A Cautionary Tale

DNA Analysis of Alleged Extraterrestrial Biological Material: Anatomy of a Molecular Forensic Investigation

National Institute for Discovery Science
4975 S. Polaris Ave, Las Vegas, NV 89118

A&S RESEARCH
253 Lombard Street, Suite B, Thousand Oaks, CA 91360

The APEG
P.O. Box W42, West Pennant Hills, NSW, 2125, Australia

Abstract

In investigating cases of alleged anomalies, particularly when physical evidence appears to corroborate the anomalous nature of an event, a tendency to jump to hasty conclusions can manifest. Here we describe a case in which multiple rounds of DNA analysis of a biological sample appeared to corroborate its reported anomalous (extraterrestrial) origins.

In September 2000, an object resembling a claw was found by a family in a bedroom of their Californian home. The family reported they were in the midst of an intense series of visitations by purported extraterrestrials. Thus, the preliminary DNA analysis on this claw began to appear to "fit in" with the multiple reports of high strangeness events as reported by the primary eyewitnesses. The linkage was made stronger because the anomalous biological sample was found in the same bedroom in which numerous visitations and intrusions by alleged extraterrestrials had occurred. Further, the sample was found during a period of intense anomalous activity in the same bedroom. Six separate rounds of subsequent DNA analysis, using different and sometimes very novel approaches, were carried out to bring this case to a conclusion. It was necessary to invent a new polymerase chain reaction using novel primers to the most conserved DNA sequences on Earth in order to finally resolve this case.

Hence, painstaking DNA analyses and the use of bioinformatics methodology over a 12 month period by highly qualified teams of experts in three countries was necessary to establish that the biological specimen found in the house was a mundane terrestrial mollusk. Mollusks, particularly snails and slugs, secrete a thick mucous that contains multiple inhibitors of many of the common enzymes that are fundamental to molecular biology and DNA analysis, including polymerase chain reaction enzymes and those used in standard molecular cloning. Further, there is a relative paucity of mollusk DNA sequences, particularly from mollusks found in California, in global DNA databanks. These two factors conspired to lead this investigation down a false path for about a year. Ultimately, however rigorous DNA analysis using a novel set of oligonucleotide primers for the polymerase chain reaction solved the puzzle. The

investigation of this case went far beyond the "business-as-usual" analyses usually afforded anomaly cases. The project evolved into a major molecular biology research project in its own right.

Independent confirmation that the sample was a dried mollusk was obtained by an expert from the Los Angeles County Natural History Museum.

The lessons from this intensive, lengthy and very expensive investigation are simple and can be generalized for all anomaly investigations: (a) always follow the scientific data, (b) resist the temptation to tie purported physical evidence with a nearby anomaly until the analysis is fully completed, (c) resist the impulse to publish (or publicize) during an ongoing investigation.

It should be noted that the original anomalies reported by the family in their house in California remain unexplained and are still being reported and investigated.

Introduction

In 2000 and 2001 a family in California contacted several investigators regarding their experience of an ongoing series of anomalous encounters. The vast majority of these encounters occurred in the family home and represented nocturnal visitations by a variety of creatures, sightings of small flying objects in the bedroom, injuries to two family members and a wide spectrum of anomalous lights, sounds and mechanical interferences. According to the family, most disturbing were multiple encounters with bizarre creatures in the parents' bedroom usually late at night. During these encounters, it was noticed that the creatures apparently entered and sometimes exited via the bedroom walk-in clothes closet. The family decided to try to obtain physical evidence of these encounters by placing an aluminum sheet underneath a towel on the floor of the closet in the hopes that the creatures would leave footprints or other physical evidence of visitation. This was one of many strategies the family followed in an effort to obtain physical evidence of these anomalies.

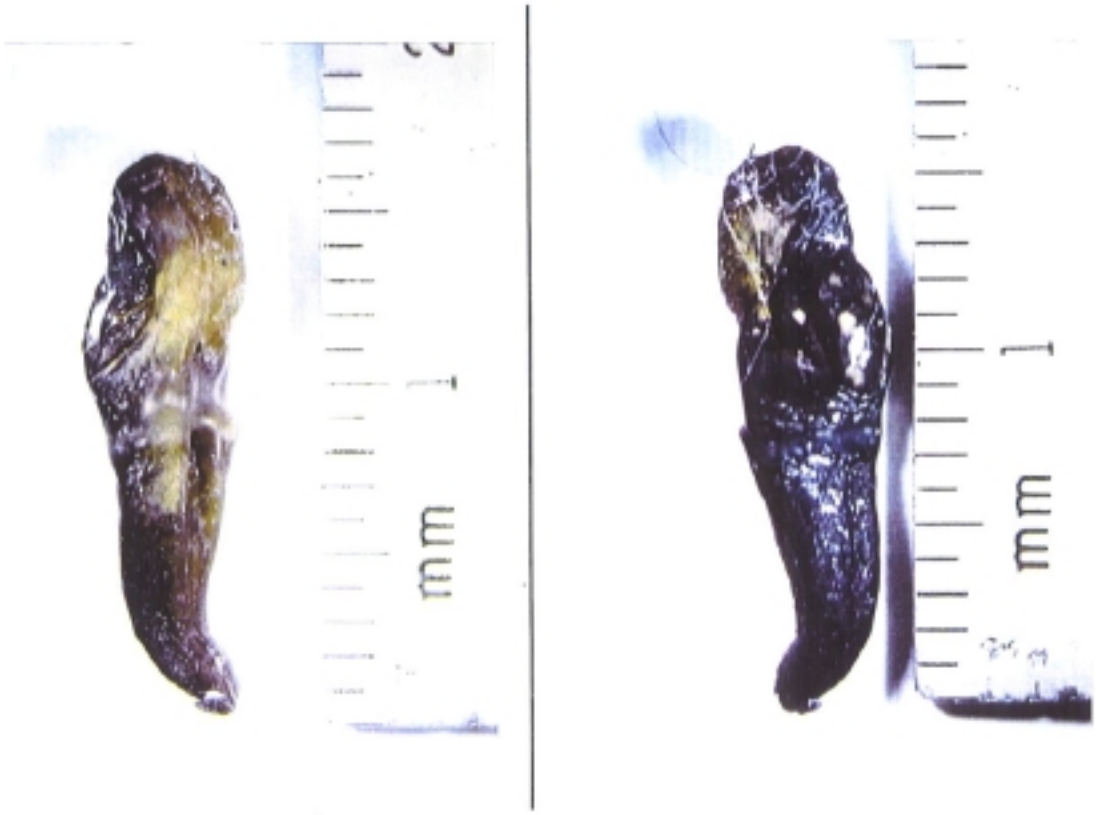
Following one visitation in September 2000 the family noticed a small claw like object that was embedded in the towel that lay on top of the aluminum sheet. The family assumed that they had found physical evidence of the visitations and began a campaign to identify the nature of the "claw". Close examination of the claw indicated what appeared to be a hair protruding from the rear of the object leading to the supposition that the claw came from a mammal.

