Esin Onal

Analysis of Genome to Identify Variants of Interest: Chromosome 7 Gene NOS3- Implications on Heart Disease

Introduction

The Human Genome Project was the first project to determine the DNA sequence of the whole human genome. Mapping the "human genome" involved sequencing a small number of people and merging the data together. The result represents not an individual's genome but multiple people's genomes assembled together. Every genome is unique, however. It is now possible to sequence and analyze individual genomes. This is called personal genomics.

Personal genomics is relevant as it can give information about which diseases a person may be likely to get, as well as how they might respond to particular diseases. Personal genomics might also give information on how they might respond to particular drugs: this is the field of pharmacogenomics, where doctors need to ensure the drug has the maximum positive effect and minimum side effects. Personal genomics can also be used to advise couples who want children, as if both parents are carriers of diseases like cystic fibrosis, options like in-vitro fertilization might be considered to allow screening of the embryo to ensure it is healthy.(1)

Thanks to advanced DNA sequencing technologies in the last decade, personal genomics is becoming more and more accessible to the public. (2) This is because the introduction of NGS added high-throughput sequencing technology while at the same time drastically lowering the cost of sequencing. Prior to NGS, the sequencing method used was Sanger's method, which was expensive. As a result, genomics could only be used to check for single genes that were linked to diseases, like Huntington's, cancer, and Cystic Fibrosis. Thanks to its capacity to map large regions of the genome, NGS is now allowing the possibility to check for more complicated diseases and traits. More rare variations are being researched and discovered, and as a result can be checked for in personal genomics tests, making the tests more diverse and informative.(3)

Exome sequencing is another way that technology has made sequencing cost cheaper, as it makes it possible to isolate and sequence only the exons, which cover only about 1-2% of the genome. Different kits are made by Agilent, Roche, and Illumina for exon capture.(3) The capture methods have to consider target choice, oligonucleotide bait length, and bait density, and are different based on the technology used.(3) The exome is enriched for functionally relevant loci. (3) Exome sequencing is usually done at high coverage (X100) in order to be reliable. (3) However, the capture of exons can be unbalanced and some exons might be missed, because of problems in the target-probe hybridization. Sanger methods can be used to trace the regions that are missed. (3) Exome sequencing is helpful for finding variants that cause monogenic diseases.(3)

This project will be looking at a genome from the Human Genome Project. The goal is to assess what kind of genetic variants the person could have that may lead to diseases. There were multiple parts to this project. Aligning was first done in order to arrange the sequence to identify regions of similarity to the human reference genome. Then genotyping was done in order to find the genotypes at each particular site in the genome. The probable genotypes

at each site can be found by looking at the nucleotides read at each site and considering the probability of seeing that nucleotide for each possible genotype. This process especially considers the SNPs and indels. The last step was annotation, which is used to add information on the variants, like which variants would be worth investigating further. Some variants may not be interesting, especially if they occur in the introns or if they are non-synonymous. The cases that occur in protein-coding regions where functional groups are important or the mutation creates a change in the amino acid would be worth investigating further. Annotation therefore adds biological information on top of our genetic data. The information added include genomic content (exonic/intronic), mutation type (synonymous/non-synonymous), RefGene ID, and metrics to see how harmful the mutation would be.

Methods

Alignment is computationally expensive, as it involves aligning millions of short nucleotide sequences to a human genome. As a result, the Burrows Wheeler algorithm was used, as it is consistent and fast at aligning short sequences to large genomic databases. The BWA program created two .sai files, which were then merged to a .bam file, after finding the best alignment position for each pair of the paired reads. The .bam file was then sorted according to the genomic coordinates of the human genome.

The genotyping algorithm was run on all the chromosomes. Calculations were weighted by allele frequency of previously found SNPs. The results were filtered by Quality Control so that the variants from well-aligned regions were kept. Those with PHRED scores of lower than 50 were removed. The genotype of the individual where they deviate from the reference genome were kept. The –q parameter was for retrimming. There was a read depth set of 100.

For annotation, the ANNOVAR tool was used to add information about the variants. The computation for annotation was not expensive.

