UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias Biológicas y Ambientales

CARACTERIZACIÓN GENÉTICA DE LA FIBROSIS QUÍSTICA EN ECUADOR, IMPLICACIONES PARA EL DIAGNÓSTICO Y TRATAMIENTO

Martina Isabella Armas Samaniego

Ingeniería en Biotecnología

Trabajo de fin de carrera presentado como requisito
para la obtención del título de
Ingeniera en Biotecnología

Quito, 12 de mayo de 2025

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias Biológicas y Ambientales

HOJA DE CALIFICACIÓN
DE TRABAJO DE FIN DE CARRERA

CARACTERIZACIÓN GENÉTICA DE LA FIBROSIS QUÍSTICA EN ECUADOR, IMPLICACIONES PARA EL DIAGNÓSTICO Y TRATAMIENTO

Martina Isabella Armas Samaniego

Nombre del profesor, Título académico **Vanessa I. Romero Aguilar, MD, PhD**

Quito, 12 de mayo de 2025

© DERECHOS DE AUTOR

Nombres y apellidos: Martina Isabella Armas Samaniego

Código: 00321194

Cédula de identidad: 1726613498

Lugar y fecha: Quito, 12 de mayo de 2025

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en http://bit.ly/COPETheses.

UNPUBLISHED DOCUMENT

Note: The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on http://bit.ly/COPETheses.

RESUMEN

En Ecuador, se estima que 1 de cada 1.252 nacidos vivos presenta fibrosis quística (FQ), la cual suele diagnosticarse tardíamente debido a la falta de un sistema eficiente de detección genética. Por ello, se realizó un estudio orientado a caracterizar genéticamente la fibrosis quística en pacientes ecuatorianos, utilizando una cohorte derivada del Hospital Baca Ortiz con sospecha clínica de la enfermedad. Este constituye el primer análisis del perfil genético de 44 pacientes procedentes de 11 provincias del país mediante secuenciación del exoma completo (WES). De los 44 casos, 26 fueron confirmados con al menos una variante clínicamente relevante en el gen CFTR. La variante G85E fue la más frecuente, detectada en el 34,62% de los pacientes confirmados, seguida por F508del y H609R, ambas con una frecuencia del 30,77%. Se identificaron siete variantes clínicamente relevantes no reportadas previamente en la población ecuatoriana: S1255L, S573C, L33_Q39del, Q237P, 579+1G>T, 2916+1G>C y L452del. Además, se detectó una variante benigna poblacional no descrita anteriormente en Ecuador, con una frecuencia del 47,73% en la muestra. El 57,69% de los alelos analizados corresponden a variantes elegibles para terapia con moduladores, lo que representa a 22 de los 26 pacientes con diagnóstico confirmado de FQ. Sin embargo, el 42,31% de los alelos incluye variantes del gen CFTR no tratables con estos moduladores, entre ellas H609R, una de las más prevalentes en esta población.

Palabras clave: Fibrosis Quística (FQ), Secuenciación del Exoma Completo (WES), variantes CFTR, población ecuatoriana, moduladores de CFTR, caracterización genética.

ABSTRACT

In Ecuador, it is estimated that 1 in every 1,252 live births is affected by cystic fibrosis (CF), which is often diagnosed late due to the lack of an efficient genetic screening system. Therefore, a study was conducted to genetically characterize cystic fibrosis in Ecuadorian patients, using a cohort derived from the Baca Ortiz Hospital with clinical presumptive of the disease. This represents the first genetic profiling of 44 patients from 11 provinces of the country using Whole Exome Sequencing (WES). Of the 44 cases, 26 were confirmed to carry at least one clinically relevant variant in the CFTR gene. The G85E variant was the most frequent, detected in 34.62% of confirmed cases, followed by F508del and H609R, both with a frequency of 30.77%. Seven clinically relevant CFTR variants not previously reported in the Ecuadorian population were identified: S1255L, S573C, L33_Q39del, Q237P, 579+1G>T, 2916+1G>C, and L452del. In addition, a previously undescribed benign population variant was found in 47.73% of the sample. 58% of the analyzed alleles correspond to variants eligible for modulator therapy, representing 22 of the 26 patients with a confirmed CF diagnosis. However, 42% of alleles include CFTR variants that are ineligible for current treament with modulators, including H609R, one of the most prevalent in Ecuadorian population.

Key words: Cystic Fibrosis (CF), Whole Exome Sequencing (WES), CFTR variants, Ecuadorian population, CFTR modulators, Genetic Characterization

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

Cystic fibrosis (CF) is an autosomal recessive genetic disease that damages the lungs, digestive system, sweat glands, and reproductive system due to defective chloride and sodium transport across secretory epithelia. This dysfunction results in thick, viscous secretions in the bronchi, biliary tract, pancreas, intestines, and reproductive system [20]. Throughout their lives, patients experience recurrent respiratory infections, most commonly caused by *Pseudomonas aeruginosa* and *Staphylococcus aureus*, which progressively lead to respiratory insufficiency and eventual failure.

Pathogenic variants located in the cystic fibrosis transmembrane conductance regulator (CFTR) cause CF [20]. CFTR protein is part of the ABC (ATP-binding cassette) family, which includes transporters for nutrients, surfactants, and multidrug-resistant proteins [2]. CFTR produces a chloride channel, which also transports bicarbonate and glutathione, and regulates other chloride and sodium channels at the cell surface. The 1480-amino acid protein consists of two groups of six membrane-spanning regions, two nucleotide-binding folds, and a highly charged "R domain" with multiple phosphorylation sites. Activation of the chloride channel requires phosphorylation of the R domain by protein kinase A and continuous ATP binding to the NBFs [21]. The phenotypic expression of CF varies significantly, primarily depending on the specific pathogenic variant(s) present. The CFTR2 databases list over 2000 variants in the CFTR gene, though not all are disease-causing [13][8].

Cystic fibrosis (CF) is the most common autosomal recessive genetic disorder in Caucasian populations, with F508del being the most prevalent pathogenic variant worldwide. This variant results from the deletion of three nucleotides encoding phenylalanine (F) at position 508 of the CFTR protein. Approximately 90% of individuals with CF carry at least one F508del allele, while around 50% are homozygous for this variant [4]. Certain CFTR variants are more frequent in specific ethnic groups; for instance, in the Ashkenazi Jewish

population, F508del, G542X, W1282X, N1303K, and 3849+10C>T account for approximately 97% of CF alleles, with a carrier frequency of 1 in 29 among healthy individuals [1]. Haplotypic studies suggest that F508del and other common variants originated between 11,000 and 34,000 years ago in a population distinct from modern European groups and subsequently spread across Europe [18]. Historically, the misconception that CF was exclusive to northern Europeans led to underdiagnosis in Latin America, a genetically diverse region shaped by Amerindian ancestry and European (primarily Spanish and Portuguese), African, Italian, and German influences [18]. Today, CF incidence in Latin America ranges from 1 in 1,600 to 1 in 14,000 live births, with Ecuador reporting an incidence of 1 in 11,110, corresponding to approximately 23 affected newborns annually [24][7]. Advances in diagnosis and treatment have improved CF outcomes in the region, and ongoing research, including studies in Ecuador, continues to enhance understanding of CFTR genetic variation in local populations [18].

The first published genetic analysis of CF in Ecuador, conducted in 1999, focused on the F508del variant and reported a frequency of 25% among CF patients [17]. To date, there are seven publications on CF genetic characterization in Ecuador, the most recent from 2019, which have identified additional variants among Ecuadorians, including F508del, H609R, G85E, and L15P [18][28][15][16][23]. A limitation of these studies is that they relied on panel-based techniques, which bias results toward previously reported pathogenic variants and may overlook novel local variants.

Next-Generation Sequencing (NGS) is the most effective approach for CF genetic analysis, with CFTR panels commonly used to identify CF in patients. The American College of Medical Genetics and Genomics (ACMG) recommends a core panel of 23 disease-causing variants to identify pathogenic CFTR variants [9]. These panels enable the efficient identification of known CFTR variants in a cost-effective way, providing quick results that help determine the CFTR variant and guide treatment decisions. Whole Exome Sequencing

(WES), however, is a broader approach that can identify novel variants not previously reported, making it a better option for populations with diverse and mixed ancestry, such as Ecuador. Although WES is more expensive and requires more complex data analysis, it offers long-term benefits in determining CFTR variants. In our study, we used WES, which allowed us to identify both previously reported variants and novel CFTR-related variants not previously documented in Ecuadorian CF patients, thereby expanding Ecuador's CF genetic database.

CF variants are classified into six classes based on their impact on the CFTR protein's function. Class I variants result in a complete absence of CFTR protein due to premature stop codons, leading to no functional protein being produced [14]. Class II variants, such as ΔF508, cause misfolding of the protein, preventing it from reaching the cell surface [14]. Class III variants affect the regulation of CFTR, preventing proper opening and closing of the chloride channel [14]. Class IV variants lead to defective conductance, where the CFTR protein reaches the surface but has reduced chloride ion transport [14]. Class V variants reduce the amount of CFTR protein, often due to splicing errors, resulting in lower protein expression on the cell surface [14]. Class VI variants cause the CFTR protein to be unstable at the cell surface, leading to increased degradation of the protein [14].

