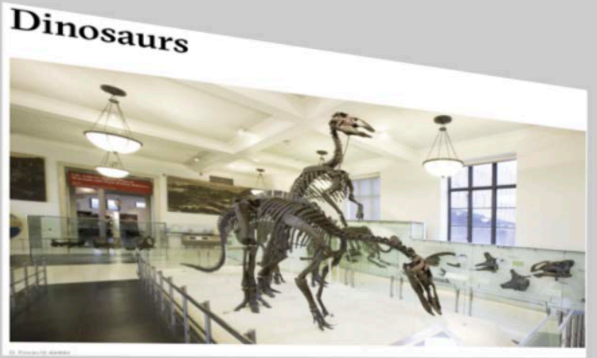
The 2020-21 AP Biology Virtual Natural History Gallery



Dinosaurs are a group of reptiles that first appeared about 245 mya. They dominated the planet until an extinction event wiped out a large swath of life on Earth about 66 mya. But, we now know that the Age of Dinosaurs never ended. Instead, a group of dinosaurs gave rise to a new, small airborne form: BIRDS

VISIT THE ROOMS TO SEE THE EXHIBITS

CLICK ON EACH EXHIBIT FOR FURTHER INFORMATION

TO ROOM 6

TO ROOM 5

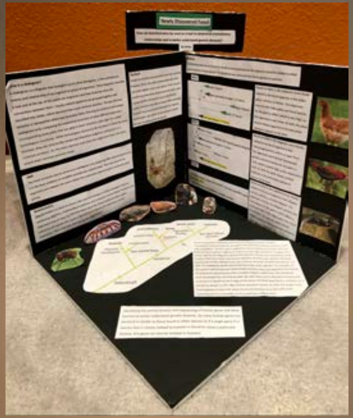
TO ROOM 2

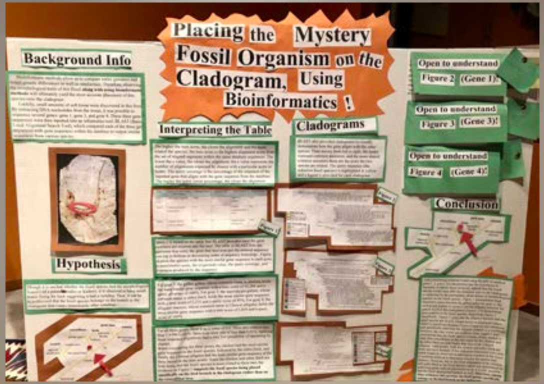
TO ROOM 4

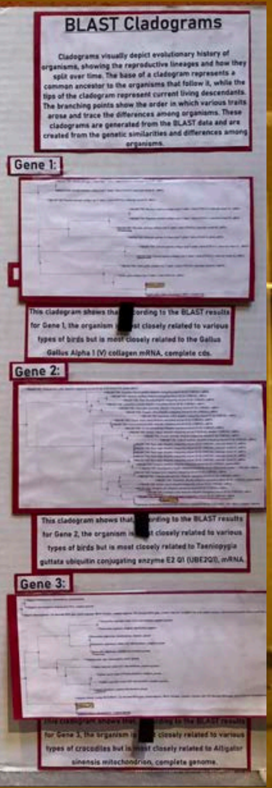
TO ROOM 3

TO ROOM 1


WELCOME
Bioinformatics for Evolution
INTRODUCTION TO SPECIES
Ants
THE FOSSIL & MORPHOLOGY
BLAST DATA
DNA NUCLEOTIDE SEQUENCES
BIOINFORMATICS
THE CURRENT CLADOGRAM
Gene 1
Gene 2
Gene 3
FINAL PLACEMENT
THANKS!
Chicken
Gallus Gallus
Zebra Finch
Taeniopygia Guttata
Chinese Alligator
Alligator Sinensis
Missing Mystery Fossil
How can bioinformatics be used as a tool to determine evolutionary relationships and to better understand genetic diseases?
great apes




Placing the Mystery Fossil Organism on the Cladogram, Using Bioinformatics !
Background Info
Hypothesis
Interpreting the Table
Cladograms
Open to understand Figure 2 (Gene 1)!
Open to understand Figure 3 (Gene 3)!
Open to understand Figure 4 (Gene 4)!
Conclusion



BLAST Cladograms

Cladograms visually depict evolutionary history of organisms, showing the reproductive lineages and how they split over time. The base of a cladogram represents a common ancestor to the organisms that follow it, while the tips of the cladogram represent current living descendants. The branching points show the order in which various traits arose and trace the differences among organisms. These cladograms are generated from the BLAST data and are created from the genetic similarities and differences among organisms.

Gene 1:

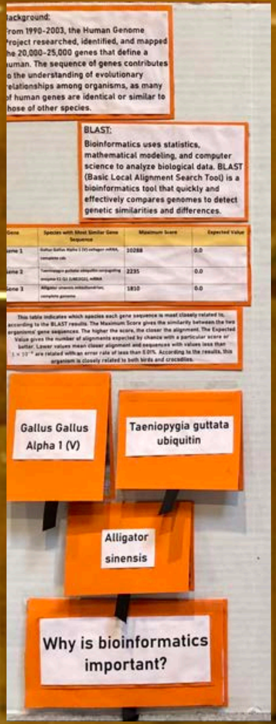
This cladogram shows that according to the BLAST results for Gene 1, the organism is most closely related to various types of birds but is most closely related to the Gallus Gallus Alpha 1 (V) collagen mRNA, complete cds.

Gene 2:

This cladogram shows that according to the BLAST results for Gene 2, the organism is most closely related to various types of birds but is most closely related to Taeniopygia guttata ubiquitin conjugating enzyme E2 Q1 (UBE2Q1), mRNA.

Gene 3:

This cladogram shows that according to the BLAST results for Gene 3, the organism is most closely related to various types of crocodiles but is most closely related to Alligator sinensis mitochondrion, complete genome.



ackground:

rom 1990-2003, the Human Genome roject researched, identified, and mapped he 20,000-25,000 genes that define a uman. The sequence of genes contributes o the understanding of evolutionary elationships among organisms, as many f human genes are identical or similar to hose of other species.

BLAST:

Bioinformatics uses statistics, mathematical modeling, and computer science to analyze biological data. BLAST (Basic Local Alignment Search Tool) is a bioinformatics tool that quickly and effectively compares genomes to detect genetic similarities and differences.

ene	Species with Most Similar Gene Sequence	Maximum Score	Expected Value
ene 1	Gallus Gallus Alpha 1 (V) collagen mRNA, complete cds	10288	0.0
ene 2	Taeniopygia guttata ubiquitin conjugating enzyme E2 Q1 (UBE2Q1), mRNA	2235	0.0
ene 3	Alligator sinensis mitochondrion, complete genome	1810	0.0

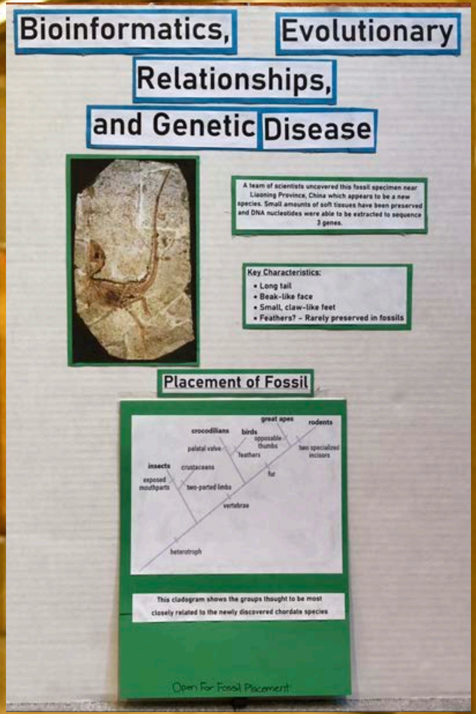
This table indicates which species each gene sequence is most closely related to, according to the BLAST results. The Maximum Score gives the similarity between the two organisms' gene sequences. The higher the score, the closer the alignment. The Expected Value gives the number of alignments expected by chance with a particular score or better. Lower values mean closer alignment and sequences with values less than 1×10^{-4} are related with an error rate of less than 0.01%. According to the results, this organism is closely related to both birds and crocodiles.

Gallus Gallus Alpha 1 (V)

Taeniopygia guttata ubiquitin

Alligator sinensis

Why is bioinformatics important?



Bioinformatics, Evolutionary Relationships, and Genetic Disease

A team of scientists uncovered this fossil specimen near Liaoning Province, China which appears to be a new species. Small amounts of soft tissues have been preserved and DNA nucleotides were able to be extracted to sequence 3 genes.

Key Characteristics:

- Long tail
- Beak-like face
- Small, claw-like feet
- Feathers? – Rarely preserved in fossils

Placement of Fossil

This cladogram shows the groups thought to be most closely related to the newly discovered chordate species

Open For Fossil Placement


START
THANK YOU
0:00 / 12:26
BLAST
Gene 1
Gene 3
Gene 4
EVOLUTIONARY LINEAGE OF A MYSTERY FOSSIL
Cladogram for Genes
The specimen is closely related to birds and crocodilians
Exploring Bioinformatics and Evolutionary Relationships
I hope you learned more about bioinformatics from the exhibit

Why Are Bioinformatics Used:

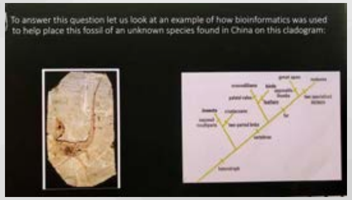
Using characteristics found by examining a fossil [illegible]
place a new species on a cladogram. [illegible]
modeling, statistics, and computer science in order to [illegible]
bioinformatic tool that is used to examine DNA [illegible]
BLAST, which stands for Basic Local Alignment Search Tool [illegible]
input a gene sequence and the program searches [illegible]
to find similarities.
While examining this fossil scientists found that there were [illegible]
preserved in the rock. They extracted DNA [illegible]
three sperate DNA sequences which were named : Gene 1, Gene 2, and Gene 3. [illegible]
sequences were then input into BLAST and scientists were able to [illegible]
fossil most closely matched.

5.)

After each gene was input into BLAST the program [illegible]
gene tests was to another known DNA sequence.

- Maximum score in column three → the highest alignment [illegible]
sequence.
- If the max score is higher that means the DNA sequences [illegible]
- The max values for each gene are [illegible]
sequences are closely related.
- Expected value in column four → the number of alignments [illegible]
- In the table above the e-value for each gene [illegible]

How can bioinformatics be used as a tool to determine evolutionary relationships and to better understand genetic diseases?