Results

The number of variants that we found were 139,605, which includes both the high and low quality variants. The number of variants that were high quality were 95,268. All future analysis will consider only the high quality variants. The variants that were exonic were 15,077. This includes both exonic and ncRNA_exonic variants. The number of SNPs were 88,746. The number of indels were 6,521. The number of variants that were synonymous were 7,376. The number of variants that were non-synonymous were 6,056. The number of variants that were frameshift were 45. The number of variants that were premature stop codons were 50.

Location of Variant	Type of Variant	Implications
Chromosome 4, Position 186236880	Non-synonymous SNV	Prekallikrein deficiency, problems related to blood clotting
Chromosome 5, Position 74685445	Non-synonymous SNV	Sandhoff disease, infantile-type

*Chromosome 5, Position 177093242	Non-synonymous SNV	Cancer progression and tumor cell motility
Chromosome 6, Position 51746853	Non-synonymous SNV	Autosomal recessive polycystic kidney disease
*Chromosome 7, Position 150999023	Non-synonymous SNV	Susceptibility to coronary artery spasm, late-onset Alzheimer disease, pregnancy-induced hypertension, hypertension resistant to conventional therapy, ischemic heart disease, ischemic stroke
Chromosome 9, Position 133436862	Non-synonymous SNV	Upshaw Shulman syndrome
Chromosome 11, Position 66560624	Stop-gain	Impairs muscle performance, muscle protein actinin-3
Chromosome 12, Position 101770477	Non-synonymous SNV	Mucolipidosis, inherited metabolic disorder
Chromosome 12, Position 121857429	Non-synonymous SNV	Maturity-onset type-3 diabetes of the young
Chromosome 16, Position 56514589	Non-synonymous SNV	Bardet-biedl syndrome
Chromosome 16, Position 69711242	Non-synonymous SNV	Susceptibility to benzene toxicity, post-chemotherapy leukemia, poor survival in post-chemotherapy breast cancer, lung cancer
Chromosome 16, Position 8811153	Non-synonymous SNV	Carbohydrate deficient glycoprotein syndrome, type 1, congenital disorder of glycosylation
*Chromosome 16, Position 27344882	Non-synonymous SNV	Susceptibility to atropy, slow progression to acquired immuno-deficiency syndrome
Chromosome 19, Position 12899706	Non-synonymous SNV	Glutaric aciduria, type 1
Chromosome 19, Position 48965830	Synonymous SNV	Sporadic abdominal aortic aneurism, neuroferritinopathy, hyperferritinemia cataract syndrome, glycogen storage disease (muscle)

*Picked for further analysis

There are 88,304 variants that are associated with disease through GWAS. The chromosomes most associated with disease through GWAS were chromosomes 1, 2, and 19, where chromosome 1 had 8583 RS numbers. The ones least associated were 18, 21, and 22, where chromosome 21 had 1304 RS numbers.

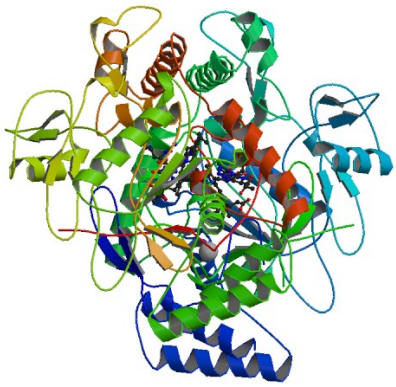
On chromosome 5 at position 177093242, we identified a non-synonymous SNP in gene FGFR4 that changed the amino acid from G to A. It was in an exomic part of the genome, and it was a pathogenic change. It is identified with cancer progression and tumor cell motility. This gene is a protein-coding gene, and plays a role in the regulation of cell proliferation, differentiation and migration, and in regulation of lipid metabolism, bile acid biosynthesis, glucose uptake, vitamin D metabolism and phosphate homeostasis. Diseases associated with this gene are Prostate Cancer and Rhabdomyosarcoma. (5) The FGFR4 SNP rs351855, which is also associated with poor prognosis, was found in the 1000 Genomes Project database at a minor allele frequency of 0.30 and was found in approximately 50% of patients with cancer.(6) The mutation alters the transmembrane-spanning segment and exposes a membrane-proximal cytoplasmic STAT3-binding site. (6) The RS ID associated with this mutation is rs351855. According to the 1000Genome study, this variant has high allele frequencies in east Asian populations, where the they are at 0.46. It least affects African populations, where they are at 0.11. (7)