The classification of CF variants is crucial for determining the type of modulator treatment a patient can receive, if eligible. Modulators are divided into potentiators, correctors, and triple therapies. Potentiators, such as Ivacaftor, enhance the opening of the CFTR chloride channel, improving ion transport in variants like G551D. Correctors, such as Lumacaftor, Tezacaftor, and Elexacaftor, help misfolded CFTR proteins, like F508del, reach the cell membrane. Triple therapy, such as Elexacaftor-Tezacaftor-Ivacaftor (ETI), combines two correctors with a potentiator, significantly improving CFTR function in a broader range of patients [29][26][22].

In Ecuador, there are significant limitations regarding the management and treatment of CF. The lack of comprehensive genetic characterization techniques has contributed to limited awareness among healthcare professionals about the true prevalence of CF in the country. As a result, many patients remain undiagnosed and without appropriate treatment. The absence of research and accessible information continues to hinder efforts to address the urgent need for broader treatment availability. Currently, only monotherapy is offered, and more effective options, such as triple therapy, are not accessible to the Ecuadorian population. There is a critical need for genetic studies that accurately reflect the frequency and spectrum of CF variants among Ecuadorians, to support improved treatment access and patient management. Therefore, considering these limitations, we performed whole-exome sequencing (WES) to analyze the genetic profiles of 44 presumptive CF patients from 11 provinces across Ecuador.

METHODS

Sample recollection

The study included 44 patients from Baca Ortiz Pediatric Hospital who were referred for genetic analysis due to a presumptive diagnosis of cystic fibrosis. Common symptoms included recurrent respiratory infections with persistent cough and thick phlegm, chronic diarrhea with steatorrhea due to malabsorption, difficulty gaining weight and growing, nutritional deficiencies, and salty-tasting sweat. Clinical tests performed included a sweat test (≥60 nmol/L) and fecal tests indicating exocrine pancreatic insufficiency (<200 µg/g). While most patients exhibited one or more of these symptoms and test abnormalities, some presented atypical or inconclusive findings.

All patients and their legal representatives were informed about the project, including the genetic analysis procedure, as well as potential risks and benefits. Informed consent was obtained from legal representatives, and assent was collected from patients based on their age, in compliance with USFQ-CEISH approval 2022-101IN.

Saliva samples were collected from the first 15 patients using the *Saliva DNA Sample Collection Kit* (ZeeSan). For the remaining patients, blood samples were collected in anticoagulant-containing tubes (purple cap) and refrigerated at -20°C. Samples from hospitalized children were collected at Baca Ortiz Hospital and delivered the same day for storage at the human genetics laboratory of USFQ. In three exceptional cases involving patients aged 1–2 years, buccal swabs were collected using the *ORACollect-Dx* kit (DNAGENOTEK) since blood extraction was deemed too invasive. In total, 44 biological samples were obtained, including blood, saliva, and buccal swabs.

DNA extraction

Genomic DNA was extracted from the samples using the *QIAamp DNA Blood Mini Kit* (ID: 51104). For each blood sample, 200 µL were used, to which 20 µL of proteinase K

was added, followed by incubation at room temperature for 5 minutes. The *QIAamp DNA Blood Mini Kit* was followed without modifications for blood samples (cat. 51104/51106). For DNA extraction from saliva and buccal swab samples, the same kit and protocol were used, with modifications in the initial phase to optimize DNA quality and extraction efficiency. In these cases, 300 µL of sample was used, with the addition of 30 µL of proteinase K, followed by incubation at 56°C for 30 minutes. From this point onward, the procedure followed the same steps as for blood samples. DNA was eluted after 5 minutes of incubation at room temperature and quantified using *NanoDrop-EPOCH*. An electrophoresis gel was run to confirm the quality of the DNA and rule out any degradation in the samples before sending them for sequencing.

Genetic Sequencing

The extracted DNA was sent for external sequencing. The company followed this sequencing workflow: After assessing DNA quality, a *gDNA* library was prepared using specialized kits with TruSeq-compatible adapters and Twist UDI index adapters (Illumina). The final purified product was quantified using TapeStation DNA screentape D1000 (Agilent). For exome capture, 1500 ng of indexed libraries per hybridization reaction were mixed with Hybridization Mix, Twist Human Core Exome probe, RefSeq probe, Blocker Solution, Universal Blocker, and Hybridization Enhancer. Sequencing was performed on the NovaSeq6000 platform (Illumina, San Diego, USA), and data were converted into FASTQ format.

Raw data processing

Upon receiving from the external company the raw sequencing data, we conducted quality control and variant analysis. Clean reads were mapped to the human reference genome (GRCh38) using the Burrows-Wheeler Aligner (BWA-MEM). Duplicate read marking, local realignment around indels, base quality score recalibration, and variant calling were performed using GATK (Sentieon) algorithms. Sequencing depth and sample coverage were calculated,

and target regions were analyzed for variations in DNA copy numbers. Expected sequencing depth was estimated using other samples from the same sequencing run as a reference. Sequence data were adjusted for guanine and cytosine variation effects. Patient samples were processed into variants using a proprietary bioinformatics pipeline.

Variant classification followed the 2015 ACMG guidelines. Likely benign and benign variants were not reported. Variants with a population frecuency >0.005 were discarded. Relevant variants were analyzed using rsID in dbSNP to assess clinical significance, gene involvement, and variant type. Protein impact was predicted using PolyPhen-2 and SIFT (D or P considered), and CADD scores >15 were prioritized. Finally, the clinical evaluation team assessed the pathogenicity of candidate variants by reviewing the patient's clinical data, relevant literature, and sequencing details to determine which variants were relevant for reporting.

Results delivery

The geneticists responsible for the patients at Baca Ortiz Pediatric Hospital delivered the reports, ensuring the results were explained and accompanied by appropriate genetic counseling.

RESULTS

Whole-exome sequencing analysis of 44 patients with presumptive cystic fibrosis of 11 provinces in Ecuador revealed that only 26 carried a pathogenic, likely pathogenic, or uncertain significance (VUS) variant in the CFTR gene (Figure 1; Figure 5). Unlike global data where F508del is the most prevalent CFTR variant, our study revealed that G85E is the most common in Ecuador. It was found in 9 patients – 34.62% of CF-positive cases –, all in a complex heterozygous state. F508del and H609R, both share the incidence of 30.77% – affecting 8 patients –, for F508del three patients were homozygous for the variant, and five patients were heterozygous, and for H609R all are found in complex heterozygosis. L15P variant was observed in 5 patients, corresponding to an incidence of 19.23%, which 2 are homozygous and 3 are complex heterozygous (Table 1; Figure 2B).

We also identified seven pathogenic and VUS variants not previously reported in the Ecuadorian population: S1255L, S573C, L33_Q39del, Q237P, 579+1G>T, 2916+1G>C, and L452del. Each of these variants had an incidence of 3.85%, with one patient affected per variant in heterozygosis, except for L452del, which was found in 15.38% of patients (present in four individuals), all in heterozygosis (Table 1; Figure 2B).

Among the variants not previously reported in the Ecuadorian population, we identified three with no associated publications: 2916+1G>C, S573C, and Q237P. The variant 2916+1G>C (rs397508456) is classified as a splice site variant located on chromosome 7 at position 117606673, involving a G>C substitution. As a splicing variant, it does not result in an amino acid change. According to ClinVar and ACMG guidelines, it is classified as pathogenic. The variant S573C (rs772223589) involves a substitution of serine for cysteine at position 573. The variant Q237P (rs1554380493) involves a substitution of glutamine for proline at position 237. Both S573C and Q237P are classified as variants of VUS

in ClinVar; however, Franklin platform, following ACMG guidelines, classified them as likely pathogenic (Table 1).

A population-level benign CFTR variant, not previously reported in Ecuador, was identified in 47.73% of the entire sample. This variant was classified as benign, with a CADD score of 0.204, a SIFT prediction of T (tolerated), and a PolyPhen-2 prediction of B (benign). It involves an asparagine-to-lysine substitution at position 417 (p.N417K). In seven patients, the N417K variant was the only CFTR variant detected, while in another 14 patients it was found alongside other pathogenic or VUS variants (Table 1; Figure 2A).

There were 18 patients that were presumptive of CF – symptoms similar to CF - , but end up negative for CFTR clinical significant variants. Of these 18 patients, 5 carried variants in genes related to Primary Ciliary Dyskinesia (PCD) [OMIM: 614679] (Figure 1). The PCD-related genes identified in the our Ecuadorian cohort included HYDIN, DNAH5, RPGR, and GAS2L2 (Table 2). While these five patients had variants only in PCD-related genes, we also found patients that carried both PCD-related genes and CFTR variants. The distribution of PCD-related variants and its frequencies among participants with and without CF is detailed in Appendix 2 and 3.

We compared the variant frequencies identified in our study with those reported in global databases such as gnomAD. In general, the frequencies observed in our cohort were higher than those reported globally. Notably, six variants identified in our study are absent from the gnomAD database, as shown in Figure 3. To facilitate comparison, frequency values were normalized by applying a base-10 logarithmic transformation using absolute values. As a result, variants with lower frequencies appear as taller bars, since their log-transformed values are of higher magnitude (Table 3; Figure 3).

Of all the variants identified in our study, only three—F508del, G85E, and L15P—are eligible for modulator therapy. A total of 22 patients carried at least one of these treatable

variants, representing 26 alleles out of 52 CF-positive alleles in the cohort. This indicates that 58% of the CFTR-affected alleles identified in our study are potentially responsive to modulator treatment. However, 21 patients carried at least one CFTR variant ineligible for modulator treatment, representing 42% of the CFTR-affected alleles identified in our study (Figure 4).

