The Long-Term Preservation and Recovery of DNA in a Borehole of the Moon

Alfie Box

I. Abstract

This report looks at the relationship between the conditions of the borehole that will be created on the Moon and the conditions required in order to preserve an archive with DNA on the Moon for long periods of time, with as little degradation of the DNA as possible. As well as this, different courses of action that can be taken in order to preserve the DNA archive in the Moon's borehole and to reconstruct a complete sequence of a genome on the Moon will be examined. The DNA being stored for the mission's archive will be carried by strands of hair. This form has a significant benefit for storing DNA with reduced degradation due to the keratin present in the hair strand. If the DNA is to be carried by strands of hair, it will be important for further use of the DNA that the information sequence can be reconstructed. This can be done with the mitochondrial DNA (mtDNA) found inside the hair shaft and nuclear DNA if the hair follicle is present and preserved. Although nuclear DNA is provides greater amounts of information than mitochondrial DNA, it is less stable.

Reliability

When conducting research, any information used will have its source referenced. Sources will be kept to published scientific journals and trusted websites to keep strong reliability for all information throughout the report and to maintain validity of the report.

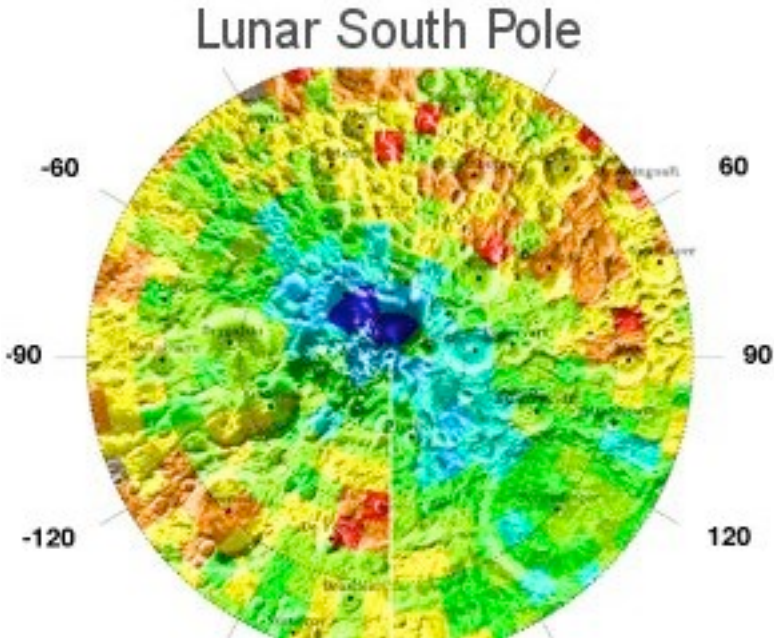
II. DNA Preservation

The borehole to be created is going to be drilled through the Shackleton crater, located on the South side of the Moon. The borehole will reach a depth of 100m into the regolith, where the capsules containing the DNA samples will be stored and preserved. The diameter of the hole will be 5cm, with about a 2.5cm core in the centre (Lunarmissionone.com, 2016). The DNA to be stored in the borehole of the

Moon is to be carried by strands of hair. In order to store DNA when it is carried by strands of hair, the roots must be attached in order for it to carry genomic DNA, as only mitochondrial DNA is found in the shaft and genomic DNA is found only in the follicle of hair strands. For it to be kept in the borehole for a long time, low temperatures are required. DNA must be stored at a temperature of -164°C in order to preserve DNA for multiple decades. The temperature on the surface of the Moon ranges, depending on where the sun shines. At areas where the sun shines, the temperature can reach 123°C, despite the Moon having no significant atmosphere to trap heat. Higher temperatures decay the bonds in DNA, thus degrading genetic information, making this area inadequate to store DNA. On, what is often called the far side of the moon, the side that always faces the dark expanse of space and where the sun never shines, temperatures drop to -153°C, close to those required to store DNA for multiple decades. The crust of the moon maintains fairly constant low temperatures, but the temperatures rise when deeper underground due to the hot core that heats the mantle and the lower and middle parts of the crust. However, this should not affect the temperature inside the borehole. When about 2 metres underground, the temperatures are at around -30 to -40°C. (Space.com, 2016) Lower temperatures will be needed in order to preserve DNA for a long period of time. Those conditions do not provide the preservation for the length of time wanted.Woolly mammoth mitochondrial DNA was found preserved, ranging in ages from 50,000 years to 12,000 years, with one of the mammoth's DNA being stored in a museum at room temperature (Live Science, 2016). Conditions similar to those present when the DNA of woolly mammoths first began preservation in their hair would be quite agreeable. Those would be similar consistent temperatures, or lower consistent temperatures for longer periods of time, for the extent of the archive. However, the temperature is only one of many factors that must be taken into account.