On Chromosome 7 at position 150999023 we identified a non-synonymous SNV that was associated with susceptibility to coronary artery spasm, late-onset Alzheimer disease, pregnancy-induced hypertension, hypertension resistant to conventional therapy, ischemic heart disease, and ischemic stroke. This change from T to G was located in the exomic part of the genome. It was in the NOS3 gene and it was a pathogenic change. The NOS3 gene is a protein-coding gene and variations in the gene are associated with susceptibility to coronary spasm. Diseases that are associated with this gene include stroke, ischemic, and Alzheimer disease. The pathways involved are the Doxorubicin Pathway (Cardiomyocyte Cell), Pharmacodynamics, and eNOS activation and regulation. (8) In patients with coronary spasm, the incidence of the allele was significantly greater than in the control group, where regression analysis showed that the most predictive independent risk factor for coronary spasm was this mutant allele. (9)The RS ID associated with this mutation is rs1799983. According to the 1000Genome study, this variant has high allele frequencies in most populations, though the frequencies for Africa seem to be higher (0.93) than Europe (0.65). (10)

On Chromosome 16 at position 27344882 we identified a non-synonymous SNV in the exomic part of the genome in gene IL4R that caused a change in the amino acid from A to G. The change was associated with a pathogenic change of susceptibility to atropy, and slow progression to acquired immuno-deficiency syndrome. Atropy is the predisposition to respond immunologically to different antigens/allergens, leading to overproduction of immunoglobulin. The result is the propensity to develop reactions to allergens. Manifestations of atopy include allergic bronchial asthma, allergic rhinitis, atopic dermatisis and food allergy. (11) Among 50 prospectively recruited adults, the mutant allele from A to G was found in 13 of 20 subjects with atopy but only in 5 of 30 without atopy. The relative

risk of atopy among those with a mutant allele was 9.3. (12) The mutant allele was associated with higher levels of expression of CD23 by interleukin-4. The IL4R gene is a protein coding gene that encodes the alpha chain of the interleukin-4 receptor, a type I transmembrane protein that can bind interleukin 4 and interleukin 13 to regulate IgE production. The gene is expressed in the lymph node. (13) The RS ID associated with this mutation is rs1805010. According to the 1000Genome study, this variant has allele frequencies, ranging around 0.42-0.51 in different populations, where the highest was in East Asian populations. (14)

Discussion: Chromosome 7 NOS3 gene T to G mutation, Implications on Heart Disease



The NOS3 gene encodes the protein nitric oxide (NO), which is implicated in vascular smooth muscle relaxation through a cGMP-mediated signal transduction pathway. It also mediates vascular endothelial growth factor (VEGF)-induced angiogenesis in coronary vessels and it promotes blood clotting through the activation of platelets.(15) NOS3 is in the NOS3 protein family. (16) NO is synthesized in endothelial cells from L-arginine by nitric oxide synthase (NOS).(17) The diagram on the right shows the structure of an endothelial 3NOS domain, which has a cyclic global symmetry.(18) The expression of the gene is largely in the spleen. The gene contains 26 exons spanning approximately 21 kb of genomic DNA and encodes an mRNA of 4,052 nucleotides. (17) The four phenotypes associated with this gene are Alzheimer disease, Hypertension, Ischemic stroke, and Preeclampsia.(16) The inheritance of Alzheimer disease, coronary artery spasm, and pregnancy-induced hypertension were autosomal dominant.(17)

The variant in the NOS3 gene is thought to have a role in heart disease. One of the earliest studies done on this was on the correlation between the variant and coronary spasm. Endothelial-derived nitric oxide (NO) has been implicated in coronary spasm: NO activities are impaired in the coronary arteries of patients with coronary spasm. A study on 113 patients with coronary spasm was made in Japan, where the diagnosis of coronary spasm was made by intracoronary injection of acetylcholine. There were 100 control subjects. A significant difference was found in the distribution of the variant; 21.2% of the coronary spasm group and 9.0% of the control group ($p = 0.014$ for dominant effect) showed the variant. (19)