DISCUSSION

CF diagnosis and research challenges in Ecuador

In Ecuador, CF affects an estimated 1 in 1,252 newborns, yet the healthcare system lacks efficient tools for early diagnosis and treatment. Our study supports the genetic characterization of patients with suspected CF, aiding accurate diagnosis and access to modulator therapy when eligible. Given a 33% rate of CFTR-negative cases, genetic testing is essential to confirm clinical presumption.

Comparison with previously reported CFTR variants in Ecuador

The most recent published genetic study on CF in Ecuador, conducted by Ruiz-Cabezas et al. (2019), analyzed a sample of 141 participants, 85 of whom were confirmed with CF. Of these, 38 were from the city of Quito [23]. In our study, we analyzed 44 samples, 21 of which were confirmed as CF. To minimize sample overlap, we prioritized patients without prior genetic analysis results or those with unreliable or inconclusive results (Appendix 4).

We found seven clinical significant variants in our study that were previously reported in Ecuadorian population, which are: F508del, G85E, L15P, H609R, N1303K, G542*, and W1098*. The identification of the other variants previously reported in Ecuadorian population gives a confirmation of the presence of these variants in the region.

The most recent publication on CF genetic characterization in Ecuador, by Ruiz-Cabezas et al. (2019), reported F508del and H609R each with a frequency of 24.7%. In our study, both variants were also identified, but with a higher frequency of 30.8%. This difference may be due to our smaller sample size (44 vs. 141 patients). Notably, G85E was the most frequent variant in our study at 34.6%, compared to 11.1% previously reported. Another variant showing a marked difference was L15P, reported at 9.4% in the previous study and at 19.2% in ours, making it the fourth most frequent variant identified. Ruiz-Cabezas et al. reported

N1303K (4.1%), W1098* (1.17%), and G542* (2.3%), while our study found slightly higher frequencies: 3.85% for each (Table 4) [23].

A key finding in our study was the identification of seven clinically relevant CFTR variants not previously reported in the Ecuadorian population: S1255L, S573C, L33_Q39del, Q237P, 579+1G>T, 2916+1G>C, and L452del. All appeared in one patient each, mostly as complex heterozygous variants (3.85% frequency), except S1255L, found in one patient in homozygosity (15.4%). Among these, four variants have been previously reported in international studies. S1255L was described by Cartault et al. (1998) in Reunion Island – France – with a 0.7% frequency in 65 patients, similar to our allele frequency of 2.3% [5]. L33_Q39del was reported by Faucz et al. (2007) in Brazil at 0.89% (1/56 patients), comparable to our 1.1% [10]. 579+1G>T has appeared in multiple studies, most recently by Banjar et al. (2022) in Saudi Arabia at 12.5% (5/40 patients), while we observed it at 1.1% [3]. L452del was reported by Stuhrmann et al. (1997) in Tyrol – Austria – with a 0.8% frequency (1/63 patients), whereas our study showed a higher allele frequency of 4.5% [25].

Three of the variants identified in our study—2916+1G>C, S573C, and Q237P—have not been previously reported in the literature. Their detection may be attributed to two main factors: first, the high genetic diversity and admixture of the Ecuadorian population, which may result in unique variant profiles; and second, the use of WES, a broad-spectrum technique capable of identifying local variants missing by previously used panel approaches [27].

WES as a superior approach for local variant detection

Our data remains valuable due to differences in the sequencing techniques used. Ruiz-Cabezas et al. employed the Ion AmpliSeq™ CFTR Panel, which targets 102 amplicons covering known CFTR variants [27]. In contrast, our study utilized WES for all participants. This distinction is significant, as WES enables a comprehensive diagnosis, including for those who were not confirmed with CF, due to its broad exon coverage. Unlike targeted panels, WES

is not limited to known CFTR variants, allowing us to identify novel variants not previously reported in Ecuadorian genetic studies of CF. The use of CFTR Panel facilitated a larger sample by the cost-effectiveness, but it introduced bias [23].

In comparison to the study by Ortiz et al. (2017), which used capillary electrophoresis for CFTR variant detection, our findings show similar incidences for F508del and H609R but higher incidences for other variants [16] (Table 4). The genetic characterization technique used by Ortiz et al. had significant limitations compared to the broader scope of WES, which enables the identification of a broader range of variants, enhancing the depth and value of genetic data.

One major advantage of WES is its ability to identify variants in genes that do not involve CFTR variants but are associated with other diseases. The diseases are crucial to consider when evaluating clinical symptoms, as these variants may cause symptoms that are often misdiagnosed as CF. In our study, nearly half of the participants initially misdiagnosed with CF had variants in PCD-related genes related or immune deficiencies (Table 2; Appendix 2; Appendix 3). Further investigation into the relationship between clinical symptoms and the genetic characterization of these alternative diseases is necessary.

Variant-based diagnosis and treatment options

Genetic analysis ensures appropriate treatment. In our study, 18 patients were misdiagnosed due to phenotypic similarities with other respiratory diseases. Incorporating genetic testing alongside clinical evaluation can prevent misdiagnoses and reduce unnecessary healthcare costs—especially in genetically diverse populations with limited access to specialized testing and healthcare.

The clinical implications of CF genetic characterization extend beyond population-level variant frequencies, as they directly influence treatment options. As of 2023, CF patients receiving modulator therapy have shown a significant increase in life expectancy, reaching up to 68 years [6]. Eligibility for these treatments depends on the specific CFTR variant present.

In our study, 58% of the total allele count of CF-positive patients – 22 out of 26 patients carried at least one of the variants F508del, G85E, or L15P – making them eligible (Figure 4).

Treatment eligibility for F508del variants depends on the zygosity and age of the patient. Heterozygotes with the F508del variant may receive a triple combination therapy (ETI or VTD) if older than 2 years. For homozygotes, both ETI and VTD are available for patients older than 2 and 6 years, respectively [12][11]. For patients younger than 2 years, dual therapy (Tez-Iva or Lum-Iva) is available [22]. The G85E and L15P variants are eligible for triple therapy, responding to class II variant that requires a corrector [19].

42% of the allele count in the CF-positive patients cohort in our study – 21 of 26 patients carried at least one non compatible modulator –have variants for which no modulators are currently available, including S1255L, H609R, Q237P, S573C, N1303K, L33_Q39del, L452del, G542*, W1098*, 579+1G>T, and 2916+1G>C. Among these, H609R is one of the most common variants in Ecuador – 30.77% –, yet no modulator is available for it (Figure 4).

Limitations and considerations

Discrepancies in variant frequencies may result from the difference in sample size; Ruiz-Cabezas et al. included 141 participants, while our study analyzed 44. The exclusive inclusion of patients with clinical suspicion of CF may have led to an overrepresentation of certain variants. The benign population variant found in 47.73% of our sample should be investigated in a broader, non-CF-biased population. Although our cohort included individuals from 11 provinces, it does not represent the entire country, as 13 provinces were not included.

Future Directions

This project remains ongoing. We are continuing to collect samples to expand the cohort and improve its representativeness in terms of both size and geographic distribution across the country. Additional studies should also explore alternative treatment options for variants that are not responsive to currently available modulators.

CONCLUSIONS

Whole Exome Sequencing (WES) has proven to be a highly effective tool for the genetic characterization of CF in Ecuador, enabling the identification of both population-specific and novel CFTR variants. In our cohort, the most common variant was G85E, present in 34.62% of CF-positive patients—surpassing F508del, which is globally the most frequent but accounted for only 30.77% in our study, tying with H609R as the second and third most common variants. We identified seven clinically relevant CFTR variants not previously reported in the Ecuadorian population: S1255L, S573C, L33_Q39del, Q237P, 579+1G>T, 2916+1G>C, and L452del. Notably, 42% of the CF-causing alleles found are not eligible for current modulator therapies, including H609R, highlighting the urgent need for local research and the development of suitable treatments. This study is the first large-scale CF genetic analysis in Ecuador using WES, emphasizing its value in uncovering novel variants and informing more precise, population-specific therapeutic strategies.

TABLES

Table 1. Cystic Fibrosis variants details and statistics

CF variants statistics							
Rs	**CFTR Variants (aa)**	**CFTR Variants (aa) (simplified nomenclature)**	**Exonic function**	**Clinsig**	**#affected patients**	**Incidence PwCF** (%)**	**Incidence sample (%)**
rs4727853	NM_000492:exon10:c.C1251A:p.N417K	p.N417K	nonsynonymous SNV	Benign/Likely_benign	21	**80,77**	47,73
rs75961395	NM_000492:exon3:c.G254A:p.G85E	p.G85E	nonsynonymous SNV	Pathogenic	9	**34,62**	20,45
rs113993960	NM_000492:exon11:c.1520_1522del:p.507_508del	p.F508del	nonframeshift deletion	Likely_pathogenic	8*	**30,77**	18,18
rs397508310	NM_000492:exon14:c.A1826G:p.H609R	p.H609R	nonsynonymous SNV	Pathogenic	8	**30,77**	18,18
rs1562876459	NM_000492:exon1:c.T44C:p.L15P	p.L15P	nonsynonymous SNV	Pathogenic	5*	**19,23**	11,36
rs397508194	CFTR:NM_000492:exon10:c.1356_1358del:p.452_453del	p.L452del	nonframeshift	VUS	4	**15,38**	9,09
rs79074685	CFTR:NM_000492:exon10:c.G1365T:p.A455A	p.A455A	synonymous SNV	Benign/Likely_benign	4	**15,38**	9,09
rs1554380493	NM_000492:exon6:c.A710C:p.Q237P	p.Q237P	nonsynonymous SNV	Pathogenic/VUS	1	**3,85**	2,27
rs397508141	NM_000492:exon2:c.98_115del:p.33_39del	p.L33_Q39del	nonframeshift deletion	Likely_pathogenic	1	**3,85**	2,27
rs397508456	NM_000492.4:c.2909-1G>C	N/A (Splice Acceptor Variant)	nonsynonymous SNV	Pathogenic	1	**3,85**	2,27
rs76649725	NM_000492:exon23:c.C3764T:p.S1255L	p.S1255L	nonsynonymous SNV	Pathogenic	1*	**3,85**	2,27