To answer this question let us look at an example of how bioinformatics was used to help place this fossil of an unknown species found in China on this cladogram:

Results gathered from the table:

The first gene most closely aligns with gallus gallus alpha, or a Red Junglefowl, which is a type of bird.

The second gene most closely aligns with Taeniopygia guttata ubiquitin, or a Zebra finch.

The third gene most closely aligns with Alligator sinensis mitochondrion, or a Chinese alligator.

BIOINFORMATICS FOR EVOLUTION

YASH MEHTA

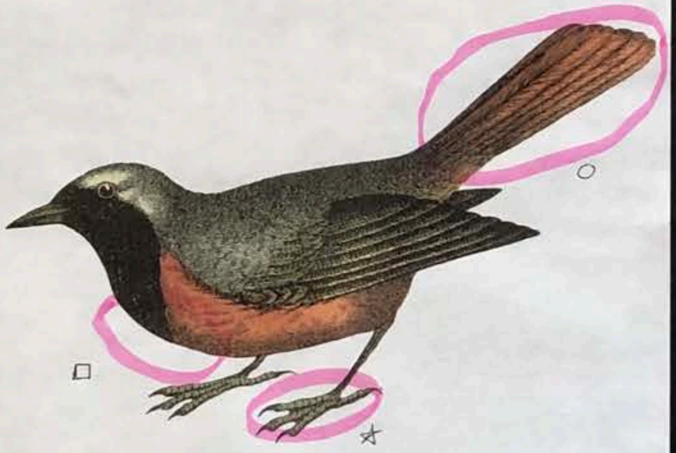


Figure 4. Drawing of a Bird.

Pink circles point to similar morphological traits between the fossil and a generic bird.

THE FOSSIL & MORPHOLOGY

On the left is Figure 3. Our team of scientists uncovered the fossil specimen in Figure 3 near the Liaoning Province in China. What observations can you make about it? Our scientists noticed a long tail, two hind legs with long feet, and a ribcage. Additionally, the head was on top of a long neck on the end of a backbone. If you noticed anything else, make sure you remember it, for it could prove useful later on. This fossil is evidence of evolution because it provides physical support for Darwin's idea of descent with modification. Organisms descend from one common ancestor, and each new generation evolves and has modified genetic and morphological traits.

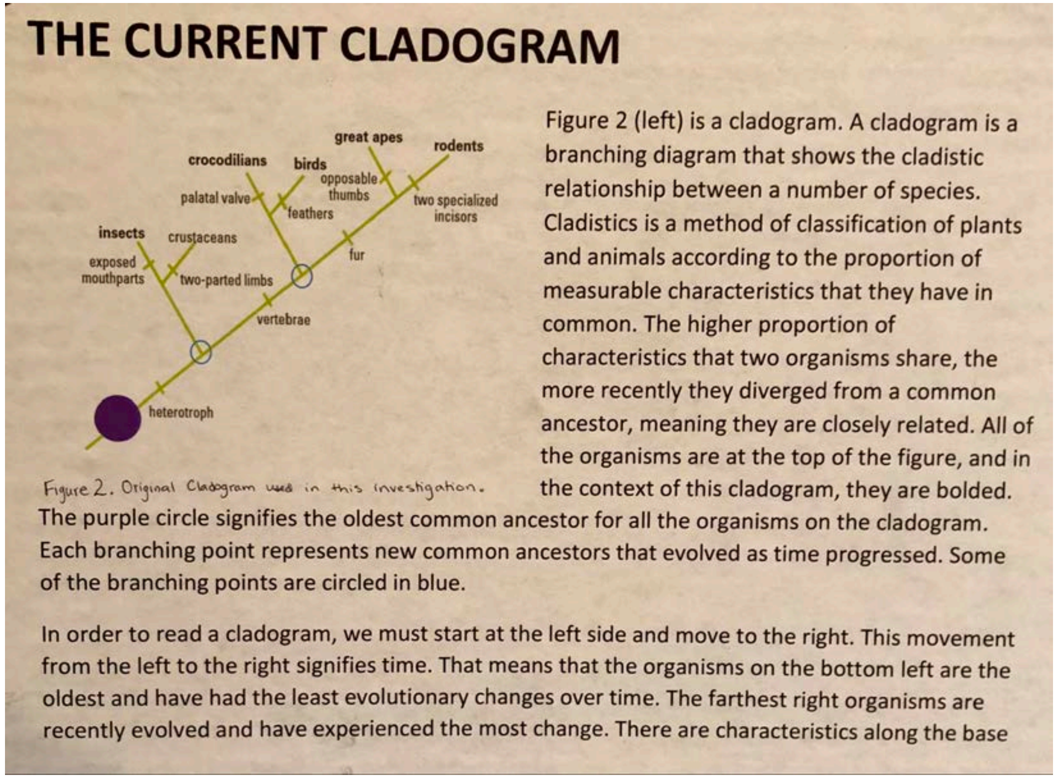
Now that we've made some observations, where would you place the organism in the fossil on the cladogram shown earlier?

Similarities in the bone structure between birds and the fossil provide further evidence of evolution. If we take a look at the cladogram mentioned earlier and try to find an animal that seems similar, it is evident that some morphological traits in this fossil place the organism on the birds branch. It more closely resembles birds than animals like apes, insects, or crustaceans. On the left is Figure 4. Shown is a generic bird that has been pulled from our database, and we can see some similarities in the tail, body structure, legs, and feet. The fossil was likely an older bird, or a recent ancestor of birds, and shows many morphological similarities. Morphological similarities are a great way to add evidence to our argument that this fossil belongs on the bird branch, however, we don't have enough evidence just yet to justify the initial claim. That's where computational biology and the BLAST program come into play.

CLICK FOR NEXT SLIDE

BIOINFORMATICS FOR EVOLUTION

YASH MEHTA

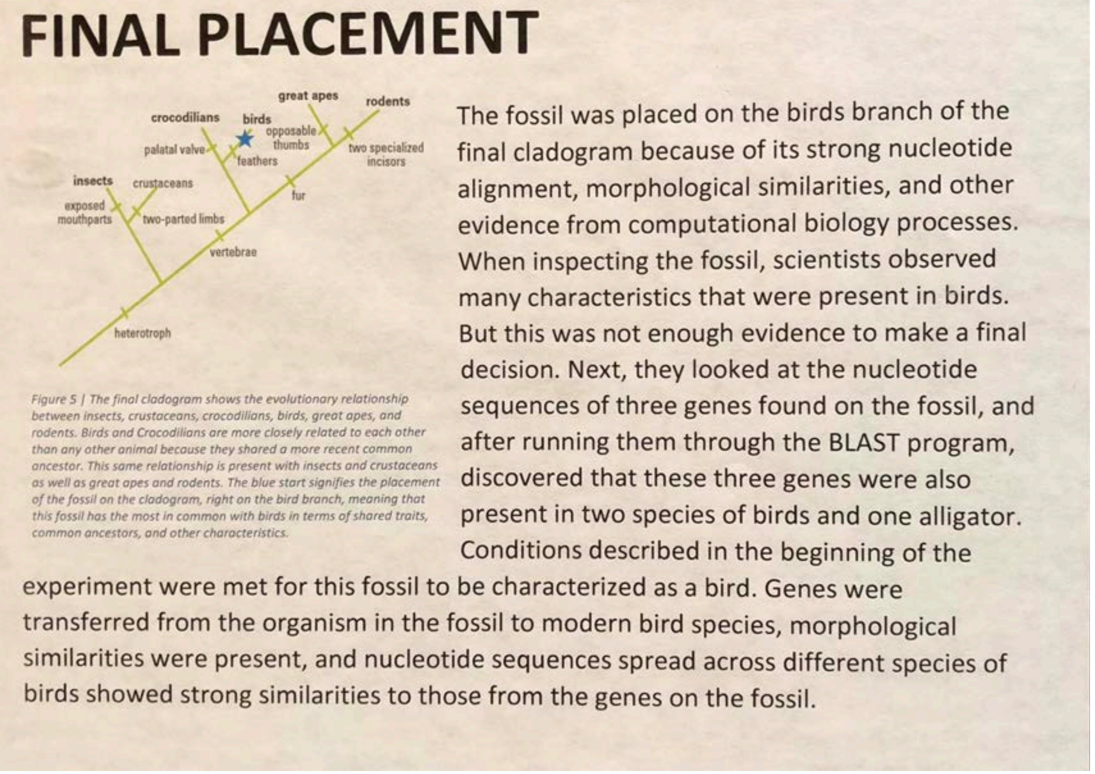


THE CURRENT CLADOGRAM

Figure 2. Original Cladogram used in this investigation.

Figure 2 (left) is a cladogram. A cladogram is a branching diagram that shows the cladistic relationship between a number of species. Cladistics is a method of classification of plants and animals according to the proportion of measurable characteristics that they have in common. The higher proportion of characteristics that two organisms share, the more recently they diverged from a common ancestor, meaning they are closely related. All of the organisms are at the top of the figure, and in the context of this cladogram, they are bolded. The purple circle signifies the oldest common ancestor for all the organisms on the cladogram. Each branching point represents new common ancestors that evolved as time progressed. Some of the branching points are circled in blue.

In order to read a cladogram, we must start at the left side and move to the right. This movement from the left to the right signifies time. That means that the organisms on the bottom left are the oldest and have had the least evolutionary changes over time. The farthest right organisms are recently evolved and have experienced the most change. There are characteristics along the base



FINAL PLACEMENT

Figure 5 | The final cladogram shows the evolutionary relationship between insects, crustaceans, crocodilians, birds, great apes, and rodents. Birds and Crocodilians are more closely related to each other than any other animal because they shared a more recent common ancestor. This same relationship is present with insects and crustaceans as well as great apes and rodents. The blue start signifies the placement of the fossil on the cladogram, right on the bird branch, meaning that this fossil has the most in common with birds in terms of shared traits, common ancestors, and other characteristics.

