Author Responses to Referees

Referee #1 (Remarks to the Author):

I must state at the outset that I am not a geneticist, and hence not fully qualified to evaluate the DNA data in proper critical fashion. Bottom line: there is obviously the backbone here of a paper that includes a lot of interesting data, but it needs a lot of work before it can be accepted.

1. My main advice (discussed further below): include better graphics (especially a similarity tree/distance tree/phylogram/cladogram of some sort).

Authors' Response: We have added 9 phylogenetic trees and added more data in the form of whole genome sequencing.

2. Remove the problematic conclusions on taxonomic status and the unnecessary section on recent hominin discoveries, and tidy up the nomenclature and wording.

-- Discussion of fossil species is probably irrelevant

The manuscript opens with a section on how recently discovered fossil hominins have changed our views on the hominin diversity of the recent past. While it seems logical to mention these taxa in passing at least somewhere in the manuscript, I think it's a bad idea to start the article off with a discussion of these forms - is this really the area of investigation that most requires review when writing about the possible existence of sasquatch? I would say no. The authors are meant to be addressing the possible existence and identity of an extant population, not adding to the

roster of fossil forms. In other words, the discussion of Neanderthals, Denisovans and so on seems like unnecessary padding. There are some basic mistakes concerning terminology and the use of binomials (e.g., the name _Homo floresiensis_ is incorrectly written with capital first letter on the species name). Neanderthals do not have the name they do "because of their skeletal morphology" (rather, they are distinguished by their skeletal morphology).

Authors'Response: We removed the taxonomic recommendation. Additionally we removed the recent hominin discoveries (Neanderthal and Denisovan).

3. It would seem far more appropriate so far as I can see to begin with a discussion of the controversy surrounding the purported existence of 'mystery hominoids', and to perhaps allude to some of the other technical studies that have claimed to find evidence for the existence of these alleged creatures.

Authors' Response: The discussion of the controversy surrounding the existence of these hominins was moved to the introduction and expounded on, including photos and video. References were added from other scientific papers and books published addressing the existence or the discussion of the existence of these hominins.

4.-- Demonstrate with better clarity what the DNA samples represent

It is stated in the discussion of the collected DNA samples that those "not consistent with _Homo sapiens sapiens_" were then evaluated further. I feel the authors must elaborate on what it was that made it clear that the samples were not part of _H. s. sapiens_ - the normal graphic way of representing this is, of course, a gene tree, distance tree or cladogram of some sort. If the authors

are saying that the samples are from a hominid and hominin, but come from a taxon outside of _H. s. sapiens_, they need to state it more clearly and provide better evidence and clearer graphics. Any diagram should also make it clear that other mammals (a useful list is included in the manuscript) were definitely out-groups relative to the group that includes the 'mystery' samples and those of definite hominins.

Authors'Response: We clarified and removed the verbiage stating "not consistent with _Homo sapiens sapiens_". We have added 6 mitochondrial phylogenetic trees which clearly show the samples had modern human maternal origins and were generated via conventional sequencing as well as phylogenetic trees extracted from next generation whole genome sequences. We provide more evidence by sequencing 3 whole genomes at the University of Texas core lab and using a subsample of extracted reads. The reads were assembled to create a consensus sequence using the human chromosome 11 as a reference. These concatemers (supercontigs) were used to find sequence homologs and generate 3 phylogenetic trees, one for each whole genome sequenced.

5. -- Terminology seems odd and needs changing

The terminology used throughout this manuscript, and the conclusions the authors reach, seem inappropriate in view of the evidence. Indeed, the title tells us that evidence for a 'new species' is presented, yet the authors actually end up naming a new 'subspecies'. I am definitely of the opinion that the naming of a new taxon seems inappropriate at this stage (it is likely to be about as accepted as Meldrum's suggested ichnotaxonomic name _Anthropoidipes ameriborealis_), and I would also add that the chosen name ('_Homo sapiens feralis_') is odd and highly problematic (use google to see what I mean). It is stated throughout the ms that the animal is of hybrid origin. If this is so, it is highly debatable as to whether or not taxonomic novelty is warranted.

Authors'Response: We have re-written and re-arranged the manuscript to use better terminology and also have removed any taxonomic references other than to call the hominins Sasquatch.

6. I would like to know exactly what is meant by those statements noting that the "paternal lineage [is] completely unknown", as the authors seems to be introducing a new layer of mystery to their conclusions. It seems radical enough that they are positing a hybrid origin for this putative animal, but are they also invoking the existence of an additional animal that was involved in the proposed hybridisation event? This all seems very peculiar and I am not convinced that the evidence presented in this manuscript explains it adequately.

