Research on Sasquatch (*Homo sapiens cognatus*) Gene MC1R.

Brent Post, Noah Knappenberger, and Lon Knappenberger
Westfield Academy and Central School, 203 East Main Street, Westfield, New York, 14787 and The Western New York Genetics in Research Partnership

University at Buffalo

Abstract

In 2013, The Sasquatch Genome project, which includes Dr. Melba Ketcham and her colleagues, published a paper in the Journal DeNova: Accelerating Science in which they claimed they had sequenced genes from a creature unknown to science, which they named *Homo sapiens cognatus*. The purpose of this project was to use the newly acquired skills of gene-act.org to test the claims of this article. A supplemental sequence titled MC1R from www.sasquatchgenomeproject.org was aligned using BLAST and the multiple sequence alignments and phylogenetic trees were constructed comparing the sequence to other primates, domesticated mammals and North American mammals to determine the source of the sequence. The sequence appears to be human and difficult to differentiate from a Sasquatch.

Introduction

On February 13, 2013, Dr. Melba Ketcham of the Sasquatch genome project (sasquatchgenomeproject.org/) and DNA Diagnostics, Inc. published her 5 year study on the DNA from a bipedal primate called Sasquatch or "Big Foot." The article titled "Novel North American Hominins, Next Generation Sequencing of Three Whole Genomes and Associated Studies" was published in a new"peer-reviewed"journal called "DeNovo: Accelerating Science" (www.denovojournal.com). According to the forty-one page document:

"One hundred-eleven samples of blood, tissue, hair, and other types of specimens were studied, characterized and hypothesized to be obtained from elusive hominins in North America commonly referred to as Sasquatch."
DNA was extracted from the samples and Sanger sequenced by a Illumina HiSeq 2000 "next generation sequencer." Dr. Ketcham and her team concluded that the sequences were from a bipedal primate unknown to science. The scientific name *Homo sapiens cagnatus* ("Blood relative of the wise people") was submitted to ZooBank, the International Commission on Zoological Nomenclature and the application was accepted.

The purpose of this project is to determine if sequence MC1R came from an unknown hominin, *Homo sapiens cognatus*, modern human *Homo sapiens* or entire different organism. Due to the fact that genes are highly conserved and that Dr. Ketcham argues that Sasquatch in a member of our genus and species, it will be almost impossible to determine if highly significant sequences came from our subspecies (*sapiens*) or the proposed Sasquatch subspecies (*cognatus*).

Figure 1: what is a theoretical image of Bigfoot. Source: https://aeon.co/ideas/why-sasquatch-and-other-crypto-beasts-haunt-our-imaginations

Methods

Websites that were used:

1) Sasquatch genome Project-Supplmental raw data http://sasquatchgenomeproject.org/sasquatch_genome_project_003.htm to obtain sequences for testing.
2) EXPASY http://web.expasy.org/translate/ to obtain 6 possible reading frame translation for sequences.
3) BLAST https://blast.ncbi.nlm.nih.gov/Blast.cgi
4) UNIPROT: http://www.uniprot.org/ to find comparable primate, domesticated and North American mammals sequences.
5) T-COFFEE http://www.ebi.ac.uk/Tools/msa/tcoffee/ to construct multiple sequence alignments and Phylogenetic Trees.

> MC1R Sequence for Sample 25/26

```
TCACTCCACCCCCACAGCCATCCCCCAGCTGGGGCTGGCTGCCAACCAGACAGGAGCCCG
GTGCCTGGAGGTGTCCATCTCTGACGGGCTCTTCCTCAGCCTGGGGCTGGTGAGCTTGGTG
GAGAACGCGCTGGTGGTGGCCACCATCGCCAAGAACCGGAACCTGCACTCACCCATGTACT
GCTTCATCTGCTGCCTGGCCTTGTCGGACCTGCTGGTGAGCGGGAGCAACGTGCTGGAGAC
GGCCGTCATCCTCCTGCTGGAGGCCGGTGCACTGGTGGCCCGGGCTGCGGTGCTGCAGCA
GCTGGACAATGTCATTGACGTGATCACCTGCAGCTCCATGCTGTCCAGCCTCTGCTTCCTGG
GCGCCATCGCCGTGGACCGCTACATCTCCATCTTCTACGCACTGCGCTACCACAGCATCGT
GACCCTGCCGCGGGCGCGGCGAGCCGTTGCGGCCATCTGGGTGGCCAGTGTCGTCTTCAG
CACGCTCTTCATCGCCTACTACGACCACGTGGCCGTCCTGCTGTGCCTCGTGGTCTTCTTCC
TGGCTATGCTGGTGCTCATGGCCGTGCTGTACGTCCACATGCTGGCCCGGGCCTGCCAGCA
CGCCCAGGGCATCGCCCGGCTCCACAAGAGGCAGCGCCCGGTCCACCAGGGCTTTGGCCT
TAAAGGCGCTGTCACCCTCACCATCCTGCTGGGCATTTTCTTCCTCTGCTGGGGCCCCTTCT
TCCTGCATCTCACACTCATCGTCCTCTGCCCCGAGCACCCCACGTGCGGCTGCATCTTCAAG
AACTTCAACCTCTTTCTCGCCCTCATCATCTGC
```