Radiation could prove to be a more considerable factor on the moon than on Earth. The surface of the Moon is exposed to large amounts of cosmic radiation which consists of high-energy particles such as protons, atomic nuclei and neutrons and the Moon has no significant atmosphere to protect against radiation. When cosmic rays hit particles on the surface of the Moon, small nuclear reactions occur which then emits further radiation in the form of neutrons. Depending on the location, the intensity of this radiation can vary. Figure 1 shows the mapping of Neutron radiation

at ground level on the south pole of the Moon. This secondary radiation and other primary radiation can cause multiple different types of damage to DNA, including the: Changing of bases, break of saccharides, rupturing of strands and formation of dimers (two similar monomers joined by covalent or intermolecular bonds). If the DNA is to be of use in the future, the exposure to radiation should be kept to a minimum, if not completely prevented, although that will prove to be very difficult to achieve if there is naturally emitting radiation from the lunar body near the DNA archive. It is possible for the majority, if not all degrading radiation, to be protected against if the DNA samples are stored at the large depth of 100m, due to the shielding of the Moon's regolith and crust.



-Figure 1 " 'Hot spots' are red; cool spots, blue. Credit: Lunar Prospector."
(Los Alamos National Laboratory, 2016)

If sources of radiation were to be located underground and even near the storage area, then the samples may not be as protected, and so may degrade more quickly, unless the storage container and material will be able to help protect the samples against this and other factors. When cosmic rays collide with the moon, they can generate a considerable amount of gamma radiation, enough to have a significant effect on Fermi/Lat data being recorded in the area (Astrobites.org, 2016). The material will need to be able to prevent the penetration of not only weak and highly ionising radiation, such as alpha particles, but other cosmic radiation, including gamma rays. Gamma rays have high penetration power which requires ~0.4m thick lead or ~2.0m thick concrete in order to be shielded against (ANS, 2016). However, the borehole is only 0.05m in diameter and there is a 0.02m cylindrical core in the centre of the hole, greatly limiting the thickness of the lining material. The space in the borehole is further limited as room for the archives must be taken into account. When taking into account the digital archive, a data center at least 5 metres underground in the Moon, would be exposed to the same level of

cosmic radiation as if it was on the Earth, assuming there to be no radioactive materials in that area or gamma rays produced by cosmic radiation collision at the borehole. At similar distances, the temperature stays at a fairly consistent temperature of about -30°C to -40°C. With these Earth-like conditions, it is possible for the environment to be recreated on Earth, and possible preservation techniques for the archive to be tested on Earth before the mission. However, if the archive is to be stored deeper into the borehole, where conditions are not known for sure, it is likely the conditions will change and thus the tested preservation apparatus and materials may not be as effective, or may even be useless. It is predicted that the temperature will reduce as the depth reaches 100m. However, the consistency of the temperature is not known. It is important for temperatures to remain consistent as frequent changes in temperature can degrade nucleic acids. However, conditions deeper underground are not known in much detail nor for certain at this time. Because of this, and the long length of time the DNA is wanted to be preserved for, it will be best to take precautions against the worst circumstances. An alternative to lead will, therefore, be needed as the borehole is too narrow to have a 0.4m protective diameter layer of lead, however it may be possible to have a multi-metre lid or cover, on top of the borehole in order to prevent the entry of cosmic rays and/or their resulting gamma rays and neutrons. Inside, a thin layer of a dense material, one with a high atomic number, will be needed to provide good shielding against radiation.

