AGATCCAGAGCACTCCGGAGCAGGTTTTCATTAAGGATGTCTCTTGACATTGCTTCATTCATCTTTCCCTCAATCCTGACTAGTCTCCCAGTTCCTGCCACTGAAAAACATTCCCACAGCATGATGCTGCGACCACCATGCTTCACTGTAGGGATGGTATTGGACAGGTGATGAGCAGTGCCTGGTTTTCTCCACACATATA
233	246	AACCTGCTACCAAGAGTCGGTCTGGCACCTGACGACGCATGATAACCTCTCAGCTCTCTGCCGATACTCAGTGACAGACTGGAGGCTATCTGCGAGTGTGTTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATGTGTGTGTCAGCACGCTGACTGACAGATTAAAACAGGGAACAACAAAATCAAACGGTGTCCCGCCCTACGGCTCTGATTTCCAATGACAAAGCAAGAAGTCTA
234	118	CGTCGAATTCTCGTTCTGCCTCGTAGAGAGGCTCGAGATTCCGTCAAAACCGACGTCTCCGGTGGGACTCGAACCCACAACCTCCGAATTGCTTTCGGATCAACCGCTAGAAGTCCAA
235	152	GATAGATAGACAGACTGATTGATTGATTGATTGATTGATTGACAGGTAGACAGACAGAAGGCAGATAGATAGATAGATAGATAGACAGACTGATTGATTGATTGATTGACAGGCAGGTAGACAGACAGAAGGCAGATAGATAGATAGACAGA

236	193	ACGATAAAGAAGTTTCTAAGGGAGGAGAGGAATCACAGGACTGGAGACACACCTGTCTCTGGTCCAGGTAACCTTATTTCTGATGCCATAAATTACGCTTTTAATAAGAATAACTCCAGTATACTGCTGCAAGCCATTTGGACCATCCCCCTTTTTACTAACTAGAGATTTTCCACACAGCACAGCTCCACCA
237	235	ATGCTTAGCAAAAGAGTATTTTGACATTTTTGTGTGCTAGTTGACCGTTTAGGTGTGAATCCGATGCCCATAGTATACTTTGCTTATGCACATTCTGCTCCGTTTCCGAGCTCACACACAAAAATCTTGTTTTCATAATCGTCGGAAGCCGAAATCAAAAATCAATTACGATTAATCGGACAGCCCTAATCAGCACAACAATGATTGTGTGGTGTTCTTGTCCAATCAATTGCTG
238	235	CAGCAATTGATTGGACAAGAACACCACACAATCATTGTTGTGCTGATTAGGGCTGTCCGATTAATCGTAATTGATTTTTGATTTCGGCTTCCGACGATTATGAAAACAAGATTTTTGTGTGTGAGCTCGGAAACGGAGCAGAATGTGCATAAGCAAAGTATACTATGGGCATCGGATTCACACCTAAACGGTCAACTAGCACACAAAAATGTCAAAATACTCTTTTGCTAAGCAT
239	136	ATCCGAACTCTAACTCTTACCCTGACCCTAACCCTAACCCTAAACATAAGCCTAAATTAAACCCTAAACATAACCCTACCCTAACCCTACCACAACACTAACCGTAACCGTAACCCTAAACCTAACCCCAACCCTA
241	129	AGCGTGACCTGCGAAGCAAAGAGACCCCAGAGAAGACGAACGTTGTGAAATCACCTGGACAGGCGGACCTGCATCCAGAGGGAGAAGACAGCAGAGAAACCGAAGAGCTTCCATACATGACAACAACCG
242	116	CCTAAATAGCCAACCCAAGACTGACGTGGGTCATAGTCCACCTGTCGTATCTGTTTTTACACTAAATATCTGTAAACATATGGTATTTCAATCAGGACCACTTCCGTCAAGTATTA