Authors'Response: By sequencing 3 whole genomes that failed to align with any animal or hominin found in NCBI, there is no other conclusion other than that the paternal nuclear DNA origins are unknown. Previous data suggested that novel sequences and high failure rate were found in the nuclear DNA sequencing and STR testing., The addition of the three genomes further supports our initial findings. The nuDNA and mtDNA origins of the Sasquatch are discordant, with mtDNA indicating human maternal lineage. Analysis of the 3 next generation whole genome sequences and analysis of preliminary phylogeny trees from the Sasquatch indicate that these individuals possesses an anomalous mosaic pattern of nuclear DNA comprising sequences that are distantly related to primates interspersed with sequences that are closely homologous to humans.

7. I would also suggest that the phrase "unknown morphology of the hair" is inappropriate: rather, the authors are reporting a morphology that is novel.

Authors' Response: We changed the phrase "unknown morphology of the hair" to "microscopic morphology of hairs classified as "novel" and "novel hairs".

8. All in all, this manuscript seems to report the discovery of a novel North American hominin lineage as determined by DNA analysis (so far as I can tell, the substantial discussion of that data is appropriate and does consider most relevant factors/avenues of investigation). This is, obviously, potentially, a significant, Nature-worthy discovery. But it is marred by poor choice of presentation (that is, the absence of a tree that immediately conveys the position of the samples to those of other taxa/populations), unnecessary discussion of fossil forms, and inappropriate, confusing and rather naive proposals concerning nomenclature and the alleged hybrid origin of the alleged animal. I would definitely like to see this manuscript 'salvaged', but it would need a thorough revision and re-organisation. I wish the authors the best of luck in their continuing efforts.

Authors' Response: We re-organized and revised the manuscript and added extra data to support our original findings.

Referee #2 (Remarks to the Author):

The authors analyse some biology samples (mainly hairs) putatively belonging to the Big Foot; they analyse the hair microscopica structre, the mtDNA, forensic STRs markers and some few

nuclear genes -including the MC1R- and conclude the results indicate the existence of a previously unknown human species, which they call Homo sapiens feralis. I am not going to go into the specific details of the results, which are quite confusing and methodologically debatable, but will explain what the average molecular biologists would do in two different scenarios related to the problem presented in this work.

1. Identify a biological sample of unknown specific origin, say, hairs, or coprolites or blood stains. This approach is widely used in zoology, for instance, to distinguish between wolf and dog after an attack to a sheep herd. Usually the people uses universal primers to amplify a diagnostic DNA fragment that could help in the identification of numerous species. 16S is probably the most widely used, although in specific situations and for trying to find out the precise match, you may want to design additional tests on cytb, nuclear SNPs, complete mtDNA, etc. This has not been done here; the use of human primers will amplify human mtDNA, as has been the case. Moreover, the fragmentary mtDNA data does not support any unknown hominin lineage, because the haplotype/haplogroup attribution fits well in what is already known of the modern human mtDNA phylogeny.

Authors'Response: The evidence samples were screened with not only published cytochrome b primers utilized for species identification but also THR/DHL which are universal mammalian primers in the literature utilized across HV1 for species identification. This is stated in the manuscript however, we have made it more prominent. We also had previously sequenced a number of whole mitochondrial genomes which are in the manuscript. We have increased that number. The initial mtDNA screening of the samples with the universal primers yielding only human results. These findings are consistent with previous attempts by other labs and scientists to validate the existence of Sasquatch through mtDNA which we have documented in the

Discussion. There was no DNA of any other species in the mtDNA in any of the samples utilized in this study. The whole mitochondrial genomes were consistent with modern human as were the samples that yielded enough mitochondrial sequence to assign a haplotype but not a whole mtDNA genome. Even the samples with very little DNA (not enough to achieve a haplotype) were screened with a short HV2 sequence and gave only human sequence. This mtDNA alone did not support an unknown hominin, however the same extractions, when sequenced on various nuclear loci and amplified with PowerPlex 16, yielded unexpected and aberrant results with some loci yielding normal human sequence along with novel sequences not found in genetic depositories. To further support our findings, we have sequenced 3 next generation whole genomes from which 3 mitochondrial DNA sequences were extracted and were homologous with human and the previous sequencing. The chromosomal sequences were discordant and novel, supporting the original findings.

2) Another scenario. We want to prove that the specimen we are studying is an undescribed, new species. Of course the definition of what a species would look like genetically is a tricky question (see for instance, the Denisovans), but you would be expected to generate a huge amount of genomic data (if possible, the complete genome), construct phylogenetic trees and show that your specimen represents a deep, undescribed clade in the current phylogeny. Even so, some might argue that genetic divergence is not equivalent to species difference, but at least you would have a good point to support your claims. This has not been done here, and the analysis of three random nuclear genes cannot be used for the purpose of defining a new species.