Silflex, by MarShield Custom Radiation Shielding products, is a flexible lead-alternative consisting of silicone pieces with tungsten. This alternative is effective against both gamma and neutron fields. It attenuates both fast and intermediate neutrons and the tungsten layer on the outside shields against both primary and secondary gamma radiation (Marshield.com, 2016). In the 5cm diameter borehole there is a ~2.5cm cylindrical core at the centre and the archive must also be stored inside this space. This does not leave much room for a protective layer; there will be ~1.25cm on either side of the core, in which the archives must also be stored. Against caesium-137, ~0.32cm Tungsten Silflex shielding has, having been tested, only a 21% shielding effectiveness and 12% against cobalt-60 (see Figure 2 in the appendix for further tests). Effectiveness of this percentage may be sufficient shielding at 100m deep if there are low amounts of radiation or infrequent radiation. However, the shielding may have to be replaced over time, especially when exposed to consistent

low temperatures required for the long-term preservation of the DNA. In use, the shielding installed can last up to ~40 years before replacement (Marshield.com, 2016), making it impractical for such long-term use. Tungsten, however, looks to be a promising material due to its high atomic number and density.

Figure 3 Image Credit: *MarShield*
Different forms of the Silflex lead-alternative radiation shielding.
(Marshield.com, 2016)

A tungsten alloy may improve the durability of the shielding while maintaining a thin layer. Tungsten alloys are often used for containers of radioactive sources in gamma radiography and also for protecting syringes during radioactive injections and so are known to be reliable. A tungsten alloy, such as those produced by Mi- Tech Metals, would provide greater resistance to more extreme temperatures. A high-density tungsten alloy provides the shame shielding as lead radiation shielding but in 1/3 of the thickness (Mi-techmetals.com, 2016). High levels of shielding can thus be achieved by thin layers with thicknesses able to fit in the borehole whilst maintaining space for the archive itself. However, tungsten alloys are fairly expensive. A more affordable alternative is polyethylene.

High-density 5% borated polyethylene provides great attenuation of neutrons due the high concentration boron and hydrogen and shielding against gamma rays because of its high density. Like Silflex and tungsten alloys, its shape can be customised into sheets, as well as cylinders. But in addition this, polyethylene can also be used in the form of polyethylene-based pellets which, although may have an overall lower density than a single solid depending the packing (Shieldwerx.com, 2016), may be easier to use to shield the borehole's archive as it can be poured around the archive without having to craft a specific shape layer.

Keratin protects DNA from external factors, which prevents degradation of the DNA, by preventing the breakdown of biopolymers. It is easier to rid hair of contamination,

such as bacterial DNA, as solutions which destroy external DNA do not affect the internal DNA. Keratin is hydrophobic (it repels water), so, unlike with bone, water does not rush into hair, carrying bacterial DNA into it. There is a hair shaft that covers and protects the DNA from damage and bacterial contamination as if the case was made of plastic. There is no nuclear DNA inside the hair shaft. However, there is mitochondrial DNA which is much smaller and more stable than nuclear DNA.

III. DNA Recovery

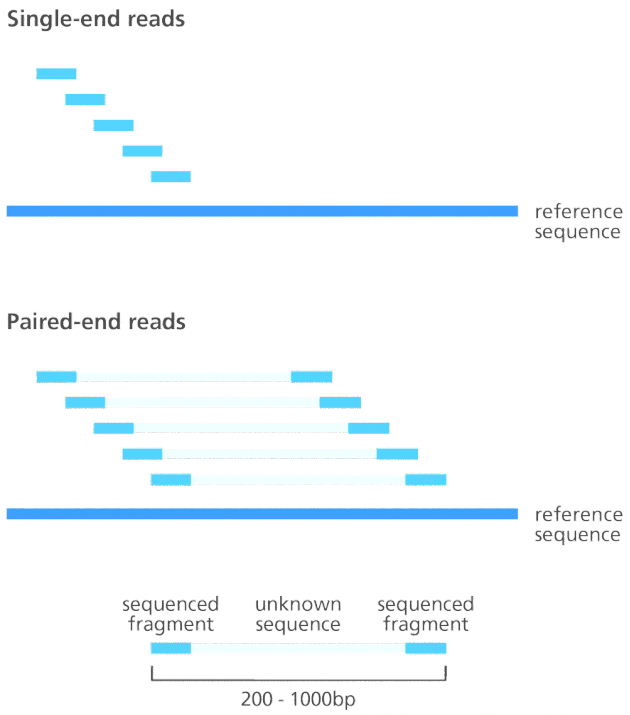
Once preserved for the wanted time on the Moon, or any other reasons, the DNA is likely to be used and so will also very likely need to be sequenced.

