Distribution of the most frequent cystic fibrosis mutations in Cuba.

Distribución de las mutaciones más frecuentes de fibrosis quística en Cuba.

Ixchel López Reyes, Teresa Collazo Mesa, Elvia Nelemi Santos González, Yaixa Piñato Roque, Manuel Gómez Martínez, Lidice Reyes Navarro, Fidel Rodríguez Cala, Teresa Casals Senent

Abstract
More than 1,900 mutations have been identified in the CFTR gene, responsible of cystic fibrosis, some of them frequent in most of the populations while others are proper of a single region or family. Due to the great multinational heterogeneity and taking into account the ethnic origin of the Cuban population, the objective of this study was to carry out the distribution of the most frequent mutations in the different regions of the country. DNA extraction was carried out to samples of peripheral blood. The detection of p.F508del, p.G542X, p.R1162X, p.R334W, p.R553X and c.3120+1G>A was carried out by means of the Polymerase Chain Reaction, followed by enzymatic digestion and electrophoresis in agarose gel. Afterwards the frequency of said mutations was calculated for each one of the country regions. From the 252 fibrocystic patients studied, 106 were from the Western region, 65 from the Central region and 81 from the Eastern region of the country. The p.F508 mutation resulted to be the most frequent one in the Central Region, p.G542X and p.R334W were more represented in the Eastern region. Nevertheless the p.R1162X and c.3120+1G>A mutations were similarly represented in the three regions, while the p.R553X was not found in the Eastern region. Taking into account the ethnic origin of our population a large molecular heterogeneity in cystic fibrosis was to be expected. The most frequent mutations detected up to now are generally distributed throughout the island, related to the impact of European, African and native-American genes, as well as depending on the migration among different regions in the country.

Keywords: Cystic fibrosis, CFTR gene, Cuba, mutation.

Resumen
Se han identificado más de 1900 mutaciones en el gen CFTR, responsable de la fibrosis quística; algunas de ellas son frecuentes en la mayoría de las poblaciones, mientras que otras caracterizan a una población o familia en particular. Dada la heterogeneidad multinacional y teniendo en cuenta los orígenes étnicos de la población cubana, el objetivo de este estudio fue caracterizar la distribución de las mutaciones más frecuentes en las diferentes regiones del país. Se extrajo ADN a partir de muestras de sangre periférica. La detección de p.F508del, p.G542X, p.R1162X, p.R334W, p.R553X y c.3120+1G>A se realizó mediante la Reacción en Cadena de la Polimerasa, seguido de una digestión enzimática y electforesis en gel de agarosa. Posteriormente se calculó la frecuencia de las mutaciones relacionadas para cada una de las regiones del país. De los 252 pacientes estudiados, 106 fueron de la región Occidental, 65 de la central y 81 de la región Este del país. Teniendo en cuenta el origen étnico de nuestra población, cabe esperar una heterogeneidad elevada de la fibrosis quística. Las mutaciones más frecuentes detectadas están distribuidas de forma general por toda la isla, en relación con el impacto de los genes europeos, africanos y nativo-americanos, así como en dependencia de la migración entre las diferentes regiones del país.

Palabras clave: Fibrosis quística, gen CFTR, Cuba, mutación.

1 Licenciada en Biología. Aspirante a Investigador. Centro Nacional de Genética Médica. La Habana, Cuba. Email: ixchel@cngen.sld.cu


3 Licenciada en Tecnología de la Salud. Centro Nacional de Genética Médica. La Habana, Cuba

4 Master y licenciada en Bioquímica. Investigador agregado. Centro Nacional de Genética Médica. La Habana, Cuba
**Introduction**

Since it was discovered that the CFTR gene was responsible of cystic fibrosis (CF) to date, 1932 mutations have been described\(^1\), some of them specific of a population while others have been identified in a single patient. Due to this heterogeneity it is important to know the ethnic origin of each diagnosed individual or the region where he comes from.\(^2\)

An analysis of the geographic distribution of more than 200 CF mutations in different European populations has allowed to detect that in the Mediterranean region there is a high CF genetic level heterogeneity. The analysis of the CFTR gene in the Spanish population confirms this high heterogeneity, with 75 identified mutations that represent from 90% to 92% of the CF alleles and only 10 mutations having frequencies greater than 1%.\(^3\)

The frequency of the F508del mutation in the worldwide population is 66%,\(^4\) with an 87,2% maximum value in Denmark, a 21,3% minimum in Turkey\(^3\) and 50% in Southern Europe.\(^6\) Latin America is mainly the combination of three populations: Native American Indian, African and Caucasian. Its incidence varies in different countries: 45% in México\(^7\), 50,8% in Brazil\(^8,9\) and between 57%\(^10\) and 60,9% in Argentina,\(^11\) while in Uruguay it is 40%.\(^12\) In Cuba it has been reported to be 37%.\(^13\) This mutation is responsible of very severe forms of the disease, i.e. those accompanied by pancreatic insufficiency. In most of the patients, the first symptoms appear during nursing, when they are less than 1 year old.\(^14\)