The most recent study on the link between the variant and heart disease was in 2007, with a much larger sample size, studying strokes. With 1,901 hospitalized stroke patients and 1,747 population controls, multiple PCR along with allele-specific hybridization was used for genotype determination. This was performed in two different regions in Germany in two separate case controlled studies. The study analyzed the contribution of 106 informative single nucleotide polymorphisms (SNPs) from 63 candidate genes for cardiovascular diseases for the risk of stroke. Permutation testing was used to narrow the SNPs down to 5 SNPs, located in the nitric oxide synthase 3, the alpha 2 integrin, the interleukin 13, the selectin P and the chemokine receptor 2 genes.

For one of the SNPs - the SNP located in NOS3 - the variant had an association with ischemic stroke. This was also replicated in the second study, and also in a combined analysis of both studies. The p value of the significance for the combined studies was significant, at p=0.0009. The analysis was done using multiple regression. The association was independent of age, gender, hypertension, diabetes and hypercholesterolemia in both studies. The variant in NOS3 was therefore found to have a role in stroke onset. (20) The table below provides the odds ratios.(20)

Table 3 Odds ratios for two SNPs in cardiovascular candidate genes and six SNPs in inflammation candidate genes and the risk of ischemic stroke in the Westphalian and Pomeranian studies, according to three models of inheritance

Cardiovascular SNP panel							
SNP (rs number)		Westphalia (1,417 cases, 996 controls)			Pomerania (484 cases, 751 controls)		
		OR[a]	95% CI	*P* (2-tailed)	OR[a]	95% CI	*P* (2-tailed)
Nitric oxide synthase 3, glu298asp (rs 1799983)							
Additive	wt	1.00	Ref.		1.00	Ref.	
	het	1.31	1.05-1.63		1.23	0.97–1.57	
	hom	1.65	1.13–2.41		1.40	0.91–2.14	
	Trend	*1.29*	*1.10–1.52*	*0.002*	*1.20*	*1.00–1.44*	*0.045*
Dominant		1.36	1.11–1.68	0.004	1.26	1.00–1.58	0.05
Recessive		1.42	1.01–2.08	0.05	1.27	0.84–1.91	0.26
Alpha 2 integrin, 873G/A (rs 1062535)							
Additive	wt	1.00	Ref.		1.00	Ref.	
	het	1.15	0.92–1.45		1.06	0.82–1.38	
	hom	1.34	0.98–1.84		1.33	0.95–1.87	
	Trend	*1.16*	*1.00–1.34*	*0.05*	*1.14*	*0.96–1.34*	*0.13*
Dominant		1.20	0.97–1.48	0.10	1.13	0.89–1.44	0.33
Recessive		1.24	0.93–1.65	0.14	1.28	0.95–1.74	0.11
Inflammation SNP panel							
SNP (rs number)		Westphalia (503 cases, 546 controls)			Pomerania (441 cases, 740 controls)		
		OR[a]	95% CI	*P* (2-tailed)	OR[a]	95% CI	*P* (2-tailed)
Interleukin 13, 4045C/T (rs 1295686)							
Additive	wt	1.00	Ref.		1.00	Ref.	
	het	1.15	0.85–1.56		1.04	0.81–1.34	
	hom	2.50	1.28–4.89		1.15	0.65–2.04	
	Trend	*1.34*	*1.05–1.69*	*0.02*	*1.06*	*0.86–1.29*	*0.61*
Dominant		1.28	0.96–1.71	0.10	1.05	0.83–1.34	0.67
Recessive		2.39	1.23–4.63	0.01	1.13	0.64–1.99	0.67
Selectin P, 40G/A (*rs 6131*)							
Additive	wt	1.00	Ref.		1.00	Ref.	
	het	1.27	0.94–1.72		1.10	0.85–1.42	
	hom	2.12	0.89–5.06		1.40	0.78–2.51	
	Trend	*1.32*	*1.02–1.72*	*0.04*	*1.13*	*0.92–1.39*	*0.24*
Dominant		1.32	0.98–1.78	0.06	1.13	0.88–1.44	0.34
Recessive		1.96	0.83–4.66	0.13	1.36	0.76–2.42	0.30
Chemokine receptor 2, 46295G/A (rs 1799864)							
Additive	wt	1.00	Ref.		1.00	Ref.	
	het	1.59	1.08–2.34		1.14	0.84–1.55	
	hom	1.68	0.36–7.90		1.32	0.29–5.98	
	Trend	*1.53*	*1.07–2.18*	*0.02*	*1.14*	*0.85–1.52*	0.38
Dominant		1.59	1.09–2.33	0.02	1.14	0.84–1.55	0.39
Recessive		1.57	0.34–7.34	0.56	1.29	0.29–5.83	0.74