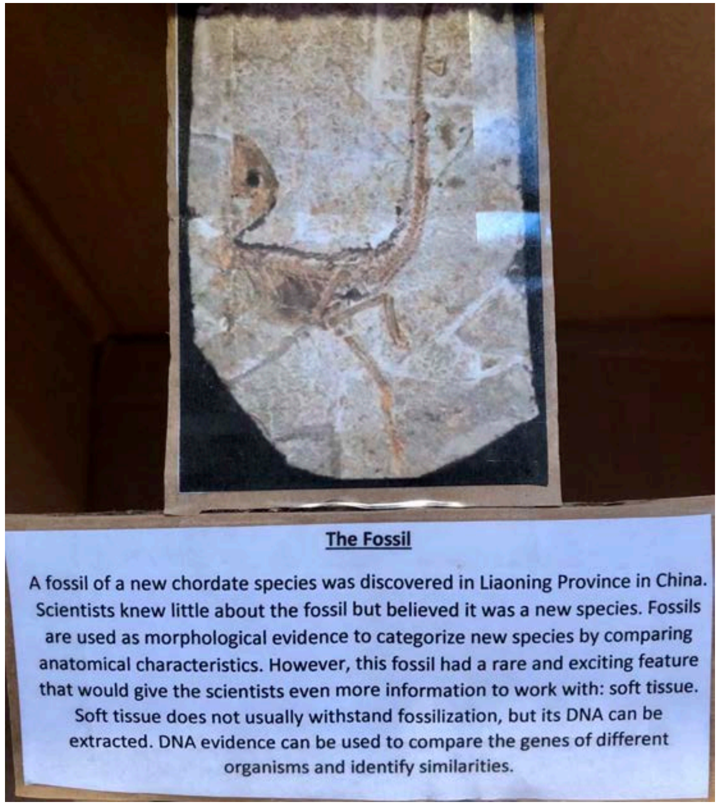
The fossil was placed on the birds branch of the final cladogram because of its strong nucleotide alignment, morphological similarities, and other evidence from computational biology processes. When inspecting the fossil, scientists observed many characteristics that were present in birds. But this was not enough evidence to make a final decision. Next, they looked at the nucleotide sequences of three genes found on the fossil, and after running them through the BLAST program, discovered that these three genes were also present in two species of birds and one alligator. Conditions described in the beginning of the experiment were met for this fossil to be characterized as a bird. Genes were transferred from the organism in the fossil to modern bird species, morphological similarities were present, and nucleotide sequences spread across different species of birds showed strong similarities to those from the genes on the fossil.

BACK TO LOBBY

BACK TO ROOM 1

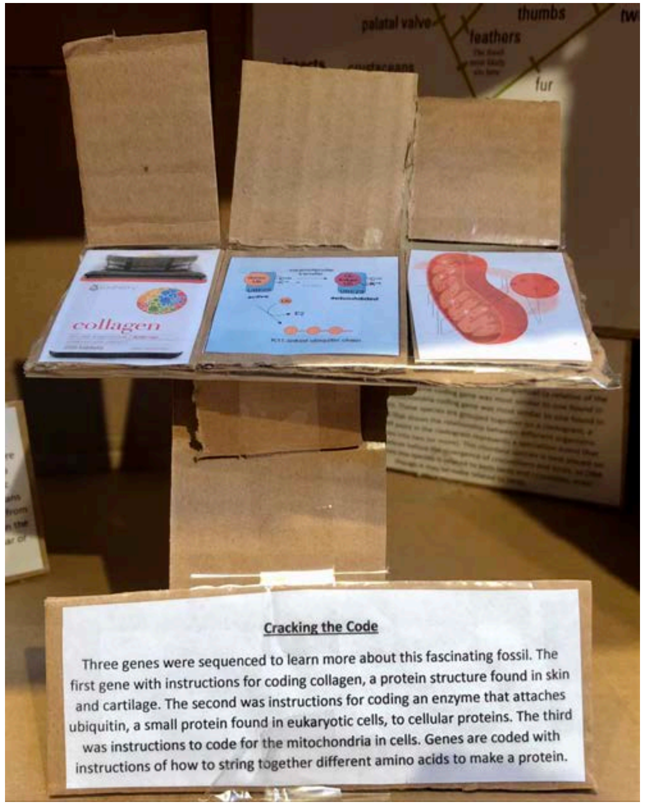
MISSING MYSTERY FOSSIL

ANGELA LANE



The Fossil

A fossil of a new chordate species was discovered in Liaoning Province in China. Scientists knew little about the fossil but believed it was a new species. Fossils are used as morphological evidence to categorize new species by comparing anatomical characteristics. However, this fossil had a rare and exciting feature that would give the scientists even more information to work with: soft tissue. Soft tissue does not usually withstand fossilization, but its DNA can be extracted. DNA evidence can be used to compare the genes of different organisms and identify similarities.



Cracking the Code

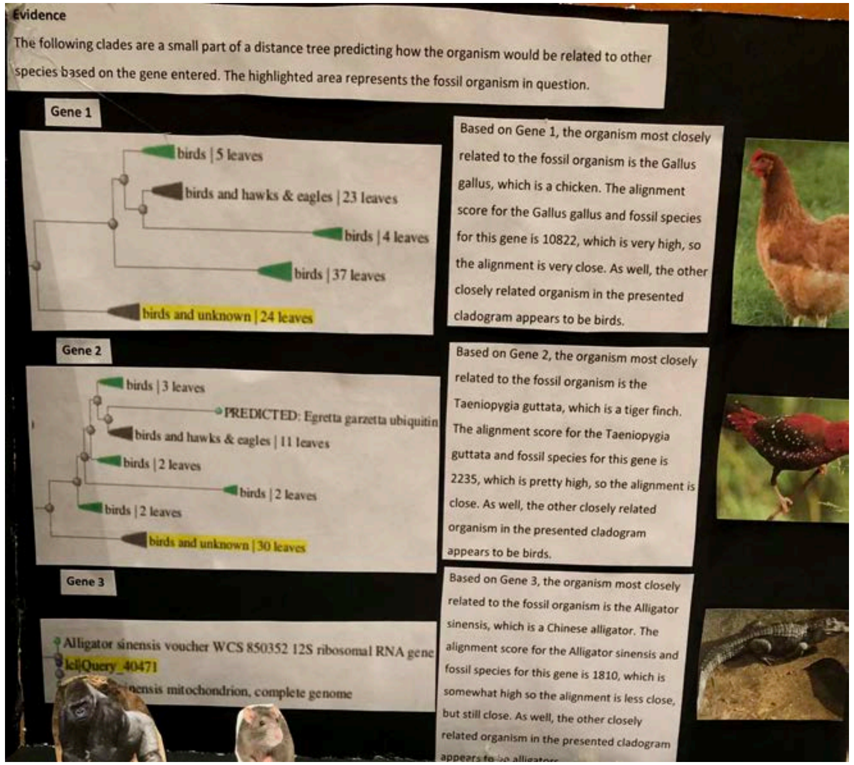
Three genes were sequenced to learn more about this fascinating fossil. The first gene with instructions for coding collagen, a protein structure found in skin and cartilage. The second was instructions for coding an enzyme that attaches ubiquitin, a small protein found in eukaryotic cells, to cellular proteins. The third was instructions to code for the mitochondria in cells. Genes are coded with instructions of how to string together different amino acids to make a protein.

BACK TO LOBBY

BACK TO ROOM 1

NEWLY DISCOVERED FOSSIL

JENNA LACY



Evidence

The following clades are a small part of a distance tree predicting how the organism would be related to other species based on the gene entered. The highlighted area represents the fossil organism in question.

Gene 1

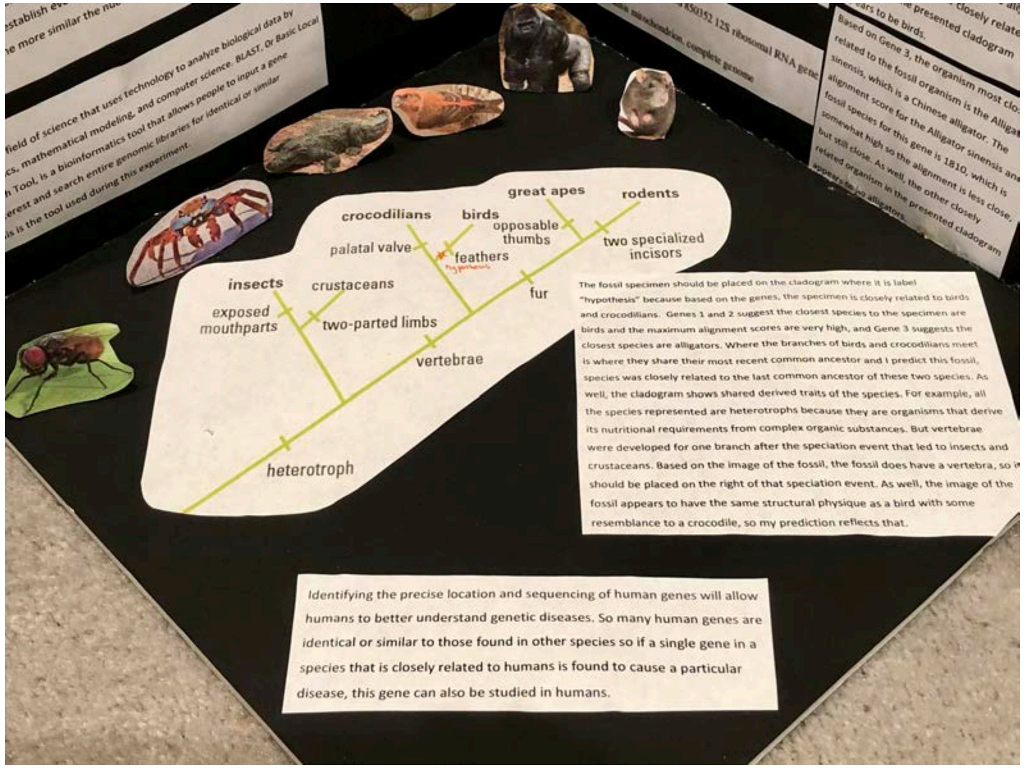
Based on Gene 1, the organism most closely related to the fossil organism is the Gallus gallus, which is a chicken. The alignment score for the Gallus gallus and fossil species for this gene is 10822, which is very high, so the alignment is very close. As well, the other closely related organism in the presented cladogram appears to be birds.

Gene 2

Based on Gene 2, the organism most closely related to the fossil organism is the Taeniopygia guttata, which is a tiger finch. The alignment score for the Taeniopygia guttata and fossil species for this gene is 2235, which is pretty high, so the alignment is close. As well, the other closely related organism in the presented cladogram appears to be birds.

Gene 3

Based on Gene 3, the organism most closely related to the fossil organism is the Alligator sinensis, which is a Chinese alligator. The alignment score for the Alligator sinensis and fossil species for this gene is 1810, which is somewhat high so the alignment is less close, but still close. As well, the other closely related organism in the presented cladogram appears to be alligators.



