Even more perfect Mischa Rayko

Polina Guseva, Anna Malykhina

March 8, 2023

Abstract

Due to genotyping chips, finding single-nucleotide polymorphisms (SNPs) became relatively fast and affordable. The next step for humanity would be editing the pathogenic mutations (or even just not favourite ones) with CRISPR-Cas9. Based on the SNPs of Mischa Rayko by 23&Me, he is a European male and got brown-green eyes. Some corrections are suggested in terms of health to make him even more perfect.

Introduction

Genotyping chips, also known as microarrays, are relatively cheap tools used in genetic research to determine the genetic variations present in an individual's DNA [1], [2], [3]. These chips contain DNA probes that can identify specific genetic variants. When a DNA sample is applied to the chip, the probes bind to any matching genetic variants and show which variants are present [1], [3]. The data obtained from genotyping chips can be used in a variety of applications, such as genetic research, medical diagnosis, and personalized medicine [2].

CRISPR-Cas9 is a gene-editing tool that can be used to make precise changes to the DNA sequence of an organism [4]. In the context of SNPs, CRISPR-Cas9 can be used to correct a single-nucleotide polymorphism (SNP) by replacing the incorrect nucleotide with the correct one. This process involves designing a guide RNA that targets the site of the SNP and the Cas9 protein that cleaves the DNA at the target site [5]. Once the DNA is cleaved, the cell's natural repair mechanisms can be harnessed to introduce the desired change to the DNA sequence [6]. This technique has the potential to be used to correct disease-causing mutations in humans, but it is still in the experimental stages and its safety and efficacy are being studied [7].

In this project, SNPs from Mischa Rayko's genome are analysed to determine attributes of his phenotype as well as some health downsides. Some suggestions are made on how to improve the genome for better human performance.

Methods

The raw data is a collection of Mischa Rayko's single nucleotide polymorphisms (SNPs) obtained using a genotyping chip (23&Me, Illumina HumanOmniExpress-24).

The `plink` program converted raw data into standard VCF format [8]. The SNPs corresponding to deletions and insertions were deleted in plink by command: `--output-chr MT --snps-only just-acgt`. Then the data were annotated by the SnpEff tool [9] using the GRCh37.75 genome as a reference with comparison to the ClinVar database. The IGV browser (https://igv.org/) was used for visualization. Maternal haplogroup was determined using MThap online tool [10], while for paternal haplogroup MorleyDna online tool was used (https://ytree.morleydna.com/extractFromAutosomal). The eye colour is specified according to [11]. By Variant Effect Predictor (VEP) [12], the risk factors are identified.

Results

Overall, 164638 SNPs are processed where are 76489 (46.5%) novel variants and (53.5%) existing ones. Most of the mutations are missense and within an intron (fig.2). However, most of these mutations are benign and have no impact on health (fig.1). According to the ClinVar database, there were 12 mutations classified as risk factors (mostly related to coronary artery disease and metabolic syndrome), 2 pathogenic mutations (related to rare diseases) and 19 mutations associated with different illnesses. Among 22 drug response mutations, 15 were for Tramadol.

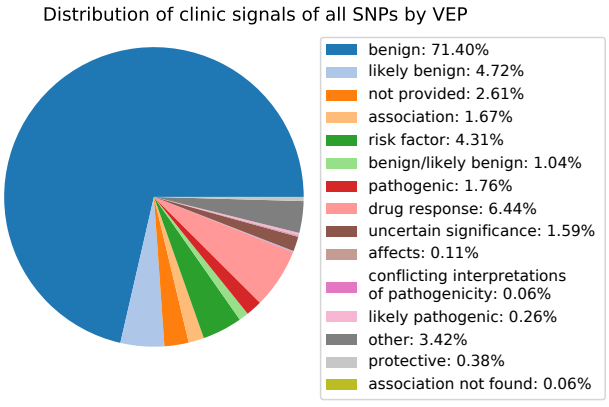

Figure 1: Clinical significance (VEP)

The maternal and paternal haplogroups of the studied genome were established as H(T152C) and R1a1a [R1a-L168 (R1a-M17, R1a-M198)] accordingly. The person is male as there is a Y chromosome in SNP data. The eye colour coding is reflected in several SNPs. Heterozygote in rs12913832 is in favour of not blue, while heterozygotes in rs16891982 and rs12203592 don't confirm either brown or green, thus the colour is in between brown and green [11]. The SNPs rs12913832, rs16891982 and rs12203592 are located in genes HERC2, SLC45A2 and IFR4 accordingly.

Discussion

According to his haplogroups, both Mischa's parents are probably European [13], [14].

Despite Mischa being perfect as hell, a few mutations could be fixed to improve some health parameters (tab.1). Sometimes all variants could have their pros and cons due to the gene multifunctionality. Therefore, the editing could be controversial, for example, for rs1801274 [15], [16], [17].