Although the purpose of the present report is to describe in some detail the DNA analysis that eventually identified the object, it is worth describing some preliminary steps in the investigation.

- The claw (see photograph below) was examined by a herpetologist at the University California, Berkeley, who said although it looked like a claw it did not look like any reptile claw he had seen before.

- A senior primatologist at San Diego zoo examined the object and opined that it resembled an extremely distorted claw from a New World monkey. Prior to the DNA analysis, the primatologist's opinion went considerably towards strengthening the hypothesis that the object was a mammalian claw.
- Thirdly, although uninformative for biological specimens, it was decided to subject the object to an SEM-EDX analysis. The elemental content of the claw appeared consistent with a biological specimen.

Therefore, the preliminary examination by the herpetologist and by the primatologist appeared to validate the hypothesis that the object was a claw, possibly of mammalian origin. At this point DNA analysis was begun.



Results and Discussion

Note: Throughout the first six phases of this investigation, much of the narrative from the scientific team who performed the analyses is preserved in order to highlight the evolution in thinking that accompanied the bizarre and sometimes confusing data that emerged from the molecular detective hunt. Secondly, the following sections are replete with molecular biology jargon and need not necessarily be focused on. The essential message of this paper lies in the Abstract.

Phase 1: The First DNA Examination

Summary and preliminary interpretation of findings from the first laboratory to conduct DNA analysis:

> A preliminary but detailed PCR analysis was carried out, in order to determine the species identity of a haired claw from an unknown large mammal. Our findings suggest that this claw may belong to a primate species not currently represented in the Genbank database.
>
> All standard forensic reactions for cytochrome b, 12S rRNA, cytochrome c or human mitochondrial HVI show only weak or non-specific products after 35-40 cycles of PCR amplification, using DNA of high quality which was extracted from a small piece of that claw. Yet under precisely the same conditions, a wide range of mammalian controls amplify strongly and specifically: namely rat, mouse, rabbit, dog or human.
>
> Any detectable PCR products were cloned into a standard plasmid vector, so that many different DNA molecules could be analyzed by automated sequencing. When studied in that way, the vast majority of clones (54 / 59) show no significant match to any sequence in Genbank. Yet a few clones (5 / 59) for 12S rRNA or cytochrome c do show a partial match to known sequences.
>
> To be precise, two clones for 12S rRNA (CS63 or CS68) show a partial match to: (a) an ancient mitochondrial insertion within the T-cell region of primates; or (b) an unknown land snail which was apparently bound within that claw when it was extracted for DNA.

So, with the benefit of hindsight, one out of fifty nine clones (CS 68) yielded a marginal homology to a land snail. Only a year later would this become significant.

Phase 2: Bioinformatics Analysis of Claw DNA

a. In order to further examine the Phase I data, a complete bioinformatics analysis was carried out on all DNA obtained to date by a high quality West Coast bioinformatics facility. See pdf file.

b. tRNAScan - nothing detected
SRP RNA Scan - nothing detected
Signal Scan - Multiple possible signals
Transfac (using MatInspector)- multiple possible signals
3rd position GC bias - inconclusive (overall GC 42 % ; fairly conserved over all sequences, appears that all the sequences could be from the same source. Some regions have significantly different 3rd position GC bias, indicating possibility of coding regions)
HMM based gene predictions - no ORFs detected.

c. Promoter analysis of DNA sequences indicated bona fide promoter sequences were present in dozens of the unknown sequences [This same analysis was

conducted on subsequent DOP-PCR sequences with positive results]. See accompanying file for promoter sequences.

Phase 3: Direct cloning from DNA, no PCR

Listed below are five new sequences from directly-cloned claw DNA (no PCR), either sample A or B. None of them match anything in Genbank, apart from a weak match in one case to a repetitive sequence on the human Y chromosome, which would be similar in primates.

Either this claw comes from a local monkey, or else from an exotic primate ancestor of man. I would have liked to have gotten 50 direct clones, however the material simply does not like to clone in E. coli, even after extensive purification.

There is one piece C left, that could be analyzed separately from A and B. More productively, we could do DOP-PCR on existing samples A and B, and gel purify to select for a size mixture of 300 bp or greater. Then some other lab could clone it easily, and sequence in great detail.

cA4.doc
GAATTCGATACCATCAATCTCACTGTAGGAACTGCC
CCAACCATCAATCTCACTGTAGGAACTGCCCCAACCATTAATCCCACTGTAG
GAACTGTCCCAACCATCAATCCCACTGT
AGGAACTGCCCCAACCATCAATCTCACTGTAGGAACTGCCCCAACCATCAA
TCCCACTGTAGGATCACTAGTGAATTC

cB8.doc
GAATTCGATTCCCAGAGTTTGCCGTGCATATGTCAC
TCTGGTTTAAGTGTCGGTCTGTTGGCCGTGCTCCATGGAGTATTGTTTATCT
CCACTTCTGGACACGGTCTGCTAGCGTC
TTCGTCTCCTTATGGGACTTAGAGGACCAGGTTTTGGCGCTCAGCTACTGAT
CACTAGTGAATTC

cA13.doc
GAATTCGATCTCCATCTCCTCTTTTTAAGTATGTCA
TTCATTGGGTCCTGTCTGGCATCCTGAAGCAGCCAGTTGTTGGATATTTTAT
TTGGCCAGAAGACTCTTAGGATTTGAAG
AAATCACTAGTGAATTC

cA20.doc
GAATTCGATGAATCAGCTTGCAAAAGTCTGATAACT
GAAAAGTATTTATCATTTTATGCATAGTTTGAAAACTCAAGATTGTCTGTAAT
CACTAGTGAATTC

cA21.doc
GAATTCGATAAGTGTTACAGCAACTGGAAAAGTTCA
AACGTCTAGAAATGCACCCTATGTACGAACGAGTATACAGATCGGGCAAAG
GCAAATTCGCCACCTCATGGATCACTAGT
GAATTC

Phase 4: Attempted Independent Corroboration of Lab Results

Methodology: The claw sample was submitted to a molecular diagnostics lab in Ontario Canada that had a high-level experience in species identification from old and degraded DNA (including successfully isolating and characterizing DNA from 100 year-old whalebone).

Summary of work carried out by Canadian lab:

1. Extracted claw DNA using Qiagen extraction protocol.
2. Quantified DNA using Pico Green (approx. 2ug)
3. Amplified 1ng of DNA with cytochrome B (approx. 250bp). There was no amplification.
4. Ran DNA (not amplified) and found most of the DNA was of low molecular weight (approx. 100-200bp.)
5. Amplified DNA with a dilution series (10ng, 1ng, 100pg and 10pg)
    a. 100pg amplified very lightly
6. Spiked 1ng of DNA with a 100pg of known ungulate DNA. Amplification of spiked product was partially inhibited.

Thus the Canadian lab concluded that the sample was heavily contaminated with inhibitors.