The fossil specimen should be placed on the cladogram where it is label "hypothesis" because based on the genes, the specimen is closely related to birds and crocodilians. Genes 1 and 2 suggest the closest species to the specimen are birds and the maximum alignment scores are very high, and Gene 3 suggests the closest species are alligators. Where the branches of birds and crocodilians meet is where they share their most recent common ancestor and I predict this fossil, species was closely related to the last common ancestor of these two species. As well, the cladogram shows shared derived traits of the species. For example, all the species represented are heterotrophs because they are organisms that derive its nutritional requirements from complex organic substances. But vertebrae were developed for one branch after the speciation event that led to insects and crustaceans. Based on the image of the fossil, the fossil does have a vertebra, so it should be placed on the right of that speciation event. As well, the image of the fossil appears to have the same structural physique as a bird with some resemblance to a crocodile, so my prediction reflects that.

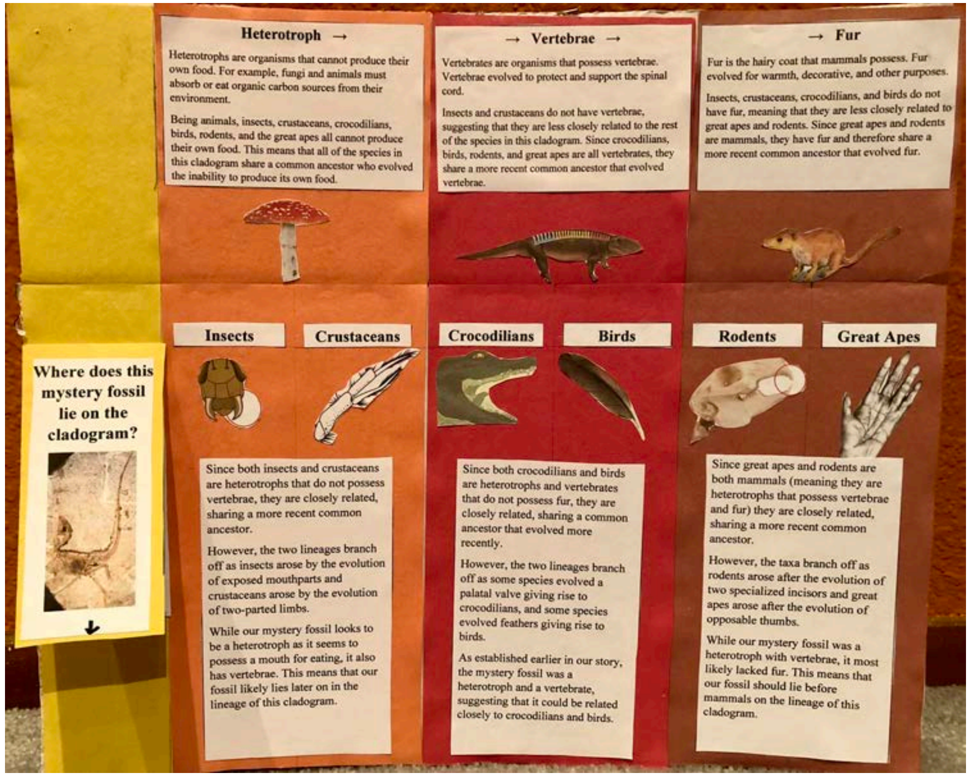
Identifying the precise location and sequencing of human genes will allow humans to better understand genetic diseases. So many human genes are identical or similar to those found in other species so if a single gene in a species that is closely related to humans is found to cause a particular disease, this gene can also be studied in humans.

BACK TO LOBBY

BACK TO ROOM 2

IDENTIFYING AN UNKNOWN FOSSIL

ANYA SHARMA



Where does this mystery fossil lie on the cladogram?

↓

Heterotroph →

Heterotrophs are organisms that cannot produce their own food. For example, fungi and animals must absorb or eat organic carbon sources from their environment.

Being animals, insects, crustaceans, crocodilians, birds, rodents, and the great apes all cannot produce their own food. This means that all of the species in this cladogram share a common ancestor who evolved the inability to produce its own food.

→ Vertebrae →

Vertebrates are organisms that possess vertebrae. Vertebrae evolved to protect and support the spinal cord.

Insects and crustaceans do not have vertebrae, suggesting that they are less closely related to the rest of the species in this cladogram. Since crocodilians, birds, rodents, and great apes are all vertebrates, they share a more recent common ancestor that evolved vertebrae.

→ Fur

Fur is the hairy coat that mammals possess. Fur evolved for warmth, decorative, and other purposes.

Insects, crustaceans, crocodilians, and birds do not have fur, meaning that they are less closely related to great apes and rodents. Since great apes and rodents are mammals, they have fur and therefore share a more recent common ancestor that evolved fur.

Insects

Crustaceans

Since both insects and crustaceans are heterotrophs that do not possess vertebrae, they are closely related, sharing a more recent common ancestor.

However, the two lineages branch off as insects arose by the evolution of exposed mouthparts and crustaceans arose by the evolution of two-parted limbs.

While our mystery fossil looks to be a heterotroph as it seems to possess a mouth for eating, it also has vertebrae. This means that our fossil likely lies later on in the lineage of this cladogram.

Crocodilians

Birds

Since both crocodilians and birds are heterotrophs and vertebrates that do not possess fur, they are closely related, sharing a common ancestor that evolved more recently.

However, the two lineages branch off as some species evolved a palatal valve giving rise to crocodilians, and some species evolved feathers giving rise to birds.

As established earlier in our story, the mystery fossil was a heterotroph and a vertebrate, suggesting that it could be related closely to crocodilians and birds.

Rodents

Great Apes

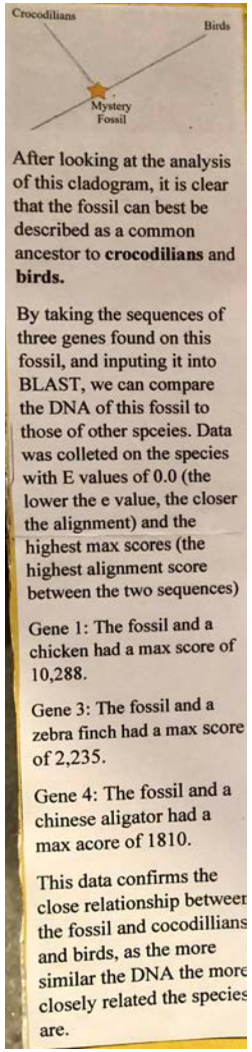
Since great apes and rodents are both mammals (meaning they are heterotrophs that possess vertebrae and fur) they are closely related, sharing a more recent common ancestor.

However, the taxa branch off as rodents arose after the evolution of two specialized incisors and great apes arose after the evolution of opposable thumbs.

While our mystery fossil was a heterotroph with vertebrae, it most likely lacked fur. This means that our fossil should lie before mammals on the lineage of this cladogram.

BACK TO LOBBY

BACK TO ROOM 2



After looking at the analysis of this cladogram, it is clear that the fossil can best be described as a common ancestor to **crocodilians** and **birds.**

By taking the sequences of three genes found on this fossil, and inputing it into BLAST, we can compare the DNA of this fossil to those of other spceies. Data was colleted on the species with E values of 0.0 (the lower the e value, the closer the alignment) and the highest max scores (the highest alignment score between the two sequences)

Gene 1: The fossil and a chicken had a max score of 10,288.

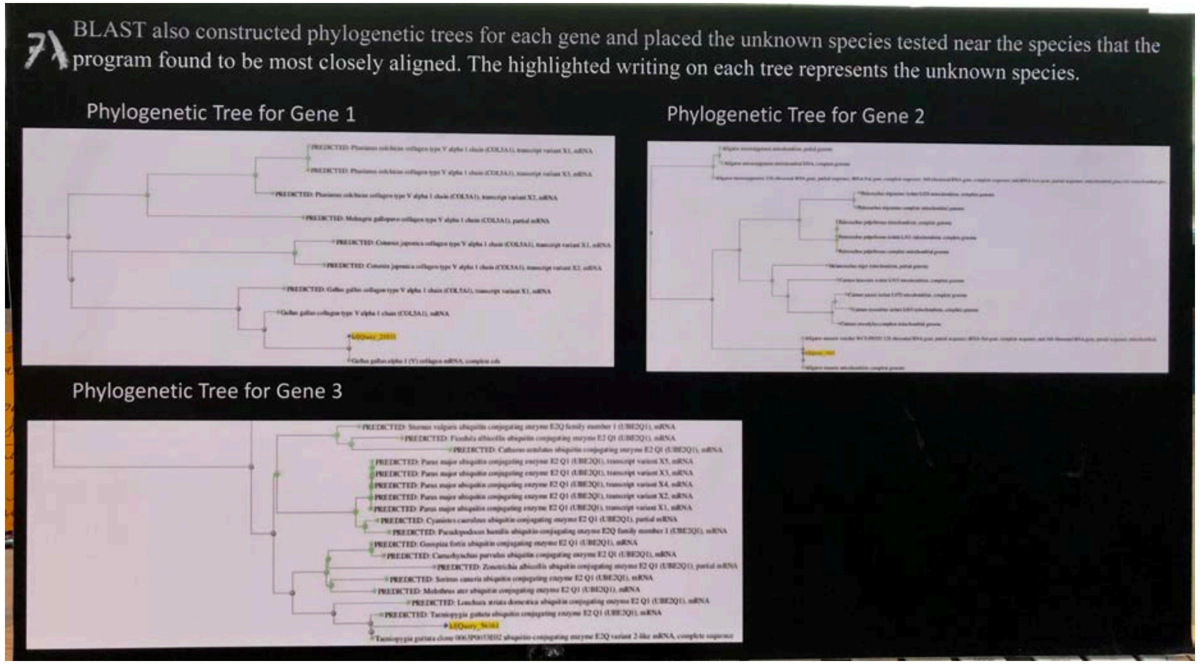
Gene 3: The fossil and a zebra finch had a max score of 2,235.

Gene 4: The fossil and a chinese aligator had a max acore of 1810.

This data confirms the close relationship between the fossil and cocodillians and birds, as the more similar the DNA the more closely related the species are.

