Project 5

1. **Downloaded and installed plink**

PLINK 1.9

https://www.cog-genomics.org/plink/1.9/

```
sudo gedit ~/.bashrc
```

```
export PATH="/home/anya/bin/plink_install:$PATH"
alias plink="/home/anya/bin/plink_install/plink"
```

```
source ~/.bashrc
```

2. **File conversion to vcf**

```
plink --23file SNP_raw_v4_Full_20170514175358.txt --recode vcf --out snps_clean --output-chr MT --snps-only 'just-acgt'
```

2. **Origins, haplogroups.**

-mtDNA Haplogroup Analysis Report

mthap version 0.19b (2015-05-11); haplogroup data version PhyloTree Build 17 (2016-02-18) raw data source (0B) ERROR: The server did not receive the file, the upload file is empty, or the upload file was too large. Please try again. Need help? First, please check the Frequently Asked Questions for guidance on how to read this report.

https://dna.jameslick.com/mthap/mthap.cgi

Just download data to the site and it will show the result

```
MorleyDNA.com Y-SNP Subclade Predictor: extract Y-DNA from an autosomal test (preprocessing stage 2 of 2)
Your data has:
63 Y-DNA position(s) lacking mutations recognised by the genetic genealogy community. These Y-DNA positions may not be very useful.
166 recognised mutation(s) with positive calls.
733 recognised mutation(s) with negative calls.
1086 recognised mutation(s) with no-calls.
```

3. **IGV Browser:**

```
sh /home/anya/bin/IGV_Linux_2.14.1/igv.sh
```

4. **Sex determination**

SNPs of SRY gene:

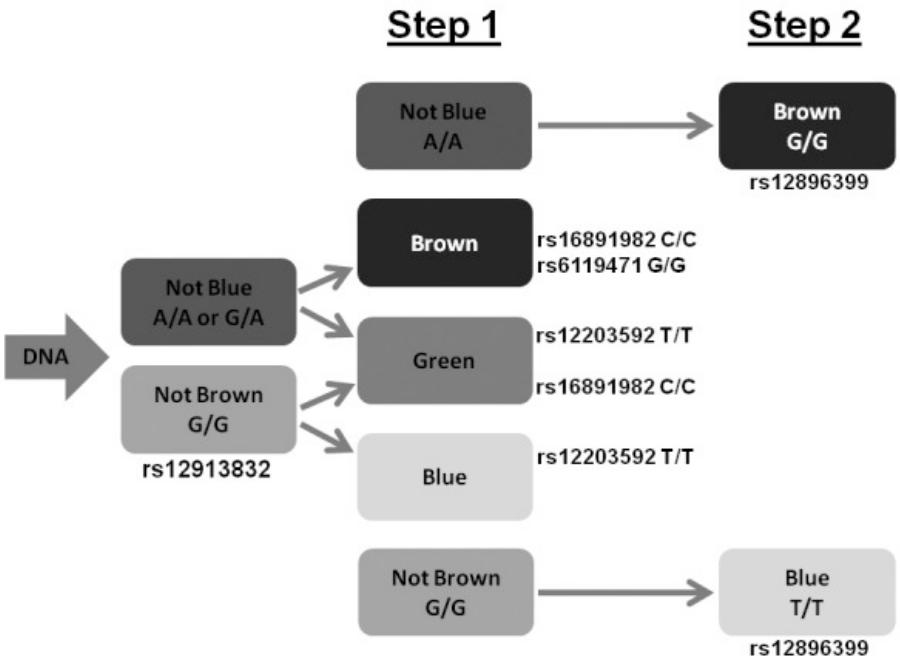
rs11575897	0	2,787,139	
rs2534636			

→ Misha is biologically male

5. Skin and eye color

> Non-dark skin color (ie, light or medium) is predicted by any two of the following alleles: G/G at rs12913832, G/G at rs16891982, A/A at rs1426654, T/T at rs1545397, or A/A at rs885479. Light skin color is predicted by more stringent conditions: G/G at rs12913832, plus G/G at rs16891982, and A/A at rs1426654. Non-light skin color (ie, medium or dark) is predicted by G/G at rs6119471.