Sequencing of DNA is working out the order of DNA bases that make up a genome, the number of bases differing greatly between species. As humans, we denote each of the four bases in DNA as adenine (A), thymine (T), cytosine (C) and guanine (G). A human genome sequence consists of other three billion of these bases. A whole genome cannot be sequenced entirely at once due to the current technology available. However, many short
sections (reads) of DNA can be handled. The genome, once broken down into small sections, is sequenced and put back together in the correct order. Depending on whether there will be a reference for the DNA sample or not, there are multiple methods to sequence a genome that can be used. If there were to be only one sample of the species to be sequenced, a version of de novo sequencing would be carried out.

De novo sequencing of DNA, as the name suggests, is when the genome of an organism is sequenced without prior knowledge of the sequence e.g. for the first time. De novo sequencing consists of two processes: Alignment and assembly. During alignment, the sample is usually compared to existing DNA sequences in order to find similarities or differences, however, as there are no existing references to be used as a template in de novo sequencing, a similar species' genome could be used as a guide if one knows that a species is very similar. In the assembly process, many reads of the DNA are taken to look for parts where they overlap with each other. One way, is to first break the genome into fairly large parts, known as clones, consisting of

about one hundred and fifty thousand base pairs (bp) each, and then use a gene map to know where each clone should be in the genome and later breaking them down into smaller pieces which overlap, for sequencing. The pieces are then put together to reconstruct the original genome. A gene map may be developed to highlight main parts in order to know where sections of DNA are, apropos to each other. Gene maps, however, can be expensive. As an alternative, data of single and paired-end reads can be used instead. Single reads consist of either a whole fragment or one end of a fragment being sequenced. These can then be put together by finding where they overlap in the sequence to form the couple sequence of DNA. Paired-end reads consist of both ends of a fragment being sequenced. The distance between the two ends are known, which makes assembling a continuous sequence of DNA from paired-end reads much easier. This is particularly useful in de novo sequencing as it provides information that would not be present without a gene map.

Figure 4 shows a visual representation of the difference between single reads and paired-end reads.



-Figure 4
Image Credit: *Genome Research Limited*, (Genome Research Limited, 2016)

The next step in de novo sequencing is comparing the reads. Those which share the same sequence of DNA are grouped together and long contiguous sequences (contigs) are then produced from the progressively larger sections. What is known as "scaffolding" then takes place. The contigs are stitched together, paying close attention to how far apart they should be. Linkages among all of the contigs form a graph. These scaffolds can then be ordered to form individual chromosomes. Annotation then takes place. This is where the start and stop of genes and other relevant parts are identified. (Yourgenome.org, 2016) It consists of two stages:

Prediction and manual annotation. The prediction process can be done by using evidence from sources other than that of the DNA sequence, or by ab initio. Using signals in the DNA sequence, a computer is used, having given instructions to find genes in the sequence. Common sequences are looked for which are known to be found at the beginning and end of genes. The sequencing data can be lined up with a reference genome and any differences can be identified. By using evidence from sources other than the DNA sequence, such as m-196RNA and protein sequences of a genome, the correct DNA sequence can be determined by working backwards through the transcription and translation processes; this is known as reverse transcription/translation.

After prediction, manual annotation can then begin. This is the process where any information from prediction is examined, e.g. to look for a specific gene. Once a sequence has been successfully annotated, it can be compared against other, similar, organisms' annotated sequences, to see differences and similarities in their DNA sequences. (Yourgenome.org, 2016)

When using a reference to sequence a DNA sequence, there isn't usually any assembly as it already has been done for the reference. Alignment is carried out and so the reads from sequencing are compared with the reference genome and aligned with the corresponding section of the reference, allowing any differences to be identified between the sample and reference sequences.

The most feasible way of Sequencing DNA on the Moon would be to use a recent, successful technology. A small sequence of genomic DNA has been successfully sequenced in a microgravity environment aboard the International Space Station (ISS). The technology used was one of *Oxford Nanopore Technologies* products known as a MinION (pronounced min-eye-on), a small, portable and plug-and-play genetic sequencer. It is a very small, light and relatively quick sequencer and looks to be the most promising method to achieve sequencing in a lunar environment having already been tested and successful. It uses disturbances in an electrical potential to identify a biological molecule, be it DNA, mRNA, miRNA, mtDNA or proteins (Nanoporetech.com, 2016). A potential is passed through the nanopores, inserted into an electrical resistant membrane, so the current flows through the aperture of the

nanopore only. A molecule, e.g. DNA strand, passed through the pore causes characteristic disturbances in the current, unique to the molecule. This is known as the nanopore signal. The molecule can be identified by measuring its unique signal. Inside the MinION, there is a flow cell, a collection of electrodes and supports, each with their own electrical channel (Nanoporetech.com, 2016). This sequencer determines the nucleotide sequence of the DNA without a required specific primer. It is powered via a USB cable to a computer