Figure 2: Raw Nucleotide Sequence for Reported "Sasquatch" gene MC1R

Results.

melanocortin 1 receptor [Homo sapiens]
Sequence ID: ALJ76158.1 Length: 317 Number of Matches: 1

Range 1: 15 to 289 GenPept Graphics Next Match

Score	Expect	Method	Identities	Positives
546 bits(1406)	0.0	Compositional matrix adjust.	274/275(99%)	275/275(100%)

```
Query  1    HSTPTAIPQLGLAANQTGARCLEVSISDGLFLSLGLVSLVENALVVATIAKNRNLHSPMY  60
            +STPTAIPQLGLAANQTGARCLEVSISDGLFLSLGLVSLVENALVVATIAKNRNLHSPMY
Sbjct  15   NSTPTAIPQLGLAANQTGARCLEVSISDGLFLSLGLVSLVENALVVATIAKNRNLHSPMY  74

Query  61   CFICCLALSDLLVSGSNVLETAVILLLEAGALVARAAVLQQLDNVIDVITCSSMLSSLCF  120
            CFICCLALSDLLVSGSNVLETAVILLLEAGALVARAAVLQQLDNVIDVITCSSMLSSLCF
Sbjct  75   CFICCLALSDLLVSGSNVLETAVILLLEAGALVARAAVLQQLDNVIDVITCSSMLSSLCF  134

Query  121  LGAIAVDRYISIFYALRYHSIVTLPRARRAVAAIWVASVVFSTLFIAYYDHVAVLLCLVV  180
            LGAIAVDRYISIFYALRYHSIVTLPRARRAVAAIWVASVVFSTLFIAYYDHVAVLLCLVV
Sbjct  135  LGAIAVDRYISIFYALRYHSIVTLPRARRAVAAIWVASVVFSTLFIAYYDHVAVLLCLVV  194

Query  181  FFLAMLVLMAVLYVHMLARACQHAQGIARLHKRQRPVHQGFGLKGAVTLTILLGIFFLCW  240
            FFLAMLVLMAVLYVHMLARACQHAQGIARLHKRQRPVHQGFGLKGAVTLTILLGIFFLCW
Sbjct  195  FFLAMLVLMAVLYVHMLARACQHAQGIARLHKRQRPVHQGFGLKGAVTLTILLGIFFLCW  254

Query  241  GPFFLHLTLIVLCPEHPTCGCIFKNFNLFLALIIC  275
            GPFFLHLTLIVLCPEHPTCGCIFKNFNLFLALIIC
Sbjct  255  GPFFLHLTLIVLCPEHPTCGCIFKNFNLFLALIIC  289
```

Figure 3: Top BLAST hit for the raw nucleotide sequence. A Human melanocortin 1 receptor (MC1R). E-value =0. melanocortin 1 receptor is one of the key proteins involved in regulating mammalian skin and hair color by making the pigment melanin (https://en.wikipedia.org/wiki/Melanocortin_1_receptor).

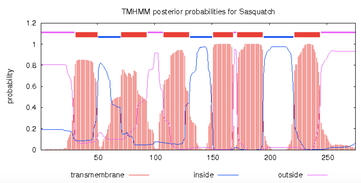

Figure 4: The Sasquatch sequence conatins 6 transmenbrane helices, which supports its function as a receptor.

melanocortin 1 receptor, partial [Sus scrofa]
Sequence ID: AAR12166.1 Length: 321 Number of Matches: 1