Phase 5: DOP-PCR Analysis of Claw DNA

Methodology Summary:

Modern DOP methods are best described in: PNAS USA, March 5, 2002, volume 99, no. 5, pp. 2942-2947, "Genome complexity reduction for SNP genotyping analysis," authors Barbara Jordan et al., Center for Cancer Research, MIT (David Housman, John Landers). That paper plus its "Supporting Methods" give the best and latest methods, which were used in this study.

Using the MIT DOP-PCR methodology, 14 x 5 ug of claw B DNA were successfully amplified—a relatively huge amount of claw genomic DNA. 3 of those 14 DOP reactions were purified on an agarose gel (no UV light), selecting sizes 400 bp or greater (0.5 ug total), and ligated 1-2% with a PCR vector.

> The first claw-A sample was quite impure (yellow) and failed to amplify. But the second claw-A Wizard sample amplified a lot to 200 bp.
>
> More significantly, both claw-B samples amplified very well to give a fat band at 200-300 bp, that trails upward towards a discrete band around 600 bp which is probably repetitive DNA, then a few products at 700-800 bp.
>
> Those size distributions match well the original DNA sizes for claw samples A and B. Any local contamination would give larger sizes, cf. controls.

The big repeat band of 600 bp for the claw matches a similar band in rat, but not in human, where the repeat is slightly larger.

Note that when using a random DOP primer, Xho-NNNNNN-ATGTGG (the standard DOP sequence), our claw DNA amplifies quite well; whereas the same claw DNA would hardly amplify at all when using specific cyt. b or 12S RNA or cytochrome c mammalian primers previously.

This result implies again that the DNA sequences within such claw DNA may be quite unusual, and fail to amplify with specific consensus primers, because the genomic sequences in those critical primer locations do not match any Earth mammalian consensus.

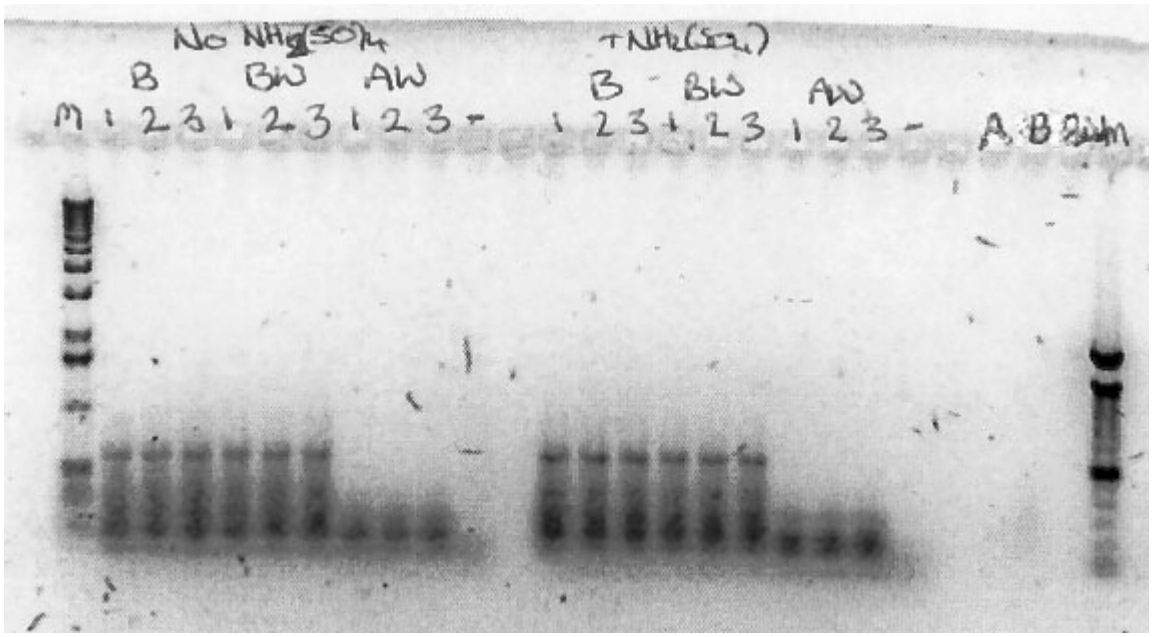
The 600 repeat band is probably centromeric alpha-satellite (my guess), which may be useful in species identification.

We will repeat again the claw A and B rxs in multiple tubes to get lots of sample (each rx yields 5 ug total), then pool and purify claw-A, claw-B over a column to remove anything less than 100 bp. One might wish to gel-select for large sizes later, 400 bp or more for example.

It looks very much like an "unknown mammal" at this point. Possibly a new species with the same biochemistry as on Earth.

Attached is a recent DOP gel for claw samples B and A, with many tubes to give large amount. Samples on the left include no ammonium sulfate, while those on the right are plus ammonium sulfate. Six B samples and three A on each side. Followed by a negative control of no DNA (-), and two neg. controls on the right without Taq enzyme.

We pooled those tubes, ran DOP-B or DOP-A over a Wizard column to remove primers, eluted in TE, and air dried in small aliquots using 250 ul tubes. Also dried and aliquotted the gel-purified claw B (400 bp or more), from which 24 clones were sent to sequence facility.



Sequence Analysis of the above DOP-PCR amplification

DOP01
GNAATTCGATTCCGACTCGAGGGGTAGATGTGGG
TGTTGATCTCACATGCCGATGTGGTCGTTGATCTCACATGCCGATGTGGGTGTTGATCTCAC
ATGCCGATGTGGGTGTTG
ATCTCACATGCCGATGTGGTCGTTGATCTCACATGCCGATGTGGGTGTTGATCTCACATGCC
GATGTGGGCGTTGATCTC
ACATGCCGATGTGGGCGTTGATCTCACATGCCGATGTGGGCGTTGATCTCGTTGCTCCACA
TCCCCGGCTCGAGTCGGAA
TCACTAGTGAATTC

DOP03
GAATTCGATTCCGACTCGAGGCAGCCGATGTGGCT
CTTTACTAGTGTGTCTGCAGGCTTTGTGCCCGTCCCGATCCTGCCGGAATGTCTGGTATCCG
GGGACCCAGAAGCTCCAG
TCCAATTTCAGGTGAGTCTCCTGAATACAGCCCAAGTCGACACGGTCCCTCGCGAATCTCTC
TTATAGCTTTCGCCTCTT
TTGGTAGACGCCTTCTGCGTTCCACTGAAGGATTCTTAGATTTCCCAGATTCTTTGCAGCCG
ATTCCACTCGTAGCCGTC
GACCACATTATGCCCTCGAGTCGGAATCACTAGTGAATTC

DOP04
GAATTCGATTCCGACTCGAGGTTAGGATGTGGTCA
TGTTTTCATCATTTCTCTGAGAAATCATCCTCAGAGATTATTTCACATTATTTTCTGTTATGTAC
TATCTAGTCCACTAA
ACAACACTGGAGAAATCTCTTCCCCTCTGCCTTCCTTCCCTATCACTATTCATTCATTCATTAA
GAACCCTGCCGACCGG
CGTTGGCGATCATGACTCTCCACATGACCCCCTCGAGTCGGAATCACTAGTGAATTC