Moreover, some additional improvements could be applied to add cool features. Thus, the CCR5-Δ32 gene mutation provides some level of resistance to HIV infection [18]. For our favourite professor, we wish him longevity (rs2542052 near the APOC3 gene [19], rs34516635 in the IGF1R gene [20], rs6873545 or rs4590183 in the d3-GHR gene [21], [22], and plus cognitive health rs3758391 in the SIRT1 gene [23]). Yet, longevity could be a controversial choice for SNPs [24].

Table 1: Inconvenient SNPs that would be profitable and good to correct

ID	allele	gene	allele 2.0	reasons to upgrade
rs4961	T/T	α-adducin 1 ADD1 gene	G/G	otherwise: 1.8x increased risk for high blood pressure [25]
rs6280	T/T	the dopamine receptor D3 DRD3 gene	C/C	better response to olanzapine, greater positive remission with schizophrenia [26]
rs1801274	G/G	Fcγ-receptors (FCGR) FCGR2A gene	C/C	better survival of severe COVID-19 [27]; lower risk of lung cancer [28]
rs231775	G/G	cytotoxic T-lymphocyte- associated protein 4 CTLA4 gene	A/A	otherwise: 2.3x risk of Hashimoto's thyroiditis, 1.47x risk of Graves' disease [29]
rs1801394	G/G	methionine synthase reductase MTRR gene	A/A	otherwise: 1.4x higher risk for meningiomas [30]

References

[1] D. Gerhold, T. Rushmore, and C. T. Caskey, "Dna chips: promising toys have become powerful tools," *Trends in biochemical sciences*, vol. 24, no. 5, pp. 168–173, 1999. doi:10.1016/S0968-0004(99)01382-1.

[2] S. J. Watson and H. Akil, "Gene chips and arrays revealed: a primer on their power and their uses," *Biological Psychiatry*, vol. 45, no. 5, pp. 533–543, 1999. doi:10.1016/S0006-3223(98)00339-4.

[3] M. Gabig-Cimińska and G. Wegrzyn, "An introduction to dna chips: principles, technology, applications and analysis," *Acta Biochimica Polonica*, vol. 48, pp. 615–622, 2001. doi:10.18388/abp.2001_3896.

[4] H. K. Ratner, T. R. Sampson, and D. S. Weiss, "Overview of crispr–cas9 biology," *Cold Spring Harbor Protocols*, vol. 2016, no. 12, pp. pdb–top088849, 2016. doi:10.1101/pdb.top088849.

[5] J. G. Doench, E. Hartenian, D. B. Graham, Z. Tothova, M. Hegde, I. Smith, M. Sullender, B. L. Ebert, R. J. Xavier, and D. E. Root, "Rational design of highly active sgrnas for crispr-cas9–mediated gene inactivation," *Nature biotechnology*, vol. 32, no. 12, pp. 1262–1267, 2014. doi:10.1038/nbt.3026.

[6] M. W. Shen, M. Arbab, J. Y. Hsu, D. Worstell, S. J. Culbertson, O. Krabbe, C. A. Cassa, D. R. Liu, D. K. Gifford, and R. I. Sherwood, "Predictable and precise template-free crispr editing of pathogenic variants," *Nature*, vol. 563, no. 7733, pp. 646–651, 2018. doi:10.1038/s41586-018-0686-x.

[7] A. Pickar-Oliver and C. A. Gersbach, "The next generation of crispr–cas technologies and applications," *Nature reviews Molecular cell biology*, vol. 20, no. 8, pp. 490–507, 2019. doi:10.1038/s41580-019-0131-5.

[8] C. C. Chang, C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee, "Second-

generation plink: rising to the challenge of larger and richer datasets," *Gigascience*, vol. 4, no. 1, pp. s13742–015, 2015. doi:10.1186/s13742-015-0047-8.

[9] P. Cingolani, A. Platts, M. Coon, T. Nguyen, L. Wang, S. Land, X. Lu, and D. Ruden, "A program for annotating and predicting the effects of single nucleotide polymorphisms, snpeff: Snps in the genome of drosophila melanogaster strain w1118; iso-2; iso-3," *Fly*, vol. 6, no. 2, pp. 80–92, 2012. doi:10.4161/fly.19695.

[10] M. Van Oven and M. Kayser, "Updated comprehensive phylogenetic tree of global human mitochondrial dna variation," *Human mutation*, vol. 30, no. 2, pp. E386–E394, 2009. doi:10.1002/humu.20921.

[11] K. L. Hart, S. L. Kimura, V. Mushailov, Z. M. Budimlija, M. Prinz, and E. Wurmbach, "Improved eye-and skin-color prediction based on 8 snps," *Croatian medical journal*, vol. 54, no. 3, pp. 248–256, 2013. doi:10.3325/cmj.2013.54.248.