243	271	GCTCGCTCGGCTCCACCAGAGACCATCATCCTTATGGCTCCAACAGGCTCCCTTGTCCCTCCGGCTACGCCTTGGTCAGACGTCACTCTGCCTGTGCCATGGACTTCTGAGCTGTCTGCTGCACTCCGTCTCTCCACCCCTTCAGGTCTGTCTGGTTCCGCCTCGGATCTGACAGTCACTGCGGCTTCGCCTCAGCCTCCAGGACCTTCGGTGTTGACCAGTCGTATCTGCTGTCCGTCTGCACATAGGGCTCCACCTCCATCAGTTTCAT
244	209	AACTAACAATCCAACTACAAGTGTCAAACCATCTGAGACAACTAACAATCCAACTACCAGTTTCAAACCATCTGTGACAACTAACAATCTACTCACCAGTTTCAAACCAACTGTGACTATAAACAATCCAACTACCAGTTGGATTGTTAGTTGAACTGGTAGTTGGATTGTTTATAGTCACAGTTGGTTTGAAACTGGTGAGTAGATTG
245	126	CCTGAATCATAAGCATCGAATATTGAATATTGCATTCTCATATTTAACAGTTCAGTAGCGGTGAGATGTGATACCTGTGGGGGTTTGTGGAGGTCTCTGAGCCTCTTGTCCTTTGCCTGCGTTAAA
246	206	CTCTGATGTTGACCCTGCCAACATAGCGTGCTTCAAAGCTGCATTCACAAAGGACCTAAACAGGCAAAAGGAGAACTCAAACCTATGGTGGTTAAAGGTGGCAACCGCTCTAGATCCTAAGTTCAAAAACCTTAGGTGCTTGCCCAGGGCAGAGATGGGAGAGGTATGGCAAAAGCTGAGTGAGATGCTGAAGGATAGAGAACCTG
247	220	CGATAGGTAACATTGGTTTTGCTTTGTTTCACAGAATATATGCTGTTAAAACACACAGTAATACCAAAACACACTTCACTTATTGATGACATCCTACCCTCGGCCGCCTGGCACATCAGATCCTGCCATAATCATTGCTGGGGTTGCGTACGTATGTGTGGGGTGGCGCTATCAAAATAGGGACGAGACCCTTTTGGGGTAGGGGCGTGTTTGTTTTGGT
248	148	TCCGCGAAAAGGCGCTGTGGTTGGAAGGCGCTGTGGTTGGAAGGCGCTGTGGTTGGAAGGCGCTGTGGTTGGAAGGCGCTGTGGTTGGAAGGCGCTGTGGCGGAGAGGCTCAGTTGCCCTGGCCGACGATGTACAGATATCCGGCGAG
249	85	CTGAGGCCGGGGAATGGCGTGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATGGCGCCACTGCACTCCAGCCTGGGCGACAG
250	92	ACAGTCACATAAAATCTAAAGTTAATAACTGAAGGGTCAAGATCCCAACTAAAGCAAACACTGAACACTATAATGGTGATACACAAATTGTG
251	172	CTTATTCTTGTCTTGGATAACCAGGCGCTTTCGGGCAAAGTAATCAGTTTTTCCCTCTCTCCTTCTCCTGAACTTCACCTGGTACCTCTTGAAGTAGGCCTTGTTCTTCACTACTTTAACGAATCCCATTTTGTAGACTTTTTAGTCTCCTAAGCCGGAGCCTGGAGTCAAA
252	128	GGCTACGTAGCTACCGCCGGCTACGGCTACGCCATGCAGCAGCCGCTGGCCACCGCCGCTCCAGGGACCGCAGCCGCAGCCGCCGCCTTCGGTCAGTACCAGCCCCAGCAGCTGCAAGCCGAACGGAT
253	103	GGAGGGTCAGCATACGGAGGGTCAGCATACGGAAGGTCAGCATACGGAAGGTCAGCATACGGAGGGTCAGCATACGGAGGGTCAGCATACGGAAGGTCAGCAT