DOP06
GNAATTCGATTCCGACTCGAGGGGATGATGTGGTC

GTTGATTTTTAGCGTCGCCTTGAGGACCTTGGTGGATAGACTGAATATGGTGTACGTGGTGT
CAGGTTTCAGGTGAGTCT
CTTGAATGCACGCCACATGCCTCGCTCGAGTCGGAATCACTAGTGAATTC

DOP07
GNAATTCGATTCCGACTCGAGGACGTCATGTGGGG
TGTCCAAGTTAACGCCTTGCCTCTCATGTGGGGTGTCCAAGTTAACGCCTTGCCTCTCATGT
GAGGTGTCCACATTCACG
TCTTGCTTCTCAAGTGGGTGTCACGTTAACGCCTTGCCTCTCATGTAGGGTGTCCACATCCA
CATCTCGAGTCGGAATCA
CTAGTGAATTC

DOP08R
GAATTCGATTCCGACTCGAGGGGGGAATGTGGCA
CATGGAAAGCAAGAATCTGATTTCTCCCAACCAGAGCGGATTTAGGACTAACATGTCAACAG
AAGACCAGGTTACCTATA
TCGCGCAAGAAATCATCGATGGCTACCACGTCCACGAGCACACAGTGGCCGTATGGGTAGA
ATGGAAAAAGCGTTCGACA
CAGTGTGGAGAGACGGGCTCAAGTCCAAACTCCTAAGCCTAGGACTTGGAGGACATATGTT
CGAGTGGATCGGCAGCTAC
CTAGAAGACCGGACAGCTAGGGTCCGAGTGCAGGGCAAAACAAGTCGAACGAACACTCGAA
CAGGGAGTGCCACAATGAG
GGGTGCTGTCCCCAACGCTCTTTCTCGCATTCGTCAACGACATTAGTACATGTTTCCCCAGA
GAGGTGCGCTCAGCCATG
TACGCGGATGATCTCGCCCTTTGGGCCAGGAGTGAGTCGATAAAAAAGGCCGAAAAAAGCT
TACAATCAGCGCTCTCCAA
GCTGGAGGACTGGACGAAGCAGTGGCTCTTAAAAATCAATGGAAAAAAGACCACGTACACG
ATATTCAGTCTGACCACCA
ATGTCATCAAGGCCACGCTAAAAATTAACGACCACATCCCCTCCTCGAGTCGGAATCACTAG
TGAATTC

DOP10
GNAATTCGATTCCGACTCGAGGGCGTTATGTGGTTGGCA
TAGGTGAGGCTGGGTCAGCCACAGGAACCTTTAAATTGAGGAAGAATTTCCTTATCTGGAGA
TCTGACCCTTGACCTAGC
CGCTGCTAAGCTGTAAGGCAACAACTTTGCTACCTTCCTAAAAAAAAAGTAAAATGCCCAGAA
ATAGTTATATCTCTACAG
TTGTACTTAAAAGTGGAAAATCTATCTACCATACAACCCCCCAGAAATTTTGTTGTTGATTCAA
TCATCTAGGTTTTGTA
CGTTGATAGTTTGCCATTGCTTTCTGAAGGGAGATTGGCATAAAATAAATTATGCGTAGCTGC
TTTCTGAAGGTAAGCTT
TTAATACCAAGATTTTTAAAAAATCCACATTGATTGATCTGCTTCCAAATTGCACATGGTGGG
ATTATATCTCAGCCCCG
TTGGGTTCACGAAGCCACATCCCAAGCTCGAGTCGGAATCACTAGTGAATTC

DOP14R
GNAATTCGATTCCGACTCGAGGGGGTCATGTGGCA
CATGGAAAGCAAAAATTTGATTCCCCCCAATCAGAGCGGATTTAGGAGTAACATGTCTACAG
AAGACCAGGTTACCTACA
TCGCGCAAGAAATCATCGATAGCTACCATGTCCACGAGCACACGCTAGCCGTATGGGTTGA
CATGGAAAAAGCGTTCGAC
ACAGTGTGGAGAGATGGGCTCAAATCAAAACTTCTAACTCTAGGAGTAGGAGGACATATGTT
CGAATGGCTCGGAAGCTA
CCTCGAAAACCGGACAGCTAGGGTCCGAGTGCAGAACAAAACGAGCCGAAAACGAACACTC
GAACAAGGCGTTCCTCAAG
GTGGAGTGCTGTCCCCAACGCTCTTTCTCGCATTTGTCAACGACATCAGTATATGTTTTCCCA

TGGAGGTGCACTCAGCC
ATGTACGCGGATGATCTCGCCCTTTGGGCCAGGAGTGGGTCGATGGAAACGGCCAAACAAA
GTGTGCAATTAGCGCTCGC
CAAGCTGGAAGACTGGGACGAAGCAGTGGCTCTTACAAATCAATGGAAATAAGACCACGTA
CACCATATTCAGTCTATCC
ACCAAGGTCCTCAAGGCGACGCTAAAAATCAACGACCACATCAAACCCTCGAGTCGGAATC
ACTAGTGAATTC

DOP16
GNAATTCGATTCCGACTCGAGGGAGTTATGTGGGC
GTTGTCAAATGTGGGCGTTGTCAAATGTAGGGGCGTTGGTAAATGTGGGCGTTGTCAAATAT
GGGCGTTGTCAAATGTGG
GCGTTGTCAAATGTGGGCGTTGTCAAATGTGGGCGTTGTCAAATATGGGCGTTGTCAAATGT
GGGCGTTGTCAAATATGG
GCGTTATCAAATGTGGGCGTTGTCAAATGAGAGCGTTGTCAAATATAGGCGTTAAAAAAACC
AAACATTCTAAAAATGAA
CTCCAGCCAAACAAGGTTTCTGAGAAGAGCCGCCATATTGTCACCACACGAGAGTCGCCAT
ATTGTCACCACACGAGAGC
CGCCATATTGTCACCACACGAGAGTCGCCATATTGTCACCACACGAGAGTCGCCATATTGTC
ACCACACGAGAGCCGCCA
TATTGTCACCACACGAGAGTCGCCACATACTCCACTCGAGTCGGAATCACTAGTGAATTC

DOP17
GNAATTCGATTCCGACTCGAGTTTAGGATGTGGTC
GACAAAGCAGTGCAGAAAGCTAGCCAACGGCAGTGCCTATTTAGAAAACTAGCCGCACCAG
CTGGGGTGCAAATCACAAT
ATTCTAAATAGAATATACACTGGTAGAGTGAGATCAGTGCTAGAATATGTCTCTGCAGCCTGT
GCCATGGCATCACATGC
ATAATTCTCCAAAGTTGGTCAAACGCAAAATCGGGCAGCACGCATTATCACAGGTGACATGA
GATCCACCCCAATAAAAT
TTCTGGAAACGATGACAGGTCTCCATCCTATGGAAGATAAGAGAGACTCAAAAGTGATACAG
CAGGCCGAGAAATACAAA
ACACTGAAATCGCACCACATGCCCCTCTCGAGTCGAATCACTAGTGAATTC

