Proceedings of the 2nd International Conference on Modern Medicine and Global Health
DOI: 10.54254/2753-8818/29/20240753

The bioinformatic analysis of CFTR in essential hypertension

Junzhen Zhu

University of Glasgow, University Ave, Glasgow G12 8QQ, UK

zhujunzhen959@foxmail.com

Abstract. The CFTR gene is associated with cystic fibrosis, a genetic disease primarily affecting the respiratory and digestive systems. Essential hypertension is a common cardiovascular disorder characterized by high blood pressure. In this bioinformatic analysis, various databases and tools were used to investigate the potential mechanisms underlying this association. Gene expression data and epigenetic analyses were used to identify the biological processes between CFTR and hypertension. Additionally, genetic variants within CFTR were analysed for potential effects on hypertension susceptibility. The results of this analysis suggest that CFTR may not play a role in hypertension. Moreover, hypertension has no special epigenetic feather. Further studies are needed to confirm the mechanisms of such a phenomenon.

Keywords: CFTR, essential hypertension, bioinformatics

1. Introduction

Essential hypertension (EH) is hypertension whose cause is unknown, and its cases account for 95% of hypertension (5% result from other diseases); EH is associated with both genes and environment [1]. Cystic fibrosis transmembrane conductance regulator (CFTR) is a type of ABC transporter, and its variants lead to cystic fibrosis, a genetic disorder in northern Europe [2]. Zhao's research shows CFTR is a negative regulator of blood pressure [3]. However, most studies of CFTR focus on cystic fibrosis. And the studies of essential hypertension limit in some key genes in hypertension, like MOV10, ULK4 and CSK [4]. Therefore, the bioinformatic research in the relation of CFTR and hypertension may be valuable to fill the gap. The aim of this research is to analyse the relationship between CFTR variants and essential hypertension from the perspective of nucleic acid sequence and methylations. The data is collected from the article database PubMed, gene database GenBank and IGSR and methylation database EWAS. The result is based on the statistical test of the data. The assignment is to explore the answers to some questions. First, is the epigenetic feature of the CFTR gene a high frequency factor in hypertension patients' chromosomes? Second, what are the highly frequent CFTR variants in hypertension patients' samples? The last one is which variant is more possible to lead to hypertension and will the syndrome be more serious if the patient has such a variant. The remainder of this assignment is structured as follows. Section 2 shows the previous studies about hypertension's gene inducibility, the relation of CFTR and hypertension in biochemistry and the bioinformatic analysis of hypertension of CFTR. Section 3 provides the method of the research. Section 4 shows the data and the statistical tests. Section 5 discusses the result of the analysis. The last section summarizes other sections, giving a conclusion and advice on the future study.

DOI: 10.54254/2753-8818/29/20240753

2. Literature Review

2.1. Introduction

Essential hypertension (EH) is hypertension whose cause is unknown, and its cases account for 95% of hypertension (5% result from other diseases); EH is associated with both genes and environment [1]. Cystic fibrosis transmembrane conductance regulator (CFTR) is a type of ABC transporter, and its variants lead to cystic fibrosis, a genetic disorder in northern Europe [2]. Zhao's research shows CFTR is a negative regulator of hypertension [3]. However, the significance of CFTR in human hypertension has not been evaluated. The aim of this literature review is to discuss and evaluate existing research on CFTR, essential hypertension and bioinformatical methods to find the CFTR's effect in human hypertension patients via bioinformatic analysis.

2.2. Theme 1: Essential hypertension and genetics factors

There is a large body of literature concerning essential hypertension and genetics. Padmanabhan et al. found more than 100 single nucleotide polymorphisms (SNPs) and proved that these SNPs are associated with blood pressure regulation [5]. The method of the research is genome wide association studies, which allow us to suppose the function of genes by genome comparison and function net association. This method is a population-based study, so it can't reflect the gene-environment interaction and the gene-gene interaction, which is important for the causes of essential hypertension. The meaning of this literature is to know how to find the SNPs in the CFTR gene and check whether the SNP is a blood pressure regulation. However, the research did not find a significant SNP for hypertension, only to find the SNP for blood pressure. Besides, in this article, they suggested another type of hypertension—monogenic forms of hypertension, the mutation of this type of hypertension accounts for Mendelian forms, and the research of this type needs family studies. And CFTR is the potential factor of polygenic hypertension, so it can also be researched by wide bioinformatic analysis.