In short, the conclusions are not supported by the methodology. I would suggest the authors to generate complete mtDNA genomes and, if possible, suficient nuclear data, even from shotgun

approaches, and build phylogenetic trees with all possible mammal mtDNA genomes and nuclear data available at genbank.

Authors' Response: We added more data to the manuscript by sequencing 3 whole genomes using next generation sequencing. Previous data suggested that novel sequences and high failure rate were found in the nuclear DNA sequencing and STR testing. The addition of the three genomes further supports our initial findings. The nuDNA and mtDNA origins of the Sasquatch are still discordant, with mtDNA indicating human maternal lineage. Analysis of the 3 next generation whole genome sequences and analysis of preliminary phylogeny trees from the Sasquatch indicate that these individuals possesses an anomalous mosaic pattern of nuclear DNA comprising sequences that are distantly related to primates interspersed with sequences that are closely homologous to humans.

Referee #3

(Remarks to the Author) I believe that among the most important abilities defining a real scientist is his/hers ability to stay open-minded and accept that the world may be radically different from common believes as long as such radical claims are supported by sufficient hard scientific evidence. Such a radical claim is presented by Ketchum et al. that basically propose the existence of new contemporary sub-species of homo in North America termed Homo sapiens feralis. The claim is based on morphological and DNA based analyses of samples such as hair and bark shavings, tissue, toenail, saliva, and blood samples taken by various people in areas where unusual bipedal hominin like creature has been visually observed. Based on these analyses the authors claim to have found evidence of a new contemporary hominin with unique hair,

strange, nuDNA, and human mtDNA. An exceptional claim such as this demands for exceptional convincing evidence - something the authors do not have:

1. First the data does not make logical sense. Mitochondrial DNA genomes being identical to that of contemporary humans can only be explained be relative recent interbreeding between this new hominin and woman of Caucasian descent. Did such interbreeding go back many thousands of years one will expect differences to modern mtDNA genomes. The mtDNA results are hardly explainable unless one believe that American woman of Caucasian descent (within the last 200-300 years as its America) runs around in the forest having sex with a undiscovered hominin and leavening the baby to their care take of the new hominin (as the rest of us have not heard about such hybrid babies yet the baby must be send of). It is also sticking that the entire mtDNA lineages of this new hominin is poorly human. One would expect at least some mtDNA genomes coming out as being accordance with this being a new hominin or if nothing els some of the mtDNA being Native American (everything being equal they have been in America more than 10,00 0 years).

Authors' Response: We now have two samples that did yield American Indian haplotypes. These were late arriving samples and were added to the manuscript. As far as when this species arrived in the United States, we do not know, however due to H haplotypes in their mitochondrial DNA, the age of these hominins is less than 15,000 years. However, they could have arrived in the United States before Native American peoples according to the Solutrean Theory now added as a reference in this manuscript. As far as the mitochondrial DNA being homologous to human, we used next generation whole genome sequencing on three samples. The mitochondrial DNA sequences were extracted from the genomes and all three were

consistent with the same human mtDNA sequence previously sequenced in the beginning of our study. The nuDNA however, was a mosaic of human sequence interspersed with novel sequence related to primate lineage. This supports the previous findings reported in the original manuscript.

2. Secondly, PCR based methods and methods used for SNP detection by the authors are known to be highly unreliable when applied to minor amounts of degraded DNA (I know that personally for the SNP detection approach and simple PCR). This is especially the case for nuDNA templates that are more prone to be affected by damage than those of mtDNA (due to copy number differences). In fact copy number differences can very well explain the DNA findings i.e. the specimens are human in origin, why the authors amplify mtDNA genomes marching contemporary Caucasians. The nuDNA , however, is too poor quality for amplifying the attempted sequence length and creating PCR artifacts. It makes good sense even for the hair sharft samples that by nature has degraded DNA (even when taken "fresh") . I am sorry but this appears to be a much more straightforward scenario than having a previously undetected (by science) hominin sub-species running around in the forest mating with Caucasian woman. I am by no means convinced that this has anything to do with a new hominin subspecies. To make a compelling case I need seeing mtDNA genomes and large numbers of nuDNA sequences that points in direction of a new hominin species i.e. ape or human like without being identical to know species.