[a] Odds ratio derived from logistic regression model with case–control status as the dependent variable and adjusted for age (5-year age groups) and gender

The strengths of this study are the large sample size, as well as the conduction of an independent replication study with identical methods that also confirms the same results. Additionally the genetic differences between the two groups are small as both the groups were from Germany. Some limitations are that populations with different genetic

backgrounds would need to be analyzed, and that an alternative explanation for the findings could be that the SNP NOS3 might be in linkage disequilibrium with one or more other SNPs within the NOS3 gene or the genes next to it. This points to a larger problem in exome sequencing, that exome sequencing does not pick up on gene-to-gene interaction, for example a variant that extends axons might help to make a case of spinal muscular atrophy much milder. (20)

There are several downsides related to exome sequencing. It is difficult to identify rare variants and to know which variants cause disease. There isn't one reference genome so based on what you pick the results can be different. Different software packages also make different assumptions about thresholds for quality scores and how sequences are mapped. The methods that are currently in use – alignment, quality assessment, checking variants and interpreting significance – are very complex and are constantly changing. There are also many different workflows that can be applied to genomes, and applying two different workflows to the same FASTQ data will give different variants. (3)

Further, there is a lot with exome sequencing that is missed. Genes are made up of more than just exons, and introns can actually play an important role in producing proteins. For example, 15% of disease-causing variants are in the introns, at the exon-intron boundary. (21) These variants would be missed in exome sequencing. Whole genome sequencing would be preferred, as soon as costs go down. Environmental factors can affect DNA, for example it can put methyl groups onto the DNA, which would that block the expression of genes, so exome sequencing would not be able to pick up certain genes.

Structural variants that move DNA but don't alter the base sequence are not picked up. It also cannot pick up triplet expansion disorders like Huntington's disease, which increase the number of trinucleotide repeats in genes, but don't change the base sequence. (22) It also does not check whether mutations come from just the mom, just the dad, or both, which can be a vital part of mechanism for some diseases, for example leading to cases of cystic fibrosis. (23) Additionally, the current understanding of the genome limits what can be done with the exomes, as the nucleotides that are not recognized to be functional will be missed. This means variants that could be vital are missed.(24) Also, variant checking is done through looking at databases, so if new variants are added but these are not reflected in the databases, this could lead to missing important variants.

Knowing these variants could be extremely helpful. Whole exome sequencing is usually considered for patients who have complex disorders that affect multiple organs or body types, when more than one disorder is suspected, and when genetic testing has already been performed and not given any results. If there is any diagnosis to be had, this could be invaluable. It is especially ideal for finding Mendelian diseases.(24) As with any

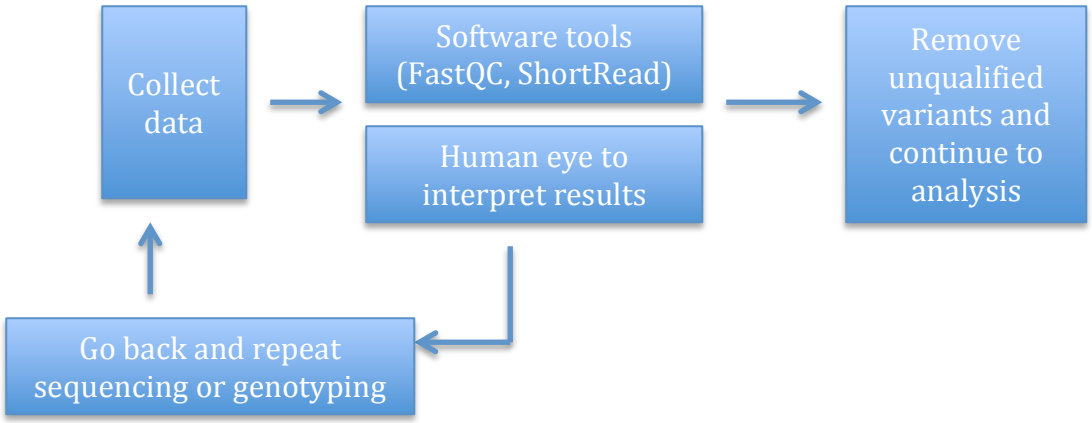
technological tool, however, it also has the potential to be interpreted incorrectly, so explaining results for the patients and addressing the ethical challenges is essential.