Wei et al. 's research is also about SNPs [6]. They succussed in finding evidence about hypertension and genes using another method—they did a case-control study on two specific genes consisting of 816 patients and 836 non-patients; then they used the SNP-environment software GMDR to find the gene-environment interaction and genetic dominance model analysis to find the gene-gene interaction. As a result, they concluded that the two genes were negatively correlated and one of them has an interaction with BMI. However, the individuals are all from an ethnic minority of China, and they can find more individuals from other nations to evaluate the interaction. Compared with the research of Padmanabhan's team, Wei's team chose the method that can provide more information about a specific gene and be more accurate. However, the former's research can find more SNPs in a wider range of populations.

Kato et al. 's research focused on the epigenetic factors rather than DNA sequence (or SNPs)[7]. They also used the genome-wide association study and human methylation bead chip to find the relationship between methylation and variation associated with blood pressure. It also shows another use of genome-wide association study. Individuals from at least three different nations were included in the study. The study's conclusion highlights the connection between genetic variation and blood pressure through DNA methylation. This research demonstrates that methylation can affect blood pressure and that epigenetic factors can contribute to abnormal occurrences in CFTR.

2.3. Hypertension and CFTR gene

A large of studies have focused on CFTR and hypertension. In Zhao's research, the team knocked down the CFTR gene in mice [3]. Then, they found that the mice whose CFTR gene was knocked down showed higher blood pressure when responding to angiotensin II (AngII). This means CFTR is a negative regulator of high blood pressure. However, the experiment didn't try to relieve the symptoms of hypertension model mice using this theory, so CFTR was still a factor concerning blood pressure not hypertension in this research. The advantage of this research is that the team found two possible pathways between CFTR and AngII by changing the calcium concentration in the cell; then they

DOI: 10.54254/2753-8818/29/20240753

discovered that CFTR affects two pathways that need calcium. This article creatively suggested that CFTR had some relationship with blood pressure, not only a chloride channel in epithelial cells in all organs, which is the core theory of this project. The conclusion drawn from the research is inadequate as the relationship between CFTR and the compound correctors was not explored. Additionally, the lack of research on humans is a major omission. To address this, bioinformatic methods can be utilized to recover the expression of CFTR and conduct further research.

Lu et al. based on the research of Zhao, focused on another aspect of blood vessels and CFTR—CFTR's important role in vascular remodelling [8]. They detected the proliferation and migration of vascular smooth muscle cells (VSMC). They found that the migration and proliferation activity is higher in CFTR overexpressed rats and lower in CFTR silenced rats. The result shows the effect of CFTR in VSMC. And this team also found two pathways associated with this phenomenon. Theoretically, the blood vessels' remodelling can be found both in the cases of essential hypertension and secondary hypertension; but here they did not talk about hypertension, they only talked about their topic 'preventing arterial restenosis. In fact, this article can also support that CFTR may have some relationship with hypertension.

Liang et al., unlike Zhao and Lu, chose the erythrocytes as the research object [9]. They found that CF patients' (whose CFTR is abnormal) erythrocytes couldn't stimulate isolated rabbit lungs to produce endogenous NO while normal erythrocytes had the opposite result. This means CFTR silencing can lead to less endogenous NO; and as a result, the patients will suffer from pulmonary hypertension. It is widely known thar NO is a factor in vasoconstriction; and in this research, it finally led to pulmonary hypertension. The change in endogenous NO can affect blood pressure. But till now, no research has found the pathway. So, this may be a gap. In this project, pathway net analysis can possibly answer it.

2.4. Theme 3: Bioinformatics in hypertension or CFTR

Much research on hypertension or CFTR has been done in recent decades. Sanders et al. collected 2,006 CFTR variants from CFTR1, CFTR2, ClinVar, TOPmed, gnomAD and COSMIC databases; then they gave a score to the pathogenicity of each variant and finally got 13 high impact variants to CF (not hypertension) [10]. Then, they searched these gene in populations from different nations. As a result, they got the possibility of these variants in different nations and found the linkage disequilibrium among them, which means they have special features in population genetic processes. The conclusion does not correspond with the need of my project as it based on the need for diagnosis of CF. However, the steps in the experiment are quite useful as they can help me screen and classify the results of bioinformatic experiments. And the databases can provide sufficient samples.