Range 1: 32 to 292 GenPept Graphics Next Match Previous Mi

Score	Expect	Method	Identities	Positives	Gaps
242 bits(617)	1e-75	Compositional matrix adjust.	194/261(74%)	200/261(76%)	0/261(0%)

```
Query  15   QPDRSPVPGGVHL*RALPQPGAGELGGERAGGGHHRQEPEPALTHVLLHLLPGLVGPAGE  74
            QPD  PVPGGVH  RALPQPGAGE  GERAGGG HRQEP+PAL HVLL LLPG VGPAGE
Sbjct  32   QPDGPPVPGGVHSRRALPQPGAGEPRGERAGGGRHRQEPQPALAHVLLRLLPGRVGPAGE  91

Query  75   REQRAGDGRHPPAGGRCTGGPGCGAAAAGQCH*RDHLQLHAVQPLLPGRERRGPLHLHLL  134
            REQRAGDGR   AGG   G PG   AAAGQCH R HL+LH VQPLLPGRERRGPL +HLL
Sbjct  92   REQRAGDGRAAAAGGGRPGRGPGRRGAAAGQCHERAHLRLHGVQPLLPGRERRGPLRVHLL  151

Query  135  RTALPQHRDPAAGAASRCGHLGGQCRLQHALHRLLRPRGRPAVPRGLLPGYAGAHGRAVR  194
            R ALPQHRD A   A   GHLGGQ  LQH LHRLL P GRPA PR LL G+AGAHG  VR
Sbjct  152  RAALPQHRDAAPRGAGHRGHLGGQRALQHPLHRLLPPHGRPAGPRQLLRGHAGAHGGTVR  211

Query  195  PHAGPGLPARPGERPAPQEAAPGPPGLWP*RRCHPHHPAGHFLPLLGPLLPASHTERPLP  254
            PHAGPGLPARP KRPAPQ+AAP PPGL P  R HPHHPAG    PLLG LLPA    RPLP
Sbjct  212  PHAGPGLPARPAERPAPQQAAPHPPGLRPQQRGHPHHPAGRLPFLLGTLLPAPLPRRPLP  271

Query  255  RAPHVRLHLQELQPLSRPHHL  275
            RAPH+RL LQE QPLS P HL
Sbjct  272  RAPHLRLRLQERQPLSGPRHL  292
```

Figure 5: Top BLAST Hit for reading frame 2 of MC1R is a tubulin beta in a human(Homo sapiens)

tubulin, beta, 4 [Homo sapiens]
Sequence ID: AAP35617.1 Length: 797 Number of Matches: 1
See 1 more title(s)

Range 1: 15 to 289 GenPept Graphics Next Match Pre

Score	Expect	Method	Identities	Positives	Gaps
551 bits(1419)	0.0	Compositional matrix adjust.	274/275(99%)	275/275(100%)	0/27

```
Query  1    HSTPTAIPQLGLAANQTGARCLEVSISDGLFLSLGLVSLVENALVVATIAKNRNLHSPMY  60
            +STPTAIPQLGLAANQTGARCLEVSISDGLFLSLGLVSLVENALVVATIAKNRNLHSPMY
Sbjct  15   NSTPTAIPQLGLAANQTGARCLEVSISDGLFLSLGLVSLVENALVVATIAKNRNLHSPMY  74

Query  61   CFICCLALSDLLVSGSNVLETAVILLLEAGALVARAAVLQQLDNVIDVITCSSMLSSLCF  120
            CFICCLALSDLLVSGSNVLETAVILLLEAGALVARAAVLQQLDNVIDVITCSSMLSSLCF
Sbjct  75   CFICCLALSDLLVSGSNVLETAVILLLEAGALVARAAVLQQLDNVIDVITCSSMLSSLCF  134

Query  121  LGAIAVDRYISIFYALRYHSIVTLPRARRAVAAIWVASVVFSTLFIAYYDHVAVLLCLVV  180
            LGAIAVDRYISIFYALRYHSIVTLPRARRAVAAIWVASVVFSTLFIAYYDHVAVLLCLVV
Sbjct  135  LGAIAVDRYISIFYALRYHSIVTLPRARRAVAAIWVASVVFSTLFIAYYDHVAVLLCLVV  194

Query  181  FFLAMLVLMAVLYVHMLARACQHAQGIARLHKRQRPVHQGFGLKGAVTLTILLGIFFLCW  240
            FFLAMLVLMAVLYVHMLARACQHAQGIARLHKRQRPVHQGFGLKGAVTLTILLGIFFLCW
Sbjct  195  FFLAMLVLMAVLYVHMLARACQHAQGIARLHKRQRPVHQGFGLKGAVTLTILLGIFFLCW  254

Query  241  GPFFLHLTLIVLCPEHPTCGCIFKNFNLFLALIIC  275
            GPFFLHLTLIVLCPEHPTCGCIFKNFNLFLALIIC
Sbjct  255  GPFFLHLTLIVLCPEHPTCGCIFKNFNLFLALIIC  289
```

Figure 6: Top BLAST Hit for reading frame 3 of MC1R is a truncated melanocortin 1 receptor in a cow (Bos taurus).

truncated melanocortin 1 receptor, partial [Bos taurus]
Sequence ID: ABX83564.1 Length: 72 Number of Matches: 1