[12] W. McLaren, L. Gil, S. E. Hunt, H. S. Riat, G. R. Ritchie, A. Thormann, P. Flicek, and F. Cunningham, "The ensembl variant effect predictor," *Genome biology*, vol. 17, no. 1, pp. 1–14, 2016. doi:10.1186/s13059-016-0974-4.

[13] I. L. Rozhanskii, A. A. Klyosov, *et al.*, "Haplogroup r1a, its subclades and branches in europe during the last 9,000 years," *Advances in Anthropology*, vol. 2, no. 03, p. 139, 2012. doi:10.4236/aa.2012.23017.

[14] P. Brotherton, W. Haak, J. Templeton, G. Brandt, J. Soubrier, C. Jane Adler, S. M. Richards, C. D. Sarkissian, R. Ganslmeier, S. Friederich, *et al.*, "Neolithic mitochondrial haplogroup h genomes and the genetic origins of europeans," *Nature communications*, vol. 4, no. 1, p. 1764, 2013. doi:10.1038/ncomms2656.

[15] H. Sato, M. Iwano, Y. Akai, T. Nishino, T. Fujimoto, H. Shiiki, and K. Dohi, "Fcγriia polymorphism in japanese patients with systemic lupus erythematosus," *Lupus*, vol. 10, no. 2, pp. 97–101, 2001. doi:10.1191/096120301677569675.

[16] D. N. Forthal, G. Landucci, J. Bream, L. P. Jacobson, T. B. Phan, and B. Montoya, "Fcγriia genotype predicts progression of hiv infection," *The Journal of Immunology*, vol. 179, no. 11, pp. 7916–7923, 2007. doi:10.4049/jimmunol.179.11.7916.

[17] S. Sinha, S. K. Mishra, S. Sharma, P. K. Patibandla, P. K. Mallick, S. K. Sharma, S. Mohanty, S. S. Pati, S. K. Mishra, B. K. Ramteke, *et al.*, "Polymorphisms of tnf-enhancer and gene for fcγriia correlate with the severity of falciparum malaria in the ethnically diverse indian population," *Malaria Journal*, vol. 7, pp. 1–11, 2008. doi:10.1186/1475-2875-7-13.

[18] W. Zapata, W. Aguilar-Jiménez, N. Pineda-Trujillo, W. Rojas, H. Estrada, and M. T. Rugeles, "Influence of ccr5 and ccr2 genetic variants in the resistance/susceptibility to hiv in serodiscordant couples from colombia," *AIDS research and human retroviruses*, vol. 29, no. 12, pp. 1594–1603, 2013. doi:10.1089/aid.2012.0299.

[19] G. Atzmon, M. Rincon, C. B. Schechter, A. R. Shuldiner, R. B. Lipton, A. Bergman, and N. Barzilai, "Lipoprotein genotype and conserved pathway for exceptional longevity in humans," *PLoS biology*, vol. 4, no. 4, p. e113, 2006. doi:10.1371/journal.pbio.0040113.

[20] G. A. Erikson, D. L. Bodian, M. Rueda, B. Molparia, E. R. Scott, A. A. Scott-Van Zeeland, S. E. Topol, N. E. Wineinger, J. E. Niederhuber, E. J. Topol, *et al.*, "Whole-genome sequencing of a healthy aging cohort," *Cell*, vol. 165, no. 4, pp. 1002–1011, 2016. doi:10.1016/j.cell.2016.03.022.

[21] D. Ben-Avraham, D. R. Govindaraju, T. Budagov, D. Fradin, P. Durda, B. Liu, S. Ott, D. Gutman, L. Sharvit, R. Kaplan, *et al.*, "The gh receptor exon 3 deletion is a marker of male-specific exceptional longevity associated with increased gh sensitivity and taller stature," *Science advances*, vol. 3, no. 6, p. e1602025, 2017. doi:10.1126/sciadv.1602025.

[22] M. Saitou, S. Resendez, A. J. Pradhan, F. Wu, N. C. Lie, N. J. Hall, Q. Zhu, L. Reinholdt, Y. Satta, L. Speidel, *et al.*, "Sex-specific phenotypic effects and evolutionary history of an ancient polymorphic deletion of the human growth hormone receptor," *Science Advances*, vol. 7, no. 39, p. eabi4476, 2021. doi:10.1126/sciadv.abi4476.

[23] M. Kuningas, M. Putters, R. G. Westendorp, P. E. Slagboom, and D. Van Heemst, "Sirt1 gene, age-related diseases, and mortality: the leiden 85-plus study," *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, vol. 62, no. 9, pp. 960–965, 2007. doi:10.1093/gerona/62.9.960.