Shi et al. did a bioinformatic analysis on CFTR variants associated with cancer[11]. Unlike Sanders' research, this research focused on the CFTR mRNA and protein expression levels, which is more quantitative. Using Oncomine they could compare the mRNA level among control individuals and all types of cancer in the databases. Using the same method, data on the mRNA level of hypertension patients and control individuals can be obtained. This method also allows the project to compare the level both in tissue level and cellular level. And finally, they use GEPIA2 and meta-analysis to explore the CFTR cancer survival landscape. The research found that CFTR expression was associated with LUAD survival. The last step is hard to learn unless there is a specific standard to score essential hypertension.

The above concerns bioinformatic analysis of CFTR; the latter two will be about the analysis of hypertension. The article written by Botzer et al. shows their steps to collect and evaluate important genes in hypertension [12]. First, collect gene participates from RAAS, endothelin, ANP/cGMP and bradykinin and the gene obtained from GWAS. Then STRING to find protein-protein interaction and PrePPI to forecast the protein-protein interaction. So, my project can find the protein-protein interaction to check if it is a hub gene for hypertension.

The research of Zheng et al. only focused on the relationship between a specific mRNA and essential hypertension [13]. They just measure the expression level of target mRNA by qRT-PCR.

DOI: 10.54254/2753-8818/29/20240753

After they found the target mRNA was differentially expressed, they performed a bioinformatic prediction of the corresponding miRNA. Zheng's method makes it easier to find the CFTR variants that the research need. But Botzer's method can help more to predict the pathway that CFTR associated with essential hypertension.

2.5. Conclusion

In the first theme, the literature about genetic factors in hypertension are reviewed. Two methods to analysis CFTR variants: one tends to population, and another tends to interaction. And the epigenetic factors can also affect the relationship between CFTR and hypertension. The second theme is the evidence that CFTR regulates blood pressure and may lead to hypertension: negative regulator for AngII; activation of the proliferation and migration of VSMC; an important factor in producing endogenous NO. And the third theme is mainly about the bioinformatic method to analyse CFTR of hypertension. The first and second methods are to evaluate the pathogenicity and predicted severity of each CFTR variant. The third and last methods are to find the target type of CFTR variant and predict its interaction with another gene.

3. Method

3.1. Introduction

The aim of this chapter is to describe and justify the overall methodology. This chapter begins with the research design, which is followed by the process of data collection and data analysis. Furthermore, the limitation of the method is stated.

3.2. Research design

This assignment is secondary research. Because there are enough easily accessible data in public databases in this field, which means they are timesaving, and the sources are adequate. In addition, most of such data are provided in 7 years, which means the risk of outdated is low. Quantitative research is conducted in this assignment. One of the reasons is that most of the original research in bioinformatics is quantitative. Another reason is that quantitative blood pressure statistics is a common method in the research of hypertension because there is no dividing line between high blood pressure and normal one [14].

3.3. Data collection

3.3.1. The collection of highly frequent CFTR variants in hypertension patients. Collect the data from the database PubMed, a common database in the field of biology. The first step is to search 'CFTR' in ClinVar of GenBank of PubMed [15]. The results show that CFTR has 588 variations till now. Only 94 of them are pathogenic, which means evidence shows these variations can cause disease. And 15 of the variations are likely-pathogenic, which means evidence can't show they are harmless. Search variants in the database IGSR (the International Genome Sample Resource) [16]. This is a database that shares the found variants around the world, and this research uses it to get the frequency of variants. Using the ID of CFTR (GRCh38.p13) to find the highly frequent variants. After excluding the intron variants and non-coding variants and limiting the global MAF to 0.05-0.5, the result is 9 variants. Check the data of the 9 variants in GWAS data of hypertension patients. Search 'essential hypertension' and 'GWAS' (genome-wide association study) in PubMed and limit the publication data to 5 years. The result is 768 articles. Two articles show the significant genes associated with hypertension.

3.3.2. The collection in the epigenetic feature. Using EWAS data hub, a database of epigenetic features [17]. Search 'CFTR' in the database. The result is 18 series of data. Click the button to show the tissue-specific hypermethylation and hypomethylation. The data for hypertension illness can also be searched by the same method, and the result is also 18 series.

DOI: 10.54254/2753-8818/29/20240753

3.3.3. The collection in blast tree. Using blast to show the evolution tree of CFTR [18]. Then, search for hypertension related genes in the result of blast.

3.4. Data analysis

3.4.1. To find the specific variants and whether the variants can lead to hypertension. Compare with the variant data from PubMed and IGSR to get the highly frequent pathogenic variant. Compare the results of normal and hypertension to find the specific variant for hypertension. Both types of result are made by the same method, so it is feasible to do so. Next step is to find whether CFTR is shown in the significant associated genes. If not, using the same formula to calculate the result.