rs772223589	NM_000492:exon13:c.C1718G:p.S573C	p.S573C	nonsynonymous SNV	Pathogenic/VUS	1	**3,85**	2,27
rs80034486	NM_000492:exon24:c.C3909G:p.N1303K	p.N1303K	nonsynonymous SNV	Pathogenic	1	**3,85**	2,27
rs397508532	NM_000492:exon20:c.G3293A:p.W1098X	p.W1098*	stopgain	Pathogenic	1	**3,85**	2,27
rs77188391	NM_000492:exon5:c.579+1G>T	splicing	splicing	Pathogenic	1	**3,85**	2,27
rs213950	CFTR:NM_000492:exon12:c.G1624T:p.G542*	p.G542*	stopgain	Pathogenic	1	**3,85**	2,27
rs1042077	CFTR:NM_000492:exon15:c.T2562G:p.T854T	p.T854T	synonymous SNV	Benign/Likely_benign	1	**3,85**	2,27
rs213950	CFTR:NM_000492:exon11:c.G1408A:p.V470M	p.V470M	nonsynonymous SNV	Benign/Likely_benign	1	**3,85**	2,27

****PwCF= patients with cystic fibrosis**

***Of the 8 patient with F508del, 5 are heterozygous and 3 are homozygous**

***Of the 5 patient with L15P, 3 are heterozygous and 2 are homozygous**

***The 1 patient with S1255L is homozygous**

Table 2. Patients without Cystic Fibrosis statistics

No CF patients statistics		
Gene affected	**#patients with variant with clinical significance**	**Incidence sample among PwNoCF(%)**
RPGR	2	18,18
HYDIN	1	9,09
DNAH5	1	9,09
GAS2L2	1	9,09
PCD-related genes	**5**	**45,45**
POLA1	1	9,09

CARD11	1	9,09
FAS*	1	9,09
TCF3*	2	18,18
PRKDC	1	9,09
OFD1	1	9,09
Otros	6	54,55
No variants	1	9,09
Total	**18**	**163,64**

***FAS and TCF3 represent a complex heterozygous variant in a single participant.**

Table 3. CFTR variant allele frequencies: global vs. this study

CFTR variants (aa)	Allele count gnomAD	Alele number gnomAD	Global Frecuency (gnomAD)	AbsLog* Global Frecuency	Allele count our study	Allele number our study	Frecuency of our study	AbsLog* Frequency of our study
F508del	19237	1612320	0,011931254	1,9233139	14	88	0,15909091	0,79835464
G85E	101	1593676	6,337549E-05	4,19807866	9	88	0,10227273	0,99024016
H606R	2	1595240	1,25373E-06	5,90179604	8	88	0,09090909	1,04139269
N1303K	252	1598060	0,000157691	3,80219254	1	88	0,01136364	1,94448267
G542*	585	1612120	0,000362876	3,4402415	1	88	0,01136364	1,94448267
S1255L	3	1613638	1,85915E-06	2	2	88	0,02272727	1,64345268
S573C	1	1603430	6,23663E-07	6,20505	1	88	0,01136364	1,94448267
579+1G>T	56	1482946	3,77627E-05	3	1	88	0,01136364	1,94448267
L33_Q39del	0	1	0	0	1	88	0,01136364	1,94448267
2916+1G>C	0	1	0	0	1	88	0,01136364	1,94448267
L452del	0	1	0	0	4	88	0,04545455	1,34242268
W1098*	0	1	0	0	1	88	0,01136364	1,94448267
Q237P	0	1	0	0	1	88	0,01136364	1,94448267
L15P	0	1	0	0	7	88	0,07954545	1,09938463

***AbsLog: represents the absolute value of the logarithm in base 10 applied to the frequencies.**

Table 4. CFTR Variants incidences compared with previous studies

Variants incidences compared with previous studies			
p.Phe508del			
Fuente	**#affected patients**	**#sample**	**Incidence (%)**
WES (our study)	8	26	30,77
Ruiz-Cabezas 2019	42	170	24,7
Ortiz 2017	13	48	20,27
p.His609Arg			
Fuente	**#affected patients**		**Incidence (%)**
WES (our study)	8	26	30,77
Ruiz-Cabezas 2019	42	170	24,7
Ortiz 2017	10	48	18,92
p.Gly85Glu			
Fuente	**#affected patients**		**Incidence (%)**
WES (our study)	9	26	34,62
Ruiz-Cabezas 2019	19	170	11,1
Ortiz 2017	5	48	8,11
p.Leu15Pro			
Fuente	**#affected patients**		**Incidence (%)**
WES (our study)	5	26	19,23
Ruiz-Cabezas 2019	16	170	9,4
Ortiz 2017	NA	NA	NA
p.Asn1303Lys			
Fuente	**#affected patients**		**Incidence (%)**
WES (our study)	1	26	3,85
Ruiz-Cabezas 2019	7	170	4,1
Ortiz 2017	1	48	1,35
p.Trp1098Ter (W1098*)			
Fuente	**#affected patients**		**Incidence (%)**
WES (our study)	1	26	3,85
Ruiz-Cabezas 2019	2	170	1,17
Ortiz 2017	2	48	2,7
p.Gly542* (G542*)			
Fuente	**#affected patients**		**Incidence (%)**
WES (our study)	1	26	58,33
Ruiz-Cabezas 2019	4	170	2,3
Ortiz 2017	NA	NA	NA

FIGURES

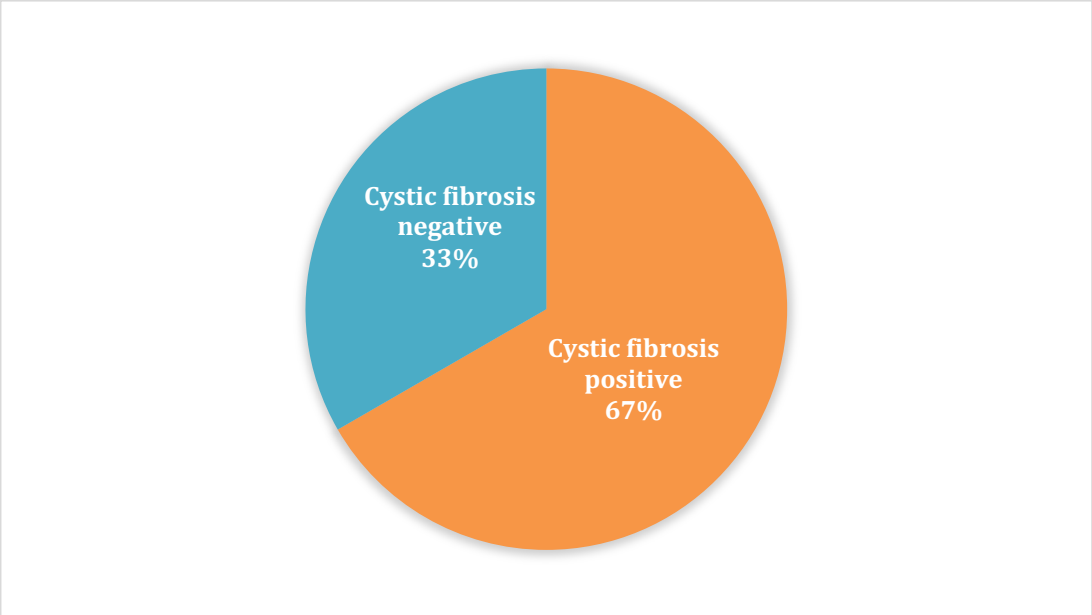
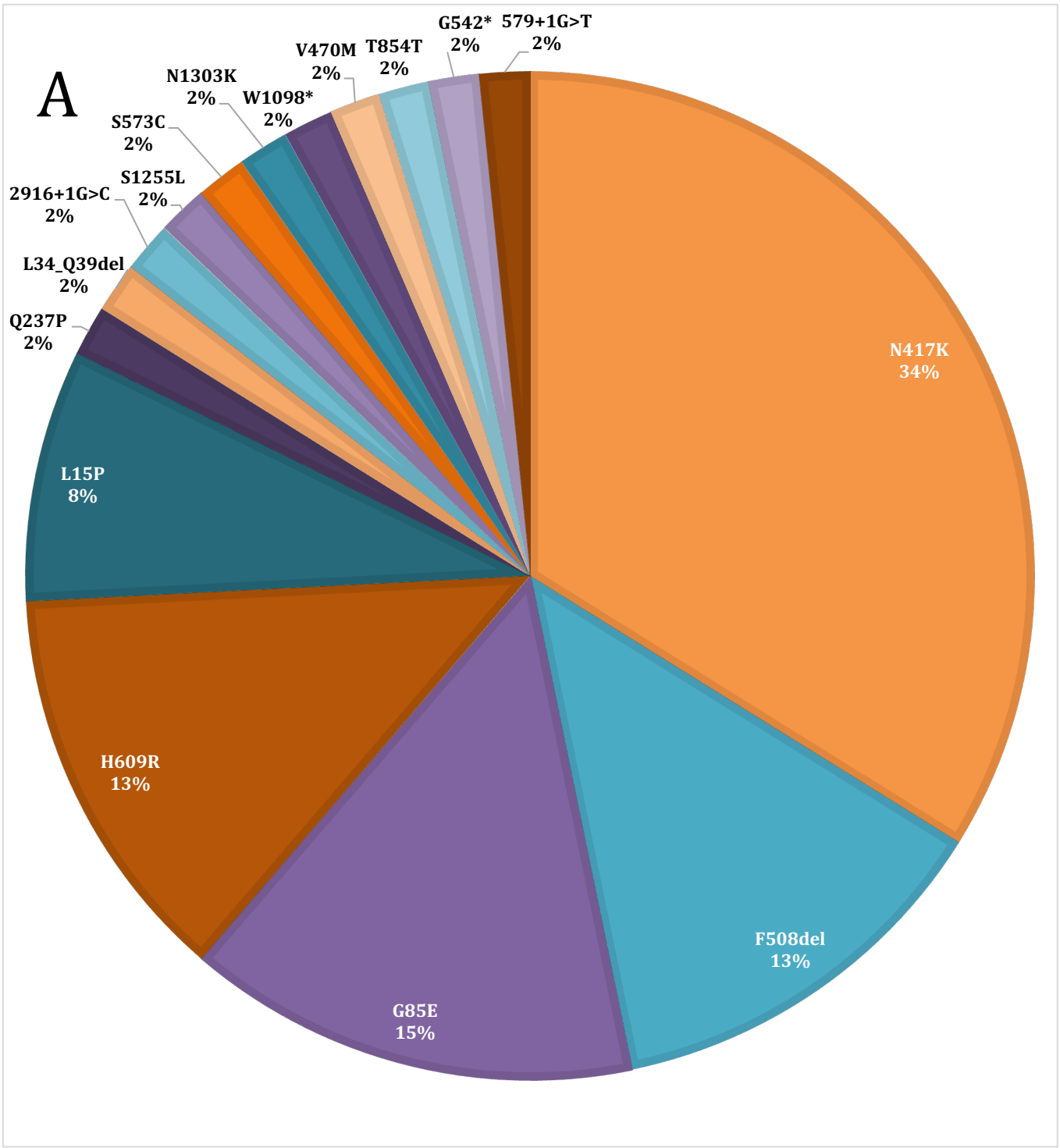