Figure 4
A MinION sequencer. Image Credit: *Oxford Nanopore Technologies* (Nanoporetech.com, 2016)
9.5 x 3.2 x 1.6 centimetres in size and under 120 grams in weight, including cable. (NASA, 2016)

Nuclear DNA is much larger than mitochondrial DNA; in humans, there are ~3.2 billion base pairs in the nuclear DNA compared to the mtDNA's ~17,000 base pairs. However, per cell, there are 2 copies of the nuclear DNA whereas there are hundreds of copies of the mtDNA per cell. Human mtDNA as been sequenced completely, for the first time in 1981 (Anderson et al) and then revised in 1999 (Andrews et al), for confirmation.

Although nuclear DNA is a more powerful sequence for information, the use of mtDNA may be more helpful and likely as it will most likely stay preserved for longer due to its stability as circular DNA. In addition, the access to mtDNA is also much greater as it can be obtained from even greatly degraded sources. This means even if radiation or heat has affected the samples, mtDNA should likely still be able to be used. This is because mtDNA has many copies and is present within an organelle with a double membrane, which further protects the mtDNA, proving more stability. When the nucleus degrades during the keratinisation of the hair shaft, the mtDNA remains intact and so it can still be analysed. However, to sequence mtDNA would require additional action in the preparation of the DNA sample. The shaft must be broken down and the DNA released. A more efficient chemical digestion of the shaft can be used rather than the more tedious physical digestion. However, a normal

process, which can involve steps of buffering, vortexing, incubation, centrifugation, along with the extraction of DNA must still be carried out, which can be difficult in an environment with microgravity (Gilbert et al., 2007). Although an information sequence has been successfully sequenced in space, the sample was prepared on Earth. The preparing of the sample for sequencing may prove to be more difficult when in an environment of microgravity, than on Earth.

III. Conclusion

Overall, the conditions at the shallow depths previously investigated on the moon will not naturally allow for the preservation of DNA for the long length of time, as aimed by the Lunar Mission One project. In the borehole created on the Moon, additional radiation shielding from tungsten alloy lead-alternative will extend the length of the integrity of the DNA samples being carried by the strands of hair, however the space for the archive in the borehole will be further limited due to the layer which may reduce the size of the archive from its originally intended size. This shielding, with a possible cover at the entrance of the borehole, should protect against the majority, if not all radiation, reducing or even preventing degradation due to radiation. If the nuclear DNA would result in being degraded and beyond use, the mitochondrial DNA should remain in a condition able to be sequenced due to its stability. The temperature in the borehole is expected to reduce to very low temperatures as it reaches greater depths near 100m. However, unlike with the start of the borehole, the conditions are not well known and so the consistency of the temperature is not known for certain. Changing in temperatures will cause the DNA to degrade, which can result in the more informative nuclear DNA found in the follicle unable to be successfully sequenced or usable. The MinION sequencer from *Oxford Nanopore Technologies* is not only extremely small and lightweight, making it easy to transport to the Moon and fuel-saving, it has already successfully sequenced DNA in a low gravity environment making it the most promising method of current technology to sequence DNA on the Moon itself.

Appendix

Cesium 137 (Cs-137)

Shielding Material	Actual Thickness	Shielding Effectiveness	HVL		TVL	
			in	cm	in	cm
1/8" Tungsten Silflex	0.115"	21%	0.36	0.92	1.21	3.07
½" Tungsten Silflex	0.500"	63%	0.36	0.92	1.21	3.07
1" Tungsten Silflex	0.980"	84%	0.36	0.92	1.21	3.07
2" Tungsten Silflex	1.960"	97%	0.36	0.92	1.21	3.07
1/8" Mixed Silflex	0.125"	18%	0.57	1.45	1.89	4.81
½" Mixed Silflex	0.500"	47%	0.57	1.45	1.89	4.81
1" Mixed Silflex	0.990"	70%	0.57	1.45	1.89	4.81
2" Mixed Silflex	1.990"	90%	0.57	1.45	1.89	4.81
1/8" Iron Silflex (Fe)	0.075"	5%	1.17	2.97	3.88	9.85
½" Iron Silflex (Fe)	0.490"	26%	1.17	2.97	3.88	9.85
1" Iron Silflex (Fe)	0.970"	44%	1.17	2.97	3.88	9.85
2" Iron Silflex (Fe)	1.970"	68%	1.17	2.97	3.88	9.85
½" Neutron Silflex	0.490"	13%	2.4	6.09	7.97	20.23
1" Neutron Silflex	0.970"	24%	2.4	6.09	7.97	20.23
2" Neutron Silflex	1.950"	43%	2.4	6.09	7.97	20.23