3.4.2. To answer whether the epigenetic feature of CFTR can affect hypertension. Find the tissue-specific epigenetic feature for the tissues related to hypertension (whole blood, kidney, liver, lung) in EWAS. After that, compare the data for normal and hypertension to show whether the epigenetic feature is different. The two groups are made by the same method, so it is feasible to do so.

3.5. Limitations

The variants are too many (more than 60,000 series), so the intron variants are excluded, non-coding variants and non-pathogenic variants. They have no relationship with cause of hypertension, but they may be the special gene only in hypertension patients. The result of collecting highly frequent variants is limited to finding the variants that participate in the process of multiple genes leading to hypertension.

4. Research findings

4.1. The specific variants for hypertension

All the 9 highly frequent variants are observed in the pathogenic or likely- pathogenic variants group. This means the nine highly frequent variants are pathogenic or likely- pathogenic variants. They are exactly what the research aims to search. The GWAS data of hypertension patients are from the research of Evangelou et al. [19]. And the frequency of the 9 variants both in normal and hypertension samples are shown. The two-sample confidence interval test shows that for a 95% confidence interval, there is a difference in the data of two variants between normal and hypertension samples. Two of the nine variants are highly observed in hypertension, which means essential hypertension patients are more likely to have such variants (Table 1).

Table 1. The frequency and the confidence interval test of each variant. Adapted from the data in the research of Evangelou et al., CFTR page of GenBank (n.d.) and IGSR (n.d.) [15, 16, 19].

ID for variant	frequency(normal)	frequency(hypertension)	P
rs1042077	0.479	0.48	0.15
rs213950	0.418	0.48	0.002
rs1042180	0.187	0.186	0.24
rs1800136	0.186	0.186	0.87
rs1800130	0.068	0.068	0.76
rs17140308	0.051	0.051	0.82
rs1800501	0.027	0.027	0.84
rs10234329	0.015	0.015	0.61
rs75789129	0.011	0.019	0.0009

DOI: 10.54254/2753-8818/29/20240753

4.2. Test for the variants

This question bases on the conclusion of the first question. The two variants in the first question are the key issues. Xu et al. (2012) showed some of the key variants in essential hypertension [4]. Their research tests the correlation between the existence of a variant and the SBP (systolic blood pressure) or DBP (diastolic blood pressure). The result of CFTR combined with part of the result of Xu et al. is shown. The standard of the key variants in Xu et al.'s research is that strong correlation ($P<0.05$) should be the result in both tests. Although rs213950 has a strong correlation with SBP, the two variants are not the key variants. Therefore, these 2 variants (rs213950 and rs75789129) exist more in hypertension patients, but they are not the variants that have the potential to cause hypertension (Table 2).

Table 2. The results of the correlation tests of some key variants and two chose CFTR variants. Adapted from Xu et al.'s research [4].

SBP		
gene name	ID for variant	P
MOV10	rs2932538	1.2*10^-9
ULK4	rs3774372	9*10^-14
NPR3	rs1173766	1.9*10^-18
CSK	rs1378942	5.7*10^-23
CFTR	rs213950	0.006
CFTR	rs75789129	0.12
DBP		
gene name	ID for variant	P
MOV10	rs2932538	1.2*10^-9
ULK4	rs3774372	9*10^-14
NPR3	rs1173766	9.1*10^-12
CSK	rs1378942	4.2*10^-8
CFTR	rs213950	0.07
CFTR	rs75789129	0.23

4.3. The relationship between the epigenetic feature of CFTR and hypertension

The data in EWAS data hub shows that the lung is the hypo-tissue of cg12124767; liver is the hyper-tissue of cg09181792 and cg25509184. No epigenetic feather has the hyper-tissue of hypo-tissue as kidney or whole blood (Table 3). And CFTR has no hypo-tissue or hyper-tissue (Table 4). And the database directly shows the P number of the two-sample test (95%) in every epigenetic point and CFTR between hypertension patients and control sample (Table 5). Therefore, the epigenetic feature of CFTR has no relationship with hypertension.

Table 3. The hyper-tissue and hypo-tissue of 18 epigenetic point. Adapted from the CFTR page of EWAS Data Hub (n.d.) [17].