DOP19
GAATTCGATTCCGACTCGAGTTGCGCATGTGGCAG
CACCCAGTTCCCTTGCATGAAAGGCTGTATAGAGACGCTGGGAGCCTTGAACGGACTGCTA
CCTTCATCAGAATCATCAA
TGTGGACGTATGAAAGCGAACGAGGAAGAGGAAGACGTCGCAGTATTTCGTTGAACATGAC
TACAGTTTACTTAGTGTTT
AGCCTGTCTATTGTTCTGCCCATCACTGTAGCTAGACATTTCAAAGAAATATCAAATTGGAGG
CCAGGAAACTAGCTACA
CTTCACGCCTTTGCTAATGTCACATAAGAGATTTCTCTCTGATAACCTCCTGCAGAGCGCATC
CTCTAGAGATTCCGGCA
ATAAAATCTCTGGAACCACATGCCGGACTCGAGTCGGAATCACTAGTGAATTC

DOP20T
GAATTCGATTCCGACTCGAGGGGTTAATGTGGCAC
ATGGAAAGCAAAAACTTGATTCCCCCCAATCAGAGCGGATTTAGGAGTAACATGTCTACAGA
AGACCAGGTTACCTACAT
CGCGCAAGAAATCATCGATAGCTACCATGTCCACGAGCACACGCTAGCCGTATGGGTTGAC
ATGGAAAAAGCGTTCGACA
CAGTGTGGAGAGATGGGCTCAAATCAAAACTTCTAACTCTAGGAGTAGGAGGACATATGTTC
GAATGGCTCGGAAGCTAC
CTCGAAAACCGGACAGCTAGGGTCCGAGTGCAGAACAAAACGAGCCGAAAACGAACACTCG
AACAAGGCGTTCCTCAAGG

TGGAGTGCTGTCCCCAACGCTCTTTCTCGCATTTGTCAACGACATCAGTACATGTTTTCCCAT
GGAGGTGCACTCAGCCA
TGTACGCGGATGATCTCGCCCTTTGGGCCAGGAGTGGGTCGATAGAAACGGCCGAAGAAA
GTGTGCAATTAGCGCTCGCC
AAGCTGGAAGACTGGACGAAGCAGTGGCTCTTAAAAATCAATGGAAATAAGACCACGTACAC
CATATTCAGTCTATCCAC
CAAAGTCCTCAAGGCGACGCTAAAAATCAACGACCACATCCGCAACTCGAGTCGGTGTGGC
NCATGGAAAGCAAAAATTT
GATTCCCCCCAATCAGAGCGGGATTTANGAGTAACATGTCTACAGAAGACCAGGTTACCTAC
ATCGNCGCNAGAAATCAT
CGNTAGCTACNTGTCCACGAGCCCACGCTANCCGNATGGGTTGANTGGNAAAAGCGTCNAA
NAGTGTGGAGAGAGGGCTC
NAATCAAACTTTAACTCTAGAGTAGGAGNNATATGTTCAATGGCTCGGGAGCTACTCCAANN
CGGACAGCTAGGNCCGAT
GCAAACAAANNAGNNAAAACAACCTCNAAANGNGTNCTCAGGGGGGGGNTGCCCCAGCNC
TTTTCCATTGNCACGNANAG
ACTGGTTNCCTGGNGTGCCTCACCTGNNCGGGTACTCCCNTTGGCCNGNGGGTCCTAAANC
GCNANAANTGCATANCCCC
CNNCGG

DOP20S
GAATTCACTAGTGATTCCGAC
TCGAGGTAGCCATGTGGTCGTTGATTTTTAGCGTCGCCTTGAGGACTTTGGTGGATAGACTG
AATATGGTGTACGTGGTC
TTATTTCCATTGATTTTTAAGAGCCACTGCTTCGTCCAGTCTTCCAGCTTGGCGAGCGCTAAT
TGCACACTTTCTTCGGC
CGTTTCTATCGACCCACTCCTGGCCCAAAGGGCGAGATCATCCGCGTACATGGCTGAGTGC
ACCTCCATGGGAAAACATG
TACTGATGTCGTTGACAAATGCGAGAAAGAGCGTTGGGGACAGCACTCCACCTTGAGGAAC
GCCTTGTTCGAGTGTTCGT
TTTCGGCTCGTTTTGTTCTGCACTCGGACCCTAGCTGTCCGGTTTTCGAGGTAGCTTCCGAG
CCATTCGAACATATGTCC
TCCTACTCCTAGAGTTAGAAGTTTTGATTTGAGCCCATCTCTCCACACTGTGTCGAACGCTTT
TTCCATGTCAACCCATA
CGGCTAGCGTGTGCTCGTGGACATGGTAGCTATCGATGATTTCTTGCGCGATGTANGTAAC
CTGGTCTTCTGTANACATG
TTACTCCTAAATCCGCTCTGATTGGGGGGGAATCAAATTTTTTGCTTTCCCATGTGCCACACC
GACTCCANTTGNCNGGA
TGTGGGNCNNTNGATTTTTANCGTNGCCTTAAGGACTTTTGGNGGANANANTGNANATGGTG
CCNNNGGNCNTNTNTNAT
TGATNNNAAAGCATGNTTCCCCCNNNCTCCANNNNGAACGCGTTANNGNNCNTTCACTCNC
TGTNNNATCANCNCNCCTG
GGCNANGGNNANNANTCGGCNNCNNGCNGNTNANCCCNCGNGANAANGGCNNGNGNCGN
TGANNATCNACANNCCCNTCG
NTNCNNNGNCCCNCNGNNCACNNGCTCNANGNGTTNACCACCCANANNNNCNNGTCTNCA
NNCCTGTATNACNAACANGN
NGCNCNCATCNCGNCGCGTCNNNNCNNTCTNNCNNNTCCNTNTANCNANNNCNNNNNNA

DOP21
GAATTCACTAGTGATTCCGACTCG
AGATGGGGATGTGGCTTGTATTCAGGAGACTTACCTGAAACCTGAGAGGAGCTTCTGGGTC
CCTGGATACCAGACATTCC
GACAGGATCGTGACGGACACAAGGGTGGCGTGCTCATCCTGGTAAAGAGCCACATGAACAA
CTCGAGTCGGAATCGAATT
C