254	144	CCCCCTACACCCAGATTTTTATTAGACATTAGTCTGATTCCTATTCCACAGGCGTTTCCCCTTCCACTTTAACACTCAGCCACACCAGAGCCAATTTTAAAAAGGCTCTTCTGCCTAGCCTGGAAAAAAAAAAAAAAACCCTTGG
255	138	CTAGCAACTAGCCATAATGCCACTCATAGCTAGCAACTAGCCATAATGCCACTCATAGCTAGCAACTAGCCATAATTCCACTCATAGCTAGCAACTAGCCATAATGCCA

		CTCATAGCTAGCAACTAGCCATAATGCCA
256	148	GTCGAGAATGCCGAGCTCGCTCTCTGCCTCGTCGACGCAGAAGACAAAGTAAAGGGTCGCGTAGTGGCGGTAGATGATTCGTAAAGGTTCTTGCGCGGTGTCGGGGTCCAGATCAAGTGCGCGACCCTTGCCCTTGAGCTGCTGCTGG
257	93	GACAAGATGCACAGCATAGCCATATTCTAGGTAGACATGAGTTTTTTCTTTAACTATATATATATATATATATATATAGTCGTAGTGGTCTTG
258	271	TGATGTTGATCTGAGAAACCCTGTTTTATTATAAAAAATAAAGCATTTTTATATATAGGGTGTGTTCACCCACCCACCACCATCCAACAAGCCATCTGATCACACATCCATTCCTTTTACTATTCATCCATCCATCCATCCATCCATCCATCTCAAATACACCTGAACAAAGCCAACAATAAGAGAGTTCAACAAGTGTGCAGGCGTCAGTCTTTAGCTCCTTCTTCCTCTCACTGTATGCATGTTAAAGAGCGTCAGTATATCATCTTAA
259	148	TTTGTCGCACAGCAGGTAGATTTCTTCACCACCGGTAACGCAACCAGCGGTACGGTCCATACGAACGATTTTCAGGTTAGAAGCGTTCGGAGCTTTAGAGTCGTAGATAGCGTCAGAAACAACCGGTTCCAGACGACGGGTGAAAGAA
260	82	GTGGTGATGGGGGTGGTGGTGGTGATGGTGGTGGAGGTGTGGGTGATGTGTGGGCAGAGGTGGAGACACGGGAGACAAAGGT
261	177	CTTGTAATTCTTGGCAAACATGAACCCACCCTCATTGATTCCTGGTGTGTTTGTGTTGTAGTTATTCAACTTTTACACCACCACCCTGCAGAATTCTTTGTGTAGACATCGTTTCATCTGTGATTTGTCCTTAAAAACAATTCTACATTTTGAAAGCTGAATACAGATGTTAAAATA
262	86	TCTTAGCGTAGCCTGAACAAAACCCTGAAAATTAATTCTTTCCTTCACCCCCCATCATCCCGTCACCTCCTCCCTCTTTATCTTCT
263	137	CATGGATTATCGGATTGTGCAAATTTAGCACCACAACCACCACAACCACCGCAACCTCCACAACCACTACTACTTATAGAGCCTCCACAGCCTCCACAACCACCGCAACCTCCACATCCACCACAGCCACCACTTGT
264	122	CATCTTCTTGCTTTTAGGCTAAAACAGAAGCACCCACAGCCTCCACAAACCTAGAGGAAGCAATTTTTTAGGTTGAATGTTGACAGATGCATATTGAACTTCATCTTCCTGCAGCACATTCT
265	205	CCTGCTCAATCTTTCCGCCTCCATACGGATGTCAATCATCTGCTGGCGGTCCTTGAAGCTTCTCGATAAAAAACGTCCCTCTGGGGTGTAGAGGGAGCGAGTAGGAGTGAGGCACCGGATGTTGCGGACCTCCTCTGAGTCACTTTCGCCACCGACGGGACCCTCACGTTTGCGATGCGGCGGGAGCCGAGAGCAGTCATCGTCG