ID for epigenetic	hyper-tissue	hypo-tissue
cg09341015	no	placenta, colon, sperm
cg21212505	no	esophagus, prostate
cg09378456	no	sperm
cg11606570	no	pancreas, skeletal muscle
cg22533025	no	no

DOI: 10.54254/2753-8818/29/20240753

Table 3. (continued).

cg12124767	no	nasopharynx, lung, spleen
cg05917537	no	placenta
cg22467052	no	no
cg26310285	sperm	placenta
cg09181792	prostate, breast, liver	sperm
cg25509184	prostate, breast, liver	no
cg17204129	prostate, breast	no
cg26635219	breast	no
cg17616554	prostate, nasopharynx, breast	no
cg09626894	nasopharynx	no
cg00735923	no	no
cg06081199	no	no
cg21461649	no	no

Table 4. CFTR's hyper-tissue and hypo-tissue. Adapted from the CFTR page of EWAS Data Hub (n.d.) [17].

gene	hyper-tissue	hypo-tissue
CFTR	no	no

Table 5. The two-sample test between normal and hypertension's CFTR epigenetic feather. Sourced from the CFTR page of EWAS Data Hub (n.d.) [17].

ID for epigenetic	p
cg09341015	0.87
cg21212505	0.69
cg09378456	0.75
cg11606570	0.44
cg22533025	0.52
cg12124767	0.68
cg05917537	0.31
cg22467052	0.52
cg26310285	0.67
cg09181792	0.82
cg25509184	0.25
cg17204129	0.24
cg26635219	0.56
cg17616554	0.25
cg09626894	0.73
cg00735923	0.46
cg06081199	0.66
cg21461649	0.68

DOI: 10.54254/2753-8818/29/20240753

4.4. The blast tree of CFTR protein

The result of blast shows that all the related proteins are CFTR in human or other animal species and there are no hypertension related proteins. This means that the function of CFTR may has no relationship with hypertension (Figure 1).

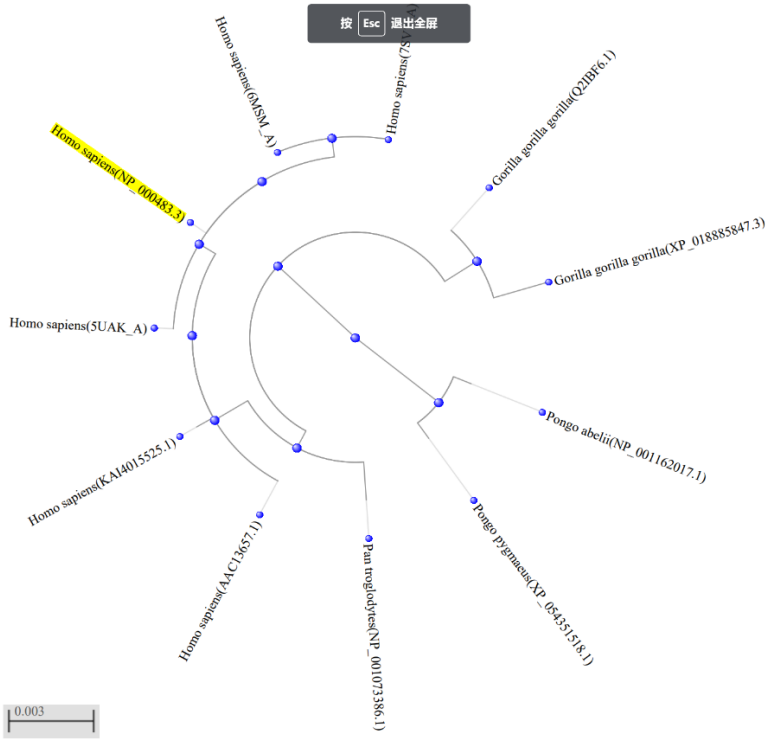

Figure 1. Blast tree of the result of CFTR blast. This tree was produced using BLAST pairwise alignments. (n.d.) [18].

5. Discussion

The process of choosing the variants gives up checking out the effect of non-pathogenic variants in fact. So, the higher existing variants may be more than two.

Although the two variants (rs213950 and rs75789129) are higher in essential hypertension, there is no evidence that they have the potential to cause hypertension. There are two possible explanations for this phenomenon. First is that logically, the two variants may be the effect of essential hypertension rather than the cause. If the cause and effect do seem like this, this phenomenon can be reasonable, but till now no evidence supports that hypertension can lead to such variants. The second is that the variants do have the potential to cause hypertension, but the effect is too slight. The SNP variants group, which contains the two variants, can only get involved in the explanation of 6% of hypertension samples [20]. If so, the test of these two variants should not test them alone. They need to be tested combined with other variants, and the interaction among the variants and the key factors of hypertension may be more complex. Going back to the process of choosing the variants, the non-pathogenic may get involved in the cause of hypertension. For example, intron can regulate the expression of exons. The sequence after the terminating signal sequence can also regulate the expression. Future researchers can check the mRNA of CFTR and the miRNA or siRNA that target these variants.