Figure 1. Distribution of CF-confirmed patients

Distribution of the 44 patients who tested positive or negative for clinically relevant variants in the CFTR gene


A
N417K
34%
F508del
13%
G85E
15%
H609R
13%
L15P
8%
Q237P
2%
L34_Q39del
2%
2916+1G>C
2%
S1255L
2%
S573C
2%
N1303K
2%
W1098*
2%
V470M
2%
T854T
2%
G542*
2%
579+1G>T
2%

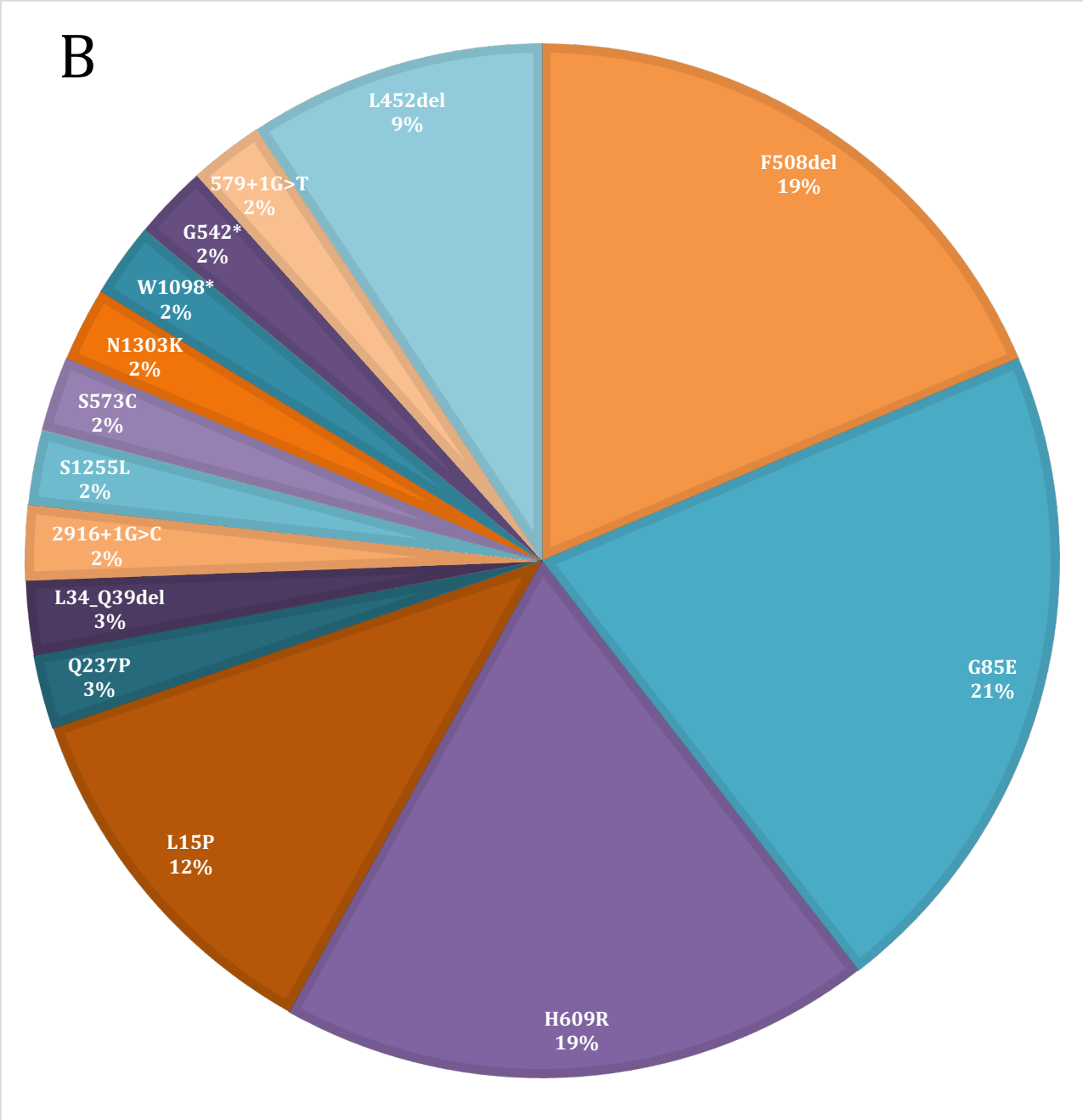


Figure 2. CFTR variants incidence

Figure 2.A Illustrates the distribution of all CFTR variants identified in the 44 participants (wholes sample). Figure 2.B presents the distribution of clinically significant CFTR variants, excluding p.N417K and other benign variants.