266	248	AATATGCAATCAGCTTGATGGTTCAGAACATAAAACACTCTAAAAGGAATGTAAAAGTGTTTATTCATTCCCCATTTCAAGATCGCTCATGCTTGTCTTTATCTTCCTGTAGGAGCTAGAAATGATGAGAAGGAGCGGCAAGTGGAAGGTATGGAGAGGAACAAGTCTCTCCAAGATTACAACGAGTATTACGATTACTCAATGCAAATGCACCGCAAACTGTTCTCTGGCCAGAAAATTGGCAGAAT
267	185	TTAAATTAGAAAACAACCAAAAATTCTGTTTAAAATTAAGACTTGGTTGTTAAATGTTAAATGGACCATGTCCGTAGTTAATGTACTACACCTGTTGTTAGGACACTTCTGTAAACTTGAGGGGAATCAACTTCATACAGCTACTGTTACACCATTTGAATAAAAATAACAGCAAAAATGACTGA
268	134	CTGCGATCAATCAGAGAGCCGCGAGTCAGTGGAATTACTGGATCGATCACTGCGATCAGTGTTTATCAATCACCCACAACAGGCCATATGGACCGGCCGGTACGTCACAAAGTCCCCGGATTTGTTGCAGGAGA
269	148	CGTCTGCAGACGGGCTACAGAAGACGGGTAGTATAACCGGACCGAACCCAGGTCCCCAGCTCTCATTACACCCGGATCAACGCATCGCTTACAGTCTGTCGGTCGGAACGAGCTCGGTTATGGGATATTAACTCAAAAGAGGATCGTG
270	180	GATGTGTTAAAAACGGACGCATGCAAGAAGACTCGTCTGGAGTTTTACATCATCAATCTGTGACACGTTGGGTGACTGTGAGCTGCATTTACATGTGGTGACTAGAAAACCAAATTTCTAGTGTTAACAGGGTGTGAGAAAAAAAAGTGAAGCGGGACTCGATTGAAATGTGAAAAAAATAG
271	110	TAAACCCCCCCAAGGGGGTTTGACTGGGGGTTTGACTGGGGATTGACTGAGGGTTTGTCTGAGGGTTGACCGAGGGTCTGACTGGTGTTTGACTAAGGGTTGACTGAGGG
272	108	ATCACCCTCACTCTCCTGTCCTCTGTGCTGTCCCTGTCTTAGGTCACGTGCCTTCAGGACTTCTTTGGGGACGATGACATCTTCCTCGCATGTGGCTCTGAGAGGTTT
273	237	CAGCAAAATAAAGTATGGCATCGCAATAAACACAATCAAACCGGACAAAATATAACCATCTAGCCATTAAAAACACGAAAAATCAAGGAAAAAACGTCAGACAGGCAAATCTCGTGGTCTCGCGAATCCAGCAGCTTCGCTTCTGAAATGCTTCTGGTGTGAGTTGACACCGCACAGAGGACGGAACAAAACCTTGTCGGAGCCGTAACACACCGCAGACGCATGCAGTCTAAACAA
274	184	GATTTGTTAACCGTCGTAATGCATTGGTTTTGTTATTCTCGTAGTCATCATCACCATCATCATCATCATCATCATCATTACTGTAACACGTGTTCCAACTCTTCCAAATGTAAGGCACATAAACTCTTCCAAGGCCACGTGACTCCAGCAGCCTTGCCTTTATCATTATCATCATCATTATCGT

275	167	AGTTTTACGTTAAAGATGATAGATGTGTTATTAAAGTCTCAATTTTTTCCTTCAGGTCGCTAATCTGATCTTGCAAAGCTTCTGTCTGTGTTTACTGGATGCTTCCAGAATCAGCCCTCCTAACAGGTACTGAAAACCTCTAAATGTAACTTTGGTTTCCTGAGATG