There are other less frequent mutations restricted to specific geographical regions or ethnic groups, such as the 3120+1 G-A one, which is present in some homozygotic patients with a severe clinical pattern of pancreatic insufficiency and respiratory symptoms and leading to an early death, similar to the F508del mutation.\(^15\) This mutation is relatively common in South African populations. Besides, 17 other mutations have frequencies between 0,1% and 0,9%, while the remaining mutations are either rare or only confined to some populations.\(^16\)

An interesting case is the R334W mutation for which the analysis of a large number of patients has made it possible to clearly define its association to a late pancreatic insufficient start with interfamilial variability as well as among different families\(^17\), while the initial clinical data, including functional studies, suggested that it was a suppressor phenotype mutation in different populations\(^18\). It has been reported in different populations, with worldwide values equal to 0,1%, while in the Spanish population the frequency is equal to 1%.

Another one of the mutations detected in the worldwide population is the R553X one, with a reported frequency value equal to 1,5% in the Cystic Fibrosis Consortium, 0,34% in the Spanish population,\(^19\) while in Chile its value is equal to 4,2%, the largest value found in the whole world, attributing the differences found in the Chilean population to a possible prevalence of American-Indian genes in this country.\(^20\)

The fibrocystic population in Cuba, as in the rest of Latin America, has shown a large mutational heterogeneity, as it was to be expected due to the ethnic origin of the population, so in this paper, an analysis regarding the frequency of the mutations studied in different regions of the country has been carried out.

**Materials and Methods**

Two hundred fifty two Cuban fibrocystic patients were studied, after giving their informed consent to participate in the study. Diagnoses were carried out following their clinical characteristics and the electrolytes in the sweat test. These patients were diagnosed from 1989 to 2012 by the National Medical Genetics Network and the Cystic Fibrosis National Commission of Cuba, both pertaining to the Ministry of Public Health. In 252 individuals the presence of F508del, G542X, R1162X, R553X, 3120+1G→A and R334W mutations was evaluated.
DNA extraction
DNA was obtained from a 10 mL peripheral blood sample using Ethylene-diamine-tetra-acetic acid (EDTA 56 mg/mL) as anticoagulant. The extraction was carried out by using the saline precipitation method, as described by Miller et al. in 1988.21 Once the DNA samples were obtained they were coded and stored until they were used.

Mutations detections
Mutations F508del, G542X and R1162X were detected by using the specific mutations refractory amplification (SMRA), described by Newton.22 In each reaction the following reactants were used: 100 ng DNA, 1 U Taq DNA polymerase (Invitrogen, USA), 0,85 pmol/μL of each primer (Biosource, USA), dNTP 0,1 mM, MgCl₂, 1,0 mM in a final volume of 25 μL. In all cases internal amplification controls were carried out. The amplification reaction consisted in denaturation during 5 min at 94°C; afterwards the Taq DNA polymerase (Promega, USA), followed by 29 cycles of repetitions with the following stages each: 1 min de denaturation at 94°C; 1 min alignment at 60°C, 1 min 30 s extension at 72°C and, finally, a last extension cycle at 72°C during 5 min. The fragments thus obtained were run at 250 V during 25 min in a 2% agarose gel containing etidium bromide and were later visualized in an ultraviolet light transilluminator.

In order to identify the 3120+1G→A, R334W and R553X mutations the polymerase chain reaction (PCR) was carried out and the amplified product was digested with the restriction enzymes corresponding to each mutation.23-25 For detecting the 3120+1G→A mutation the following conditions were established: 100 ng DNA, 1 U Taq DNA polymerase (Invitrogen, USA), per reaction, each primer with 6,5 pmol/μL (Biosource, USA), dNTP 0,1 mM, MgCl₂ 2,5 mM, in a 25 μL final reaction volume. The amplification reaction consisted in the following steps: denaturation during 5 min, followed by 35 repeated cycles with the following stages each: 30 s denaturation at 94°C, 30 s hybridization at 55°C, 1 min extension at 72°C and finally a last extension cycle during 5 min at 72°C. The amplified product was digested during 3 h at 37°C, using 25 U of the restriction Hinc II (Bio Labs, USA), to obtain a final volume equal to 35 μL. The digested product was run in a 3% agarose gel and was visualized in an ultraviolet light transilluminator.

After detecting the six mutations, the results were separated by provinces and the frequency of each mutation was calculated for each province.