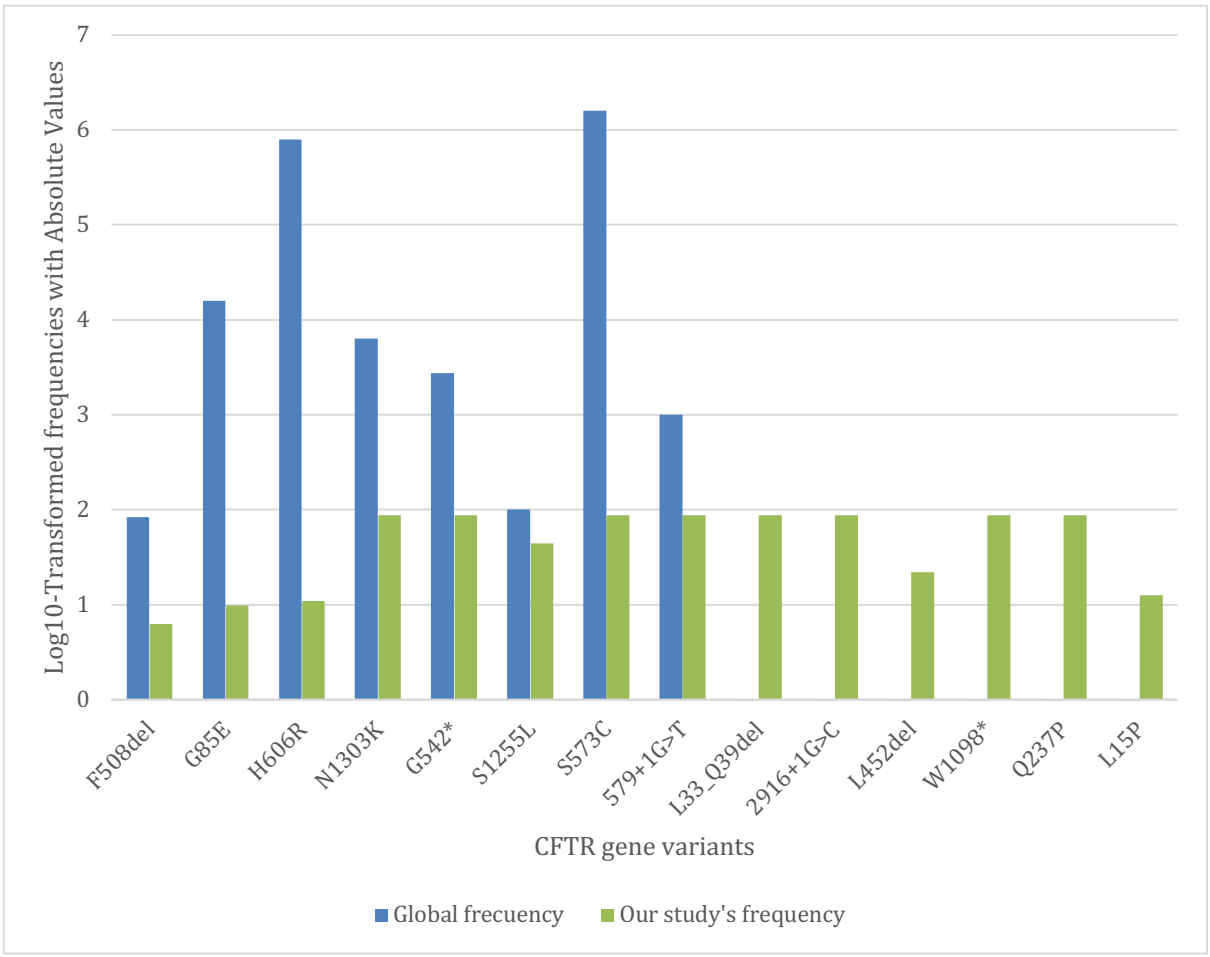


Figure 3. Log₁₀-Transformed and Absolute Allele Frequencies: Global vs. Our Study

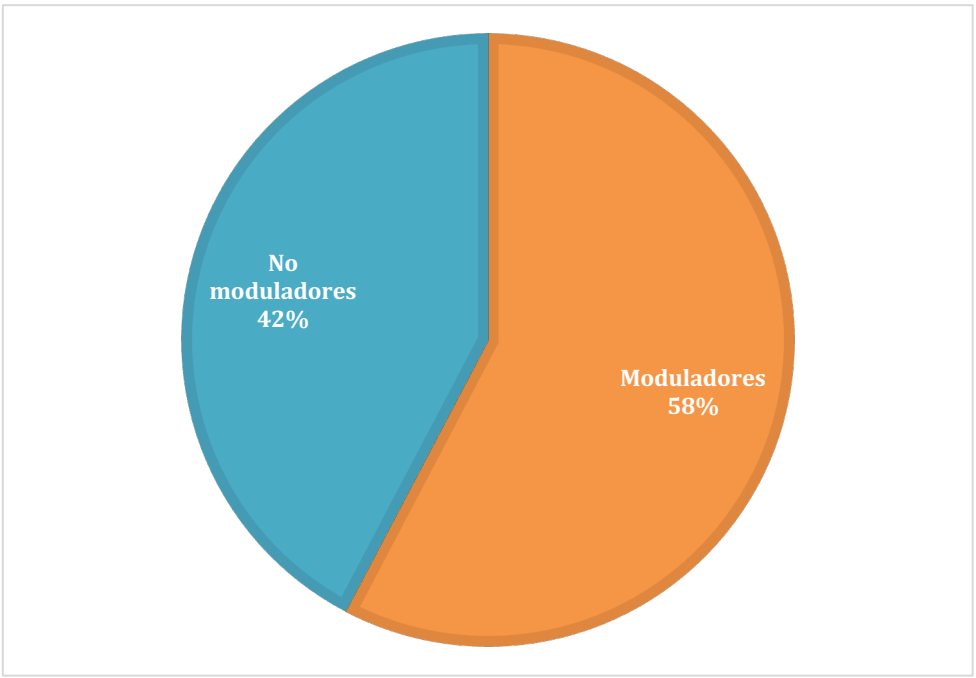


Figure 4. Allele distribution by modulator treatment eligibility

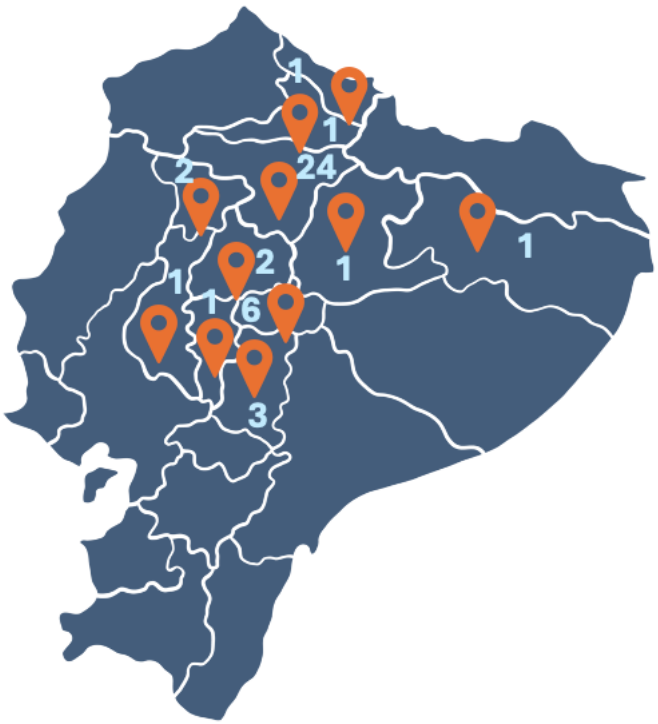


Figure 5. Geographic distribution of patients across the provinces of Ecuador

Distribution of 44 samples from patients with presumptive cystic fibrosis across 11 of Ecuador's 24 provinces

REFERENCES

[1] Abeliovich, D., Lavon, I. P., Lerer, I., Cohen, T., Springer, C., Avital, A., & Cutting, G. R. (1992). Screening for five mutations detects 97% of cystic fibrosis (CF) chromosomes and predicts a carrier frequency of 1:29 in the Jewish Ashkenazi population. *American journal of human genetics, 51*(5), 951–956.

[2] Anderson, M.P., Berger, H., Rich, D., Richard, G., Smith, A., Welsh, M. (1991). Nucleoside triphosphates are required to open the CFTR chloride channel. *Cell, 67*(4), 775-784. https://doi.org/10.1016/0092-8674(91)90072-7

[3] Banjar, H., Qeretli, R., Ramadan, A., Al-Ibraheem, A., & Bnatig, F. (2022). The first report on CFTR mutations of meconium ileus in cystic fibrosis population in Saudi Arabia: A single center review. *International journal of pediatrics & adolescent medicine*, *9*(1), 32–35. https://doi.org/10.1016/j.ijpam.2021.03.008

[4] Boyle, M. P., & De Boeck, K. (2013). A new era in the treatment of cystic fibrosis: correction of the underlying CFTR defect. The Lancet. *Respiratory medicine, 1*(2), 158–163. https://doi.org/10.1016/S2213-2600(12)70057-7

[5] Cartault, F., Steffann, J., Vidaud, D., Bousquet, S., Lesure, F., Renouil, M., McDonell, N., Feingold, J., Beldjord, C., & Bienvenu, T. (1998). Detection of more than 91% cystic fibrosis mutations in a sample of the population from Reunion Island and identification of two novel mutations (A309G, S1255L) and one novel polymorphism (L49L). *Clinical genetics*, *54*(5), 437–439.

[6] CFF. (2023). *Patient Registry.* Cystic Fibrosis Foundation. https://www.cff.org/medical-professionals/patient-registry

[7] Cuji-Gutiérrez, Genesis Monserrath, & Philco-Toaza, Priscila Elizabeth. (2023). La fibrosis quística en niños en el Ecuador. *Revista Arbitrada Interdisciplinaria de Ciencias de la Salud. Salud y Vida, 7*(14), 4-16. Epub 30 de agosto de 2023.https://doi.org/10.35381/s.v.v7i14.2510

[8] Cystic Fibrosis Genotype-Phenotype Consortium (1993). Correlation between genotype and phenotype in patients with cystic fibrosis. *The New England journal of medicine, 329*(18), 1308–1313. https://doi.org/10.1056/NEJM199310283291804

[9] Deignan, J. L., Astbury, C., Cutting, G. R., Del Gaudio, D., Gregg, A. R., Grody, W. W., Monaghan, K. G., Richards, S., & ACMG Laboratory Quality Assurance Committee (2020). CFTR variant testing: a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genetics in medicine : official journal of the American College of Medical Genetics, 22*(8), 1288–1295. https://doi.org/10.1038/s41436-020-0822-5

[10] Faucz, F. R., Gimenez, J., Ramos, M. D., Pereira-Ferrari, L., Estivill, X., Raskin, S., Casals, T., & Culpi, L. (2007). Cystic fibrosis in a southern Brazilian population: characteristics of 90% of the alleles. *Clinical genetics*, *72*(3), 218–223. https://doi.org/10.1111/j.1399-0004.2007.00854.x

[11] Manufacturer's prescribing information for ALYFTREK (vanzacaftor- tezacaftor-Deutivacaftor), 12/2024. Available at: https://pi.vrtx.com/files/uspi_vanzacaftor_tezacaftor_deutivacaftor.pdf (Accessed on January 01, 2025).

[12] Manufacturer's prescribing information for TRIKAFTA (elexacaftor-tezacaftor-ivacaftor), 12/20/2024. Available at: https://pi.vrtx.com/files/uspi_elexacaftor_tezacaftor_ivacaftor.pdf (Accessed on December 31, 2024).