HOW CAN BIOINFORMATICS BE USED AS A TOOL TO DETERMINE EVOLUTIONARY RELATIONSHIPS

SMARANDA LUPU

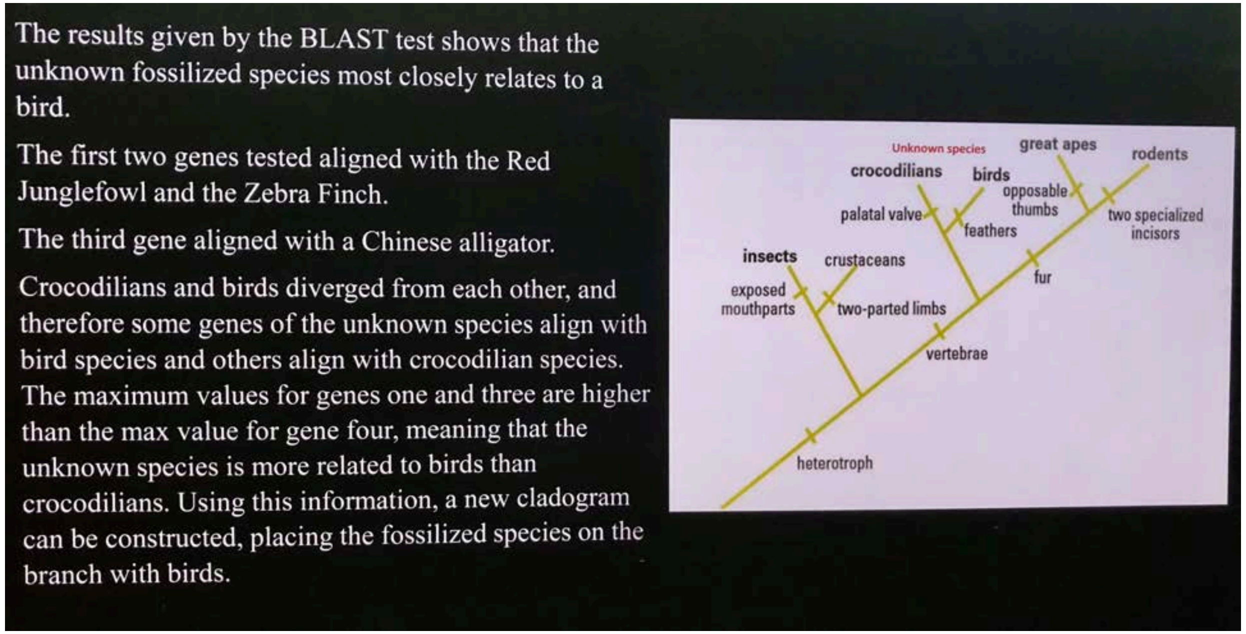


7) BLAST also constructed phylogenetic trees for each gene and placed the unknown species tested near the species that the program found to be most closely aligned. The highlighted writing on each tree represents the unknown species.

Phylogenetic Tree for Gene 1

Phylogenetic Tree for Gene 2

Phylogenetic Tree for Gene 3



The results given by the BLAST test shows that the unknown fossilized species most closely relates to a bird.

The first two genes tested aligned with the Red Junglefowl and the Zebra Finch.

The third gene aligned with a Chinese alligator.

Crocodilians and birds diverged from each other, and therefore some genes of the unknown species align with bird species and others align with crocodilian species. The maximum values for genes one and three are higher than the max value for gene four, meaning that the unknown species is more related to birds than crocodilians. Using this information, a new cladogram can be constructed, placing the fossilized species on the branch with birds.

BACK TO LOBBY

BACK TO ROOM 6

EXPLORING BIOINFORMATICS AND EVOLUTIONARY RELATIONSHIPS

TANVI MATHRUBOOTHAM

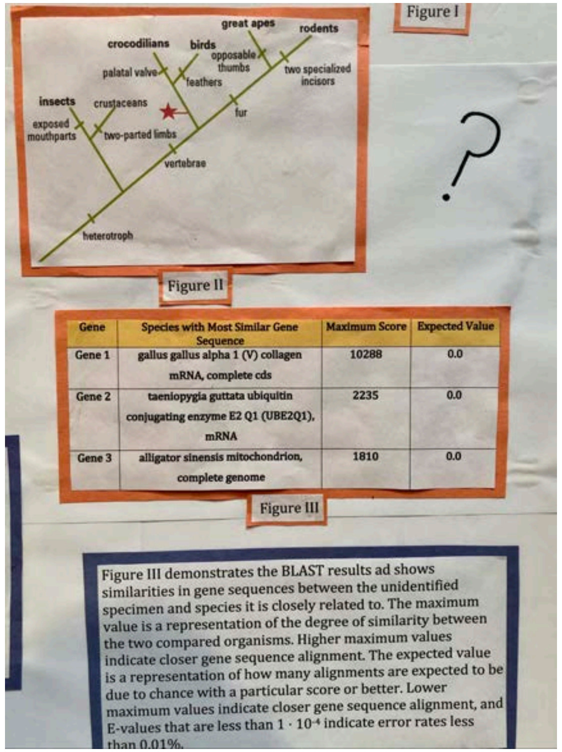


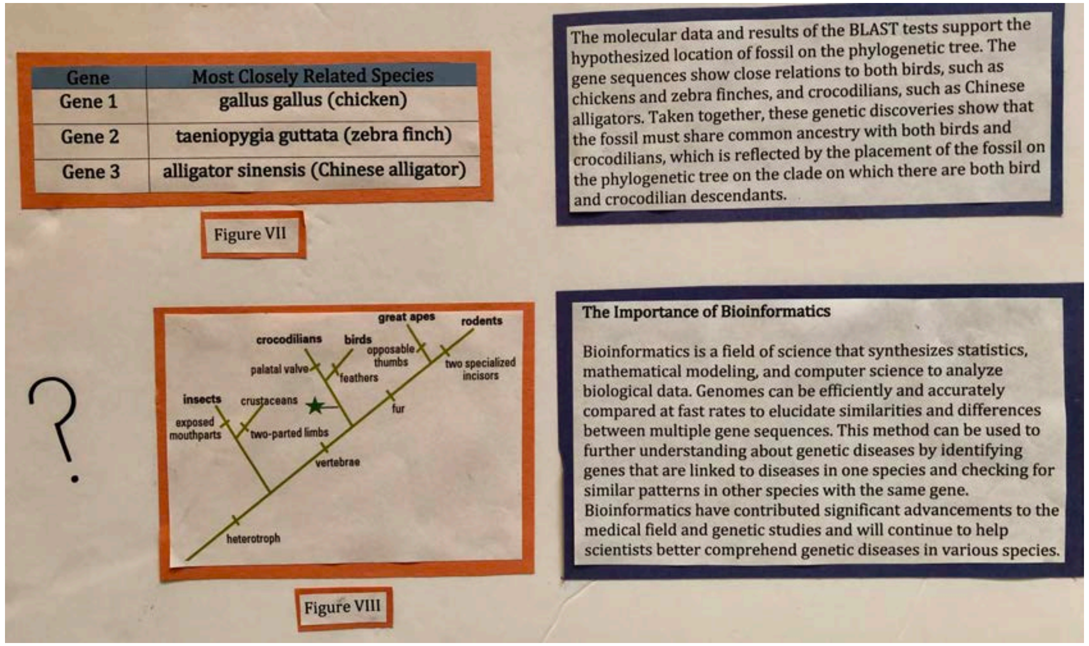
Figure I

Figure II

Gene	Species with Most Similar Gene Sequence	Maximum Score	Expected Value
Gene 1	gallus gallus alpha 1 (V) collagen mRNA, complete cds	10288	0.0
Gene 2	taeniopygia guttata ubiquitin conjugating enzyme E2 Q1 (UBE2Q1), mRNA	2235	0.0
Gene 3	alligator sinensis mitochondrion, complete genome	1810	0.0

Figure III

Figure III demonstrates the BLAST results ad shows similarities in gene sequences between the unidentified specimen and species it is closely related to. The maximum value is a representation of the degree of similarity between the two compared organisms. Higher maximum values indicate closer gene sequence alignment. The expected value is a representation of how many alignments are expected to be due to chance with a particular score or better. Lower maximum values indicate closer gene sequence alignment, and E-values that are less than $1 \cdot 10^{-4}$ indicate error rates less than 0.01%.



Gene	Most Closely Related Species
Gene 1	gallus gallus (chicken)
Gene 2	taeniopygia guttata (zebra finch)
Gene 3	alligator sinensis (Chinese alligator)

Figure VII

The molecular data and results of the BLAST tests support the hypothesized location of fossil on the phylogenetic tree. The gene sequences show close relations to both birds, such as chickens and zebra finches, and crocodilians, such as Chinese alligators. Taken together, these genetic discoveries show that the fossil must share common ancestry with both birds and crocodilians, which is reflected by the placement of the fossil on the phylogenetic tree on the clade on which there are both bird and crocodilian descendants.

Figure VIII

The Importance of Bioinformatics

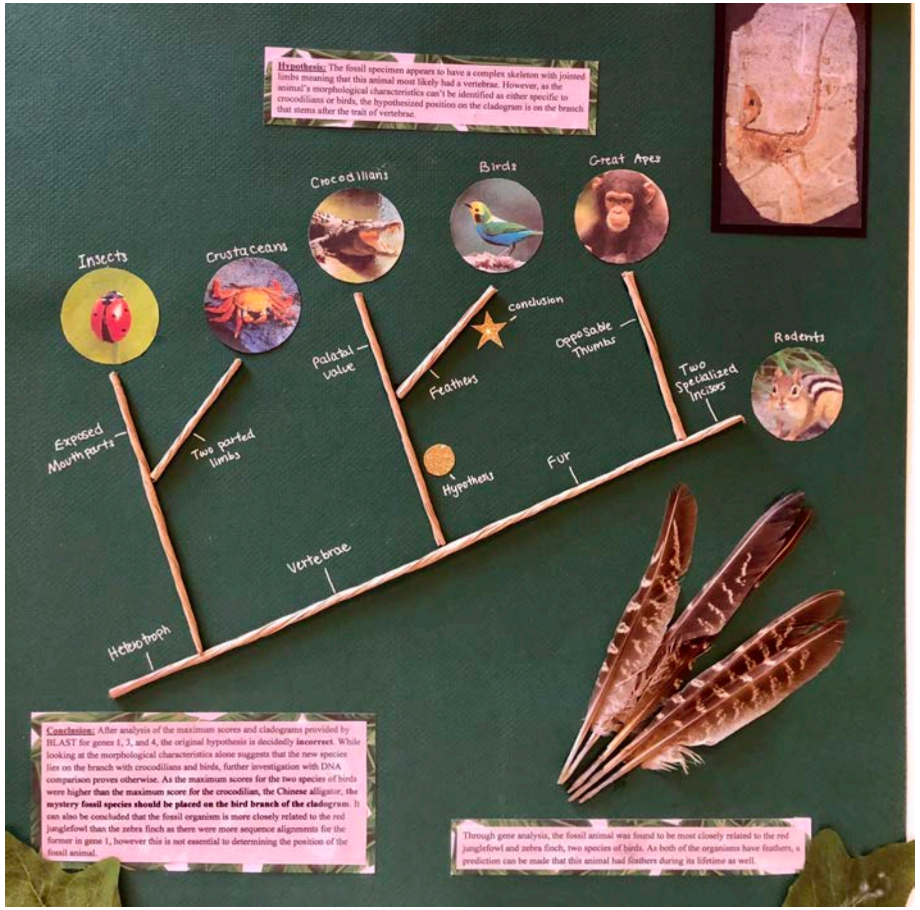
Bioinformatics is a field of science that synthesizes statistics, mathematical modeling, and computer science to analyze biological data. Genomes can be efficiently and accurately compared at fast rates to elucidate similarities and differences between multiple gene sequences. This method can be used to further understanding about genetic diseases by identifying genes that are linked to diseases in one species and checking for similar patterns in other species with the same gene. Bioinformatics have contributed significant advancements to the medical field and genetic studies and will continue to help scientists better comprehend genetic diseases in various species.