DOP23R
GAATTCACTAGTGATTCCGACTC
GAGGTGCGCATGTGGTCGTTGATTTTTAGCGTCGCCTTGAGGACCTTGGTGGATAGACTGA
ATATGGTGTACGTGGTCTT
ATTTTCATTGATTTTTAAGAGCCACTGCTTCGTCCAGTCTTCCAGCTTGGCGAGCGCTAATTG
CACACTTTCTTCGGCCG
TTTCTATCGACCCACTCCTGGCCCAAAGGGCGAGATCATCCGCGTACATGGCTGAGTGCAC
CTCCGTGGGAAAACATGTA
CTGATGTCGTTGACAAATGCGAGAAAGAGCGTTGGGGACAGCACTCCACCTTGAGGAACGC
CTTGTTCGAGTGTTCGTTT
TCGGCTCGTTTTGTTCTGCACTCGGACCCTAGCTGTCCGGTTTTCGAGGTAGCTTCCGAGCC
ATTCGAACATATGTCCTC
CTACTCCTAGAGTTAGAAGTTTTGATTTGAGCCCATCTCTCCACACTGTGTCGAACGCTTTTT
CCATGTCAACCCATACG
GCTAGCGTGTGCTCGTGGACATGGTAGCTATCGATGATTTCTTGCGCGATGTAGGTAACCTG
GTCTTCTGTAGACATGTT
ACTCCTAAATCCGCTCTGATCGGGGGGGAATCAAATTTTTGCTTTCCATGTGCCACATCAACTC
CTCGAGTCGGAATCGAA
ATTC

DOP27
GAATTCACTAGTGATTCCGACTCG
AGGGGTTCATGTGGAACACGTATAGTCATATTCATGCATAAACCAACCCTCGGCCACATGGG
GAACACGTATAGTCATAT
TCATGCATAAACCAACCCTCGGCCACATGGGGAACACGTATAGTCATATTCATGCATAAACC
AACCCTCGGCCACATGGG
GAACACGTATAGTCATATTCAAGCATAAAACAGCCCCCGGCCACATTGCTCCCTCGAGTCGG
AATCGAATTC

DOP31R
GAATTCACTAGTGATTCCGACTC
GAGGCTGTCATGTGGTCGTTGATTTTTAGCGTCGCCTTGAGGACCTTGGTGGATAGACTGAA
TATGGTGTACGTGGTCTT
ATTTTCATTGATTTTTAAGAGCCACTGCTTCGTCCAGTCTTCCAGCTTGGCGAGCGCCAATTG
CACACTTTCTTCGGCCG
TTTCTATCGACCCACTCCTGGCCCAAAGGGCGAGATCATCCGCGTACATGGCTGAGTGCAC
CTCCATGGGAAAACATGTA
CTGATGTCGTTGACAAATGCGAGAAAGAGCGTTGGGGACAGCACTCCTCCTTGAGGGACGC
CTTGTTCGAGTGTTCGTTT
TCGGCTCGTTTTGTTCTTCACTCGGACCCTAGCTGTCCGGTTTTCGAGGTAGCTTCCGAGCC
ATTCGAACATATGTCCTC
CTACTCCTAGAGTTAGAAGTTTTGATTTGAGCCCATCTCTCCACACTGTGTTGAACGCTTTTT
CCATGTCAACCCATACG
GCTAGCGTGTGCTCGTGGACATGGTAGCTATCGATGATTTCTTGCGCGATGTAGGTAACCTG
GTCTTCTGTAGACATGTT
ACTCCTAAATCCGCTCTGGTTGGGAGGAATCAAATTTTTGCTTTCCATGTGCCACATCCAGCT
CTCGAGTCGGAATCGAA
ATTC

DOP34
GAATTCACTAGTGATCCGACTCGAG
GGGCATATGTGGCTGTTTACCAGTATGAGCATGCCACCCTTGTGCCCATCTCGATCTGGTCG
GAATGTCTAGTATCCGGG
GACCCAGAAGCTCCTCTCAGGTTTCAGGTGAGTCCCTTGAATACAAGCCACATCCCACCCTC
GAGTCGGAATCGAATTC

DOP37
GAATNCACTAGTGATTCCGACT
CGAGTTGGGGATGTGGAGCAATGAGATCAACGCCCACATCGGCATGTGAGATCAACGCCCA
CATCGGCATGTGAGATCAA
CGCCCACATCGGCATGTGAGACCAACGCCCACATCGGCATGTGAGACCAACGCCCACATCG
GCATGTGAGATCAACGCCC
ACATCGGCATGTGAGATCAACGCCCACATCGGCATGTGAGATCAACGCCCACATCGGCATG
TGAGATCAACGCCCACATC
GGCATGTGAGATCAACGCCCACATGCCCGCCTCGAGTCGGAATCGAATTC

DOP38R
GAATTCACTAGTGATTCCGACTCG
AGTGGGGAATGTGGTCGTTGATTTTTAGCGTCGCCTTGATGACCTTGGTGGACAGACTGAAT
ATCCTGTACGTGGTCTTA
TTTCCATTGATTTTTAAGAGCCACTGCTTCGTCCAGTCCTCCAGCTTGGCGAGCGTTGATTG
AAAGCATTGTTCGGCCGT
TTCTATCGACCCACTCCTGGCCCAAAGGGCGAGATCATCCGCGTACATGGCAGAGCGCACC
TCTCTGGGAAAACATGTAC
TAATGTCGTTGACAAATGCGAGAAAGAGCGTTGGGGACAGCACCCCTCCTTGGGGGACTCC
CTGTTCGAGTGTTTGTTTT
CGGCTCGTTTTGTTCTGCACTCGGACCCTAGCCGTCCGGTTTTCGAGGTAGCTGCCGATCC
ATTCGAACATATGTCCTCC
GACTCCTAGACTTAGGAGTTTGGACTTGAGCCCATCTCTCCACACTGTGTCGAACGATTTTT
CCATGTCAACCCATACAG
CCAGTGTATGCTCGTGGACGTGGTAGCCATCGATGATTTCTTGCGCGATGTAGATTACCTGG
TCTTCAGTTGACATGTTA
CTCCTAAATCCGCTCTGGTTGGGAGAAATCAGATTCTTGCTTTCTATGTGCCACATCCCGCA
CTCGAGTCGGAATCGAAT
TC

DOP43
GAATTCACTAGTGATTCCGACTC
GAGGCGTTGATGTGGGGCTCCCACATGGGTCCCTCCCAGCAAACGTTATGCACGACATCAG
TACAGCAACTTGACAAGTT
TCAGTAAGAGTCGCAGGCATCGCGCGCAAAGAACAGATGTGCAGAAAAGTCGCGACTCTAA
CATCCACCCGTCCGCCACA
TAACAGCCTCGAGTCGGAATCGAATTC