276	113	CACAGAACTTATTTTACTCTTAAATTCAAGGAACTCATTAAACCCCATTAAGAAAATCTTGAGGGAACCAGGGGTGGGGTTTTGACATCAGACTTCCTCCACTCTATTTCTTA
277	65	AGTACTAACCAGGCCCGACTCTGCTTAGCTTCCGAGATCAGACGAGATCGGGCGTGCTCAGAGTG
278	117	TAGATGGTTCGATTAGTCTTTCGCCCCTATACCCAGGTCGGACGACCGATTTGCACGTCAGGACCGCTACGGACCTCCACCAGAGTTTCCTCTGGCTTCGCCCTGCCCAGGCATAGT
279	113	CTGCTGTCTATATCAACCAACACCTTTTCTGGGGTCTGATGAGCGTCGGCATCGGGCGCCTTAACCCGGCGTTCGGTTCATCCCGCAGCGCCAGTTCTGCTTACCAAAAGTGG
280	100	CATGCGGACCAAGAAATAATCAAAATCTCTTCAGGTGAAAATGTTAACACAGAGTATGTATTGCTACATGATCATGAGGGGGTCTCAATGTTTCCTCCAA
281	73	TGTAATAAAATACCTCACTCACCGGTCAAATAATAAAAACTCCATCATCACCGCATCACTTTTACAACATTTT

Non Actinopterygii Sequences

282	149	ATTCGGCAAATTGGATCCGTAACTTCGGGATAAGGATTGGCTCTAAGGGTTGGGTACATCGGGCCCTGGTTGGAAGCCGCTGATGCTGGCTTGGACTGCTGCGTGGGAACATGCGGTGGACCGAGCCGGCGTCGGCGTGTGGACGGCCA
283	150	TTGTGCAACAGCGATAAAGTCGGTAGGAACAAGACGGGTGCCCTGATGCAGTAGCTGGTAGAAAGCGTGCTGACCATTAGTTCCAGGTTCACCCCAATAGATTTCTCCGGTGGAGGTAGTAACGGGGCTGCCATCGACACGAACAGACTT
284	121	ACGCAAGGCCTGAATTAGTGACACCATTTTGTTGTGCAGTCCAGGAACAGGTCGTCAAGACAAGGAAGGTTTGTCTGTGACGAAGCAGAAGACGGAGGACACACGATGACGTCACACCTTC
285	150	TTGTTGGTAAAGGCTCTTTTTTTAAATAAGTCCCAAATGGGTTTGTGATCGGTATGCACATTAACCTTATGACCTAAAAGTATGTGTTTAAATTTCTTTAGACCCCCATATAAAGCTAAACACTCTCTTTCTGTTGTACTATAGTTTCTT
286	100	GTGAAAAAGCATTCAGTTTAGAGATATCAGCTACTTAGGATATGAATAGTCAAACAGAAGAACGGATTTCAGTACCTGGTATTTCTAATATTCATGACTT
287	150	ACGCAAGGCCTGAATTAGTGACACCATTTTGTTGTGCAGTCCAGGAACAGGTCGTCAAGACAAGGAAGGTTTGTCTGTGACGAAGCAGAAGACGGAGGACACACGATGACGTCACACCTACAAAACGAAAAGCCCCCCGCAGGACCACAC
288	150	ATACAACACAATTTAGATATATTTTTTTATTTTTATGATCTTTATATGTATGTAAAAATAACTATGCACAAACATTCTACATGCATGCATCTACTTATCAATATATATATATATATATATATATATATATATATATATATAATATTTTAGATAT
289	109	CAGCGGCTGCCCGGGCGGCGCAAGGAGGCCTGCCGACGCAGCGGCCTCGTCCAACTCCTTCCCGTTGTATCCCACGGCTCCTTCGACCTCGGACTCCGCCCCGGCGCCC