BACK TO LOBBY

BACK TO ROOM 5

EVOLUTIONARY LINEAGE OF A MYSTERY FOSSIL

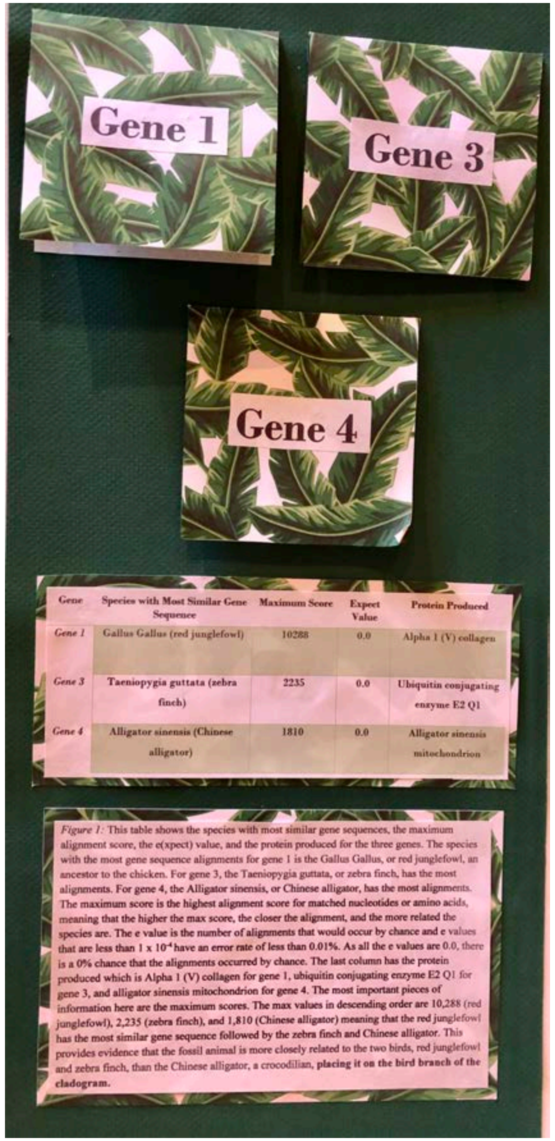
AYA SUGOI-MUNSON



Hypothesis: The fossil specimen appears to have a complex skeleton with jointed limbs meaning that this animal most likely had a vertebrae. However, as the animal's morphological characteristics can't be identified as either specific to crocodilians or birds, the hypothesized position on the cladogram is on the branch that stems after the trait of vertebrae.

Conclusion: After analysis of the maximum scores and cladograms provided by BLAST for genes 1, 3, and 4, the original hypothesis is decidedly **incorrect**. While looking at the morphological characteristics alone suggests that the new species lies on the branch with crocodilians and birds, further investigation with DNA comparison proves otherwise. As the maximum scores for the two species of birds were higher than the maximum score for the crocodilian, the Chinese alligator, the **mystery fossil species should be placed on the bird branch of the cladogram**. It can also be concluded that the fossil organism is more closely related to the red junglefowl than the zebra finch as there were more sequence alignments for the former in gene 1, however this is not essential to determining the position of the fossil animal.

Through gene analysis, the fossil animal was found to be most closely related to the red junglefowl and zebra finch, two species of birds. As both of the organisms have feathers, a prediction can be made that this animal had feathers during its lifetime as well.



Gene	Species with Most Similar Gene Sequence	Maximum Score	Expect Value	Protein Produced
Gene 1	Gallus Gallus (red junglefowl)	10288	0.0	Alpha 1 (V) collagen
Gene 3	Taeniopygia guttata (zebra finch)	2235	0.0	Ubiquitin conjugating enzyme E2 Q1
Gene 4	Alligator sinensis (Chinese alligator)	1810	0.0	Alligator sinensis mitochondrion

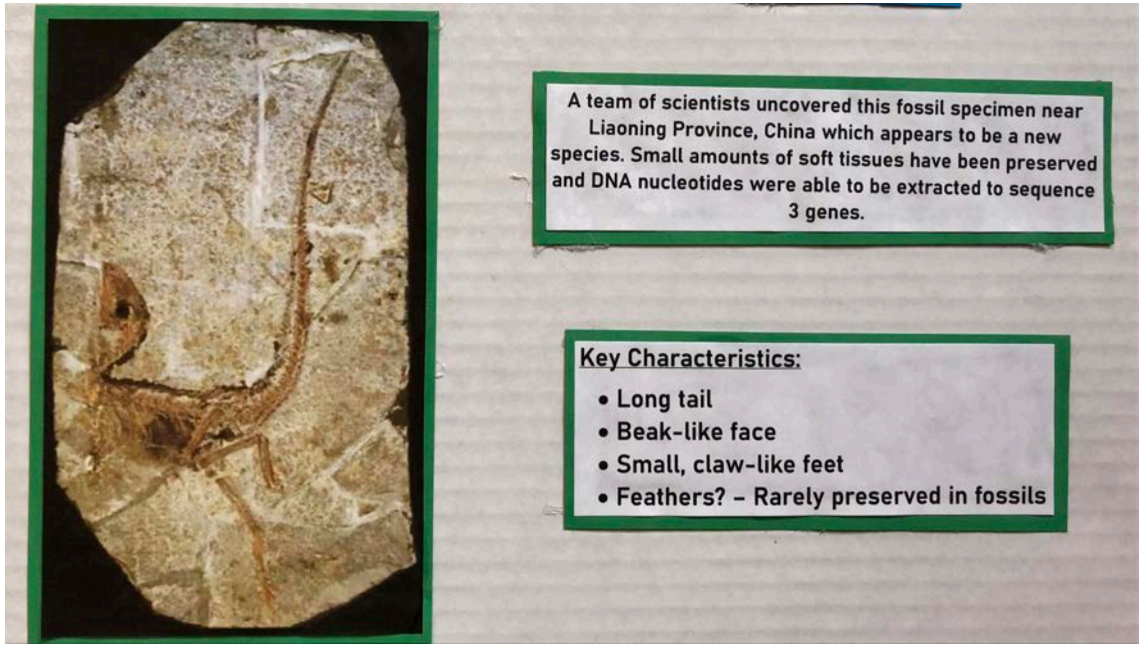
Figure 1: This table shows the species with most similar gene sequences, the maximum alignment score, the e(xpect) value, and the protein produced for the three genes. The species with the most gene sequence alignments for gene 1 is the Gallus Gallus, or red junglefowl, an ancestor to the chicken. For gene 3, the Taeniopygia guttata, or zebra finch, has the most alignments. For gene 4, the Alligator sinensis, or Chinese alligator, has the most alignments. The maximum score is the highest alignment score for matched nucleotides or amino acids, meaning that the higher the max score, the closer the alignment, and the more related the species are. The e value is the number of alignments that would occur by chance and e values that are less than 1 x 10^{-4} have an error rate of less than 0.01%. As all the e values are 0.0, there is a 0% chance that the alignments occurred by chance. The last column has the protein produced which is Alpha 1 (V) collagen for gene 1, ubiquitin conjugating enzyme E2 Q1 for gene 3, and alligator sinensis mitochondrion for gene 4. The most important pieces of information here are the maximum scores. The max values in descending order are 10,288 (red junglefowl), 2,235 (zebra finch), and 1,810 (Chinese alligator) meaning that the red junglefowl has the most similar gene sequence followed by the zebra finch and Chinese alligator. This provides evidence that the fossil animal is more closely related to the two birds, red junglefowl and zebra finch, than the Chinese alligator, a crocodilian, **placing it on the bird branch of the cladogram.**

BACK TO LOBBY

BACK TO ROOM 5

BIOINFORMATICS AND EVOLUTIONARY RELATIONSHIPS

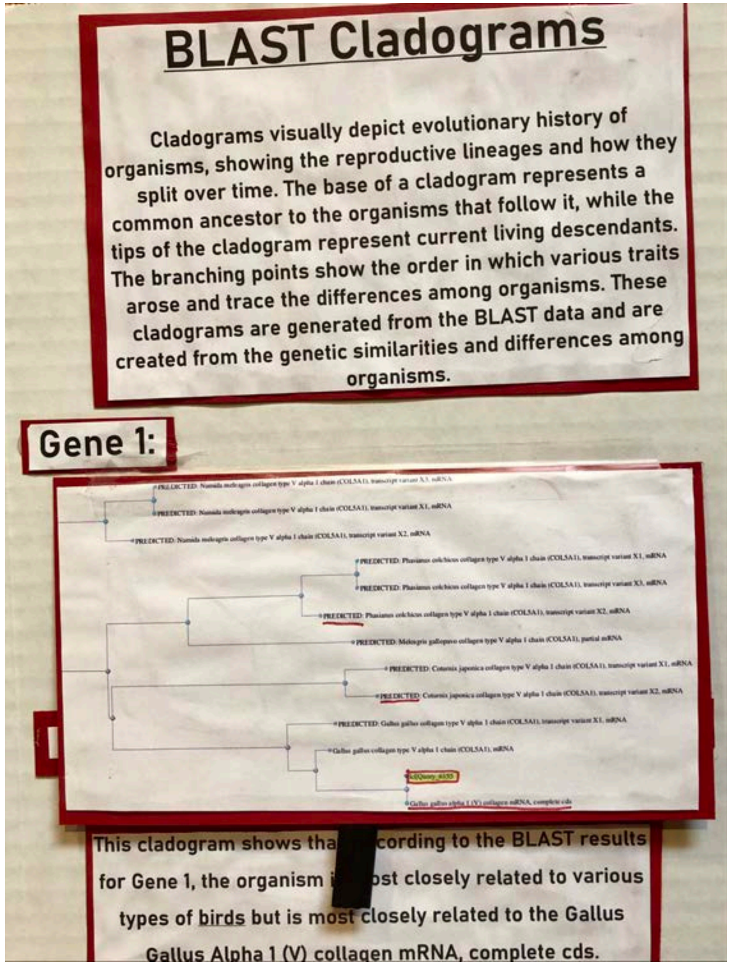
OLIVIA OKAMOTO



A team of scientists uncovered this fossil specimen near Liaoning Province, China which appears to be a new species. Small amounts of soft tissues have been preserved and DNA nucleotides were able to be extracted to sequence 3 genes.