DOP46
GAATTCACTAGTGATTCCGACTC
GAGGTGTGATGTGGCCCAGGTTGAAAGTCAGATGTCGACAAAGATCACCCGAAGCTGATAA
CCCGACAGCACCAGAGGTC
ACGTCAAATCTACAATCTCTGCAGACTTCACATTAATCCCAGAAATGCTACCGCAGCTTTGAG
GCATTCGGACATGAGAA
AAATCAACACACGTGAGAGGTCGCGCGGGCTTGCGTACACCAGGTTTGTGCGTGTCTATGA
CGATGGTAGTCTCACGAGG
ACTCCCTCTAGCGCGATGGCGAAAGCGCTCCTGAATGTCGAGGTGTGTCTGCTGACAGATC
CTGATCGCCTCCGTGTTCT
CCTGGGCAGGCTGATTCCCTGGTTCTGGGAACTAGCCGTCTCTGGGACATTCCCACCTTCC
AGCCAACACTTGATTTACC
TTTCTTTCTAGTCTTCCTCACATCGACGTCTGCCTGTCCACATACACCACAGCTGCCTGTCCA
CATACACCACAGCTGCC
TGTCCACATACACCACAGCTGCCTGTCCACATGCGCCGCTCGAGTAATCGAATTC

DOP47R
GAATTCACTAGTGATTCCGAC

TCGAGGGGGGAATGTGGTCGTTGATTTTTAGCGTCGCCTTGAGGACCTTGGTGGATAGACT
GAATATGGTGTACGTGGTC
TTATTTCCATTGATTTTTAAGAGCCACTGCTTCGTCCAGTCTTCCAGCTTGGCGAGCGCTAAT
TGCACACTTTGTTCGGC
CGTTTCTATCGACCCACTCCTGGCCCAAAGGGCGAGATCATCCGCGTACATGGCTGAGTGC
ACCTCCATGGGAAAACATG
TACTGATGTCGTTGACAAATGCGAGAAAGAGCGTTGGGGACAGCACTCCTCCTTGAGGGAC
GCCTTGTTCGAGTGTTCGT
TTTCGGCTCGTTTTGTTCTTCACTCGGACCCTAGCTGTCCGGTTTTCGAGGTAGCTTCCGAG
CCATTCGAACATATGTCC
TCCTACTCCTAGAGTTAGAAGTTTTGATTTGAGCCCATCTCTCCACACTGTGTCGAACGCTTT
TTCCATGTCAACCCATA
CGGCTAGCGTGTGCTCGTGGACATGGTAGCTATCGATGATTTCTTGCGCGATGTAGGTAAC
CTGGTCTTCTGTAGACATG
TTACTCCTAAATCCCGCTCTGGTTGGGAGGAATCAAATTTTTGCTTTCCATGTGCCACATGAA
ACTCTCGAGTCNGAATC
GAANTTC

All DNA sequences were put through both nucleotide and translation BLAST algorithms. The major repeat of 700 bp had homology to an unknown but highly-repetitive LINE element, that in all R (repeat) or S (SP6 primer) or T (T7 primer) clones, contain a long ORF for an unknown but insect-like variety of reverse transcriptase proteins.

> Put any of the R-S-T seqs. through translation BLAST, and you will be amazed. There is almost no nucleotide homology however to anything in the database. A clear pattern is seen only when the bases are conceptually translated into protein. A few other ORFs show up in non-R sequences, weakly related to lower mammals such as mouse or rabbit (i.e. proteoglycans).

Thus, even following the DOP-PCR analysis, the claw DNA appeared to retain the anomalous (extraterrestrial?) features. As in the previous PCR analysis, most of the DNA did not match anything in Genbank.

Phase 6: MDA Analysis

In an extraordinary decision, one of the labs involved in the project decided to try yet another technique that was an alternative form of DNA amplification but completely different from the polymerase chain reaction (or from DOP-PCR). The technique called multiple displacement amplification (MDA) is considered very effective for whole genome amplification. MDA uses Φ29 DNA polymerase and random exonuclease-resistant primers. DNA is amplified in a 30°C reaction not requiring thermal cycling. This is made possible in part by the great processivity of Φ29 DNA polymerase, which synthesizes DNA strands 70 kb in length.

A fuller description of the power and versatility of the MDA technique can be found in: "Comprehensive human genome amplification using multiple displacement amplification**.** Proc Natl Acad Sci U S A. 2002 Apr 16;99(8):5261-6.

The following is a narrative summary from the lab that attempted the MDA of the claw DNA.

> Here are 28 randomly-primed claw "B" MDA seqs (for reasons of space the actual DNA code is not shown), just received from seq. facility an hour ago. Two clones 34 and 42 labeled "y" are probably cloning artifacts, and can be discarded. Eight other clones labeled "x" seem to be repetitive DNA, and hence may not be very informative. The other 18 will be analyzed as soon as BLAST for DNA and protein can be done. All good clones should start with Eco RI GAATTC, and also end with Eco RI GAATTC, as an internal test of sequencing quality.
>
> Have just analyzed the new MDA sequences attached below. There are currently 26 good ones. MD34y and MD42y were cloning artifacts and have been deleted. Of those 26 good clones, only two (MD17 and MD53) code with certainty for any sort of plausible protein. A translation-BLAST search shows that those two clones code for different parts of the non-LTR reverse transcriptase found earlier in DOP PCR work. MD17 codes for amino acids 149-228, while clone MD53 codes for amino acids 262-298.
>
> The other 24 MDA clones contain various kinds of repetitive DNA, which do not match anything in the Genbank database. Hence by this fully random, representative MDA method, claw genome B appears to contain 24/26 = 90% junk DNA, 2/26 = 8% retrotransposons, and less than 4% protein-coding genes, none of which were detected in 26 clones. Clone MD07 could possibly code for TFIID protein, but this is not certain.
>
> These efforts should conclude our analysis of claw DNA as studied here. Further cloning and sequencing may be done on DOP PCR samples A, B, B+, or the new MDA sample (once Taq-modified), at another lab if so desired.
>
> In summary, we have performed DNA amplification by three means: (1) specific PCR primers for cytochrome b, cytochrome c, or 12S RNA; (2)semi-random DOP PCR primers; and (3) purely random MDA primers. The vast majority of the 114 clones so obtained, show no match at either a nucleotide or translated protein level to Genbank.
>
> Of all clones (1), we found just one strange cytochrome c gene (CY57), two partial genes for a strange splicing factor II (CY07, CY59), and two strange 12S sequences (CS63, CS68). Of all clones (2), we found a variety of non-LTR reverse transcriptase proteins (DOPxxR), which make up a prominent 700 bp genomic repeat. Of all clones (3), we found just two non-LTR proteins (MD17, MD53) of the kind seen in (2), with a lot of repetitive DNA.
>
> As a control for the specificity of amplification (1), we found the same internal DNA sequence in duplicate clones CY07 and CY59, despite different end-primers. We also found high fidelity of protein coding over 220 amino acids in amplification (2); plus accurate amplification of a

> human control in another (2). Finally, we saw no PCR amplification of the 12S gene in MDA-amplified (3), meaning that possible impurities which could inhibit such amplification in the raw sample may be ruled out. It does not amplify in MDA-pure form with standard primers, because it contains no standard genes.

Further studies of claw sample A may confirm this picture of claw sample B in greater detail. There should be in total: 20 sequences CY01 to CY60 (cytochrome c primers); 5 directly cloned sequences A4 to B8; 23 sequences CS29 to CS95 (12S RNA primers); 16 sequences CB01 to CB45 (cytochrome b primers); 24 sequences DOP01 to DOP47R (semi-random DOP primers); and 26 sequences MD04 to MD58 (fully random MDA primers).