Results
The detection of F508del, G542X, R1162X, 3120+1G→A, R334W y R553X mutations was carried out in the 252 fibrocystic patients of the different country regions. Out of these, 106 were from the Western region, 65 from the Central region and 81 from the Eastern region, as represented in Figure 1.

Figure 1. Fibrocystic patients distribution in the Western, Central and Eastern regions of Cuba.
Distribution of the most frequent cystic fibrosis mutations in Cuba.

Table 1. Detection of mutations F508del, G542X, R1162X, 3120+1G→A, R334W and R553X in fibrocystic patients in different regions of the country.

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Western region</th>
<th>Central region</th>
<th>Eastern region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chrom no.</td>
<td>%</td>
<td>Chrom no.</td>
</tr>
<tr>
<td>F508del</td>
<td>57</td>
<td>26.8</td>
<td>44</td>
</tr>
<tr>
<td>G542X</td>
<td>7</td>
<td>3.3</td>
<td>6</td>
</tr>
<tr>
<td>R1162X</td>
<td>5</td>
<td>2.3</td>
<td>3</td>
</tr>
<tr>
<td>3120+1G→A</td>
<td>2</td>
<td>0.9</td>
<td>2</td>
</tr>
<tr>
<td>R334W</td>
<td>5</td>
<td>2.3</td>
<td>3</td>
</tr>
<tr>
<td>R553X</td>
<td>3</td>
<td>1.4</td>
<td>3</td>
</tr>
</tbody>
</table>

The most frequent mutation worldwide is the F508del one, which was represented in all provinces in the country, with a greater frequency in patients from the Central region, with a frequency value equal to 33.8% and less frequent in 19% of the cases in the Eastern region. The provinces with the greater number of F508del diagnosed cases were Sancti Spíritus, Artemisa, Cienfuegos and Villa Clara, most of them located in the central region. This mutation was less frequently found in the Santiago de Cuba, Las Tunas and Matanzas provinces.

Mutation G542X, also frequent in the population worldwide, was not detected in the provinces Granma, Cienfuegos and Island of Youth. It turned out to be more frequent in the Eastern region with a frequency equal to 5.5% and less frequent on the Western one (3.3%), resulting in a greater number of positive cases with this mutation in the Sancti Spíritus, Holguín, Guantánamo and Pinar del Rio provinces.

Regarding the R1162X mutation, it was detected in patients from Santiago de Cuba, Villa Clara, Camagüey, Pinar del Rio, Las Tunas, Ciego de Ávila and La Habana provinces; most frequently in Las Tunas and Ciego de Ávila provinces with a very similar frequency value in the three country regions.

Like mutation R1162X, the 3120+1G→A one, frequent in the African origin population, was found to have similar frequency values the three regions of the country.

The R334W mutation, which presented the major frequency in the Cuban fibrocystic population, was present in the Santiago de Cuba, Holguín, Granma, Camagüey, Matanzas, Las Tunas, Ciego de Ávila, La Habana and Mayabeque provinces, with the higher frequency value equal to 15.7% in the Holguín province and the Eastern region the one with the higher frequency value equal to 5.5%.

In the Eastern region the R553X mutation is not present, not being detected in any of the provinces of this region, unlike the rest of the mutations, and its frequency is very similar in the Western and Central regions of the country. It was only found in Villa Clara, Mayabeque, Cienfuegos and La Habana provinces with similar frequency values, while it was not detected in the remaining provinces.

Discussion

Cuba is located in the Caribbean Sea, at the entrance of the Gulf of Mexico, the closest lands being the following ones: Key West to the north, the Yucatan Peninsula in Mexico to the west, Jamaica to the south and Haiti to the east. Cuba is divided into 15 provinces and the Island of Youth special municipality. The most populated Cuban provinces are La Habana, Santiago de Cuba, Holguín, Granma, Villa Clara, and Camagüey, correspondingly being the provinces with the greater number of fibrocystic patients diagnosed in the country, as has been represented in Figure 2.
The ancestry of white Cubans (65.05%) is mainly originated by Spaniards. During the XVIIIth, XIXth and mainly the first part of the XXth century, Canarian, Galician, Asturian and Catalan emigrants arrived in large numbers from Spain to Cuba. Other European nationalities that also emigrated include British, among them the Scottish, as well as Russian, Polish, Portuguese, Rumanian, Italian, Greek, French, German and Irish ones. There is also a small remainder of a Jewish community and significant ethnical affluences of diverse populations of the Middle East, specially Lebanese, Palestinian, Turkish and Syrian populations.

Afro-Cubans constitute from 10.08% to 23.84% of the population, mainly with Congo origins, a Central African population. Asian origin Cubans represent 1% of the population, mainly from Chinese, Japanese or Korean origin. The genetic influence of the Taínos is a very little one. It is estimated that their influence is only present in 1.0% of the Cuban population. Some American Indians from the United States established in Cuba in the XIX Century, but there are not exact figures regarding their present relatives.