Key Characteristics:

- Long tail
- Beak-like face
- Small, claw-like feet
- Feathers? – Rarely preserved in fossils

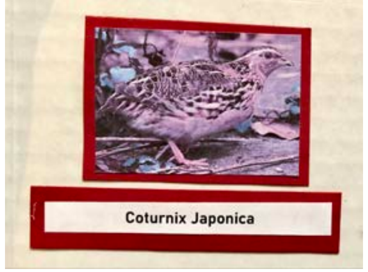


BLAST Cladograms

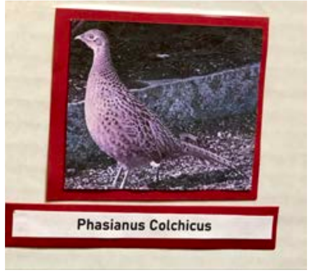
Cladograms visually depict evolutionary history of organisms, showing the reproductive lineages and how they split over time. The base of a cladogram represents a common ancestor to the organisms that follow it, while the tips of the cladogram represent current living descendants. The branching points show the order in which various traits arose and trace the differences among organisms. These cladograms are generated from the BLAST data and are created from the genetic similarities and differences among organisms.

Gene 1:

This cladogram shows tha cording to the BLAST results for Gene 1, the organism i ost closely related to various types of birds but is most closely related to the Gallus Gallus Alpha 1 (V) collagen mRNA, complete cds.



Coturnix Japonica



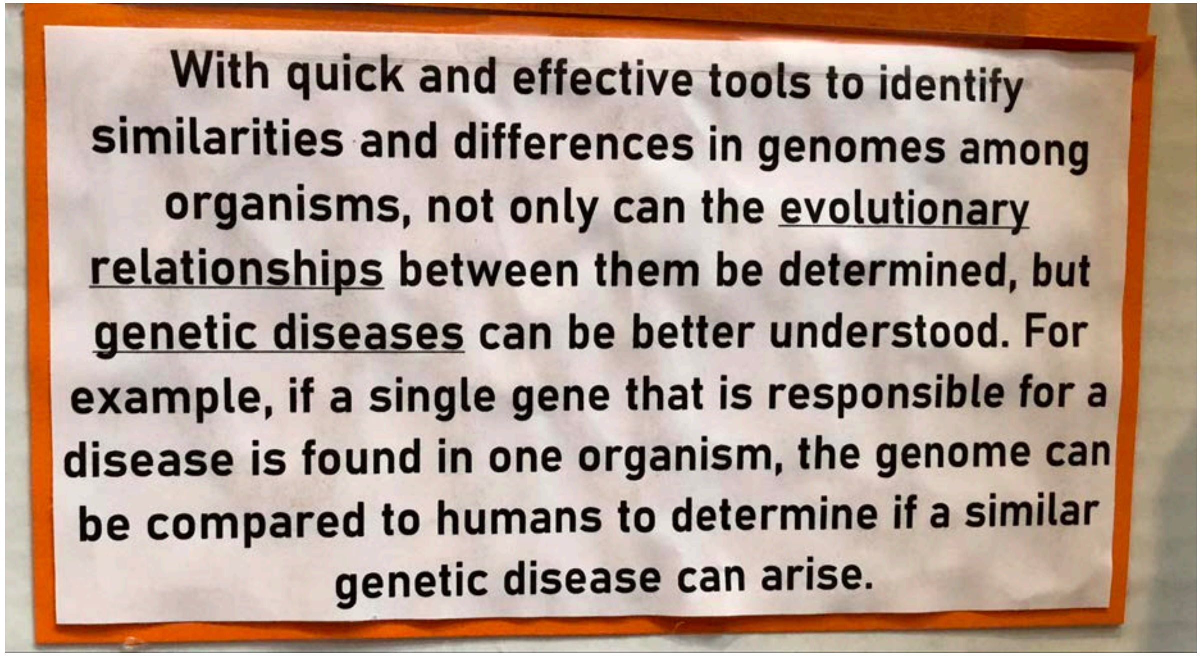
Phasianus Colchicus

CLICK FOR NEXT SLIDE

BIOINFORMATICS AND EVOLUTIONARY RELATIONSHIPS

OLIVIA OKAMOTO

Why is bioinformatics important?



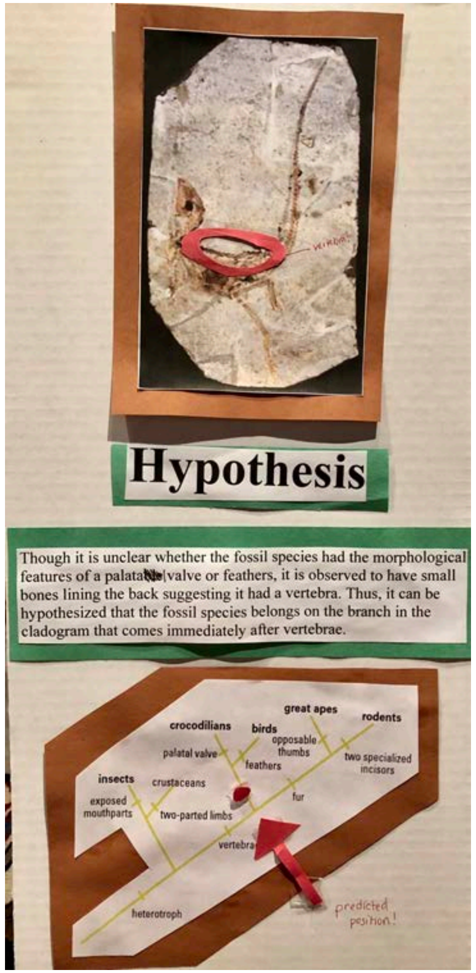
With quick and effective tools to identify similarities and differences in genomes among organisms, not only can the evolutionary relationships between them be determined, but genetic diseases can be better understood. For example, if a single gene that is responsible for a disease is found in one organism, the genome can be compared to humans to determine if a similar genetic disease can arise.

BACK TO LOBBY

BACK TO ROOM 4

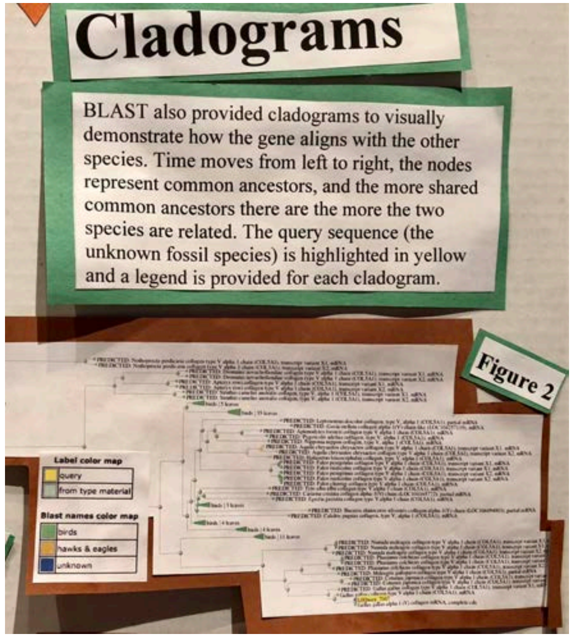
PLACING THE MYSTERY FOSSIL ORGANISM ON A CLADOGRAM

ELIZABETH CHEN



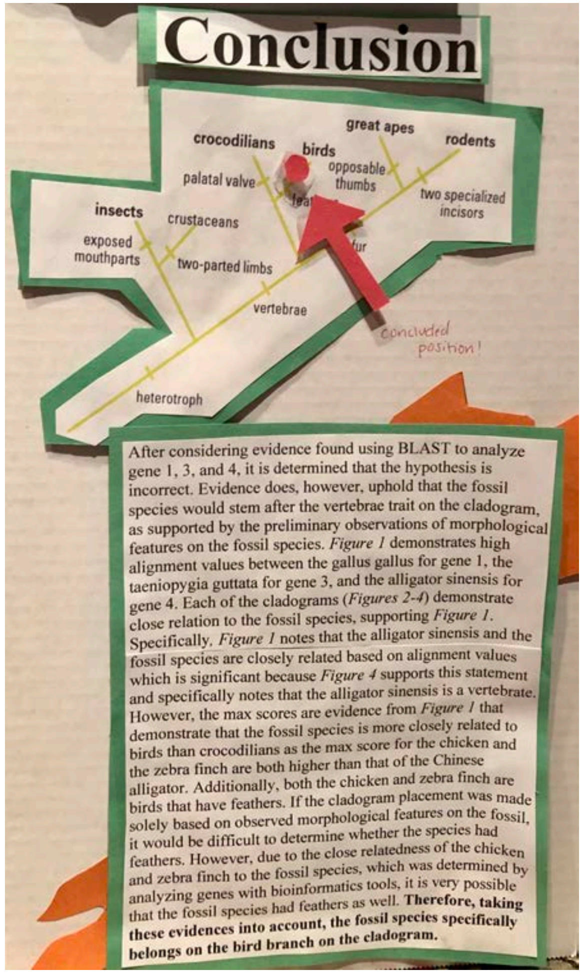
Hypothesis

Though it is unclear whether the fossil species had the morphological features of a palata[illegible] valve or feathers, it is observed to have small bones lining the back suggesting it had a vertebra. Thus, it can be hypothesized that the fossil species belongs on the branch in the cladogram that comes immediately after vertebrae.