The epigenetic features of CFTR in normal samples and hypertension patients are almost the same. In fact, the silence of CFTR is associated with suppressing apoptosis and activating cancer-related

DOI: 10.54254/2753-8818/29/20240753

genes in cancer of the intestine, lungs, breast, head and neck (Shin et al., 2020). The results of the research show that the path of essential hypertension does not involve these mechanisms in cancer.

Blast is a way to predict the function of a protein. Therefore, the result of the blast is just for reference. Indeed, the function of CFTR has been already systematically analysed.

6. Conclusion

The aim of this project was to find the variants of CFTR that especially or highly exist in patients who suffer from essential hypertension and then find the relationship among the variants, the epigenetic feature and essential hypertension. The research found that there are two variants which are specifically highly frequent in the patients' DNA. However, there's no evidence that these two variants have a relationship with the cause of essential hypertension. The result of the epigenetic feature test showed that essential hypertension has no special epigenetic feature in CFTR. In addition, CFTR's epigenetic feature is stable and approximate in most types of tissue in the human body. The conclusion of the research helps to make the prediction of essential hypertension more accurate. They can predict hypertension by combining it with other prediction factors. The result of the research needs more tests to verify the mechanization. Future research in this field is recommended to focus on the non-coding variants and proteins of CFTR.

References

[1] Aderhold J, Davydov V Yu, Fedler F, Klausing H, Mistele D, Rotter T, Semchinova O, Stemmer J and Graul J 2001 *J. Cryst. Growth* 222 701
[2] Williams B et al. 2018 *European.Heart.J* 39 3021-104
[3] Liu F, Zhang Z, Csanády L, Gadsby DC and Chen J 2017 *Cell* 169 85-95
[4] Zhao L, Yuan F, Pan N, Yu Y, Yang H, Liu Y, Wang R, Zhang B and Wang G 2021 *Frontiers in Bioscience-Landmark* 26 1396-410
[5] Xu RW, Yan WL 2012 *Yi Chuan= Hereditas* 34 793-809
[6] Padmanabhan S, Aman A and Dominiczak AF 2017 *Pharmacological Research* 121 219-29
[7] Wei BL, Yin RX, Liu CX, Deng GX, Guan YZ and Zheng PF 2021 *Frontiers in Cardiovascular Medicine* 8 720884
[8] Kato N et al. 2015 *Nature genetics* 47 1282-93
[9] Lu LY et al. 2022 *J.Cardiovascular Pharmacology* 79 914
[10] Liang G, Stephenson AH, Lonigro AJ and Sprague RS 2005 *A. J. Physiology-Heart and Circulatory Physiology* 288 H1580-5
[11] Sanders M et al. 2021 *Human genetics* 140 423-39
[12] Shi X, Kou M, Dong X, Zhai J, Liu X, Lu D, Ni Z, Jiang J and Cai K 2022 *Genomics* 114 110279
[13] Botzer A, Grossman E, Moult J and Unger R 2018 *J. hypertension* 36 1094
[14] Zheng S, He X, Sun J, Li Q, Zhang T and Zhang L 2020 *J. clinical laboratory analysis* 34 e23339
[15] Padmanabhan S, Joe B 2017 *Physiological reviews* 97 1469-528
[16] GenBank (n.d.), accessed 02/10/23 from URL: ncbi.nlm.nih.gov/genbank/
[17] International Genome Sample Resource (IGSR) (n.d.), accessed 02/10/23 from URL: internationalgenome.org
[18] EWAS Data Hub (n.d.) CFTR Genes & Probe, accessed 02/10/23 from URL: ngdc.cncb.ac.cn/ewas/datahub
[19] BLAST (n.d.), accessed 02/10/24 from URL: blast.ncbi.nlm.nih.gov
[20] Evangelou E et al. 2018 *Nature genetics* 50 1412-25
[21] Huang S, Wang J, Liu N, Li P, Wu S, Qi L and Xia L 2023 *Frontiers in Genetics* 14 1114174
[22] Shin Y, Kim M, Won J, Kim J, Oh SB, Lee JH and Park K 2020 *Journal of clinical medicine* 9 734