Range 1: 30 to 71 GenPept Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps
46.6 bits(109)	0.001	Compositional matrix adjust.	31/42(74%)	31/42(73%)	0/42(0%)

```
Query  99   CSSWTMSLT*SPAAPCCPASASWAPSPWTATSPSSTHCATTA  140
            CSSWTMS T S A P CPASASW P  WTATSPSST C TT
Sbjct  30   CSSWTMSSTCSSADPWCPASASWVPLLWTATSPSSTPCGTTV  71
```

Figure 7: Top BLAST Hit for reading frame 5 of MC1R is a hypothetical protein in a chimney swift (Chaetura pelagica), which is a bird.

hypothetical protein M959_05470, partial [Chaetura pelagica]
Sequence ID: KFU85774.1 Length: 176 Number of Matches: 1

Range 1: 3 to 148 GenPept Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps
137 bits(346)	8e-37	Compositional matrix adjust.	89/146(61%)	98/146(67%)	0/146(0%)

```
Query  114  AEDDTGHPDGRNGSPRPRQGNDAVVAQCVEDGDVAVHGDGAQEAEAQGHAAGDHVNDIV  173
            A D  G PD  +G   P QGNDAVVAQ VEDGDVAVHGD  QE E G   GAA +HV+DIV
Sbjct  3    ADGDAGQPDAGNGHHSPLQGNDAVVAQSVEDGDVAVHGDDPQEGEGGHDGAADEHVDDIV  62

Query  174  QLLQHRSPGHQCTGLQQEDDGRLQHVAPAHQQVRQGQAADEAVHG*VQVPVTLGDGCHHQR  233
            +     HQ    LQQ+ +  L+ VA AHQ VR GQAADE VHG VQVPVL DG HHQ+
Sbjct  63   HVADDGGADHQHAVLQQQQEEGLAQVADAHQHVRDQQAADEVVHGKVQVPVLQQGSHHQQ  122

Query  234  VLHQAHQPQAEEEPVKDGHLQAPGSC  259
            VLHQAHQPQ EE+ V D   APGS
Sbjct  123  VLHQAHQPQREEQLVGDVESLAPGSA  148
```

Figure 8: The chimney swift (Chaetura pelagica), not a primate. source: https://ixquick-proxy.com

Phylogenetic tree with primates.

Figure 9: Phylogenetic tree of a T-COFFEE of the MC1R gene in primates. The sasquatch gene is on the same branch as humans.

Phylogenetic tree with domesticated animals.

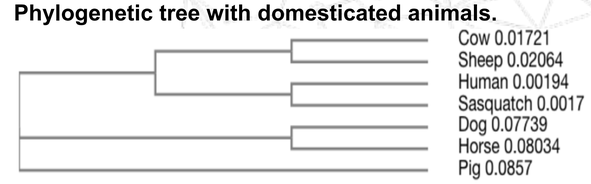

Figure 10: Phylogenetic tree of a T-COFFEE of the MC1R gene in domesticated animals. The sasquatch gene is on the same branch as humans.

Phylogenetic tree with North American animals.

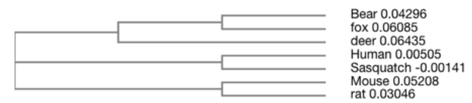

Figure 11: Phylogenetic tree of a T-COFFEE of the MC1R gene in North American animals. The sasquatch gene is on the same branch as humans.

Conclusion

The top BLAST hit for the raw nucleotide sequence is a Human melanocortin 1 receptor with an e-value of zero. However, the subject alignment starts at position 15 and differs by one amino acid. Bit score is 546. The top BLASTN hit for Reading frame 1 of the sequence is a MC1R is a Melanocortin receptor in a pig (Sus scofa) with an e-value of 1e-75. The subject alignment does not start to position 32 and Query postion 15. Bit Score is 242. The top BLASTN hit for Reading frame 2 of the sequence is a MC1R is a tubulin beta in a human with an e-value of 0.001. The subject alignment does not start to position 15. Bit Score is 551.

The Phylogenetic trees for the reported "Sasquatch" gene place it next to a human. If Sasquatch is a *Homo sapiens cognatus*, it will be nearly impossible to separate its genome from humans (*Homo sapiens sapiens*), who also spend time in the woods and contaminate samples.

References

1. http://sasquatchgenomeproject.org/
2. http://sasquatchgenomeproject.org/linked/novel-north-american-hominins-final-pdf-download.pdf

Acknowledgments

Supported by an NSF Innovative Technology Experiences for Students and Teachers (ITEST) Award - 1311902

www.buffalo.edu