Cladograms

BLAST also provided cladograms to visually demonstrate how the gene aligns with the other species. Time moves from left to right, the nodes represent common ancestors, and the more shared common ancestors there are the more the two species are related. The query sequence (the unknown fossil species) is highlighted in yellow and a legend is provided for each cladogram.



Conclusion

After considering evidence found using BLAST to analyze gene 1, 3, and 4, it is determined that the hypothesis is incorrect. Evidence does, however, uphold that the fossil species would stem after the vertebrae trait on the cladogram, as supported by the preliminary observations of morphological features on the fossil species. *Figure 1* demonstrates high alignment values between the gallus gallus for gene 1, the taeniopygia guttata for gene 3, and the alligator sinensis for gene 4. Each of the cladograms (*Figures 2-4*) demonstrate close relation to the fossil species, supporting *Figure 1*. Specifically, *Figure 1* notes that the alligator sinensis and the fossil species are closely related based on alignment values which is significant because *Figure 4* supports this statement and specifically notes that the alligator sinensis is a vertebrate. However, the max scores are evidence from *Figure 1* that demonstrate that the fossil species is more closely related to birds than crocodilians as the max score for the chicken and the zebra finch are both higher than that of the Chinese alligator. Additionally, both the chicken and zebra finch are birds that have feathers. If the cladogram placement was made solely based on observed morphological features on the fossil, it would be difficult to determine whether the species had feathers. However, due to the close relatedness of the chicken and zebra finch to the fossil species, which was determined by analyzing genes with bioinformatics tools, it is very possible that the fossil species had feathers as well. **Therefore, taking these evidences into account, the fossil species specifically belongs on the bird branch on the cladogram.**

BACK TO LOBBY

BACK TO ROOM 3

BLAST INVESTIGATION INTO A MYSTERY FOSSIL

IVY HAIGHT

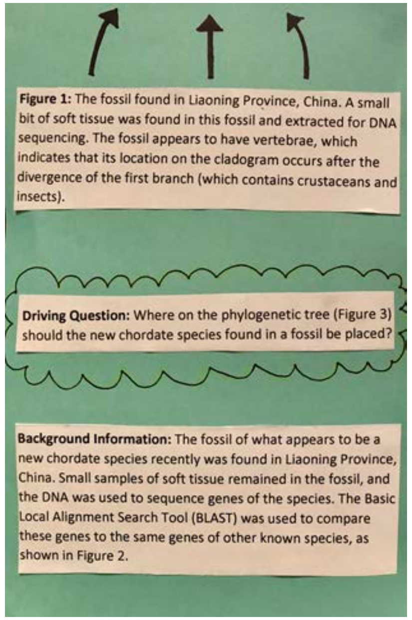
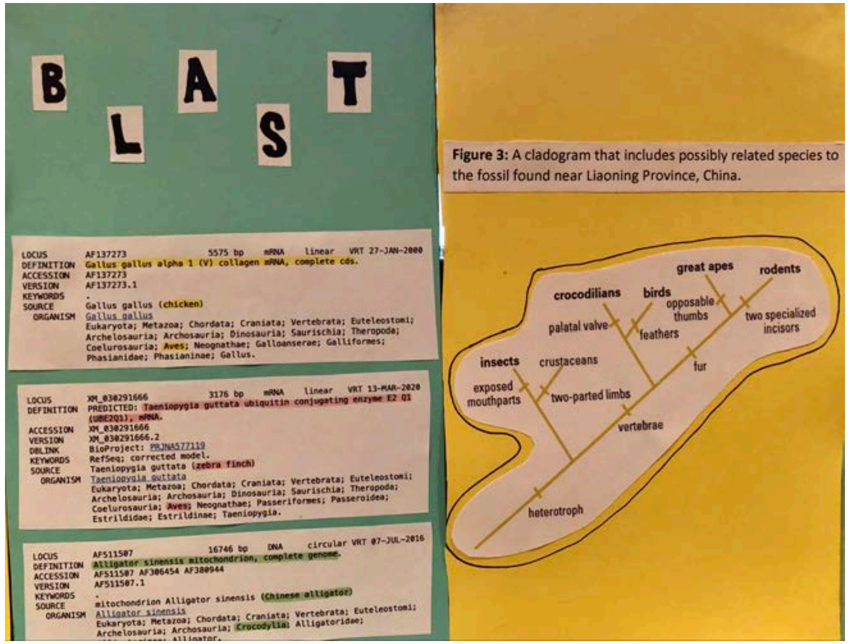


Figure 1: The fossil found in Liaoning Province, China. A small bit of soft tissue was found in this fossil and extracted for DNA sequencing. The fossil appears to have vertebrae, which indicates that its location on the cladogram occurs after the divergence of the first branch (which contains crustaceans and insects).

Driving Question: Where on the phylogenetic tree (Figure 3) should the new chordate species found in a fossil be placed?

Background Information: The fossil of what appears to be a new chordate species recently was found in Liaoning Province, China. Small samples of soft tissue remained in the fossil, and the DNA was used to sequence genes of the species. The Basic Local Alignment Search Tool (BLAST) was used to compare these genes to the same genes of other known species, as shown in Figure 2.

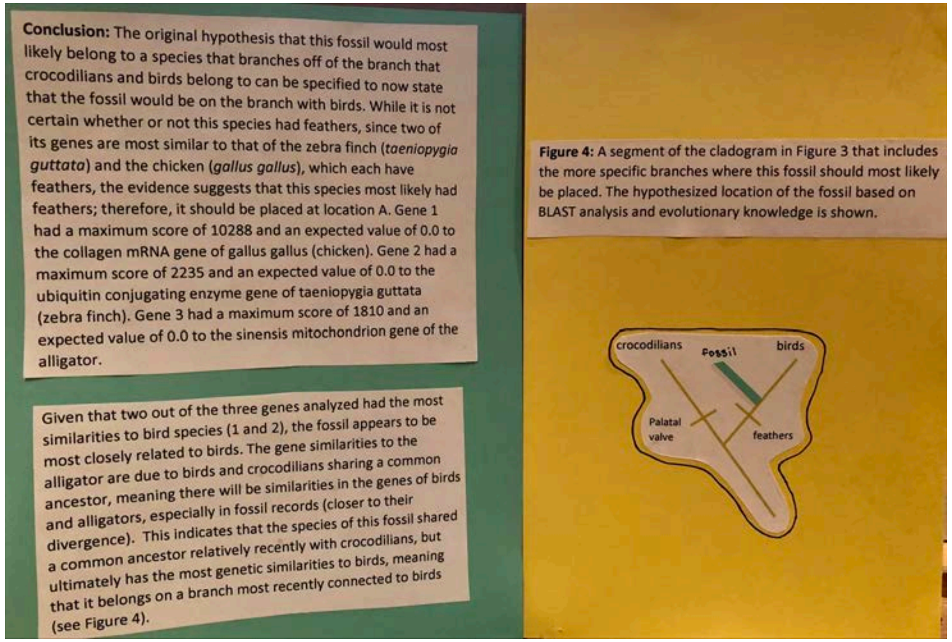


```
LOCUS       AF137273                5575 bp    mRNA    linear   VRT 27-JAN-2000
DEFINITION  Gallus gallus alpha 1 (V) collagen mRNA, complete cds.
ACCESSION   AF137273
VERSION     AF137273.1
KEYWORDS    .
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archelosauria; Archosauria; Dinosauria; Saurischia; Theropoda;
            Coelurosauria; Aves; Neognathae; Galloanserae; Galliformes;
            Phasianidae; Phasianinae; Gallus.
```

```
LOCUS       XM_030291666            3176 bp    mRNA    linear   VRT 13-MAR-2020
DEFINITION  PREDICTED: Taeniopygia guttata ubiquitin conjugating enzyme E2 Q1
            (UBE2Q1), mRNA.
ACCESSION   XM_030291666
VERSION     XM_030291666.2
DBLINK      BioProject: PRJNA577119
KEYWORDS    RefSeq; corrected model.
SOURCE      Taeniopygia guttata (zebra finch)
  ORGANISM  Taeniopygia guttata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archelosauria; Archosauria; Dinosauria; Saurischia; Theropoda;
            Coelurosauria; Aves; Neognathae; Passeriformes; Passeroidea;
            Estrildidae; Estrildinae; Taeniopygia.
```

```
LOCUS       AF511507               16746 bp    DNA     circular VRT 07-JUL-2016
DEFINITION  Alligator sinensis mitochondrion, complete genome.
ACCESSION   AF511507 AF306454 AF380944
VERSION     AF511507.1
KEYWORDS    .
SOURCE      mitochondrion Alligator sinensis (Chinese alligator)
  ORGANISM  Alligator sinensis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archelosauria; Archosauria; Crocodylia; Alligatoridae;
```

Figure 3: A cladogram that includes possibly related species to the fossil found near Liaoning Province, China.



Conclusion: The original hypothesis that this fossil would most likely belong to a species that branches off of the branch that crocodilians and birds belong to can be specified to now state that the fossil would be on the branch with birds. While it is not certain whether or not this species had feathers, since two of its genes are most similar to that of the zebra finch (*taeniopygia guttata*) and the chicken (*gallus gallus*), which each have feathers, the evidence suggests that this species most likely had feathers; therefore, it should be placed at location A. Gene 1 had a maximum score of 10288 and an expected value of 0.0 to the collagen mRNA gene of gallus gallus (chicken). Gene 2 had a maximum score of 2235 and an expected value of 0.0 to the ubiquitin conjugating enzyme gene of taeniopygia guttata (zebra finch). Gene 3 had a maximum score of 1810 and an expected value of 0.0 to the sinensis mitochondrion gene of the alligator.

Given that two out of the three genes analyzed had the most similarities to bird species (1 and 2), the fossil appears to be most closely related to birds. The gene similarities to the alligator are due to birds and crocodilians sharing a common ancestor, meaning there will be similarities in the genes of birds and alligators, especially in fossil records (closer to their divergence). This indicates that the species of this fossil shared a common ancestor relatively recently with crocodilians, but ultimately has the most genetic similarities to birds, meaning that it belongs on a branch most recently connected to birds (see Figure 4).

Figure 4: A segment of the cladogram in Figure 3 that includes the more specific branches where this fossil should most likely be placed. The hypothesized location of the fossil based on BLAST analysis and evolutionary knowledge is shown.

BACK TO LOBBY

BACK TO ROOM 3