**H+, or how to build a perfect human.**

Alekseeva L., Kureev N., Ovsyannikova A.

**Abstract**

Genetic variations, especially SNPs, have a lot of influence on our lives. Normally they appear every day in different organisms due to various mutagens. And people are no exception. Many of polymorphisms can be repaired by special system in our cells. However, sometimes people need these SNPs as a cure. In future we could treat many patients using gene therapy. Also we could change some parts of our genome to upgrade our organisms and maybe control our evolution. In this project we present possible solutions for this problem.

**Introduction**

Today CRISPR/Cas9 system is a most perspective technology for gene therapy. It is important that this structure has prokaryotic origin and plays immune system role. It consists of two main parts: a group of CRISPR palindromic repeats (24 - 28 b.p.) and a cas gene cluster. CRISPR groups have many spacers, related to various exogenous DNAs. It's like a library with records about target sequences. Cas proteins are necessary for introduction of spacers in a DNA sequence, processing CRISPR transcript for targeting final complex and destroying DNA or RNA molecules. In 2012 a way of custom molecule destruction was offered. A chemically synthesized CRISPR and programmed spacer could work together to cut DNA or RNA molecules in a specific place. This allows to modify many genes at one attempt [1].

**Materials and Methods**

In this project was analysed raw 23andMe data. Plink was used to create a vcf file from the 23andMe data.

Also we used ClinVar vcf file as a reference for disease-related SNPs checking. SnpEff and SNpSift were used for annotation of obtained SNPs.

To establish ethnicity we used GENOtation Ancestry with default parameters. Optimal results was received on 10 000 SNPs as data.

**Results**

The resulting vcs file contained all the analyzed SNPs, but was only interested in variable positions. At the same time, were removed all SNPs corresponding to deletions and insertions.

By filtering sample genotype using ClinVar file we adjust many SNPs which was not responsible for any diseases. Using annotation we 4 SNPs, 2 of which were without an indication of the disease. Another 2 SNPs were related to:

CLNDN=Generalized\_epilepsy\_with\_febrile\_seizures\_plus (rs1374591;652920, 2 chromosome, C > T);

CLNDN=Hereditary\_cancer-predisposing\_syndrome (rs2707761;826468, 5 chromosome, A > G);

Searching for mtDNA haplogroups we received M6a haplogroups as a best match, Also we obtained following markers:

**Matches(29): 73G 263G 461T 489C 750G 1438G 2706G 3537G 4769G 5082C 5301G 7028T 8701G 8860G 9540C 10398G 10400T 10640C 10873C 11719A 12705T 14766T 14783C 15043A 15301A 15326G 16223T 16231C 16362C**

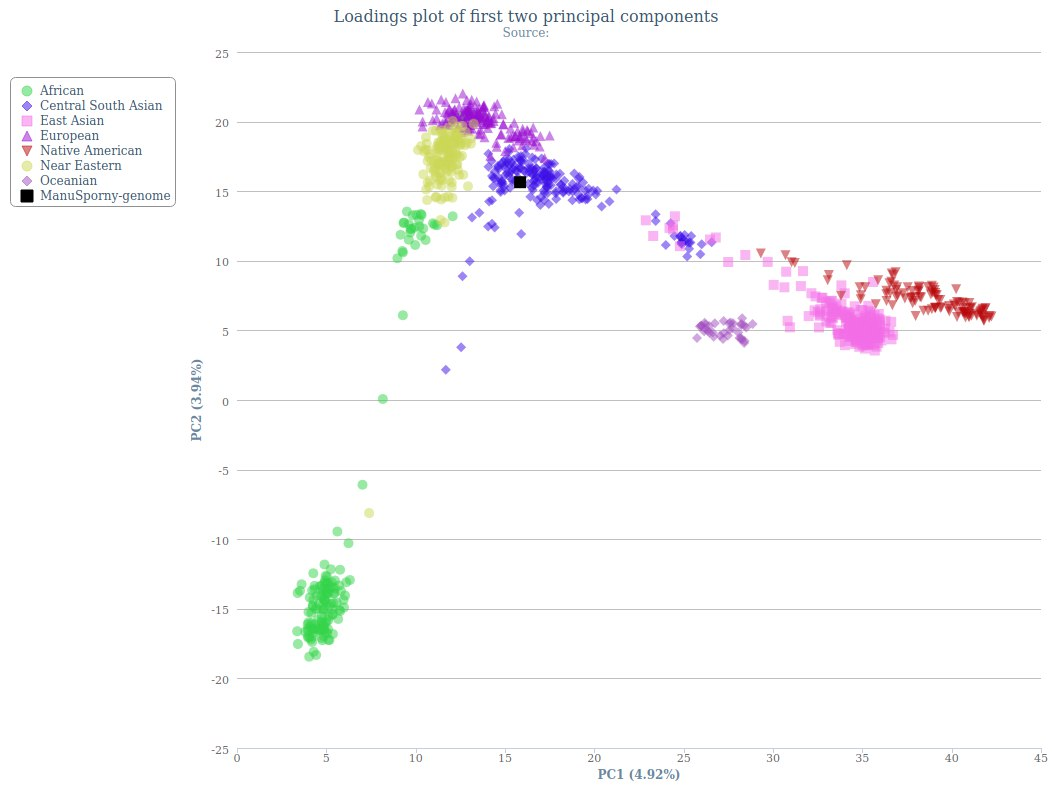
**Extras(3): 8563G 16111T 16320T (16519C)**

**No-Calls(1): 14128G**

**Untested(1): 5558**

There were found 2708 markers at 2558 positions covering 15.4% of mtDNA.

Using GENOtation we received ethnicity attachment of our genotype.



**Figure 1. GENOtation visualization using 10000 SNPs**

As we can see on Figure 1 it was identified that genotype we used as our research data can be attributed to Central South Asian ethnicity group.

**Discussion**

From mtDNA analysis we understood that our genotype sample related to M6a haplogroup. According to phylogenetic researches found out that this haplogroup is restricted to Indian region[2]. This is confirmed by mitochondrial DNA sequencing.

By using GENOtation for visualization of ethnicity attachment we obtain graph, which gave us an information that researched genotype related to Central South Asian ethnicity, which is correlated to our previous result using mtDNA.

Also we’ve got information about possible association with diseases. 2 SNPs were correlated to hereditary cancer-predisposing syndrome and generalized epilepsy with febrile seizures plus. Former is a syndrome occurring due to inherited mutation. The affected genes concerned usually have a controlling function on the cell cycle or the repair of DNA damage. The probability of the disease increases under the influence of stress factors [3]. Individuals with GEFS+ present with a range of epilepsy phenotypes. These include febrile seizures that end by age 6 and development of myoclonic seizures and myoclonic astatic seizures [4].

**References**

1. <https://postnauka.ru/faq/59807>
2. Chandrasekar A, Kumar S, Sreenath J, et al. Updating phylogeny of mitochondrial DNA macrohaplogroup m in India: dispersal of modern human in South Asian corridor. *PLoS One*. 2009;4(10):e7447. Published 2009 Oct 13. doi:10.1371/journal.pone.0007447
3. Rahner N, Steinke V. Hereditary cancer syndromes. *Dtsch Arztebl Int*. 2008;105(41):706–714. doi:10.3238/arztebl.2008.0706
4. Scheffer I, Berkovic S (1997). "Generalized epilepsy with febrile seizures plus. A genetic disorder with heterogeneous clinical phenotypes". Brain. 120 (3): 479–90.