Currently using exome sequencing in a clinical setting only presents a diagnosis for 25% of patients, where the rest of patients receive negative results. Since the majority of patients who do exome sequencing have likely exhaustive considered other options to learn more about their disease, 25% is a good amount. Nonetheless, it still means the majority of patients in a clinical setting are not receiving any information about their disease. Additionally, interpreting the negative results correctly is also extremely important. These negative results could mean that there actually wasn't a diagnosis to be had, but it could also mean that for the current moment in time there was no match for the genome given the databases, as clinical databases for variants may be incomplete on what counts as a clinically significant finding for a diagnosis. (25) It's important to keep the databases up to date and to properly inform patients on the meaning of negative diagnoses.

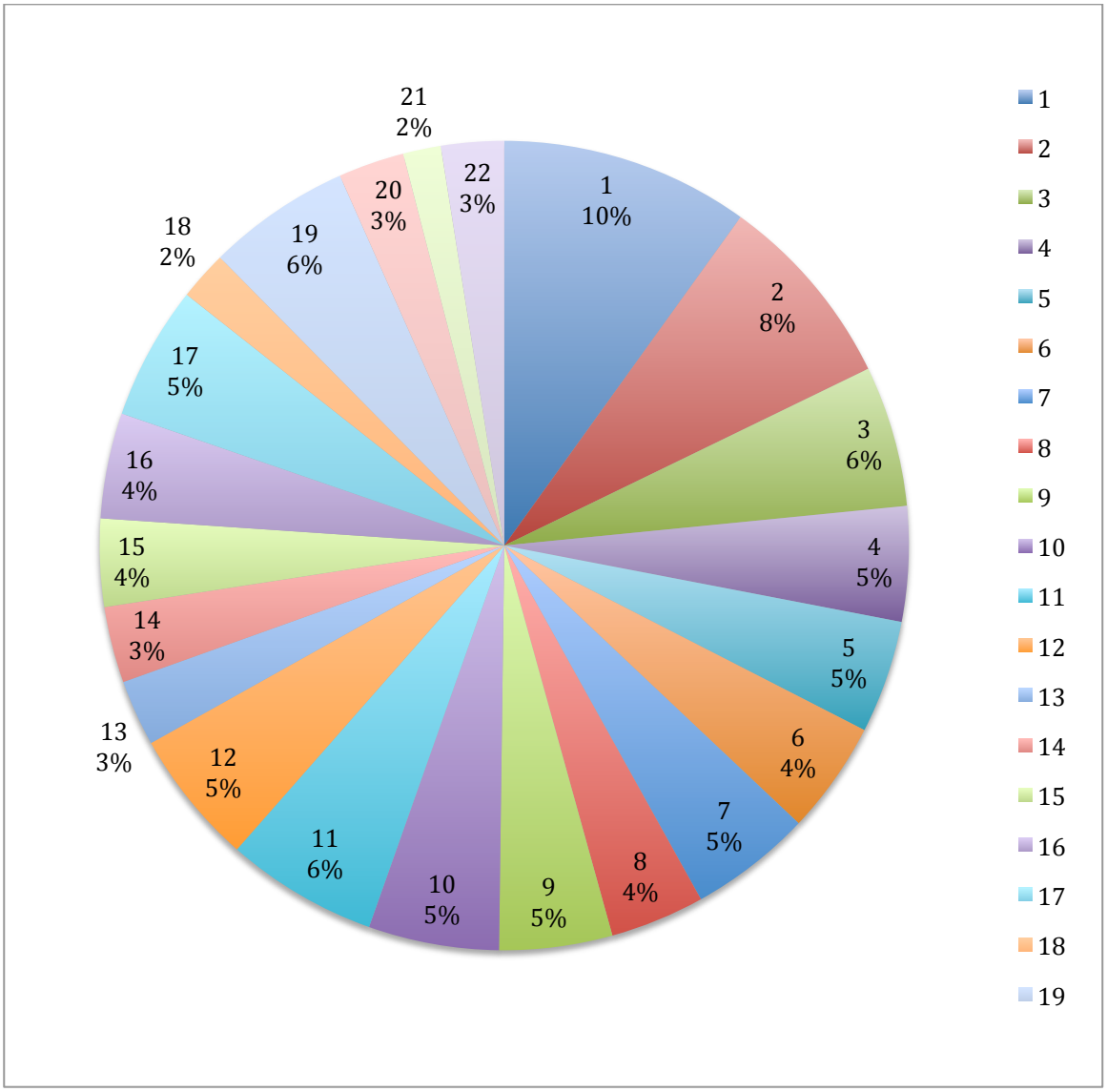
Some caveats of personal genomics are the control of data over the genome. In a clinical setting, there is the issue of what may happen with the data, if it is possible to choose between advancement of science and the best interests of the patient. In direct-to-consumer tests, it is unclear what users expect from the tests. Most seem to perceive them as having the power of predicting outcomes for complex disorders, however, the predictive power of these tests are weak, and users may be left unsatisfied. This is especially the case as individuals may not get access to their data, and it is up to the discretion of each company what they do with it. (26)