Phase 7: 18S, 5.8S and 5S Ribosomal DNA analysis

So far all of the techniques had yielded very puzzling results that were still consistent with an anomalous (extraterrestrial?) biological sample. In an attempt to definitely determine the DNA sequence, it was decided to amplify regions of DNA that were probably the most highly conserved on Earth.

18S RNA, 5.8S RNA, and 5S RNA are present in the nucleus but not in the mitochondrion. Those three multicopy genes are not usually studied for Earth evolution, because they do not vary enough to be useful (at least 5.8S and 5S). The primer sequences which were used to get nice bands from the claw (34 cycles, 1.5 mM MgCl2, anneal 50 C) were as follows: The two sets of 18S primers were taken straight from the Small Subunit Ribosomal RNA Database, whereas 5.8S and 5S primers were designed de novo, using compilations of known sequences. In effect, this was the design of an entirely novel PCR reaction that was employed using arguably the most conserved DNA on Earth.

> Twenty clones were sent to the sequencing facility this morning: 4 each of 5 genetic regions (18S left, claw A and B separately; 18S right, claw A and B combined; 5.8S, claw A and B combined; 5S, claw B only, since A gave no band). Another 16 clones are spare in the freezer if we need them. So far everything has gone okay. There is a chance that some clones will not sequence owing to lots of secondary structure in these molecules; in which case they can be re-submitted using a different sequencing primer on the other end.

Completely Unexpected Results Obtained from this new analysis

Various mollusks for 18S-left-A, 18S-left-B, 5.8S-AB and 5S-B. However 18S-right-AB shows no match anywhere. However, for the first time an unambiguous result emerged from two separate claw DNA preps ie homology with mollusk sequences.

Those mollusks secrete a defensive fluid through their hyperbranchial glands, cf. duct in sample.

> The rRNA clones came out well technically, however the sequences obtained show a variety of forms, all of which match somewhat the rRNA of various common snails or slugs. Not the same snail or slug for each kind of rRNA, but that family in general. Typically Gastropoda as the class of marine snails, some of which are found in California.

This completely surprising result emerged from the ribosomal analysis, arguably the most conserved sequences on Earth. The results although surprising were unambiguous. The evidence finally pointed in a consistent direction. The object was possibly a dried mollusk—a slug or snail. Given the large series of previous DNA analyses that had all pointed towards an anomalous identity, why was this not picked up by standard Mt DNA PCR, by DOP-PCR or by the MDA technique? The answer was revealed in Phase 8.

Phase 8: Investigation of the evidence for Mollusk DNA

(A) Consultation with an expert in mollusk molecular biology

> Most molluscs are very 'mucousy.' The mucous is a major problem in molecular work, as it inhibits most enzymes, as it binds to proteins indiscriminately. The solution I have found is to use cesium chloride centrifugation after phenol extraction. Obviously, you need a fair amount of starting material (several animals) for this. The extraction is conducted in the presence of heparin, which seems to prevent the mucous from binding.
>
> Also, I use liquid nitrogen to freeze the slug, and then crush it in a pestle and mortar, this also prevents mucous production. The second option is to use the CTAB technique, which was originally developed for plant DNA extractions, which also suffer from (heavy) polysaccharide contamination. This will help remove some of the mucous and doesn't require much starting material. Thirdly, I came across this reference, I haven't used it yet, but it might be a useful procedure, again it doesn't require much starting material.

Returning to the Phase 4 work carried out by the Canadian lab. The above statements were consistent with a high level of inhibitors in the claw sample as found by the Ontario lab. It should also be noted that one of the original PCR sequences in Phase 1 showed some homology to a member of the snail family.

(B) Consultation with Los Angeles County Natural History Museum. A senior official from this institution replied as follows to our request for identification of the biological sample:

> After careful examination of the photographs, and the results of the SEM-EDS, absorption spectroscopy, and DNA analyses, I have concluded that the specimen you refer to me in your letter of May 28, 2003 is a common California slug of the family Limacidae.
>
> The photographs are self-evident and clearly illustrate a slug that has dried out. I can recognize all the parts of a slug anatomy, including the foot, the mantle with longitudinal ridges, and the dorsal hump in which a pneumostone is visible. The hairs attached to it have probably been pasted to the specimen when it was alive and then become permanently attached when the mucous of the mollusk solidified.
>
> The absorption spectroscopy confirms the presence of proteins related to collagen, which are the main components of slug's epidermis. The SEM-EDS analysis results show high contents of elements expected to be found in a living organism.
>
> One of the strongest evidences that point towards the slug identification is the DNA analysis. The results you sent to me show that the closest relatives of which gene sequence data are stored in Gene Bank are all mollusks. For instance the BLAST of the 5S rRNA gene sequence shows that this sequence is very similar to that of Arion rufus (the common European slug) and Helix pomatia (the common garden snail, introduced in the US). The 18S sequence, show high affinity values with a number of mollusks, including Limax maximus (the giant garden slug, introduced in the US), Acusta despecta steboldii (a Japanese snail), and Cepaea nemoralis (a European snail), as well as many other species. Because of the small number of California species of slugs that have been sequenced and stored in Gene Bank it is not possible to determine what species this is exactly.
>
> It has been my pleasure to assist you to identify this specimen and do not hesitate to contact me if you need further information.

Therefore, the combination of DNA analysis and positive identification by a mollusk expert has established beyond doubt that the putative extraterrestrial biological sample was in fact a dried mollusk.

Conclusions and Lessons Learned

There is little doubt that the sample found embedded in the towel in California in September 2000 is a dried up mollusk. In addition to identifying the nature of the object, we have described the DNA of a previously unknown mollusk. The DNA can and will be added to the GenBank global DNA database. Further, part of this investigation involved the invention of a new PCR for nuclear DNA analysis.

It could be argued that the above-described project is the most comprehensive and wide-ranging study of slug/snail DNA ever conducted. It is also the most comprehensive molecular investigation into a putative extraterrestrial biological sample ever conducted.

The use of cutting-edge molecular techniques were necessary to resolve this puzzle. Thus the investigation of this anomaly using scientific methodology went far beyond any routine DNA analyses conducted previously in anomaly research.

Throughout this investigation there were several branch points when suggestions were made to publish the data, even in preliminary form, as a DNA analysis of an extraterrestrial biological sample. However, caution ruled and it was decided by all team members to maintain a relatively low profile until a definitive result could be obtained. In retrospect, this decision was fortunate.

The lessons from this intensive, lengthy and very expensive investigation are simple and can be generalized for all anomaly investigations: (a) always follow the scientific data, (b) resist the temptation to tie purported physical evidence with a nearby anomaly until the analysis is fully completed, (c) resist the impulse to publish (or publicize) during an ongoing investigation.

It should be noted that the original anomalies reported by the family in their house in California remain unexplained and are still being reported and investigated.