In a recent study by Marcheco et al., using autosomal molecular markers, they concluded that the Cuban genome is formed by 72% of European origin genes, 20% of African origin genes and 8% of native American ones. Besides, in the mitochondrial DNA information and the Y chromosome, the analysis of the mitochondrial DNA revealed that 38.8% of the haploid types are of African origin ones, 34.5% native American and 26.7% Are Eurasian origin ones. The analysis of the Y chromosome markers specified that 81.8% of the haploid types are Eurasian, 17.7% have an African origin and 0.5% are native American origin ones. That is to say that according to the uniparental markers, we come from European fathers and mostly African and native American mothers.

Considering the ethnical origin of our population, of course it was to be expected in cystic fibrosis a great molecular heterogeneity, like in other populations of Southern Europe as Spain, to which correspond the greatest percent of our genome, as it also happens in other Latin American countries. Casals et al. pointed out that in Spain 75 mutations would be responsible of 90.3% of all alterations. Out of them all, only 3 present a frequency greater than 2%: F508del (53.2%), G542X (8.4%) and N1303K (2.6%). The distribution of these mutations is quite unequal in the different areas of the peninsula. So, for instance, in the north F508del happens to be much more frequent (80% in Asturias and 73% in the Basque Country).

The F508del, G542X, R1162X, 3120+1G→A, R334W and R553X mutations, like in the Spanish population and other Latin American ones as Mexico, Chile, and Argentina are represented with a frequency greater than 1%. In Cuba, the detection of these six mutations allows diagnosing 55% of the fibrocystic chromosomes. In Latin American countries, F508del varies from 22.98% in Costa Rica to 59.5% in Argentina. The impact found in Argentina and Brazil, is not comparable to the one corresponding to the rest of Latin America, due to the different ethnical origin of its populations; these results agree with the fact that Argentina is one of the Latin American populations which had the greater Caucasian contribution. In Cuba it is represented in...
Distribution of the most frequent cystic fibrosis mutations in Cuba.

all provinces, but it must be pointed out that it is less frequent in the Eastern region, where the black race impact is greater, likewise it is the region where the frequency of sickle-cell disease is greater.

Regarding the G542X mutation, is not represented in all provinces as does F508del, although it is present in all regions. It is more represented in some of the Eastern provinces, as Guantánamo (13,6%) and Holguín (10,5%), while in the Central region it is found with a value equal to 14,2% and up to this moment it is the second most frequent mutation in the affected population in Cuba. In Spain, for instance, it is most frequent in the Mediterranean basin, having a value equal to 11,8%.

In the Cuban fibrocystic population we found the R334W mutation in a proportion greater than in other populations (4,9%), in the Spanish one the reported frequency is 1,78%, while the value reported in Argentina, Brazil, Chile, Colombia and Uruguay equals 0,9%. It is interesting to stand out that in Holguín the frequency value found was greater than in any other region (15,7%), which might be explained as due to the effect of a founder effect.

Mutations R1162X, R553X andy 3120+1G→A are represented in few provinces of the country.

We determined that R553X mutation was present in 2,6% of the studied cases. This value surpasses the frequency described in most of the studies; for instance, in the Spanish population it is 0.4%. Ríos et al. detected in the Santiago region in Chile, a frequency of 4,2%, which has been explained by the contribution of Indian-American genes to this population. In 2001 Reppetto et al. reported a frequency for R553X in Chile equal to 1% and in Mexico, Brazil and Argentina values of 0,5%, 0,6 % and 0,2% respectively. In Cuba, this mutation is represented in only 4 provinces, all of them in the Western and Central regions of the country, although the provinces with a higher reported percent of native-American origin genes are Granma (15%), Holguín (12%) and Las Tunas (12%), corresponding to the Eastern region, but there exists a large migration from this region to the remaining two regions of the country, which might explain the frequency of this mutation in regions where the native-American genes percent has been less.

In the case of the 3120+1G→A mutation, frequent in African population, the greater percent found in Cuba corresponded to Santiago de Cuba (3,2%), a province where the African contribution to our genome has been greater, Santiago de Cuba is reported to have 40 % of African genes. In the La Habana province all studied mutations are represented, which can be accounted for by the migration from the whole island to the capital.

Conclusions

Cystic fibrosis in Cuba is molecularly very heterogeneous and the most frequent mutations detected to date are distributed in the whole island, related to the impact of European, African and native American genes, as well as depending of the population migration among regions.

References


27. Key History. «SEMINOLO INDIANS» (en inglés). Consultado el 7-5-2010.

28. Personnal communication, 2014. Beatriz Marcheco Teruel DSc, National Center of Medical Genetics, Cuba


Distribution of the most frequent cystic fibrosis mutations in Cuba.