EC

There are several things one could do to improve the quality control methods in alignment and genotyping. There are several software packages like FastQC and ShortRead that can be used to better understand the quality of the sequencing. They show base per sequence quality (y-axis=quality, x-axis=base position), per sequence quality score, overrepresented sequences – which can be saved into a separate file for BLAT or BLASTN, and k-mer content (which shows data for a sequence of 5 sets of 5 nucleotides) plotted by relative enrichment and position in read. Errors should also be visually analyzed, including a quality control plot, a histogram, boxplots, and an MA plot. Outliers in the QC step have to be analyzed and regions in an array that have defects (scratches, hybridization errors), should be identified. (3) A fuller analysis of errors, using both software and a human eye, is needed before continuing immediately onto the analysis of the variants step. A decision would then be made about going forward or going back and collecting data again, maybe in a different way. A workflow is given below.



The below is a pie chart showing all mutations on different chromosomes.



Conclusion

The current growth of the field, especially since NGS technologies have been invented, is promising. The limitations of exome sequencing at the moment, which include inability to check for gene-to-gene interactions, and inability to give diagnosis to most patients, are likely to improve in the future. This person who has been studied has several pathogenic variants that may be likely to make them more susceptible to disease, some of which include heart disease, atopy, and cancer progression. Knowing their variants could be a great way for them to practice preventative approaches towards the diseases they might be more predisposed towards, and to have some information on what has worked for others with similar genotypes. This approach towards personalized medicine could be an incredibly illuminating way to treat patients.

References

(1) "Personal Genomics: the Future of Healthcare?" Stories, The Public Engagement Team at the Wellcome Genome Campus, 14 June 2016, https://www.yourgenome.org/stories/personal-genomics-the-future-of-healthcare.

(2) Riposati, Andrea. "Whole Genome Sequencing Cost." Dante Labs, Dante Labs, 26 Mar. 2018, https://us.dantelabs.com/blogs/news/whole-genome-sequencing-cost.

(3) Pevsner, Jonathan. Bioinformatics and Functional Genomics. Wiley Blackwell, 2016.

(5) Database, Gene. "FGFR4 Gene (Protein Coding)." GeneCards, https://www.genecards.org/cgi-bin/carddisp.pl?gene=FGFR4.

(6) "FIBROBLAST GROWTH FACTOR RECEPTOR 4; FGFR4." OMIM, https://www.omim.org/entry/134935#0001.

(7) "rs351855 RefSNP Report - DbSNP - NCBI." National Center for Biotechnology Information, U.S. National Library of Medicine, https://www.ncbi.nlm.nih.gov/snp/rs351855#frequency_tab.