[24] C. Caruso, M. E. Ligotti, G. Accardi, A. Aiello, G. Duro, D. Galimberti, and G. Candore, "How important are genes to achieve longevity?," *International Journal of Molecular Sciences*, vol. 23, no. 10, p. 5635, 2022. doi:10.3390/ijms23105635.

[25] D. Cusi, C. Barlassina, T. Azzani, G. Casari, L. Citterio, M. Devoto, N. Glorioso, C. Lanzani, P. Manunta, M. Righetti, *et al.*, "Polymorphisms of α-adducin and salt sensitivity in patients with essential hypertension," *The*

Lancet, vol. 349, no. 9062, pp. 1353–1357, 1997. doi:10.1016/S0140-6736(97)01029-5.

[26] D. H. Adams, S. Close, M. Farmen, A. M. Downing, A. Breier, and J. P. Houston, "Dopamine receptor d3 genotype association with greater acute positive symptom remission with olanzapine therapy in predominately caucasian patients with chronic schizophrenia or schizoaffective disorder," *Human Psychopharmacology: Clinical and Experimental*, vol. 23, no. 4, pp. 267–274, 2008. doi:10.1002/hup.930.

[27] R. López-Martínez, G. M. Albaiceta, L. Amado-Rodríguez, E. Cuesta-Llavona, J. Gómez, M. García-Clemente, D. Vázquez-Coto, V. Alvarez, and E. Coto, "The fcgr2a rs1801274 polymorphism was associated with the risk of death among covid-19 patients," *Clinical Immunology*, vol. 236, p. 108954, 2022. doi:10.1016/j.clim.2022.108954.

[28] J. He, L. Yu, Z. Qiao, B. Yu, Y. Liu, and H. Ren, "Genetic polymorphisms of fcgr2a, orai1 and cd40 are associated with risk of lung cancer," *European Journal of Cancer Prevention*, vol. 31, no. 1, pp. 7–13, 2021. doi:10.1097/CEJ.0000000000000671.

[29] W.-H. Ting, M.-N. Chien, F.-S. Lo, C.-H. Wang, C.-Y. Huang, C.-L. Lin, W.-S. Lin, T.-Y. Chang, H.-W. Yang, W.-F. Chen, *et al.*, "Association of cytotoxic t-lymphocyte-associated protein 4 (ctla4) gene polymorphisms with autoimmune thyroid disease in children and adults: case-control study," *PloS one*, vol. 11, no. 4, p. e0154394, 2016. doi:10.1371/journal.pone.0154394.

[30] L. Bethke, E. Webb, A. Murray, M. Schoemaker, M. Feychting, S. Lönn, A. Ahlbom, B. Malmer, R. Henriksson, A. Auvinen, *et al.*, "Functional polymorphisms in folate metabolism genes influence the risk of meningioma and glioma," *Cancer Epidemiology Biomarkers & Prevention*, vol. 17, no. 5, pp. 1195–1202, 2008. doi:10.1158/1055-9965.EPI-07-2733.

Supplementary materials

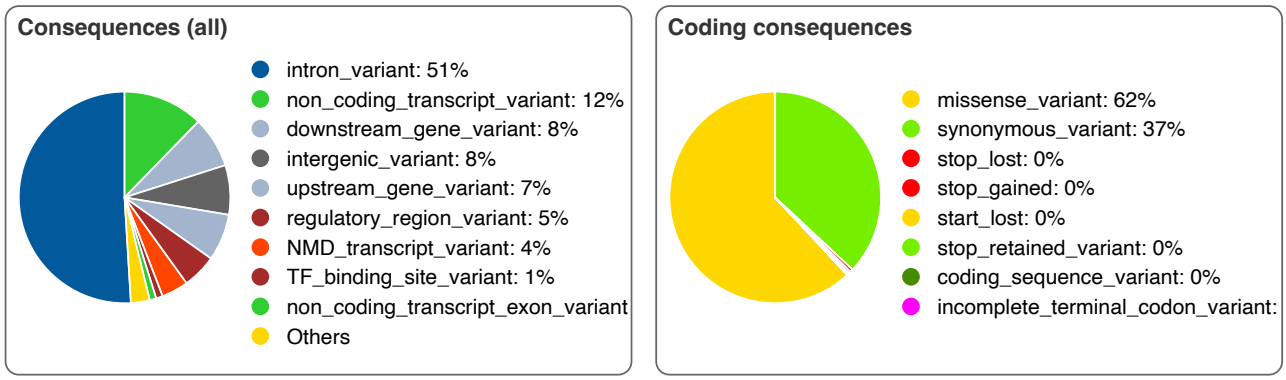

Figure 2: Distribution of parameters in the VEP output