[13] Mickle, J. E., & Cutting, G. R. (2000). Genotype-phenotype relationships in cystic fibrosis. *The Medical clinics of North America, 84*(3), 597–607. https://doi.org/10.1016/s0025-7125(05)70243-1

[14] Moskowitz, S. M., Chmiel, J. F., Sternen, D. L., Cheng, E., Gibson, R. L., Marshall, S. G., & Cutting, G. R. (2008). Clinical practice and genetic counseling for cystic fibrosis and CFTR-related disorders. *Genetics in medicine : official journal of the American College of Medical Genetics, 10(*12), 851–868. https://doi.org/10.1097/GIM.0b013e31818e55a2

[15] Moya-Quiles, M. R., Glover, G., Mondéjar-López, P., Pastor-Vivero, M. D., Fernández-Sánchez, A., & Sánchez-Solís, M. (2009). CFTR H609R mutation in Ecuadorian patients with cystic fibrosis. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society, 8*(4), 280–281. https://doi.org/10.1016/j.jcf.2009.05.001

[16] Ortiz, S. C., Aguirre, S. J., Flores, S., Maldonado, C., Mejía, J., & Salinas, L. (2017). Spectrum of CFTR gene mutations in Ecuadorian cystic fibrosis patients: the second report of the p.H609R mutation. *Molecular genetics & genomic medicine, 5*(6), 751–757. https://doi.org/10.1002/mgg3.337

[17] Paz-y-Miño, C., Pérez, J. C., Burgos, R., Dávalos, M.V, & Leone, P. E. (1999). The ΔF508 mutation in Ecuador, South America. *Human Mutation, 14*(4), 348–350. https://doi.org/10.1002/(SICI)1098-1004(199910)14:4<348::AID-HUMU11>3.0.CO;2-8

[18] Pérez, M. M., Luna, M. C., Pivetta, O. H., & Keyeux, G. (2007). CFTR gene analysis in Latin American CF patients: heterogeneous origin and distribution of mutations across the continent. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society, 6(*3), 194–208. https://doi.org/10.1016/j.jcf.2006.07.004

[19] Playford, M. P., Nurminen, E., Pentikäinen, O. T., Milgram, S. L., Hartwig, J. H., Stossel, T. P., & Nakamura, F. (2010). Cystic fibrosis transmembrane conductance regulator interacts with multiple immunoglobulin domains of filamin A. *The Journal of biological chemistry, 285*(22), 17156–17165. https://doi.org/10.1074/jbc.M109.080523

[20] Ratjen, F., & Döring, G. (2003). Cystic fibrosis. *Lancet (London, England), 361*(9358), 681–689. https://doi.org/10.1016/S0140-6736(03)12567-6

[21] Rich, D. P., Gregory, R. J., Anderson, M. P., Manavalan, P., Smith, A. E., & Welsh, M. J. (1991). Effect of deleting the R domain on CFTR-generated chloride channels. *Science (New York, N.Y.), 253(*5016), 205–207. https://doi.org/10.1126/science.1712985

[22] Rose, N. R., Chalamalla, A. R., Garcia, B. A., Krick, S., Bergeron, J., Sadeghi, H., Schellhase, D. E., Ryan, K. J., Dowell, A. E., Acosta, E. P., & Guimbellot, J. S. (2024). Pharmacokinetic variability of CFTR modulators from standard and alternative regimens. *Pulmonary pharmacology & therapeutics,* 86, 102301. https://doi.org/10.1016/j.pupt.2024.102301

[23] Ruiz-Cabezas, J. C., Barros, F., Sobrino, B., García, G., Burgos, R., Farhat, C., Castro, A., Muñoz, L., Zambrano, A. K., Martínez, M., Montalván, M., & Paz-Y-Miño, C. (2019). Mutational analysis of CFTR in the Ecuadorian population using next-generation sequencing. *Gene, 696*, 28–32. https://doi.org/10.1016/j.gene.2019.02.015

[24] Silva Filho, L. V., Castaños, C., & Ruíz, H. H. (2016). Cystic fibrosis in Latin America-Improving the awareness. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society, 15(*6), 791–793. https://doi.org/10.1016/j.jcf.2016.05.007

[25] Stuhrmann, M., Dörk, T., Frühwirth, M., Golla, A., Skawran, B., Antonin, W., Ebhardt, M., Loos, A., Ellemunter, H., & Schmidtke, J. (1997). Detection of 100% of the CFTR mutations in 63 CF families from Tyrol. *Clinical genetics*, *52*(4), 240–246. https://doi.org/10.1111/j.1399-0004.1997.tb02555.x

[26] Taylor-Cousar, J. L., Munck, A., McKone, E. F., van der Ent, C. K., Moeller, A., Simard, C., Wang, L. T., Ingenito, E. P., McKee, C., Lu, Y., Lekstrom-Himes, J., & Elborn, J. S. (2017). Tezacaftor-Ivacaftor in Patients with Cystic Fibrosis Homozygous for Phe508del. *The New England journal of medicine, 377*(21), 2013–2023. https://doi.org/10.1056/NEJMoa1709846

[27] Trujillano, D., Weiss, M. E., Köster, J., Papachristos, E. B., Werber, M., Kandaswamy, K. K., Marais, A., Eichler, S., Creed, J., Baysal, E., Jaber, I. Y., Mehaney, D. A., Farra, C., & Rolfs, A. (2015). Validation of a semiconductor next-generation sequencing assay for the clinical genetic screening of CFTR. *Molecular genetics & genomic medicine, 3*(5), 396–403. https://doi.org/10.1002/mgg3.149

[28] Valle, E. P., Burgos, R. I., Valle, J. R., Béjar, D. E., & Ruiz-Cabezas, J.C. (2007). Analysis of CFTR gene mutations and Cystic Fibrosis incidence in the Ecuatorian population. *Investigación Clínica, 48*(1), 91-98. http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0535-51332007000100009&lng=es&tlng=en.

[29] Van Goor, F., Hadida, S., Grootenhuis, P. D., Burton, B., Cao, D., Neuberger, T., Turnbull, A., Singh, A., Joubran, J., Hazlewood, A., Zhou, J., McCartney, J., Arumugam, V., Decker, C., Yang, J., Young, C., Olson, E. R., Wine, J. J., Frizzell, R. A., Ashlock, M., … Negulescu, P. (2009). Rescue of CF airway epithelial cell function in vitro by a CFTR potentiator, VX-770. *Proceedings of the National*

Academy of Sciences of the United States of America, 106(44), 18825–18830. https://doi.org/10.1073/pnas.0904709106

APPENDICES

APPENDIX 1. Variant details of participants without cystic fibrosis

No CF participants's variants details

Gene	Rs	Variant	Exonic function	Clinsig	Zygosity
DNAH5	rs1258331272	NM_001369:exon63:c.A10766C:p.Q3589P	nonsynonymous SNV	pathogenic	heterozygosity
RPGR	rs756933299	NM_001034853:exon15:c.G2557A:p.E853K	nonsynonymous SNV	Uncertain_significance	homozygosity
RPGR*	rs187844918	NM_001034853:exon15:c.A2808T:p.E936D	nonsynonymous SNV	benign	heterozygosity
RPGR*	rs1399249421	NM_001034853:exon15:c.2799_2804del:p.933_935del	nonframeshift deletion	not reported	heterozygosity
POLA1	rs141903278	NM_001330360:exon5:c.G361A:p.A121T,NM_016937:exon5:c.G343A:p.A115T	nonsynonymous SNV	VUS	homozygosity
CARD11	rs1780131821	NM_032415:exon8:c.A1084G:p.K362E,NM_001324281:exon9:c.A1084G:p.K362E	nonsynonymous SNV	VUS	heterozygosity
FAS	rs563551720	NM_000043:exon5:c.C488T:p.T163I,NM_001320619:exon5:c.C488T:p.T163I,NM_152871:exon5:c.C488T:p.T163I,NM_152872:exon5:c.C488T:p.T163I	nonsynonymous SNV	VUS	heterozygosity
TCF3	rs535113552	NM_001136139:exon14:c.G1147A:p.D383N,NM_001351778:exon14:c.G1147A:p.D383N,NM_001351779:exon14:c.G1147A:p.D383N,NM_003200:exon14:c.G1147A:p.D383N	nonsynonymous SNV	not reported	heterozygosity
GAS2L2	rs587633197	NM_139285:exon5:c.887_890del:p.V296fs	frameshift deletion	Pathogenic	homozygosity
PRKDC	rs587777687	NM_006904.7:c.1777-710_1777-711ins	frameshift insertion	Pathogenic	homozygosity
OFD1	rs1473408729	NM_001330209:exon2:c.61delA:p.K21fs,NM_003611:exon2:c.61delA:p.K21fs	frameshift deletion	VUS	heterozygosity
HYDIN	rs783893	NM_001270974:exon81:c.C13913T:p.T4638M	nonsynonymous SNV	not reported	heterozygosity

***both RPGR variants represent a complex heterozygous variant in a single participant.**

APPENDIX 2. PCD-related gene variants and their incidence among participants with and without cystic fibrosis