(8) Database, Gene. "NOS3 Gene (Protein Coding)." GeneCards, https://www.genecards.org/cgi-bin/carddisp.pl?gene=NOS3.

(9) "NITRIC OXIDE SYNTHASE 3; NOS3." OMIM, https://omim.org/entry/163729.

(10) "rs1799983 RefSNP Report - DbSNP - NCBI." National Center for Biotechnology Information, U.S. National Library of Medicine, https://www.ncbi.nlm.nih.gov/snp/rs1799983#frequency_tab.

(11) Vaillant, Angel A. Justiz. "Atopy." StatPearls [Internet]., U.S. National Library of Medicine, 6 Nov. 2019, https://www.ncbi.nlm.nih.gov/books/NBK542187/.

(12)“INTERLEUKIN 4 RECEPTOR; IL4R.” OMIM, https://www.omim.org/entry/147781#allelicVariants.

(13) “IL4R Interleukin 4 Receptor [Homo Sapiens (Human)] - Gene - NCBI.” National Center for Biotechnology Information, U.S. National Library of Medicine, https://www.ncbi.nlm.nih.gov/gene/3566#gene-expression.

(14) “rs1805010 RefSNP Report - DbSNP - NCBI.” National Center for Biotechnology Information, U.S. National Library of Medicine, https://www.ncbi.nlm.nih.gov/snp/rs1805010#frequency_tab.

(15) UniProt ConsortiumEuropean Bioinformatics InstituteProtein Information ResourceSIB Swiss Institute of Bioinformatics. “Nitric Oxide Synthase, Endothelial.” UniProt ConsortiumEuropean Bioinformatics InstituteProtein Information ResourceSIB Swiss Institute of Bioinformatics, 13 Nov. 2019, https://www.uniprot.org/uniprot/P29474.

(17) “NITRIC OXIDE SYNTHASE 3; NOS3.” OMIM, https://omim.org/entry/163729.
(18) Niu, et al. “Structural Characterization of Nitric Oxide Synthase Isoforms Reveals Striking Active-Site Conservation.” Nat.Struct.Mol.Biol., http://www.rcsb.org/structure/3NOS.

(19) Yoshimura, M, et al. “A Missense Glu298Asp Variant in the Endothelial Nitric Oxide Synthase Gene Is Associated with Coronary Spasm in the Japanese.” Human Genetics, U.S. National Library of Medicine, July 1998, https://www.ncbi.nlm.nih.gov/pubmed/9737779.

(20) Berger, Klaus, et al. “The glu298asp Polymorphism in the Nitric Oxide Synthase 3 Gene Is Associated with the Risk of Ischemic Stroke in Two Large Independent Case-Control Studies.” Human Genetics, U.S. National Library of Medicine, Apr. 2007, https://www.ncbi.nlm.nih.gov/pubmed/17165044.

(21) “6 Things You Might Not Know about Whole Exome Sequencing.” Blueprint Genetics, https://blueprintgenetics.com/resources/6-things-you-might-not-know-about-whole-exome-sequencing/.

(22) Spence, J E, et al. “Uniparental Disomy as a Mechanism for Human Genetic Disease.” American Journal of Human Genetics, U.S. National Library of Medicine, Feb. 1988, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1715272/.

(23) Lewis, Ricki. “10 Things Exome Sequencing Can t Do-but Why It s Still Powerful.” Scientific American Blog Network, Scientific American, 16 May 2012, https://blogs.scientificamerican.com/guest-blog/10-things-exome-sequencing-cant-do-but-why-its-still-powerful/.

(24) Biesecker, Leslie G, et al. “Exome Sequencing: the Expert View.” Genome Biology, BioMed Central, 14 Sept. 2011, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3308041/.

(25) Skinner, Debra, et al. “The Nuanced Negative: Meanings of a Negative Diagnostic Result in Clinical Exome Sequencing.” Sociology of Health & Illness, U.S. National Library of Medicine, Nov. 2016, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5089912/.

(26) Angrist, Misha. “Personal Genomics: Where Are We Now?” Applied & Translational Genomics, Elsevier, 1 Feb. 2016, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4796805/.

Citations created with EasyBib.