SNP			eye	skin	result
rs12913832	eye + skin	AG	not blue	not light/medium, not light	not blue eyes + not light skin
rs12203592	eye	CT	not green		not green eyes
rs12896399	eye	GG	brown if rs12913832 = AA → not brown		not brown eyes
rs16891982	eye + skin	CG	not brown	not light/medium not light	not brown eyes, not light skin
rs1426654	skin	AA		light/medium, light	light skin
rs885479	skin	GG		not light/medium,	not light skin
rs1545397	skin	—	—	—	
rs6119471	eye + skin	—	—	—	



6. **VEP**

Results of VEP analysis:

	A	B	C	D
1	rs1024611	17:32579788-32579788	G	also known as the -2578A>G SNP due to its position in the promoter of the monocyte chemoattractant protein-1 MCP-1 CCL2 gene, influences the production of its corresponding protein, a chemokine involved in inflammatory responses. discussed in the 23andMe blog as being relevant to HIV
2	rs1049296	3:133494354-133494354	T	Involved in the transport of iron, transferrin and its subtypes have been linked at times to various conditions, perhaps most notably Alzheimer's disease.
3	rs10757274	9:22096055-22096055	G	~1.3x increased risk for heart disease
4	rs1169288	12:121416650-121416650	C	The genetic susceptibility to type 2 diabetes ???
5	rs12150220	17:5485367-5485367	T	slightly increased risk for several autoimmune diseases
6	rs13266634	8:118184783-118184783	T	increased risk for type-2 diabetes
7	rs1801197	7:93055753-93055753	G	osteoporosis ???
8	rs1801274	1:161479745-161479745	G	не наш случай, тк замена на Г
9	rs1801275	16:27374400-27374400	G	1.4x higher risk for meningiomas (рак головного мозга)
10	rs1801394	5:7870973-7870973	G	normal
11	rs1801968	9:132580901-132580901	G	носинг интерестинг
12	rs2004640	7:128578301-128578301	T	1.4x increased risk for SLE (волчанка)
13	rs2073658	1:161010762-161010762	T	is associated with a modestly increased risk to develop type 2 diabetes
14	rs2184026	9:101304348-101304348	T	носинг интерестинг
15	rs2239704	6:31540141-31540141	C	носинг интерестинг
16	rs2241880	2:234183368-234183368	G	strongly associated with ileal Crohn's disease
17	rs2281845	1:201081943-201081943	T	носинг интерестинг
18	rs231775	2:204732714-204732714	G	2.3x risk of Hashimoto's thyroiditis, 1.47x risk of Graves' disease
19	rs4402960	3:185511687-185511687	T	1.2x increased risk for type-2 diabetes, ~1x risk for gestational diabetes
20	rs4880	6:160113872-160113872	G	не наш случай, тк замена на Г
21	rs4961	4:2906707-2906707	T	1.8x increased risk for high blood pressure
22	rs4977574	9:22098574-22098574	G	Most studies find a somewhat elevated (~1.5x) risk for myocardial infarction
23	rs5174	1:53712727-53712727	T	не наш случай, тк замена на Т
24	rs5186	3:148459988-148459988	C	~1.4x increased risk of hypertension
25	rs61747071	16:53720436-53720436	T	uncertain-significance,
26	rs6265	11:27679916-27679916	T	Slightly increased risk for ADHD or depression; somewhat quicker mental decline in Alzheimer patients
27	rs6280	3:113890815-113890815	T - normal	rs6280, also known as Ser9Gly, is a SNP in the dopamine receptor D3 DRD3 gene, those who were rs6280(C;C) homozygotes had greater positive symptom remission for schizophrenia
28	rs6504649	17:48437456-48437456	G	nothing
29	rs699	1:230845794-230845794	G	increased risk of hypertension
30	rs763110	1:172627498-172627498	T	~0.80x reduced cancer risk
31	rs7794745	7:146489606-146489606	T	slightly increased risk for autism
32	rs909253	6:31540313-31540313	G	Myocardial infarction Psoriatic arthritis
33				