Authors' Response: We have now included Figures 7 and 13 that show the quality of the DNA on yield gels. There was little to no smearing and the DNA was pristine. This supports the fact

that the DNA was not degraded enough to mar the results. We sequenced 30 mitochondrial whole and partial genomes and they were all homologous with modern human mtDNA. We also sequenced several nuclear loci from the same extractions which encompassed a number of long sequences up to 900 bases. It is difficult to amplify and sequence long amplicons with degraded DNA. In order to ascertain if the novel nuclear DNA was an anomaly, we used next generation sequencing to generate 3 whole genomes to determine if the nuclear DNA was indeed novel. If the DNA was badly degraded, the sequencing of whole genomes would have been impossible. The 3 extractions utilized for next generation whole genome sequencing passed all quality controls in our laboratory and the university core laboratory prior to sequencing. The libraries generated from these samples also passed stringent quality control measures in the core laboratory prior to the next generation sequencing. If the libraries had not passed QC, the sequencing would not have been performed. We have furnished 9 phylogenetic trees (mitochondrial and nuclear) to support the results of this study.

3. I'm not an expert on hair morphology but I expect that identification mistakes are made.

Authors' Response: The hair expert that examined the hair is a forensic trace evidence supervisor in a large forensic laboratory. He performs hair analysis and testifies about his findings in court on a daily basis. He also testifies concerning and examines hair from many species of animals as well as human. Furthermore, he had a wide collection of animal hair standards with which to compare these hair samples utilized in this manuscript.

Referee #4 (Remarks to the Author):

The authors seek to identify the species of 130 unknown, though purportedly hominin, hair and tissue samples. The samples are subjected to multiple forensic and genetic tests, including hair analysis, mitochondrial and nuclear genome sequencing and electron microscopy. They postulate that, while the mitochondrial genomes of all tested samples are conclusively human, discrepancies in Y-chromosome STR and amelogenin amplification, lack of sequence homology to known species, as well as structural abnormalities in DNA viewed by electron microscopy are indicative of an unusual hominin source. They identify this source as a new species, which they call Homo sapiens feralis.

Extraordinary claims require extraordinary evidence. At no point do the authors provide adequate evidence to support their outrageous claims. The paper suffers from a myriad of faults, some of the most egregious being:

1. Results are not documented in any adequate detail. For example, the DNA sequences determined are not given, rendering any coherent understanding of the results impossible.

Authors' Response: Sequences not shown in mutation reports have been added to the Supplemental Data. Mutation reports are now added to the manuscript in the Supplemental Data allowing transparency of sequences previously reported as well as new sequencing that has been added.

2. Failures in tests such as SNP analysis, amplification and electrophoresis are taken as evidence for differentiation, when they are likely explained by DNA degradation or contamination. In one instance the authors attempt to replicate DNA degradation by leaving a blood sample at room temperature for 4 days and using the sample as a positive control in further testing. They fail to assess the extent of degradation quantitatively, making this positive control uninformative.

Authors' Response: We have added Figure 13 to show the level of degradation of the human control sample in comparison with some of the Sasquatch samples in the study. We have added 3 whole genomes that were successfully sequenced using next generation sequencing technology. These genomes supported our previous data and were novel. Furthermore, we have added photos of the raw extracted DNA on agarose to support the quality of the DNA.

3. When an unknown DNA sequence is amplified from a sample this is taken as evidence that it comes from the purported unknown hominin when in fact mispriming from DNA of an unknown microorganism is a much more plausible scenario.

Authors'Response: We used next generation sequencing to sequence three unique whole genomes. The findings from these genomes support our previous findings. These samples were high quality and yielded outstanding genomes. The quality control prior to sequencing ruled out any large degree of bacterial contamination. We also provided new Figures 7, 8 and 13 to address any levels of degradation and contamination. Figure 13 and 8 as well as the histopathology report on Sample 26 which was added intentionally to show that that sample 26, which was one of the whole genomes sequenced, was neither degraded nor had a high

concentration of bacteria. Figure 7 was a yield gel showing some of the DNA utilized in this study. Note that the samples are not smeared.

4. Genetic differentiation based on electron microscopy is improbable. On this scale, one cannot distinguish the differences between any two species. The observed structural differences, if legitimate, are at best indicative of DNA damage.

Authors'Response: The electron microscopy was not intended to provide species identification but was included as supporting evidence for the unusual behavior of the amplified DNA. We have further addressed this in the revised manuscript.

5. There is a minimum of statistical analysis. One would like to see, for instance, a phylogeny based on the mitochondrial genome or HVI regions, or even the pair-wise distances between the samples and other humans.

Authors'Response: We have provided six mtDNA phylogenetic trees and three nuclear phylogenetic trees derived from the 3 whole genomes sequenced with next generation sequencing and have shown the pair-wise distances.

6. The authors do not follow the proper protocol for naming a new species.

Authors'Response: We have removed any taxonomic references and have chosen to call the hominins Sasquatch.