Primary Ciliary Dyskinesia (PCD)									
Gene	**Rs**	**Variant**	**Exonic function**	**Zygosity**	**Clinsig**	**#affected patients**	**#affected PwCF***	**Incidence sample (%)**	**Incidence PwCF* (%)**
RPGR	rs756933299	NM_001034853:exon15:c.G2557A:p.E853K	nonsynonymous SNV	homozygosity	Uncertain_significance	2	1	6.25	50
RPGR	rs187844918	NM_001034853:exon15:c.A2808T:p.E936D	nonsynonymous SNV	heterozygosity	benign	2	1	6.25	50
RPGR	rs1399249421	NM_001034853:exon15:c.2799_2804del:p.933_935del	nonframeshift	heterozygosity	not reported	2	1	6.25	50
RPGR	rs201134185	NM_001034853:exon15:c.3074_3085del:p.1025_1029del	nonframeshift	heterozygosity	likely benign	1	0	3.13	0
RPGR	NA (ChrX:38317316)	NM_000328:exon6:c.A619G:p.T207A, NM_001034853:exon6:c.A619G:p.T207A, NM_001367245:exon6:c.A619G:p.N207D,NM_001367246:exon6:c.A619G:p.T207A, NM_001367247:exon6:c.A619G:p.T207A, NM_001367248:exon6:c.A649G:p.T217A, NM_001367249:exon6:c.A616G:p.T206A, NM_001367250:exon6:c.A619G:p.N207D, NM_001367251:exon6:c.A619G:p.T207A	nonsynonymous SNV	homozygosity	not reported	1	0	3.13	0
RPGR	**Total**	**-**	**-**	**-**	**-**	**6**	**3**	**18.75**	**50.0**
HYDIN	rs2258307	NM_001270974:exon43:c.A6725G:p.Q2242R	nonsynonymous SNV	heterozygosity	not reported	3	2	9.38	66.67

HYDIN	rs78964247	NM_001198542:exon2:c.C86A:p.T29K, NM_001198543:exon2:c.C56A:p.T19K, NM_001270974:exon2:c.C5A:p.T2K, NM_017558:exon2:c.C5A:p.T2K	nonsynonymous SNV	heterozygosity	not reported	6	4	18.75	66.67
HYDIN	rs568317137	NM_001270974:exon75:c.C12746T:p.T4249I	nonsynonymous SNV	heterozygosity	not reported	1	1	3.13	100
HYDIN	rs75357093	NM_001198542:exon2:c.A178G:p.S60G, NM_001198543:exon2:c.A148G:p.S50G, NM_001270974:exon2:c.A97G:p.S33G, NM_017558:exon2:c.A97G:p.S33G	nonsynonymous SNV	heterozygosity	not reported	5	4	15.63	80
HYDIN	rs1770434	NM_001270974:exon77:c.C13088T:p.S4363F	nonsynonymous SNV	heterozygosity	not reported	1	1	3.13	100
HYDIN	rs545493400	NM_001270974:exon46:c.G7604A:p.R2535H	nonsynonymous SNV	heterozygosity	not reported	1	1	3.13	100
HYDIN	NA (Chr16:70903761)	NM_001270974:exon52:c.G8713A:p.V2905I	nonsynonymous SNV	heterozygosity	not reported	1	0	3.13	0
HYDIN	rs566213056	NM_001270974:exon70:c.G11926A:p.A3976T	nonsynonymous SNV	heterozygosity	not reported	1	1	3.13	100
HYDIN	NA (Chr16:70860795)	NM_001270974:exon70:c.G11884A:p.G3962S	nonsynonymous SNV	heterozygosity	not reported	1	1	3.13	100
HYDIN	rs2076362302	NM_001270974:exon53:c.C8956G:p.Q2986E	nonsynonymous SNV	heterozygosity	not reported	1	0	3.13	0
HYDIN	rs783893	NM_001270974:exon81:c.C13913T:p.T4638M	nonsynonymous SNV	heterozygosity	not reported	1	0	3.13	0
HYDIN	**Total**	**-**	**-**	**-**	**-**	**12**	**7**	**37.50**	**58.3**
DNAH5	rs1258331272	NM_001369:exon63:c.A10766C:p.Q3589P	nonsynonymous SNV	heterozygosity	pathogenic	1	0	3.13	0
DNAH5	rs142036266	NM_001369:exon79:c.C13778T:p.T4593M	nonsynonymous SNV	heterozygosity	Uncertain_significance	1	1	3.13	100
DNAH5	rs754976621	NM_001369:exon38:c.G6382A:p.V2128I	nonsynonymous SNV	heterozygosity	not reported	1	0	3.13	0
DNAH5	rs115075057	NM_001369:exon75:c.A12923G:p.Y4308C	nonsynonymous SNV	heterozygosity	benign	1	1	3.13	100

DNAH5	rs767156018	NM_001369:exon15:c.A2066G:p.H689R	nonsynonymous SNV	heterozygosity	not reported	1	1	3.13	100
DNAH5	rs188638970	NM_001369:exon20:c.G3021T:p.L1007F	nonsynonymous SNV	heterozygosity	benign	1	1	3.13	100
DNAH5	chr5:13700775	NM_001369:exon78:c.A13588G:p.T4530A	nonsynonymous SNV	heterozygosity	not reported	1	1	3.13	100
DNAH5	rs115004914	NM_001369:exon15:c.C2253A:p.N751K	nonsynonymous SNV	heterozygosity	benign	1	1	3.13	100
DNAH5	rs151336435	NM_001369:exon38:c.A6416G:p.K2139R	nonsynonymous SNV	heterozygosity	Uncertain_significance	1	0	3.13	0
DNAH5	chr5:137355199	NM_001369:exon68:c.T11693G:p.L3898W	nonsynonymous SNV	heterozygosity	not reported	1	1	3.13	100
DNAH5	**Total**	**-**	**-**	**-**	**-**	**9**	**6**	**28.13**	**66.7**
GAS2L2	rs112674267	NM_139285:exon6:c.G1946A:p.G649D	nonsynonymous SNV	heterozygosity	not reported	1	1	3.13	100
GAS2L2	rs587633197	NM_139285:exon5:c.887_890del:p.V296fs	frameshift deletion	homozygosity	Pathogenic	1	0	3.13	0
GAS2L2	**Total**	**-**	**-**	**-**	**-**	**2**	**1**	**6.25**	**50.0**
RPGR HYDIN DNAH5 GAS2L2	**Total**	**-**	**-**	**-**	**-**	**22**	**13**	**68.75**	**59.1**

Note: consider that each participant can have multiple variants of different genes
***PwCF = patients with cystic fibrosis**

APPENDIX 3. Participants age and geografic distribution

Código	FQ-CFTR mut (Si/No)	Edad	Provincia
FC001	NO	13	Pichincha
FC002	SI	16	Pichincha
FC003	NO	10	Tungurahua
FC004	SI	14	Tungurahua
FC005	SI	10	Pichincha
FC006	SI	15	Orellana
FC007	SI	15	Pichincha
FC008	NO	13	Bolivar
FC009	SI	7	Chimborazo
FC010	NO	4	Pichincha
FC011	SI	19	Cotopaxi

FC012	SI	9	Pichincha
FC013	SI	7	Pichincha
FC014	SI	8	Santo Domingo
FC015	NO	12	Pichincha
FC016	SI	12	Tungurahua
FC017	NO	8	Santo Domingo
FC018	SI	9	Tungurahua
FC019	SI	7	Pichincha
FC020	SI	12	Pichincha
FC023	SI	2	Pichincha
FC024	NO	10	Pichincha
FC025	SI		Los Rios
FC026	NO	14	Pichincha
FC029	SI	8	Carchi
FC031	NO	8	Pichincha
FC032	NO		Tungurahua
FC034	SI	11	Imbabura
FC035	SI	14	Pichincha
FC036	NO	9	Cotopaxi
FC037	SI	12	Pichincha
FC038	SI	12	Tungurahua
FC039	NO	9	Pichincha
FC040	SI	2	Pichincha
FC041	SI	11	Pichincha
FC042	SI	1	Pichincha
FC043	NO	12	Pichincha
FC045	NO	12	Napo
FC046	NO	13	Pichincha
FC047	NO	9	Chimborazo
FC048	SI	7	Pichincha
FC049	SI	8	Cotopaxi
FC050	NO	10	Pichincha

APPENDIX 4. CFTR variants combinations

A.

	N417K	F508del	G85E	H609R	L15P	Q237P	L34_Q39del	Splice Accep. Chr7:117606673:G>C)	S1255L	S573C	N1303K	W1098*
N417K	3	1A|1B|1C	2|1C|1D	1A|1D|1E	1|1E					1	1B	
F508del		3		1				1				
G85E			2			1						
H609R				1			1					
L15P												1
Q237P												
L34_Q39del												
Splice Accep. Chr7:117606673:G>C)												
S1255L									1			
S573C												
N1303K												
W1098*												

Note: Variants in format: "1A", is a combination of 3 variants, among the 3 variants with the same "1A" symbol. If a number is without a letter then is just a 2 variant combination

B.

Variants	Combination	Repetition
N417K	Alone	3
N417K + G85E	2 combination	2
N417K + L15P	2 combination	1
N417K + S573C	2 combination	1
N417K + F508del + H609R	3 combination	1
N417K + F508del + N1303K	3 combination	1
N417K + F508del + G85E	3 combination	1
N417K + G85E + H609R	3 combination	1
N417K + G85E + L15P	3 combination	1

Appendix 5A, is the CFTR variants combination (2 or 3 variants) in one patient and its frequencies. Appendix 5B, is the combination of N417K variant against the rest of CFTR-variants