Cobalt 60 (Co-60)

Shielding Material	Actual Thickness	Shielding Effectiveness	HVL		TVL	
			in	cm	in	cm
1/8" Tungsten Silflex	0.115"	12%	0.63	1.60	2.09	5.32
½" Tungsten Silflex	0.500"	44%	0.63	1.60	2.09	5.32
1" Tungsten Silflex	0.980"	65%	0.63	1.60	2.09	5.32
2" Tungsten Silflex	1.960"	87%	0.63	1.60	2.09	5.32
1/8" Mixed Silflex	0.125"	11%	0.94	2.40	3.14	7.97
½" Mixed Silflex	0.500"	32%	0.94	2.40	3.14	7.97
1" Mixed Silflex	0.990"	52%	0.57	1.45	1.89	4.81
2" Mixed Silflex	1.990"	76%	0.57	1.45	1.89	4.81
1/8" Iron Silflex (Fe)	0.075"	3%	1.66	4.21	5.51	13.99
½" Iron Silflex (Fe)	0.490"	18%	1.66	4.21	5.51	13.99
1" Iron Silflex (Fe)	0.970"	34%	1.66	4.21	5.51	13.99
2" Iron Silflex (Fe)	1.970"	56%	1.66	4.21	5.51	13.99
½" Neutron Silflex	0.490"	10%	3.38	8.59	11.24	28.55
1" Neutron Silflex	0.970"	18%	3.38	8.59	11.24	28.55
2" Neutron Silflex	1.950"	33%	3.38	8.59	11.24	28.55

Figure 2. Shielding Effectiveness
SRR Test Results (Marshield.com, 2016)

Bibliography

Yourgenome.org. (2016). *How do you put a genome back together after sequencing?*. [online] Available at: http://www.yourgenome.org/facts/how-do-you-put-a-genome-back-together-after-sequencing [Accessed 1 Aug. 2016].

Science, L. (2007). *Woolly Mammoth Hair Yields 'Fantastic' DNA*. [online] Live Science. Available at: http://www.livescience.com/9530-woolly-mammoth-hair-yields-fantastic-dna.html [Accessed 3 Aug. 2016].

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K., Li, S., Yang, H., Wang, J. and Wang, J. (2009). De novo assembly of human genomes with massively parallel short read sequencing. Genome Research, [online] 20(2), pp.265-272. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2813482/ [Accessed 14 Aug. 2016].

Genomenewsnetwork.org. (2016). What's a Genome?. [online] Available at: http://www.genomenewsnetwork.org/resources/whats_a_genome/Chp2_1.shtml#chp2#4 [Accessed 15 Aug. 2016].

cience.nasa.gov. (2016). *Radioactive Moon - NASA Science*. [online] Available at: http://science.nasa.gov/science-news/science-at-nasa/2005/08sep_radioactivemoon/ [Accessed 29 Aug. 2016].

Los Alamos National Laboratory, (2016). *Neutron Data Map North Pole*. [image] Available at: http://lunar.arc.nasa.gov/results/neures.htm [Accessed 29 Aug. 2016].

ANS. (2016). *Protecting Against Exposure - ANS*. [online] Available at: http://nuclearconnect.org/know-nuclear/science/protecting [Accessed 8 Sep. 2016].

Lunarmissionone.com. (2016). *Lunar Mission One - Reach for the MoonThe Technology | Lunar Mission One*. [online] Available at: https://lunarmissionone.com/lunar-mission-one/the-technology [Accessed 8 Sep. 2016].

Marshield.com. (2016). *Lead Alternative | Silflex | Non-Lead Radiation Shielding | MarShield Custom Radiation Shielding Products*. [online] Available at: http://marshield.com/nuclear-shielding/silflex/ [Accessed 9 Sep. 2016].

Live Science. (2016). Woolly Mammoth Hair Yields 'Fantastic' DNA. [online] Available at: http://www.livescience.com/9530-woolly-mammoth-hair-yields-fantastic-dna.html [Accessed 11 Sep. 2016]